Damien Desfontaines and Balázs Pejó

SoK: Differential privacies

Abstract: Shortly after it was first introduced in 2006, *differential privacy* became the flagship data privacy definition. Since then, numerous variants and extensions were proposed to adapt it to different scenarios and attacker models. In this work, we propose a systematic taxonomy of these variants and extensions. We list all data privacy definitions based on differential privacy, and partition them into seven categories, depending on which aspect of the original definition is modified.

These categories act like dimensions: variants from the same category cannot be combined, but variants from different categories can be combined to form new definitions. We also establish a partial ordering of relative strength between these notions by summarizing existing results. Furthermore, we list which of these definitions satisfy some desirable properties, like composition, post-processing, and convexity by either providing a novel proof or collecting existing ones.

Keywords: Differential privacy, Data privacy, Survey, Systematization of knowledge

DOI Editor to enter DOI

Received ..; revised ..; accepted ...

1 Introduction

What does it mean for data to be anonymized? Samarati and Sweeney discovered that removing explicit identifiers from dataset records was not enough to prevent information from being re-identified [143, 152], and they proposed the first definition of anonymization. This notion, called *k-anonymity*, is a property of a dataset: each combination of re-identifying fields must be present at least k times. In the following decade, further research showed that sensitive information about individuals could still be leaked when releasing k-anonymous datasets, and many variants and definitions were proposed such as *l-diversity* [122], *t-closeness* [113], and *n-confusion* [150].

A shortcoming of these approaches is that they define anonymity as a property of the *dataset*: without knowing how the dataset is generated, arbitrary information can be leaked. This approach was changed when Dwork and McSherry introduced *differential privacy* (DP) [48, 53]: rather than being a property of the sanitized dataset, anonymity was defined as a property of the *process*. It was inspired by Dalenius' privacy goal that "Anything about an individual that can be learned from the dataset can also be learned without access to the dataset" [33], a goal similar to one already used in probabilistic encryption [145].

Thanks to its useful properties, DP quickly became the flagship of data privacy definitions. Many algorithms and statistical processes were adapted to satisfy DP and were adopted by organizations like the US Census Bureau [5, 70], Google [61], Apple [154], and Microsoft [37].

Since the original introduction of DP, many relaxations have been proposed to adapt it to different contexts or assumptions. These new definitions enable practitioners to get privacy guarantees, even in cases that the original DP definition does not cover well. This happens in a variety of scenarios, e.g., the noise mandated by DP can be too large and force the data custodian to consider a weaker alternative or the attacker model might require the data owner to consider correlations in the data explicitly to make stronger statements on what information the privacy mechanism reveals.

Figure 1 shows the prevalence of this phenomenon: approximately *200* different notions, inspired by DP, were defined in the last 15 years.[1] These DP definitions can be *extensions* or *variants* of DP. An extension encompasses the original DP notion as a special case, while a variant changes some aspect, typically to weaken or strengthen the original definition.

With so many definitions, it is difficult for new practitioners to get an overview of this research area. Many definitions have similar goals, so it is also challenging to understand which it is appropriate to use in which context. These difficulties also affect experts: several definitions listed in this work have been defined independently

Damien Desfontaines: ETH Zürich / Google, E-mail: damien@desfontain.es

Balázs Pejó: CrySyS Lab, Dept. of Networked Systems and Services, Budapest University of Technology and Economics, E-mail: pejo@crysys.hu

1 We count all the definitions which are presented as "new", e.g., definitions which may appeared earlier, but are not cited.

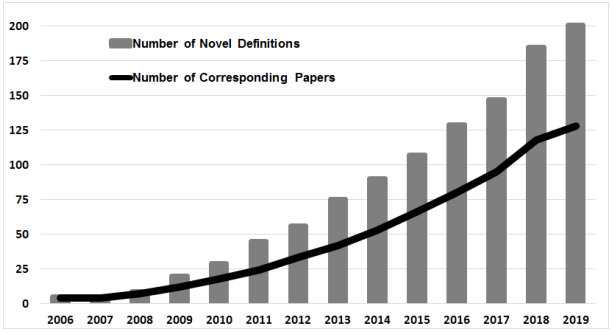

Fig. 1. Accumulated number of papers which are introducing new DP notions (line) and the exact number of these definitions (bar).

multiple times (occasionally with identical meaning but different names or identical names but different meanings). Finally, variants are often introduced without a comparison to related notions.

This work attempts to solve these problems. It is a taxonomy of various and extensions of DP, providing short explanations of the intuition, use cases and basic properties of each. By categorizing these definitions, we also hope to simplify the understanding of existing variants and extensions, and of the relations between them.

Contributions and organization

We systematize the scientific literature on variants and extensions of DP, and propose a unified and comprehensive taxonomy of these definitions. We define seven *dimensions*: these are ways in which the original definition of DP can be modified. Moreover, we highlight representative definitions for each dimension, and we enlist whether they satisfy Kifer et al.'s *privacy axioms* [103, 104], (post-processing and convexity), and whether they are composable. Our survey is organized as follows.

In Section 2, we recall the original definition of DP and introduce our dimensions along which DP can be modified. Moreover, we present the basic properties of DP, and define how definitions can relate to each other. In the following 7 sections (Sections 3 to 9), we introduce our dimensions and list the corresponding definitions. In Section 10, we review the methodology and scope of this work and mention the relevant works from the literature. In Section 11, we conclude the paper.

2 Differential Privacy

In this section, we recall the original definition of DP and we introduce our seven dimensions. We also enlist desirable properties of data privacy definitions, and define how two definitions can relate to each other. Table 1 summarizes the notations used throughout the paper.

Notation	Description
$\mathcal{T}$	Set of possible records
$t \in \mathcal{T}$	A possible record
$\mathcal{D} = \mathcal{T}^*$	Set of possible datasets (sequences of records)
$D \in \mathcal{D}$	Dataset (we also use $D', D_1, D_2, \dots$)
$D(i)$	i-th record of the dataset ($i \leq \lvert D \rvert$)
D_{-i}	Dataset D, with its i-th record removed
$\mathcal{M}$	Privacy mechanism (probabilistic)
$\mathcal{M}(D)$	The distribution (or an instance of this distribution) of the outputs of $\mathcal{M}$ given input D
$\mathcal{O}$	Set of possible outputs of the mechanism
$S \subseteq \mathcal{O}$	Subset of possible outputs
$O \in \mathcal{O}$	Output of the privacy mechanism
π	Probability distribution on $\mathcal{T}$
Θ	Family of probability distributions on $\mathcal{D}$
$\theta \in \Theta$	Probability distribution on $\mathcal{D}$

Table 1. Notations used through the paper.

The first DP mechanism was proposed in 1965 [165], and data privacy definitions that are a property of a mechanism and not of the output dataset were proposed as early as 2003 [62]. However, DP and the related notion of *ε-indistinguishability*[2] were first formally defined in 2006 [48, 53, 54].

Definition 1 (ε-indistinguishability [54])**.** *Two random variables A and B are ε-indistinguishable, denoted $A \approx_\varepsilon B$, if for all measurable sets X of possible events, $\mathbb{P}[A \in X] \leq e^\varepsilon \cdot \mathbb{P}[B \in X]$ and $\mathbb{P}[B \in X] \leq e^\varepsilon \cdot \mathbb{P}[A \in X]$.*

Informally, A and B are ε-indistinguishable if their distributions are "close". This notion is used to define DP.

Definition 2 (ε-differential privacy [48])**.** *A privacy mechanism $\mathcal{M}$ is ε-differential private (or ε-DP) if for all datasets D_1 and D_2 that differ only in one record, $\mathcal{M}(D_1) \approx_\varepsilon \mathcal{M}(D_2)$.*

2 This notion originates from the cryptographic notion of indistinguishability [77]. A similar notion, $(1, \varepsilon)$-*privacy*, is defined in [25], where $(1 + \varepsilon)$ used in place of e^ε.

2.1 Dimensions

Variants and extensions of DP modify the original definition in various ways. To establish a comprehensive taxonomy, a natural approach is to partition them into *categories*, depending on which aspect of the definition they change. Unfortunately, this approach fails for privacy definitions: many of them modify several aspects at once, so it is impossible to have a categorization such that every definition falls neatly into only one category.

The approach we take is to define *dimensions* along which the original definition can be modified. Each variant or extension of DP can be seen as a point in a multidimensional space, where each coordinate corresponds to one possible way of changing the definition along a particular dimension. To make this representation possible, our dimensions need to satisfy two properties:

- **Mutual compatibility**: definitions that vary along different dimensions can be combined to form a new, meaningful definition.
- **Inner exclusivity**: definitions in the same dimension cannot be combined to form a new, meaningful definition (but they can be pairwise comparable).

In addition, each dimension should be *motivatable*: there should be an intuitive explanation of what it means to modify DP along each dimension. Moreover, each possible choice within a dimension should be similarly understandable, to allow new practitioners to determine quickly which kind of definition they should use or study, depending on their use case.

We introduce our dimensions by reformulating the guarantee offered by DP, highlighting aspects that have been modified by its variants or extensions. Each dimension is attributed a letter, and we note the dimension letter corresponding to each highlight. This formulation considers the point of view of an attacker, trying to find out some sensitive information about some input data using the output of a mechanism.

An attacker with **perfect background knowledge** (**B**) and **unbounded computation power** (**C**) is **unable** (**R**) to **distinguish** (**F**) **anything about an individual** (**N**), **uniformly across users** (**V**) even in the **worst-case scenario** (**Q**).

This informal definition of DP with the seven highlighted aspects give us seven distinct dimensions. We denote each one by a letter and summarize them in Table 2; each is introduced in its corresponding section.

Note that the interpretation of DP is subject to some debate. In [157], the authors summarize this debate, and show that DP can be interpreted under two possible lenses: it can be seen as an *associative* property, or as a *causal* property. The difference between the two interpretations is particularly clear when one supposes that the input dataset is modeled as being generated by a probability distribution.

- In the associative view, this probability distribution is *conditioned* upon the value of one record. If the distribution has correlations, this change can affect other records as well.
- In the causal view, the dataset is first generated, and the value of one record is then *changed* before computing the result of the mechanism.

While the causal view does not require any additional assumption to capture the intuition behind DP, the associative view requires that either all records are independent in the original probability distribution (the *independence assumption*), or the adversary must know all data points except one (the *strong adversary assumption*, which we picked in the reformulation above).

Dimension	Description	Usual Motivations
Quantification of Privacy Loss	How is the privacy loss quantified across outputs?	Averaging risk, having better composition properties
Neighborhood Definition	Which properties are protected from the attacker?	Protecting specific values or multiple individuals
Variation of Privacy Loss	Can the privacy loss vary across inputs?	Modeling users with different privacy requirements
Background Knowledge	How much prior knowledge does the attacker have?	Using less noise in the mechanism
Formalism change	How to formalize the attacker's knowledge gain?	Exploring other intuitive notions of privacy
Relativization of Knowledge Gain	What is the knowledge gain relative to?	Guaranteeing privacy for correlated data
Computational Power	How much computational power can the attacker use?	Combining cryptography techniques with DP

Table 2. The seven dimensions and their usual motivation.

These considerations can have a significant impact on DP variants and extensions, either leading to distinct variants that attempt to capture the same intuition, or to the same variant being interpreted in different ways.

2.2 Properties

Two important properties of data privacy notions are called *privacy axioms*. They were proposed in [103, 104]. These are not "axioms" in a sense that they are assumed to be true, but rather, they are consistency checks: properties that, if not satisfied by a data privacy definition, indicate a flaw in the definition.

Definition 3 (Privacy axioms [103, 104])**.**

1. ***Post-processing***[3] *(or* transformation invariance*): A privacy definition* Def *satisfies the post-processing axiom if, for any mechanism* $\mathcal{M}$ *satisfying* Def *and any probabilistic function* f*, the mechanism* $D \to f(\mathcal{M}(D))$ *also satisfies* Def*.*
2. ***Convexity*** *(or* privacy axiom of choice*): A privacy definition* Def *satisfies the convexity axiom if, for any two mechanisms* $\mathcal{M}_1$ *and* $\mathcal{M}_2$ *satisfying* Def*, the mechanism* $\mathcal{M}$ *defined by* $\mathcal{M}(D) = \mathcal{M}_1(D)$ *with fixed probability* p *and* $\mathcal{M}(D) = \mathcal{M}_2(D)$ *with probability* $1 - p$ *also satisfies* Def*.*

A third important property is *composability*. It guarantees that the output of two mechanisms satisfying a privacy definition still satisfies the definition, typically with a change in parameters. There are several types of composition: *parallel composition* (the mechanisms are applied to disjoint subsets of a larger dataset), *sequential composition* (the mechanisms are applied on the entire dataset), and *adaptive composition* (each mechanism can access dataset and the output of the previous mechanisms). In this work, we only consider sequential composition.

Definition 4 (Composability)**.** *A privacy definition* Def *with parameter* α *is* composable *if for any two mechanisms* $\mathcal{M}_1$ *and* $\mathcal{M}_2$ *satisfying respectively* α_1-Def *and* α_2-Def*, the mechanism* $\mathcal{M}(D) = (\mathcal{M}_1(D), \mathcal{M}_2(D))$ *satisfies* α-Def *for some (non-trivial)* α*.*

3 This definition must be slightly adapted for some variants, see for example Proposition 12 in Appendix A.

2.3 Relations

When learning about a new data privacy notion, it is often useful to know what are the known relations between this notion and other definitions. However, definitions have parameters that often have different meanings, and whose value is not directly comparable. To claim that a definition is stronger than another, we utilize the concept of ordering established in [31] using α and β as tuples, encoding multiple parameters.

Definition 5 (Relative strength)**.** *Let* α-Def_1 *and* β-Def_2 *be privacy definitions. We say that* Def_1 *is* stronger than Def_2*, and denote it* $\mathrm{Def}_1 \succ \mathrm{Def}_2$*, if:*

1. *for all* α*, there is a* β *such that* α-$\mathrm{Def}_1 \implies \beta$-$\mathrm{Def}_2$*;*
2. *for all* β*, there is an* α *such that* α-$\mathrm{Def}_1 \implies \beta$-$\mathrm{Def}_2$*.*

If $\mathrm{Def}_1 \succ \mathrm{Def}_2$ *and* $\mathrm{Def}_2 \succ \mathrm{Def}_1$*, we say that the two definitions are* equivalent*, and denote it* $\mathrm{Def}_1 \sim \mathrm{Def}_2$*.*

Relative strength implies a partial ordering on the space of possible definitions. It is useful to classify variants but does not capture extensions well. Thus, we introduce a second notion to represent when a definition can be seen as a special case of another.

Definition 6 (Extensions)**.** *Let* Def_1 *and* Def_2 *be privacy definitions with respective parameters* α *and* β*. We say that* Def_1 *is* extended by Def_2*, and denote is as* $\mathrm{Def}_1 \subset \mathrm{Def}_2$*, if for all* α*, there is a* β *such that* α-Def_1 *is identical (i.e., provides the same privacy guarantee) to* β-Def_2*.*

Note that the original definition of relative strength only required the second condition to hold; which would classify any extension as a stronger variant.

Both relations are reflexive and transitive; and we define the symmetric counterpart of these relations as well (i.e., $\prec$ and $\supset$). Moreover, for brevity, we combine these two concepts in a single notation: if $\mathrm{Def}_1 \subset \mathrm{Def}_2$ and $\mathrm{Def}_1 \succ \mathrm{Def}_2$, we say that Def_2 is a weaker extension of Def_1, and denote it $\mathrm{Def}_1 \subset^{\succ} \mathrm{Def}_2$.

In the following sections, we detail the different ways in which researchers changed the original DP definition. Due to space constrains, only a handful of definitions are formally defined while the majority of the definitions are presented with greater detail in the full version of this work. A summarizing table is presented at the end of this work, where for each definition, we also highlight its dimensions and its relation to other notions. In Table 4 we also specify whether these notions satisfy the privacy

axioms and the composability property (✓: yes, ✗: no, ?: currently unknown); in Appendix A and B we either provide a reference or a novel proof for each of these claims.

3 Quantification of privacy loss (Q)

DP and its associated risk model is a *worst-case* property: it quantifies not only over all possible neighboring datasets but also over all possible outputs. However, in a typical real-life risk assessment, events with vanishingly small probability are ignored, or their risk is weighted according to their probability. It is natural to consider analogous relaxations, especially since these relaxations often have better composition properties, and enable natural mechanisms like the Gaussian mechanism to be considered private [57].

Most of the definitions within this section can be expressed using the *privacy loss random variable*[4], so we first introduce this concept. Roughly speaking, it measures how much information is revealed by the output of a mechanism.

Definition 7 (Privacy loss random variable [39])**.** *Let $\mathcal{M}$ be a mechanism, and D_1 and D_2 two datasets. The* privacy loss random variable between $\mathcal{M}(D_1)$ and $\mathcal{M}(D_2)$ *is defined as:*

$$\mathcal{L}_{D_1/D_2}(O) = \ln\left(\frac{\mathbb{P}[\mathcal{M}(D_1) = O]}{\mathbb{P}[\mathcal{M}(D_2) = O]}\right)$$

if neither $\mathbb{P}[\mathcal{M}(D_1) = O]$ nor $\mathbb{P}[\mathcal{M}(D_2) = O]$ is 0; in case only $\mathbb{P}[\mathcal{M}(D_2) = O]$ is zero then $\mathcal{L}_{D_1/D_2}(O) = \infty$, otherwise $\mathcal{L}_{D_1/D_2}(O) = -\infty$.

3.1 Allowing a small probability of error

The first option, whose introduction is commonly attributed to [52], relaxes the definition of ε-indistinguishability by allowing an additional small density of probability on which the upper ε bound does not hold. This small density δ can be used to compensate for outputs for which the privacy loss is larger than e^{ε}. This led to the definition of *approximate DP* [52], also called (ε, δ)-DP. As of today, this is the most commonly used relaxation in the literature.

The δ in (ε, δ)-DP is sometimes explained as the probability that the privacy loss of the output is larger than e^{ε}. However, this intuition corresponds to a different definition, called *probabilistic DP* [20, 124, 127]. These two definitions can be combined to form *relaxed DP* [176], requiring approximate DP with probability < 1.

3.2 Averaging the privacy loss

As ε-DP corresponds to a *worst-case* risk model, it is natural to consider relaxations to allow for larger privacy loss for some outputs. It is also natural to consider *average-case* risk models: allowing larger privacy loss values only if lower values compensate it in other cases. One such relaxation is called *Kullback-Leibler privacy* [9, 31]: it considers the *arithmetic* mean of the privacy loss random variable, which measures how much information is revealed when the output of a private algorithm is observed.

Rényi DP [128] extends this idea by adding a parameter $\alpha \geq 1$ which allows controlling the choice of averaging function.

Definition 8 ((α, ε)-Rényi DP [128])**.** *Given $\alpha > 1$, a privacy mechanism $\mathcal{M}$ is (α, ε)-Rényi DP if for all pairs of neighboring datasets D_1 and D_2:*

$$\mathbb{E}_{O\sim\mathcal{M}(D_1)}\left[e^{(\alpha-1)\mathcal{L}_{D_1/D_2}(O)}\right] \leq e^{(\alpha-1)\varepsilon}$$

The property required by Rényi DP can be reformulated as $D_\alpha(\mathcal{M}(D_1)\|\mathcal{M}(D_2)) \leq \varepsilon$, where D_α is the Rényi divergence. It is possible to use other divergence functions to obtain other relaxations, such as *binary-$|\chi|^\alpha$-* and *tenary-$|\chi|^\alpha$ DP* [163], *total variation privacy* [9] or *quantum DP* [30].

Another possibility to average the privacy loss is to use *mutual information* to formalize the intuition that any individual record should not "give out too much information" on the output of the mechanism (or vice-versa). This is captured by *mutual-information DP* [31], which guarantees that the mutual information between $\mathcal{M}(D)$ and $D(i)$ conditioned on D_{-i} is under a certain threshold. The bound is taken over all possible priors on D, which avoids having to reason about the attacker's background knowledge.

4 First defined in [39] as the *adversary's confidence gain*.

3.3 Controlling the tail distribution of the privacy loss

Some definitions go further than simply considering a worst-case bound on the privacy loss, or averaging it across the distribution. They try to obtain the benefits of (ε, δ)-DP with a smaller ε which holds in most cases, but control the behavior of the bad cases better than (ε, δ)-DP, which allows for catastrophic privacy loss in rare cases.

The first attempt to formalize this idea was proposed in [58], where authors introduce *mean concentrated DP*. In this definition, a parameter controls the privacy loss variable globally, and another parameter allows for some outputs to have a greater privacy loss; while still requiring that the difference is smaller than a Gaussian distribution. In [19] the authors show that it does not verify the post-processing axiom, and proposed another formalization of this idea called *zero-concentrated DP*, which requires that the privacy loss random variable is concentrated around zero.

Definition 9 ((ξ, ρ)-zero-concentrated DP [19])**.** *A mechanism $\mathcal{M}$ is (ξ, ρ)-zero-concentrated DP if for all pairs of neighboring datasets D_1 and D_2 and all $\alpha > 1$:*

$$\mathbb{E}_{O\sim\mathcal{M}(D_1)}\left[e^{(\alpha-1)\mathcal{L}_{D_1/D_2}(O)}\right] \leq e^{(\alpha-1)(\xi+\rho\alpha)}$$

Four more variants of concentrated DP exist: *approximate zero concentrated DP* [19], *Collinson-concentrated DP*[5] [30], *bounded zero concentrated DP* [19] and *truncated concentrated DP* [18]. The first takes the Rényi divergence on events with high enough probability instead of on the full distributions, the second requires all the Rényi divergences to be smaller than a threshold, while the last two requires this only for some Rényi divergences. While we present the connections between the most widely used relaxations in Table 4, we list the exact connection between all these notions in the full version of this work.

3.4 Extensions

Most definitions of this section can be seen as bounding the divergence between $\mathcal{M}(D_1)$ and $\mathcal{M}(D_2)$, for different possible divergence functions. In [9], the authors use this fact to generalize them and define *divergence DP*, which takes an f-divergence as a parameter. Note that [46] also defined divergence DP, using a special set of f-divergences. A similar idea was explored in [24] as *capacity bounded DP* which uses "restricted divergences".

Finally, approximate DP, probabilistic DP and Rényi DP can be extended to use a *family* of parameters rather than a single pair. As shown in [148] (Theorem 2), finding the tightest possible family of parameters (for either definition) for a given mechanism is equivalent to specifying the behavior of its privacy loss random variable entirely.

3.5 Multidimensional definitions

Allowing a small probability of error δ by using the same concept as in (ε, δ)-DP is very common; a lot of new DP definitions were proposed in the literature with such a parameter. Unless it creates a particularly notable effect, we do not mention it explicitly and present the definitions without this parameter.

Definitions in this section can be used as standalone concepts: (ε, δ)-DP is omnipresent in the literature, and the principle of averaging risk is natural enough for Rényi privacy to be used in practical settings, like posterior sampling [74] or resistance to adversarial inputs in machine learning [137]. Most variants in this section, however, are only used as technical tools to get better results on composition or privacy amplification [57, 66, 111, 163].

4 Neighborhood definition (N)

The original DP definition considers datasets differing in one record. Thus, the datasets can differ in two possible ways: either they have the same size and differ only on one record, or one is a copy of the other with one extra record. These two options do not protect the same thing: the former protects the *value* of the records while the latter also protects their *presence* in the data: together, they protect any property about a single individual.

In many scenarios, it makes sense to protect a different property about their dataset, e.g., the value of a specific sensitive field, or entire groups of individuals. It is straightforward to adapt DP to protect different sensitive properties: all one has to do is change the definition of neighborhood in the original definition.

5 Originally called truncated concentrated DP, we rename it here to avoid a name collision.

4.1 Changing the sensitive property

The original definition states that DP should hold for "any datasets D_1 and D_2 that differ only in one record". Modifying the set of pairs (D_1, D_2) such that $\mathcal{M}(D_1) \approx_\varepsilon \mathcal{M}(D_2)$ is equivalent to changing the protected sensitive property.

In DP, the difference between D_1 and D_2 is sometimes interpreted as "one record value is different", or "one record has been added or removed". In [105], the authors formalize these two options as *bounded DP*[6] and *unbounded DP*. They also introduced *attribute DP* and *bit DP*, for smaller changes within the differing record.

More restrictive definitions are also possible: in [126] the authors defined *client/participant DP* which covers the case when the same person can contribute multiple times to a dataset. (c, ε)-*group privacy* [49] considers datasets that do not differ in one record, but possibly several, to protect multiple individuals. This can also be interpreted as taking correlations into account when using DP: *DP under correlation* [26] uses an extra parameter to describe the maximum number of records that the change of one individual can influence.

These two definitions are formally equivalent; but the implicit interpretation of DP behind them is different. (c, ε)-group privacy is compatible with the associative view under the strong adversary assumption (the adversary knows all records except c) or the causal view (c records are changed after the data is generated). Meanwhile, DP under correlation implicitly considers the associative view with the independence assumption; and tries to relax that assumption. This last approach was further developed via *dependent DP* [116], which uses "dependence relationships" to describe how much the variation in one record can influence the other records. Equivalents to this definition also appear in [169, 170] as *correlated DP*, and in [173] as *bayesian DP*. [7]

The strongest possible variant is considered in [105] where the authors define *free lunch privacy* in which the attacker must be unable to distinguish between any two datasets, even if they are completely different. This guarantee is a reformulation of Dalenius' privacy goal [33]; as such, all mechanisms that satisfy free lunch privacy have a near-total lack of utility.

6 Called *Per-Person DP* in [66].

7 Note that another notion with the same name was defined in [112], which we introduce in Section 6.

Another way to modify the neighborhood definition in DP is to consider that only certain types of information are sensitive. For example, if the attacker learns that their target has cancer, this is more problematic than if they learn that their target does *not* have cancer. This idea is captured in *one-sided DP* [42]: the neighbors of a dataset D are obtained by replacing a single sensitive record with any other record (sensitive or not). The idea of sensitivity is formalized by a "policy", which specifies which records are sensitive.

Definition 10 ((P, ε)-one-sided differential privacy [42])**.** *Given a policy $P \subseteq \mathcal{T}$, a privacy mechanism $\mathcal{M}$ is (P, ε)-one-sided DP iff for all datasets D_1 and D_2, where D_2 has been obtained by replacing a record $t \in D_1 \cap P$ by any other record and for all $S \subseteq \mathcal{O}$:*

$$\mathbb{P}[\mathcal{M}(D_1) \in S] \leq e^\varepsilon \cdot \mathbb{P}[\mathcal{M}(D_2) \in S]$$

A similar ideas were proposed in [100] as *protected DP*, which adopts DP for graphs and guarantees that no observer can learn much about the set of edges corresponding to any protected node while offering no guarantees for the rest. Another similar definitions is *sensitive privacy* [7], which determines sensitive records after quantifying the database (instead of assuming that being an outlier/inlier is independent of the database). Another close definition was introduced in [16] as *anomaly-restricted DP*, which assumes that there is only one outlier (which is not protected). Finally, in [130] a more sophisticated formalization of these ideas was presented as *utility-optimized DP*.

4.2 Limiting the scope of the definition

Redefining the neighborhood property can also be used to reduce the scope of the definitions. In [149], the authors note that DP requires ε-indistinguishability of results between any pair of neighboring data sets, but in practice, the data custodian has only *one* data set D they want to protect. Thus, they only require ε-indistinguishability between this data set D and all its neighbors, calling the resulting definition *individual DP* [149].

Definition 11 ((D, ε)-individual differential privacy [149])**.** *Given a dataset $D \in \mathcal{D}$, a privacy mechanism $\mathcal{M}$ satisfies (D, ε)-individual DP if for any data set D' that differs in at most one record from D, $\mathcal{M}(D) \approx_\varepsilon \mathcal{M}(D')$.*

This was further restricted in *per-instance DP* [162], where besides fixing a dataset D, a record t was also fixed.

4.3 Applying the definition to other types of input

Many adaptations of DP simply change the neighborhood definition to protect different types of input data than datasets.

DP was adopted to graph-structured data in [38, 85, 136, 140, 144, 151, 153], to streaming data in [51, 55, 56, 64, 102], to symbolic control systems in [93], to text vectors in [175], to set operations in [172], to images in [174], to genomic data in [147], to recommendation systems in [80], to machine learning in [119], to location data in [28], to outsourced database systems in [101], to bandit algorithms in [12, 156], to RAMs in [3, 21, 158] and to Private Information Retrieval in [135, 155]. We detail the corresponding definitions in the full version of this work.

4.4 Extensions

It is natural to generalize the variants of this section to arbitrary neighboring relationships. One example is mentioned in [105], under the name *generic DP*, where the neighboring relation is entirely captured by a *relation* $\mathcal{R}$ between datasets.

Other definitions use different formalizations to also generalize the concept of changing the neighborhood relationship. For example pufferfish privacy[8] uses "pairs of predicate" (ϕ_1, ϕ_2) that D_1 and D_2 must respectively satisfy to be neighbors [106] while coupled-worlds privacy[8] use "private functions", denoted priv, and define neighbors to be datasets D_1 and D_2 such as $\mathsf{priv}(D_1) \neq \mathsf{priv}(D_2)$ [11].

Finally, *blowfish privacy* [84, 86], use a "policy graph" specifying which pairs of tuple values must be protected. Others use a "distance function" between datasets, and neighbors are defined as datasets a distance lower than a given threshold; this is the case for *DP under a Δ-neighborhood*, introduced in [63] and *adjacent DP*, introduced in [108]. This distance can also be defined as the sensitivity of the mechanism, like in *sensitivity induced DP* [142], or implicitly defined by a set of constraints, like in *induced DP* [105].

8 Introduced later.

4.5 Multidimensional definitions

Modifying the protected property is orthogonal to modifying the risk model implied by the quantification of privacy loss: it is straightforward to combine these two dimensions. Indeed, many definitions mentioned in this section were actually introduced with a δ parameter allowing for a small probability of error. One particularly general example is *adjacency relation divergence DP* [97], which combines an arbitrary neighborhood definition (like in generic DP) with an arbitrary divergence function (like in divergence DP).

As the examples in Section 4.3 show, it is very common to change the definition of neighborhood in practical contexts to adapt what aspect of the data is protected. Further, local DP mechanisms like RAPPOR [61] implicitly use bounded DP: the participation of one individual is not secret, only the value of their record is protected. Variants that limit the scope of the definition to one particular database or user, however, provide few formal guarantees and do not seem to be used in practice.

5 Variation of privacy loss (V)

In DP, the privacy parameter ε is *uniform*: the level of protection is the same for all protected users or attributes, or equivalently, only the level of risk for the most at-risk user is considered. In practice, some users might require a higher level of protection than others or a data custodian might want to consider the level of risk across all users, rather than only considering the worst case. Some definitions take this into account by allowing the privacy loss to vary across inputs, either explicitly (by associating each user to an acceptable level of risk), or implicitly (by allowing some users to be at risk, or averaging the risk across users).

5.1 Varying the privacy level across inputs

In Section 4, we show how changing the definition of the neighborhood allows us to adapt the definition of DP to protect different aspects of the input data. However, the privacy protection in those variants is binary: either a given property is protected, or it is not. A possible option to generalize this idea further is to allow the privacy level to vary across inputs.

One natural example is to consider that some users might have higher privacy requirements than others, and make the ε vary according to which user differs

between the two datasets. This is done in [59, 76, 94, 118, 133] via *personalized DP*, also defined in [2] as *heterogeneous DP* and in [34] as *personalized location DP*.

Definition 12 (Ψ-personalized DP [94])**.** *A privacy mechanism $\mathcal{M}$ provides Ψ-personalized DP if for every pair of neighboring datasets (D, D_{-i}) and for all sets of outputs $S \subseteq \mathcal{O}$:*

$$\mathbb{P}\left[\mathcal{M}\left(D_{-i}\right) \in S\right] \leq e^{\Psi(D(i))}\mathbb{P}\left[\mathcal{M}\left(D\right) \in S\right]$$

where Ψ is a privacy specification: $\Psi : \mathcal{T} \to \mathbb{R}^+$ maps the records to personal privacy preferences and $\Psi\left(D(i)\right)$ denotes the privacy preference of the i-th record.

This definition can be seen as a refinement of the intuition behind one-sided DP, which separated records into sensitive and non-sensitive ones. In [120], the authors define *tailored DP*, which generalizes this further: the privacy level depends on the entire database, not only in the differing record.

This concept can be applied to strengthen or weaken the privacy requirement for a record depending on whether they are an outlier in the database. In [120], the authors formalize this idea and introduce *outlier privacy*, which tailors an individual's protection level to their "outlierness". Other refinements (instances of tailored DP) include *simple outlier privacy*, *simple outlier DP*, and *staircase outlier privacy*.

Finally, varying the privacy level across inputs also makes sense in *continuous* scenarios, where the neighborhood relationship between two datasets is not binary, but quantified, like *ε-geo-indistinguishability* [6].

5.2 Randomizing the variation of privacy levels

Varying the privacy level across inputs can also be done in a randomized way, by guaranteeing that some random fraction of users have a certain privacy level. One example is proposed in [83] as *random DP*: the authors note that rather than requiring DP to hold for any possible datasets, it is natural to only consider *realistic datasets*, and allow "edge-case" datasets to not be protected. This is captured by generating the data randomly, and allowing a small proportion γ of cases not to satisfy the ε-indistinguishability property.

Definition 13 ($(\pi, \gamma, \varepsilon)$-random DP [83])**.** *Let π be a probability distribution on $\mathcal{T}$, D_1 a dataset generated by drawing n i.i.d. elements in π, and D_2 the same dataset as D_1, except one element was changed to a new element drawn from π. A mechanism $\mathcal{M}$ is $(\pi, \gamma, \varepsilon)$-random DP if $\mathcal{M}\left(D_1\right) \approx_\varepsilon \mathcal{M}\left(D_2\right)$, with probability at least $1 - \gamma$* on the choice of D_1 and D_2.

This definition was also introduced as *predictive DP* in [82] and as *model-specific DP* in [125], with slightly different ways of defining the randomness which models this notion of "realistic" datasets.

This relaxation looks similar to probabilistic DP, but is actually different: both have a small probability that the risk is unbounded, but while random DP this probability is computed across inputs of the mechanism (i.e., users or datasets), for probabilistic DP it is computed across mechanism outputs. Also similarly to probabilistic DP, excluding some cases altogether creates definitional issues: random DP does not satisfy the convexity axiom (see Proposition 10 in Appendix A). We postulate that using a different mechanism to allow some inputs to not satisfy the mechanism, similar to Rényi DP, could solve this problem.

Usually, data-generating distributions are used for other purposes: they typically model an adversary with partial knowledge. However, definitions in this section still compare the outputs of the mechanisms given fixed neighboring datasets: the only randomness in the indistinguishability property comes from the mechanism. By contrast, definitions of Section 6 compare the output of the mechanism on a random dataset, so the randomness comes both from the data-generating distribution and the mechanism.

5.3 Multidimensional definitions

The definitions described in Section 4 (e.g., generic DP or blowfish privacy), have the same privacy constraint for all neighboring datasets. Thus, they cannot capture definitions that vary the privacy level across inputs. *$d_{\mathcal{D}}$-privacy* is introduced in [22] to capture both ideas of varying the neighborhood definition and varying the privacy levels across inputs: the function $d_{\mathcal{D}}$ takes both databases as input, so it can both capture arbitrary neighborhood definitions and return different values depending on the difference between the two.

Definition 14 ($d_{\mathcal{D}}$-privacy [22])**.** *Let $d_{\mathcal{D}} : \mathcal{D}^2 \to \mathbb{R}_\infty$. A privacy mechanism $\mathcal{M}$ satisfies $d_{\mathcal{D}}$-privacy if for all pairs of datasets D_1, D_2 and all sets of outputs $S \subseteq \mathcal{O}$:*

$$\mathbb{P}\left[\mathcal{M}\left(D_1\right) \in S\right] \leq e^{d_{\mathcal{D}}(D_1, D_2)} \cdot \mathbb{P}\left[\mathcal{M}\left(D_2\right) \in S\right]$$

Equivalent definitions appeared in [60] as *l-privacy*, and in [98] as *extended DP*. Several other definitions, such as *weighted DP* [138], *smooth DP* [9] and *earth mover's privacy* [68] can be seen as instantiations of $d_{\mathcal{D}}$*-privacy* for some distance functions d.

Random DP can also be combined with changing the neighborhood definition: in [171], the authors define *DP on a δ-location set*, for which the neighborhood is defined by a set of "plausible" locations around the true location of a user. In [177], the authors define *constrained DP*, in which two neighboring datasets are also required to satisfy a utility related metric constraint. Furthermore, this idea was combined with only imposing the privacy property on a large density of datasets in *distributional privacy* [141, 177].

Different risk models, like the definitions in Section 3, are also compatible with varying the privacy parameters across inputs. For example, in [107], the author proposes *endogeneous DP*, which is a combination of (ε, δ)-DP and personalized DP. Similarly, *pseudo-metric DP*, defined in [36], is a combination of $d_{\mathcal{D}}$-privacy and (ε, δ)-DP; while *extended divergence DP* [97] combines $d_{\mathcal{D}}$-privacy with divergence DP.

Randomly limiting the scope of the definition can also be combined with ideas from the previous sections. For example, in [10], authors introduce *typical privacy*, which combines random DP with approximate DP. In [164], the authors introduce *on average KL privacy*, which uses KL-divergence as quantification metric, but only requires the property to hold for an "average dataset", like random DP. A similar notion appears in [67] as *average leave-one-out KL stability*.

In [103], the authors introduce *general DP*[9], which goes further and generalizes the intuition from generic DP, by abstracting the indistinguishability condition entirely: the privacy relation $\mathcal{R}$ is still the generalization of the neighborhood and the privacy predicate is the generalization of ε-indistinguishability to arbitrary functions. This definition was further extended via *abstract DP*, however, that definition does not satisfy the privacy axioms in general.

Definitions in this section are particularly used in the context of local DP[10] and in particular for applications to location privacy: various metrics have been discussed to quantify how indistinguishable different places should be to provide users of a local DP mechanism with meaningful privacy protection [23].

6 Background knowledge (B)

In DP, the attacker is implicitly assumed to have full knowledge of the dataset: their only uncertainty is whether the target belongs in the dataset or not. This implicit assumption is also present for the definitions of the previous dimensions: indeed, the attacker has to distinguish between two *fixed* datasets D_1 and D_2. The only source of randomness in ε-indistinguishability comes from the mechanism itself. In many cases, this assumption is unrealistic, and it is natural to consider weaker adversaries, who do not have full background knowledge. One of the main motivations to do so is to use significantly less noise in the mechanism [44].

The typical way to represent this uncertainty formally is to assume that the input data comes from a certain probability distribution (named "data evolution scenario" in [106]): the randomness of this distribution models the attacker's uncertainty. Using a probability distribution to generate the input data means that the ε-indistinguishability property cannot be expressed between two fixed datasets. Instead, one natural way to express it is to condition this distribution on some sensitive property such as in *noiseless privacy* [14, 44][11]

Definition 15 ((Θ, ε)-noiseless privacy [14, 44])**.** *Given a family Θ of probability distribution on $\mathcal{D}$, a mechanism $\mathcal{M}$ is (Θ, ε)-noiseless private if for all $\theta \in \Theta$, all i and all $t, t' \in \mathcal{T}$:*

$$\mathcal{M}(D)_{|D\sim\theta, D(i)=t} \approx_{\varepsilon} \mathcal{M}(D)_{|D\sim\theta, D(i)=t'}$$

This definition follows naturally from considering the associative view of DP with the strong adversary assumption, and attempting to relax this assumption. The exact way to model the adversary's uncertainty can be changed; for example *DP under sampling* [115], an instance of noiseless privacy, models it using random sampling.

In [11] however, the authors argue that in the presence of correlations in the data, noiseless privacy can be too strong. Indeed, if one record can have an arbitrarily large influence on the rest of the data, conditioning on the value of this record can lead to very distin-

9 Originally called generic DP; we rename it here to avoid a name collision.

10 For details, see Section 10.3.

11 Not to be confused with the definition with the same name introduced in [65], mentioned in Section 10.2.

guishable outputs even if the mechanism only depends on global properties of the data. To fix this problem, they propose *distributional DP*, an alternative definition that only considers the influence of one user once the database has already been randomly picked from the data-generating distribution. In this definition, one record is changed *after* the dataset has been generated, an approach closer to the causal interpretation of DP.

6.1 Multidimensional definitions

Modifying the risk model while limiting the attacker's background knowledge has interesting consequences. In [35], the authors show that two options are possible: either consider the background knowledge as additional information given to the attacker or let the attacker *influence* the background knowledge. This distinction between an *active* and a *passive* attacker does not matter if only the worst-case scenario is considered, like in noiseless privacy. However, under different risk models, such as allowing a small probability of error, they lead to two different definitions.

The first, *active partial knowledge DP*, quantifies over all possible values of the background knowledge. It was introduced in [11, 14] and reformulated in [35] to clarify that it implicitly assumes an active attacker.

The second definition, *passive partial knowledge DP* [35], is strictly weaker: it models a passive attacker, who cannot choose their background knowledge, and thus cannot influence the data. In this context, δ does not only apply to the output of the mechanism, but also to the value of the background knowledge.

Modifying the neighborhood definition is simpler: it is clearly orthogonal to the dimensions introduced in this section. In both noiseless privacy and distributional DP, the two possibilities between which the adversary must distinguish are similar to bounded DP. This can easily be changed to choose other properties to protect from the attacker. This is done in *pufferfish privacy* [106], which extends the concept of neighboring datasets to neighboring *distributions* of datasets.

Definition 16 ($(\Theta, \Phi, \varepsilon)$-pufferfish privacy [106])**.** *Given a family of probability distributions Θ on $\mathcal{D}$, and a family of pairs of predicates Φ on datasets, a mechanism $\mathcal{M}$ verifies $(\Theta, \Phi, \varepsilon)$-pufferfish privacy if for all distributions $\theta \in \Theta$ and all pairs of predicates $(\phi_1, \phi_2) \in \Phi$:*

$$\mathcal{M}(D)_{|D\sim\theta,\phi_1(D)} \approx_\varepsilon \mathcal{M}(D)_{|D\sim\theta,\phi_2(D)}$$

Pufferfish privacy starts with a set of data-generating distributions, then conditions them on sensitive attributes. This notion extends noiseless privacy, as well as other definitions like *bayesian DP* [112], in which neighboring records only have a fraction of elements in common, and some are generated randomly.

The same idea can be formalized by comparing pairs of distributions directly. This is done in [98] via *distribution privacy*, instantiated in [73] via *profile-based DP*, in which the attacker tries to distinguish between different probabilistic user profiles.

Further relaxations encompassing the introduced dimensions are *probabilistic distribution privacy*[12] (combination of distribution privacy and probabilistic DP), *extended distribution privacy* [98] (combination of distribution privacy and $d_{\mathcal{D}}$-privacy), *divergence distribution privacy* [97] (combination of distribution privacy and divergence DP), *extended divergence distribution privacy* [97] (combination of the latter two options), and *divergence distribution privacy with auxiliary inputs* [97] which considers the setting where the attacker uses full or approximate knowledge rather than perfect knowledge on the input probability distributions (as in profile-based privacy).

Definitions of this section are an active area of research; a typical question is to quantify in which conditions deterministic mechanisms can provide some level privacy. However, they are not used a lot in practice, likely because of their fragility: if the assumptions about the attacker's limited background knowledge are wrong in practice, then the definitions do not provide any guarantee of protection.

7 Change in formalism (F)

ε-indistinguishability compares the distribution of outputs given two neighboring inputs. This is not the only way to encompass the idea that a Bayesian attacker should not be able to gain too much information on the dataset. One such a formalism reformulates DP in terms of hypothesis testing [95, 166] by limiting the type I and the type II error of the hypothesis that the output O of a mechanism originates from D_1 (instead of D_2). This interpretation was also used in [40] in *f-DP* (a reformulation of generic DP) and in *Gaussian DP* (a specific instance of f-DP), which can capture both (ε, δ)-DP and Rényi DP.

12 Appears in the extended ArXiv version of [98].

Other formalisms model the attacker *explicitly*, by formalizing their prior belief as a probability distribution over all possible datasets. This can be done in two distinct ways. Some variants consider a specific prior (or family of possible priors) of the attacker, implicitly assuming a limited background knowledge, like in Section 6. We show that these variants can be interpreted as changing the prior-posterior bounds of the attacker. Another possibility compares two posteriors, quantifying over all possible priors. In practice, these definitions are mostly useful in that comparing them to DP leads to a better understanding of the guarantees that DP provides.

7.1 Changing the shape of the prior-posterior bounds

DP can be interpreted as giving a bound on the posterior of a Bayesian attacker as a function of their prior. This is exactly the case in *indistinguishable privacy*, an equivalent reformulation of DP defined in [117]: suppose that the attacker is trying to distinguish between two options $D = D_1$ and $D = D_2$, where D_1 corresponds to the option "$t \in D$" and D_2 to "$t \notin D$". Initially, they associate a certain prior probability $\mathbb{P}[t \in D]$ to the first option. When they observe the output of the algorithm, this becomes the posterior probability $\mathbb{P}[t \in D|\mathcal{M}(D) = O]$. From Definition 2, we have:

$$\frac{\mathbb{P}[t \in D|\mathcal{M}(D) = O]}{\mathbb{P}[t \notin D|\mathcal{M}(D) = O]} \leq e^{\varepsilon} \cdot \frac{\mathbb{P}[t \in D]}{\mathbb{P}[t \notin D]} \Rightarrow$$
$$\mathbb{P}[t \in D|\mathcal{M}(D) = O] \leq \frac{e^{\varepsilon} \cdot \mathbb{P}[t \in D]}{1 + (e^{\varepsilon} - 1)\mathbb{P}[t \in D]}$$

A similar, symmetric lower bound can be obtained. Hence, DP can be interpreted as bounding the posterior level of certainty of a Bayesian attacker as a function of its prior. We visualize these bounds in the top left side of Figure 2.

Some variants of DP use this idea in their formalism, rather than obtaining the posterior bound as a corollary to the classical DP definition. For example, *positive membership privacy* [114] requires that the posterior does not increase too much compared to the prior. Like noiseless privacy [14], it assumes an attacker with limited background knowledge.

Definition 17 ((Θ, ε)-positive membership privacy [114])**.** *A privacy mechanism $\mathcal{M}$ provides (Θ, ε)-positive membership privacy if for any distribution $\theta \in \Theta$, any record $t \in \mathcal{D}$ and any $S \subseteq \mathcal{O}$:*

$$\mathbb{P}_{D\sim\theta}[t \in D|\mathcal{M}(D) \in S] \leq e^{\varepsilon}\mathbb{P}_{D\sim\theta}[t \in D] \text{ and}$$
$$\mathbb{P}_{D\sim\theta}[t \notin D|\mathcal{M}(D) \in S] \geq e^{-\varepsilon}\mathbb{P}_{D\sim\theta}[t \notin D]$$

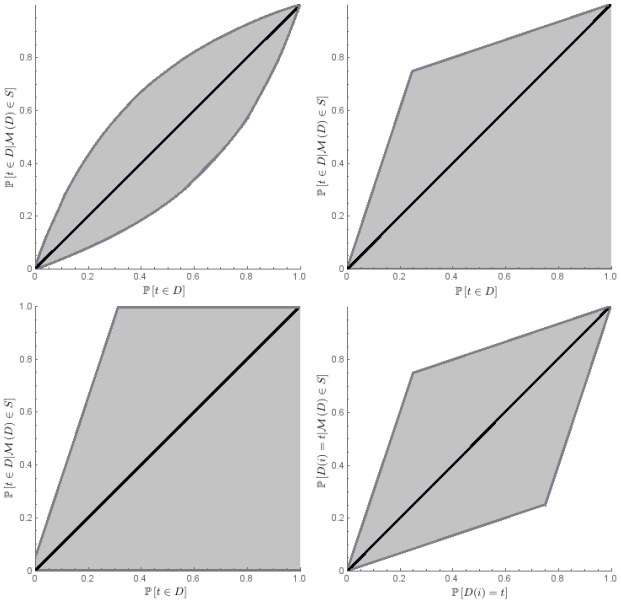

Fig. 2. From top left to bottom right, using $\varepsilon = \ln 3$: posterior-prior bounds in the original DP, positive membership privacy, adversarial privacy (with $\delta = 0.05$) and aposteriori noiseless privacy.

Note that this definition is *asymmetric*: the posterior is bounded from above, but not from below. It is visualized the top right part of Figure 2. In the same paper, the authors also define *negative membership privacy*, which provides the symmetric lower bound, and *membership privacy*, which is the conjunction of the two. Additionally, the same idea is captured in *ε-DP location obfuscation* [41] in the context of location privacy.

A previous attempt at formalizing the same idea was presented in [139] as *adversarial privacy*. This definition is similar to positive membership privacy, except only the first relation is used, and there is a small additive δ as in approximate DP. We visualize the corresponding bounds on the bottom left of Figure 2. Moreover, in [14] *aposteriori noiseless privacy* is introduced with the corresponding bounds seen in the bottom right side of Figure 2.

Further relaxations from [139] are *tuple indistinguishability*, *relaxed indistinguishability* and *bounded adversarial privacy*. Finally, the idea behind distribution privacy is reformulated in Bayesian terms in [90].

7.2 Comparing two posteriors

In [96], the authors propose an approach that captures an intuitive idea proposed by Dwork in [48]: "any conclusions drawn from the output of a private algorithm must be similar whether or not an individual's data is present in the input or not". They define *semantic privacy*: instead of comparing the posterior with the prior belief like in DP, this bounds the difference between two posterior belief distributions, depending on which database was secretly chosen. The distance chosen to represent the idea that those two posterior belief distributions are close is the statistical distance. One important difference between the definitions in the previous subsection is that semantic privacy quantifies over all possible priors: like in DP, the attacker is assumed to have arbitrary background knowledge.

Definition 18 (ε-semantic privacy [69, 96])**.** *A mechanism $\mathcal{M}$ is ε-semantically private if for any distribution over datasets θ, any index i, any $S \subseteq \mathcal{O}$, and any set of datasets $X \subseteq \mathcal{D}$:*

$$|\mathbb{P}[D \in X \mid \mathcal{M}(D) \in S] - \mathbb{P}[D \in X \mid \mathcal{M}(D_{-i}) \in S]| \leq \varepsilon$$

where D is chosen randomly from θ.

Other definitions comparing posteriors directly are *posteriori DP* [160] (although the prior is not made explicit), *inferential privacy* [75] (a reformulation of noiseless privacy), and *range-bounded privacy* [47] (a technical definition, equivalent to DP up to a change in parameters).

7.3 Multidimensional definitions

Definitions that limit the background knowledge of the adversary explicitly formulate it as a probability distribution. As such, they are natural candidates for Bayesian reformulations. In [168], the authors introduce *identity DP*, which is an equivalent Bayesian reformulation of noiseless privacy.

Another example is *inference-based distributional DP* [11], which relates to distributional DP the same way as noiseless privacy and its aposteriori version: they are equivalent, however, when a small additive error is introduced to the definitions, the inference and aposteriori based versions become weaker [11].

Further, it is possible to modify the neighborhood definition. In [43], the authors introduce *information privacy*, which can be seen as a posteriori noiseless privacy combined with free lunch privacy: rather than considering the knowledge gain of the adversary on one particular user, it considers its knowledge gain about any possible value of the database.

Definitions in this section mostly appear in theoretical research papers, to provide a deeper understanding of guarantees offered by DP and its alternatives. They do not seem to be used in practical applications.

8 Relativization of the knowledge gain (R)

When using the associative interpretation with the independence assumption, it is unclear how to adapt DP to correlated datasets like social networks: data about someone's friends might reveal sensitive information about this person. The causal interpretation of DP does not suffer from this problem, but how to adapt the associative view to such correlated contexts? Changing the definition of the neighborhood is one possibility (see Section 4.1), but it requires knowing in advance the exact impact of someone on other records. A more robust option is to impose that the information released does not contain more information than the result of some predefined algorithms on the data, without the individual in question. The method for formalizing this intuition borrows ideas from *zero-knowledge proofs* [78].

In a data privacy context, instead of imposing that the result of the mechanism is roughly *the same* on neighboring datasets D_1 and D_2, it is possible to impose that the result of the mechanism on D_1 can be *simulated* using only some information about D_2. The corresponding definition, called *zero-knowledge privacy* and introduced in [72], captures the idea that the mechanism does not leak more information on a given target than a certain class of aggregate metrics (called *model of aggregate information*).

Definition 19 ($(\mathsf{Agg}, \varepsilon)$-zero-knowledge privacy [72])**.** *Let Agg be a family of (possibly randomized) algorithms agg. A privacy mechanism $\mathcal{M}$ is $(\mathsf{Agg}, \varepsilon)$-zero-knowledge private if there exists an algorithm $\mathsf{agg} \in \mathsf{Agg}$ and a simulator Sim such as for all pairs of neighboring datasets D_1 and D_2, $\mathcal{M}(D_1) \approx_\varepsilon \mathsf{Sim}(\mathsf{agg}(D_2))$.*

8.1 Multidimensional definitions

Using a simulator allows making statements of the type "this mechanism does not leak more information on a given target than a certain class of aggregate metrics".

Similarly to pufferfish privacy, we can vary the neighborhood definitions (to protect other types of information than the presence and characteristics of individuals), and explicitly limit the attacker's background knowledge using a probability distribution. This is done in [11] as *coupled-worlds privacy*, a generalization of distributional privacy, where a family of functions priv represents the protected attribute.

Definition 20 ($(\Theta, \Gamma, \varepsilon)$-coupled-worlds privacy [11])**.** *Let Γ be a family of pairs of functions $(\mathsf{agg}, \mathsf{priv})$. A mechanism $\mathcal{M}$ satisfies $(\Theta, \Gamma, \varepsilon)$-coupled-worlds privacy if there is a simulator Sim such that for all distributions $\theta \in \Theta$, all $(\mathsf{agg}, \mathsf{priv}) \in \Gamma$, and all possible values p:*

$$\mathcal{M}(D)_{|D\sim\theta, \mathsf{priv}(D)=p} \approx_\varepsilon \mathsf{Sim}(\mathsf{agg}(D))_{|D\sim\theta, \mathsf{priv}(D)=p}$$

This definition is a good example of combining different dimensions: it changes several aspects of the original definition according to **N**, **B** and **R**. Moreover, **Q** and **F** can easily be integrated with this definition by using (ε, δ)-indistinguishability with a Bayesian reformulation. This is done explicitly in *inference-based coupled-worlds privacy* [11].

We did not find any evidence that these variants and extensions are used in practice.

9 Computational power (C)

The ε-indistinguishability property in DP is *information-theoretic*: the attacker is implicitly assumed to have infinite computing power. This is unrealistic in practice, so it is natural to consider definitions where the attacker only has polynomial computing power. Changing this assumption leads to weaker data privacy definitions. In [129], two approaches have been proposed to formalize this idea.

The first approach is to model the distinguisher explicitly as a polynomial Turing machine: this is *indistinguishability-based computational DP*.

Definition 21 (ε_κ-IndCDP [129])**.** *A family $(\mathcal{M}_\kappa)_{\kappa\in\mathbb{N}}$ of privacy mechanisms $\mathcal{M}_\kappa$ provides ε_κ-IndCDP if there exists a negligible function neg such that for all non-uniform probabilistic polynomial-time Turing machines Ω (the distinguisher), all polynomials $p(\cdot)$, all sufficiently large $\kappa \in \mathbb{N}$, and all datasets $D_1, D_2 \in \mathcal{D}$ of size at most $p(\kappa)$ that differ only one one record:*

$$\mathbb{P}[\Omega(\mathcal{M}(D_1)) = 1] \leq e^{\varepsilon_\kappa} \cdot \mathbb{P}[\Omega(\mathcal{M}(D_2)) = 1] + \mathsf{neg}(\kappa)$$

where neg is a function that converges to zero asymptotically faster than the reciprocal of any polynomial.

One instantiation of this is *output-constrained DP*, presented in [87]: the definition is adapted to a two-party computation setting, where each party has their own set of privacy parameters.

The second approach is to require that the mechanism "looks like" a truly differentially private mechanism, at least to a computationally bounded distinguisher. In [129], the authors introduce SimCDP, short for *simulation-based computational DP*.

Definition 22 (ε_κ-SimCDP [129])**.** *A family $(\mathcal{M}_\kappa)_{\kappa\in\mathbb{N}}$ of privacy mechanisms $\mathcal{M}_\kappa$ provides ε_κ-SimCDP if there exists a family $(\mathcal{M}'_\kappa)_{\kappa\in\mathbb{N}}$ of ε_κ-DP and a negligible function neg such that for all non-uniform probabilistic polynomial-time Turing machines Ω, all polynomials $p(\cdot)$, all sufficiently large $\kappa \in \mathbb{N}$, and all datasets $D \in \mathcal{D}$ of size at most $p(\kappa)$:*

$$\mathbb{P}[\Omega(\mathcal{M}(D)) = 1] - \mathbb{P}\left[\Omega(\mathcal{M}'(D)) = 1\right] \leq \mathsf{neg}(\kappa)$$

where neg is a function that converges to zero asymptotically faster than the reciprocal of any polynomial.

In [129], the authors show that ε_κ-SimCDP implies ε_κ-IndCDP.

9.1 Multidimensional Definitions

Some DP variants which explicitly model an adversary with a simulator can relatively easily be adapted to model a computationally bounded adversary, simply by imposing that the simulator must be polynomial. This is done explicitly in [72], where the authors define *computational zero-knowledge privacy*, which could also be adapted to e.g., the two coupled-worlds privacy definitions.

Modeling a computationally bounded adversary is orthogonal to changing the type of input data, as well as considering an adversary with limited background knowledge: in [8], the authors define *differential indistinguishability*, which prevents a polynomial adversary from distinguishing between two Turing machines with random input.

Limiting the computational power of the attacker is a reasonable assumption, but for a large class of queries, it cannot provide significant benefits over classical DP in the typical client-server setting [79]. Thus, existing work using it focuses on multi-party settings [13].

10 Scope and related work

In this section, we list related works and existing surveys in the field of data privacy, we detail our criteria for excluding particular data privacy definition from our work, and we list some relevant definitions that were excluded by this criteria.

10.1 Methodology

Whether a data privacy definition fits our description is not always obvious, so we use the following criterion: the attacker's capabilities must be clearly defined, and the definition must prevent this attacker from learning about a protected property. Consequently, we do not consider definitions which are a property of the output data and not of the mechanism, variants of technical notions that are not data privacy properties (like different types of sensitivity), nor definitions whose only difference with DP is in the context and not in the formal property (like the distinction between local and global models).

To find a comprehensive list of DP notions, besides the definitions we were aware of or suggested by experts, we conducted a wide literature review using two research datasets: BASE[13] and Google Scholar[14]. The exact queries were run twice: first on October 31th, 2018, and second on August 1st, 2019. The corresponding result counts are summarized in Table 3; the number in parentheses corresponds to the first run.

Query (BASE)	Hits
"differential privacy" relax year:[2000 to *]	115 (99)
"differential privacy" variant -relax year:[2000 to *]	110 (87)
Query (Google Scholar)	**Hits**
"differential privacy" "new notion"	206 (162)
"differential privacy" "new definition" -"new notion"	171 (129)

Table 3. Queries for the literature review.

First, we manually reviewed each abstract and filter out papers until we had only papers which either contained a new definition or were applying DP in a new setting. All papers which defined a variant or extension of DP are cited in this work.

13 https://www.base-search.net/
14 https://scholar.google.com/

10.2 Out of scope definitions

Besides syntactic DP definitions, some definitions 1) do not provide a clear privacy guarantee or 2) are only used as a tool in order to prove links between existing definitions. As such, we did not include them in our survey (more details can be found in the full version of this work).

Examples for the former include *ε-privacy* [121] (the first attempt at formalizing an adversary with restricted background knowledge, whose formulation did not have the same interpretation as noiseless privacy), *differential identifiability* [110] (bounds the probability that a given individual's information is included in the input datasets, *noiseless privacy* [65] but does not measure the *change* in probabilities between the two alternatives), *crowd-blending privacy* [71] (combines DP with k-anonymity), and *(k, ε)-anonymity* [89] (performs k-anonymisation on a subset of the quasi identifiers and then ε-DP on the remaining quasi-identifiers with different settings for each equivalence class).

Examples for the latter include further DP relaxations were created by changing the *sensitivity* of the function what the mechanism protects. There are many variants to the initial concept of global sensitivity [54]: local sensitivity [134], smooth sensitivity [134], restricted sensitivity [17], empirical sensitivity [27], recommendation-aware sensitivity [178], record and correlated sensitivity [179], dependence sensitivity [116], per-instance sensitivity [162], individual sensitivity [32], elastic sensitivity [92] and derivative sensitivity [109]. We did not consider these notions as these do not modify the actual definition of DP.

10.3 Local model

In this work we focused on DP variants/extensions typically used in the *global model*, in which a central entity has access to the whole dataset. It is also possible to use DP in other contexts, without formally changing the definition. The main alternative is the *local model*, where each individual randomizes their own data before sending it to an aggregator. This model is used e.g. by Google [61], Apple [154], or Microsoft [37]. Another option is the *shuffled model* [15], which falls in-between the local and global models.

Some definitions we listed actually were presented in the local model, such as $d_{\mathcal{D}}$-privacy [22], geo-indistinguishability [6], earth mover's Pr [68], location Pr [60], profile-based DP [73], divergence DP and smooth DP from [9], utility-optimized DP [97], and ex-

tended DP, distribution Pr, and extended distribution Pr from [98]. Besides these, a handful of other local DP definitions are mentioned below; a longer discussion can be found in the full version of this work (soon to be uploaded to ArXiv[15]).

The idea behind *local DP* [45] was proposed in [146] as *distributed DP*, where authors additionally assume that only a portion of participants is honest. *Joint DP* [99] models a game in which each player cannot learn the data from other players. In *multiparty DP* [167], the view of each subgroup of players is differentially private with respect to other players. Some variants introduced in this work were also considered in the local setting; examples include *localized information privacy* [91], *one-sided local DP* [130], *personalized local DP* [132], and $d_{\mathcal{D}}$*-local DP* [4, 81] (called *condensed local DP* in [81]).

10.4 Related work

The relation between the main syntactic models of anonymity and DP was studied in [29], in which the authors claim that the former is designed for privacy-preserving data publishing (PPDP), while DP is more suitable for privacy preserving data mining (PPDM).

In [161], authors establish connections between differential privacy (seen as the additional disclosure of an individual's information due to the release of the data), *identifiability* (seen as the posteriors of recovering the original data from the released data), and *mutual-information privacy* (which measures the average amount of information about the original dataset contained in the released data).

Some of the earliest surveys focusing on DP were written by Dwork [49, 50], and summarize algorithms achieving DP and applications. The more detailed "privacy book" [57] presents an in-depth discussion about the fundamentals of DP, techniques for achieving it, and applications to query-release mechanisms, distributed computations or data streams. Another recent survey focuses on the release of histograms and synthetic data with DP [131].

In [88], the authors classify different enhancing technologies (PETs) into 7 complementary dimensions. Indistinguishability falls into the *Aim* dimension, but within this category, only k-anonymity and oblivious transfer are considered. In [1], the authors survey privacy concerns, measurements and techniques used in online social networks and recommender systems. They classify privacy into 5 categories; DP falls into *Privacy-preserving models*. In [159], the authors classify 80+ privacy metrics into 8 categories based on the output of the privacy mechanism. One of their classes is *Indistinguishability*, which contains DP as well as several variants. Some variants are classified into other categories; e.g., Rényi DP is classified into *Uncertainty* and mutual-information DP into *Information gain/loss*. The authors list 8 DP definitions; our taxonomy can be seen as an extension of the contents of their work (and in particular of the *Indistinguishability* category).

Finally, some surveys focus on location privacy. In [123], authors highlight privacy concerns in this context and list mechanisms with formal provable privacy guarantees, while in [23], authors analyze different kinds of privacy breaches and compare metrics that have been proposed to protect location data.

11 Conclusion

We proposed a classification of DP variants and extensions using the concept of dimensions. When possible, we compared definitions from the same dimension, and we showed that definitions from the different dimensions can be combined to form new, meaningful definitions. In theory, it means that even if there were only three possible ways to change a dimension, this would result in $3^7 = 2187$ possible definitions: the ≈ 200 existing definitions shown in Figure 1 are only scratching the surface of the space of possible notions. Using these dimensions, we unified and simplified the different notions proposed in the literature. We highlighted their properties such as composability and whether they satisfy the privacy axioms by either collecting the existing results or creating new proofs, and whenever possible, we showed their relative relations to one another. We hope that this work will make the field of data privacy more organized and easier to navigate, especially for new practitioners.

12 Acknowledgments

The authors would like to thank Alex Kulesza, Esfandiar Mohammadi, David Basin, and the anonymous reviewers for their helpful comments. This work was partially funded by Google, and done while Balázs Pejó was at University of Luxembourg.

15 https://arxiv.org/abs/1906.01337

Name & references	Dimensions[24]	Axioms P.P.[18]	Axioms Cv.[19]	Cp.[20]	Relations
(ε,δ)-approximate DP [52]	**Q**	✓[1]	✓[1]	✓[13]	(ε,δ)-**DP** $\supset_{\prec}$ ε-DP
(ε,δ)-probabilistic DP [20, 124, 127]	**Q**	✗[2]	✗[3]	✓[13]	(ε,δ)-DP $\prec$ (ε,δ)-**ProDP** $\supset_{\prec}$ ε-DP
ε-Kullback-Leiber Pr [9, 31]	**Q**	✓[1]	✓[1]	✓[13]	(ε,δ)-DP $\prec$ ε-**KLPr** $\prec$ ε-DP
(α,ε)-Rényi DP [128]	**Q**	✓[1]	✓[1]	✓[13]	ε-KLPr $\subset_{\prec}$ (α,ε)-**Rényi DP** $\supset_{\prec}$ ε-DP
ε-mutual-information DP [31]	**Q**	✓[1]	✓[1]	✓[13]	(ε,δ)-DP $\prec$ ε-**MIDP** $\prec$ ε-KLPr
(μ,τ)-mean concentrated DP [58]	**Q**	✗[2]	?	✓[13]	(ε,δ)-DP $\prec$ (μ,τ)-**mCoDP** $\prec$ ε-DP
(ξ,ρ)-zero concentrated DP [19]	**Q**	✓[1]	✓[1]	✓[13]	(ξ,ρ)-**zCoDP** $\sim$ (μ,τ)-mCoDP
(f,ε)-divergence DP [9]	**Q**	✓[1]	✓[1]	?	most definitions in **Q** $\subset$ (f,ε)-**DivDP**
ε-unbounded DP [105]	**N**	✓[4]	✓[4]	✓[14]	ε-DP $\sim$ ε-**uBoDP** $\subset^{\sim}$ (c,ε)-GrDP
ε-bounded/attribute/bit DP [105]	**N**	✓[4]	✓[4]	✓[14]	ε-**BitDP** $\prec$ ε-**AttDP** $\prec$ ε-**BoDP** $\prec$ ε-DP
(c,ε)-group DP [49]	**N**	✓[4]	✓[4]	✓[14]	ε-DP $\subset^{\sim}$ (c,ε)-**GrDP**
ε-free lunch Pr [105]	**N**	✓[4]	✓[4]	✓[14]	all definitions in **N** $\prec$ ε-**FLPr**
(R,c,ε)-dependent DP [116]	**N**	✓[4]	✓[4]	✓[14]	(c,ε)-GrDP $\subset$ (R,c,ε)-**DepDP**
(P,ε)-one-sided DP [42]	**N**	✓[4]	✓[4]	✓[14]	(P,ε)-**OnSDP** $\supset_{\prec}$ ε-BoDP
(D,ε)-individual DP [149]	**N**	✓[4]	✓[4]	✓[14]	(D,ε)-**InDP** $\prec$ ε-DP
(D,t,ε)-per-instance DP [162]	**N**	✓[4]	✓[4]	✓[14]	(D,t,ε)-**PIDP** $\prec$ (D,ε)-InDP
$(\mathcal{R},\varepsilon)$-generic DP [105]	**N**	✓[4]	✓[4]	✓[14]	most definitions in **N** $\subset$ $(\mathcal{R},\varepsilon)$-**GcDP**
$(G,\mathcal{I}_Q,\varepsilon)$-blowfish Pr [84, 86]	**N**	✓[4]	✓[4]	✓[14]	$(G,\mathcal{I}_Q,\varepsilon)$-**BFPr** $\subset$ $(\mathcal{R},\varepsilon)$-GcDP
ε-adjacency-relation div. DP [97]	**Q,N**	✓[1,5]	✓[1,5]	?	$(\mathcal{R},\varepsilon)$-GcDP $\subset$ $(\mathcal{R},f,\varepsilon)$-**ARDDP** $\supset$ (f,ε)-DivDP
Ψ-personalized DP [59, 76, 94, 118]	**V**	✓[8]	✓[8]	✓[14]	ε-DP $\subset$ Ψ-**PerDP**
Ψ-tailored DP/$\varepsilon(\cdot)$-outlier Pr [120]	**V**	✓[8]	✓[8]	✓[14]	Ψ-PerDP $\subset$ Ψ-**TaiDP** $\supset$ $\varepsilon(\cdot)$-**OutPr**
(π,γ,ε)-random DP [83]	**V**	✓[9]	✗[10]	✓[16]	(π,γ,ε)-**RanDP** $\supset_{\prec}$ ε-DP
$d_{\mathcal{D}}$-Pr [22]	**N,V**	✓[8]	✓[8]	✓[14]	ε-**DP** $\subset$ $d_{\mathcal{D}}$-**Pr**
(ε,γ)-distributional Pr [141, 177]	**N,V**	?[21]	?[21]	?	ε-FLPr $\subset$ (ε,γ)-**DlPr** $\supset$ (π,γ,ε)-RanDP
$(\varepsilon(\cdot),\delta(\cdot))$-endogenous DP [107]	**Q,V**	✓[8]	✓[8]	✓[14]	(ε,δ)-DP $\subset$ $(\varepsilon(\cdot),\delta(\cdot))$-**EndDP** $\supset_{\prec}$ Ψ-PerDP
$(d_{\mathcal{D}},\varepsilon,\delta)$-pseudo-metric DP [36]	**Q,N,V**	?	?	✓[15]	(ε,δ)-DP $\subset$ $(d_{\mathcal{D}},\varepsilon,\delta)$-**PsDP** $\supset_{\prec}$ $d_{\mathcal{D}}$-Pr
$(\theta,\varepsilon,\gamma,\delta)$-typical Pr [10]	**Q,V**	✓[9]	✗[10]	✓[14]	(ε,δ)-DP $\subset_{\succ}$ $(\theta,\varepsilon,\gamma,\delta)$-**TypPr** $\supset_{\prec}$ (π,γ,ε)-RanDP
(Θ,ε)-on average KL Pr [164]	**Q,V**	✓[1,9]	?	?[22]	ε-KL Pr $\subset$ (Θ,ε)-**avgKL Pr** $\supset$ $(\Theta,\gamma,\varepsilon)$-Ran DP
(f,d,ε)-extended divergence DP [97]	**Q,N,V**	✓[8]	✓[8]	?	$d_{\mathcal{D}}$-Pr $\subset$ (f,d,ε)-**EDiv DP** $\supset$ (f,ε)-Div DP
$(\mathcal{R},M)$-general DP [103]	**Q,N,V**	✓[5]	✓[5]	?	$(\mathcal{R},M)$-**Gl DP** $\supset$ (ε,δ)-DP
(Θ,ε)-noiseless Pr [14, 44]	**B**	✓[4]	✓[4]	✗[17]	ε-DP $\subset$ (Θ,ε)-**NPr**
(Θ,ε)-distributional DP [11, 35]	**B**	✓[6]	✓[6]	✗[17]	(Θ,ε)-**DistDP** $\supset_{\prec}$ ε-DP
$(\Theta,\varepsilon,\delta)$-active PK DP [11, 14, 35]	**Q,B**	✓[7]	✓[7]	✗[17]	$(\Theta,\varepsilon,\delta)$-**APKDP** $\supset_{\prec}$ (Θ,ε)-NPr
$(\Theta,\varepsilon,\delta)$-passive PK DP [35]	**Q,B**	✓[7]	✓[7]	✗[17]	$(\Theta,\varepsilon,\delta)$-APKDP $\succ$ $(\Theta,\varepsilon,\delta)$-**PPKDP** $\supset_{\prec}$ (Θ,ε)-NPr
$(\Theta,\Phi,\varepsilon)$-pufferfish Pr [106]	**N,B**	✓[4]	✓[4]	✗[17]	(Θ,ε)-NPr $\subset$ $(\Theta,\Phi,\varepsilon)$-**PFPr** $\supset$ $(\mathcal{R},\varepsilon)$-GcDP
$(\Theta,\varepsilon,\delta)$-distribution Pr [98]	**Q,N,B**	✓[8]	✓[8]	✗[17]	$(\Theta,\varepsilon,\delta)$-APKDP $\subset$ $(\Theta,\varepsilon,\delta)$-**DnPr**
(d,Θ,ε)-extended DnPr [98]	**N,V,B**	✓[8]	✓[8]	✗[17]	$d_{\mathcal{D}}$-Pr $\subset$ (d,Θ,ε)-**EDnPr** $\supset$ (Θ,ε)-DnPr
(f,Θ,ε)-divergence DnPr [97]	**Q,N,B**	✓[8]	✓[8]	✗[17]	(f,ε)-DP $\subset$ (f,Θ,ε)-**DDnPr** $\supset$ (Θ,ε)-DnPr

Name & references	Dimensions[24]	Axioms P.P.[18]	Axioms Cv.[19]	Cp.[20]	Relations
$(d, f, \Theta, \varepsilon)$-ext. div. DnPr [97]	**Q,N,V,B**	✓[8]	✓[8]	✗[17]	(f, Θ, ε)-DDPr $\subset$ $(d, f, \Theta, \varepsilon)$-**EDDnPr** $\supset$ (d, Θ, ε)-EDnPr
(Θ, ε)-positive membership Pr [114]	**B,D**	✓[11]	✓[11]	✗[17]	ε-BoDP $\subset$ (Θ, ε)-**PMPr**
$(\Theta, \varepsilon, \delta)$-adversarial Pr [139]	**F**	✓[11]	✓[11]	✗[17]	(ε, δ)-DP $\subset$ $(\Theta, \varepsilon, \delta)$-**AdvPr** $\prec$ (Θ, ε)-PMPr
(Θ, ε)-aposteriori noiseless Pr [14]	**B,D**	✓[11]	✓[11]	?	(Θ, ε)-**ANPr** $\sim$ (Θ, ε)-NPr
ε-semantic Pr [69, 96]	**F**	?	?	?	ε-**SemPr** $\sim$ ε-DP
$(\mathsf{Agg}, \varepsilon)$-zero-knowledge Pr [72]	**R**	✓[6]	✓[6]	?[22]	ε-DP $\prec$ $(\mathbf{Agg}, \varepsilon)$-**ZKPr**
$(\Theta, \Gamma, \varepsilon)$-coupled-worlds Pr [11]	**N,B,R**	✓[6]	✓[6]	✗[8]	(Θ, ε)-DistDP $\subset$ $(\Theta, \Gamma, \varepsilon)$-**CWPr**
$(\Theta, \Gamma, \varepsilon, \delta)$-inference-based CW Pr [11]	**Q,N,B,D,R**	?	?	✗[8]	$(\Theta, \Gamma, \varepsilon)$-CWPr $\prec$ $(\Theta, \Gamma, \varepsilon, \delta)$-**IBCWPr**
ε_κ-SIM-computational DP [129]	**C**	✓[12]	✓[12]	✓[23]	ε_κ-**SimCDP** $\prec$ ε-DP
ε_κ-IND-computational DP [129]	**C**	✓[12]	✓[12]	✓[23]	ε_κ-**IndCDP** $\prec$ ε_κ-SimCDP
$(\mathsf{Agg}, \varepsilon)$-computational ZK Pr [72]	**R,C**	✓[12]	✓[12]	?	$(\mathbf{Agg}, \varepsilon)$-**CZKPr** $\supset^{\prec}$ $(\mathsf{Agg}, \varepsilon)$-ZKPr

Table 4. Summary of variants/extensions of DP representing the main options in each combination of dimensions.

Notes

1. See Prop. 1.
2. See Prop. 2.
3. See Prop. 3.
4. See Prop. 4.
5. See Prop. 5.
6. See Prop. 6.
7. See Prop. 7.
8. See Prop. 8.
9. See Prop. 9.
10. See Prop. 10.
11. See Prop. 11.
12. See Prop. 12.
13. See Prop. 13.
14. See Prop. 14.
15. See Prop. 15.
16. See Prop. 16.
17. See Prop. 17.
18. Post-processing
19. Convexity
20. Composition
21. A modified definition was presented in [106] which is an instance of PF Pr.
22. A proof for a restricted scenario appears in the paper introducing the definition.
23. This claim appears in [128], its proof is in the unpublished full version.
24. Abbreviations for dimensions:
    - **Q**: Quantification of privacy loss
    - **N**: Neighborhood definition
    - **V**: Variation of privacy loss
    - **B**: Background knowledge
    - **F**: Formalism of privacy loss
    - **R**: Relativization of knowledge gain
    - **C**: Computational power

Appendix

A Proofs for the axioms

Many variants in **Q** satisfy both axioms, as shown by the following generic result.

Proposition 1. *All instantiations of Div DP satisfy both privacy axioms.*

Proof. The post-processing axiom follows directly from the monotonicity property of the f-divergence. The convexity axiom follows directly from the joint convexity property of the f-divergence. □

As a direct corollary, approximate DP, MI Pr, KL Pr, Rényi Pr, and zCo DP satisfy both axioms.

Proposition 2 (Th. 1&2 in [127], App. A in [19])**.** *Pro DP and mCo DP do not satisfy the post-processing axiom.*

Proposition 3. *Pro DP do not satisfy the convexity axiom.*

Proof. Consider the following mechanisms $\mathcal{M}_1$ and $\mathcal{M}_2$, with input and output in $\{0, 1\}$.

- $\mathcal{M}_1(0) = 0$, $\mathcal{M}_1(1) = 1$ with probability δ, and $\mathcal{M}_1(1) = 0$ with probability $1 - \delta$.
- $\mathcal{M}_2(0) = \mathcal{M}_2(1) = 1$.

Both mechanisms are $\left(\frac{1}{1-\delta}, \delta\right)$-Pro DP. Now, consider the mechanism $\mathcal{M}$ which applies $\mathcal{M}_1$ with probability $1 - 2\delta$ and $\mathcal{M}_2$ with probability 2δ. $\mathcal{M}$ is a convex combination of $\mathcal{M}_1$ and $\mathcal{M}_2$, but the reader can verify that it is not $\left(\frac{1}{1-\delta}, \delta\right)$-Pro DP. □

To show that definitions in **N**, **B** and **R** satisfy privacy axioms, we use a few generic results.

Proposition 4 (Th. 5.1 in [106])**.** *All instantiations of PF Pr satisfy both privacy axioms. As an immediate corollary, all definitions which combine notions in* ***N*** *and N Pr also satisfy both axioms.*

Proposition 5 (Sec. 2.1 [103])**.** *All instantiations of Gl DP satisfy both privacy axioms.*

Proposition 6 (Th. 3&4 in [11])**.** *All instantiations of CW Pr satisfy both privacy axioms. As an immediate corollary, Dist DP and ZK Pr also satisfy both axioms.*

Proposition 7 (Prop. 5 in [35])**.** *PPK-DP and APK-DP satisfy both privacy axioms.*

We can extend this result to definitions in **V** which simply vary the privacy parameters across inputs.

Proposition 8. *$d_{\mathcal{D}}$-Pr satisfies both privacy axioms. Further, EDiv DP also satisfies both privacy axioms.*

Proof. The proof of Proposition 4 (Appendix B in [106]) is a proof by case analysis on every possible protected property. The fact that ε is the same for every protected property has no influence on the proof, so we can directly adapt the proof to $d_{\mathcal{D}}$-Pr, and its combination with PF Pr. Similarly, the proof can be extended to arbitrary divergence functions, like in Proposition 1. □

However, for the definitions in **V** which only consider a random subset of dataset pairs, the post-processing axiom is satisfied but not the convexity axiom.

Proposition 9 (Lemma 2.4 in [10])**.** *Ran DP satisfies the post-processing axiom.*

Proposition 10. *Ran DP does not satisfy the convexity axiom.*

Proof. Let π be the uniform distribution on $\{0,1\}$, let D_1 be generated by picking 10 records according to π, and D_2 by flipping one record at random. Let $\mathcal{M}_0$ return 0 if all records are 0, and $\perp$ otherwise. Let $\mathcal{M}_1$ return 1 if all records are 1, and $\perp$ otherwise. Both mechanisms are $(\pi, 2^{-9}, 0)$-Ran DP: with probability $> 2 \cdot 2^{-10}$, neither D_1 nor D_2 have all their records set to 0 or 1, so the mechanism returns $\perp$, which does not leak anything. However, the mechanism $\mathcal{M}$ obtained by applying either $\mathcal{M}_1$ or $\mathcal{M}_2$ uniformly randomly does not satisfy $(\pi, 2^{-9}, 0)$-Ran DP: the ε-indistinguishability property does not hold if D_1 or D_2 have all their records set to *either* 0 or 1, which happens twice as often as either option alone. □

The definitions in **F** which change the shape of the prior-posterior bounds also satisfy both axioms.

Proposition 11. *All variants of M Pr, Adv Pr, and AN Pr satisfy both axioms.*

Proof. We prove it for PM Pr. A mechanism $\mathcal{M}$ satisfies (Θ, ε)-PM Pr if for all $t \in \mathcal{T}$, $\theta \in \Theta$, and $S \subseteq \mathcal{O}$, $\mathbb{P}_{D\sim\theta}[t \in D \mid \mathcal{M}(D) \in S] \leq e^{\varepsilon} \cdot \mathbb{P}_{D\sim\theta}[t \in D]$. We first prove that it satisfies the convexity axiom. Suppose $\mathcal{M}$ is a convex combination of $\mathcal{M}_1$ and $\mathcal{M}_2$. Simplifying $\mathbb{P}_{D\sim\theta}[\dots]$ into $\mathbb{P}[\dots]$, we have the following where $X_i = \mathbb{P}[\mathcal{M}(D) \in S \text{ and } \mathcal{M} = \mathcal{M}_i]$ for $i \in \{1,2\}$:

$$\begin{aligned}
&\mathbb{P}[t \in D \mid \mathcal{M}(D) \in S] \\
&= \frac{\mathbb{P}[t \in D \text{ and } \mathcal{M}(D) \in S \text{ and } \mathcal{M} = \mathcal{M}_1]}{\mathbb{P}[\mathcal{M}(D) \in S]} \\
&\quad + \frac{\mathbb{P}[t \in D \text{ and } \mathcal{M}(D) \in S \text{ and } \mathcal{M} = \mathcal{M}_2]}{\mathbb{P}[\mathcal{M}(D) \in S]} = \\
&\frac{X_1 \cdot \mathbb{P}[t \in D \mid \mathcal{M}_1(D) \in S]}{X_1 + X_2} + \frac{X_2 \cdot \mathbb{P}[t \in D \mid \mathcal{M}_2(D) \in S]}{X_1 + X_2} \\
&\leq \frac{X_1 (e^{\varepsilon} \cdot \mathbb{P}[t \in D])}{X_1 + X_2} + \frac{X_2 (e^{\varepsilon} \cdot \mathbb{P}[t \in D])}{X_1 + X_2} \leq e^{\varepsilon} \cdot \mathbb{P}[t \in D]
\end{aligned}$$

The proof for the post-processing axiom is similar, summing over all possible outputs $\mathcal{M}(D)$. It is straightforward to adapt the proof to all other definitions which change the shape of the prior-posterior bounds. □

Finally, CDP also satisfies both axioms.

Proposition 12. *Both versions of CDP satisfy both privacy axioms; where the post-processing axiom is modified to only allow post-processing with functions computable on a probabilistic polynomial time Turing machine. As a direct corollary of Proposition 6, CZK Pr also satisfies both privacy axioms.*

Proof. For Ind-CDP and the post-processing axiom, the proof is straightforward: if post-processing the output could break the ε-indistinguishability property, the attacker could do this on the original output and break the ε-indistinguishability property of the original definition.

For Ind-CDP and the convexity axiom, without loss of generality, we can assume that the sets of possible outputs of both mechanisms are disjoint (otherwise, this give strictly less information to the attacker).

The proof is then the same as for the post-processing axiom.

For Sim-CDP applying the same post-processing function to the "true" differentially private mechanism immediately leads to the result, since DP satisfies post-processing. The same reasoning holds for convexity. □

B Proofs for composition

In this section, if $\mathcal{M}_1$ and $\mathcal{M}_2$ are two mechanisms, we denote $\mathcal{M}_{1+2}$ the mechanism defined by $\mathcal{M}_{1+2}(D) = (\mathcal{M}_1(D), \mathcal{M}_2(D))$.

Almost all definitions in **Q** are composable; the only one for which we could not find any result in the literature is (f, ε)-Div DP.

Proposition 13. *If $\mathcal{M}_1$ and $\mathcal{M}_2$ are respectively...*

1. *$(\varepsilon_1, \delta_1)$-DP and $(\varepsilon_2, \delta_2)$-DP then $\mathcal{M}_{1+2}$ is $(\varepsilon_1 + \varepsilon_2, \delta_1 + \delta_2)$-DP.*
2. *$(\varepsilon_1, \delta_1)$-Pro DP and $(\varepsilon_2, \delta_2)$-Pro DP then $\mathcal{M}_{1+2}$ is $(\varepsilon_1 + \varepsilon_2, \delta_1 + \delta_2)$-Pro DP.*
3. *ε_1-MI DP and ε_2-MI DP then $\mathcal{M}_{1+2}$ is $(\varepsilon_1 + \varepsilon_2)$-MI DP.*
4. *ε_1-KL DP and ε_2-KL DP then $\mathcal{M}_{1+2}$ is $(\varepsilon_1 + \varepsilon_2)$-KL DP.*
5. *$(\alpha_1, \varepsilon_1)$-Rényi DP and $(\alpha_2, \varepsilon_2)$-Rényi DP then $\mathcal{M}_{1+2}$ is $(\max(\alpha_1, \alpha_2), \varepsilon_1 + \varepsilon_2)$-Rényi DP.*
6. *(μ_1, τ_1)-mCo DP and (μ_2, τ_2)-mCo DP then $\mathcal{M}_{1+2}$ is $\left(\mu_1 + \mu_2, \sqrt{{\mu_1}^2 + {\mu_2}^2}\right)$-mCo DP.*
7. *(ξ_1, ρ_1)-zCo DP and (ξ_2, ρ_2)-zCo DP then $\mathcal{M}_{1+2}$ is $(\xi_1 + \xi_2, \xi_1 + \xi_2)$-zCo DP.*

Further, all definitions that are combinations of definitions in **N** and **V** are composable.

Proposition 14. *If $\mathcal{M}_1$ is $d^1_{\mathcal{D}}$-Pr and $\mathcal{M}_2$ is $d^2_{\mathcal{D}}$-Pr, then $\mathcal{M}_{1+2}$ is $d^{1+2}_{\mathcal{D}}$-Pr, where $d^{1+2}_{\mathcal{D}}(D_1, D_2) = d^1_{\mathcal{D}}(D_1, D_2) + d^2_{\mathcal{D}}(D_1, D_2)$.*

Proof. The proof is essentially the same as for ε-DP. $\mathcal{M}_1$'s randomness is independent from $\mathcal{M}_2$'s, so:

$$
\begin{aligned}
&\mathbb{P}\left[\mathcal{M}_1(D_1) = O_1 \text{ and } \mathcal{M}_2(D_1) = O_2\right] \\
&= \mathbb{P}\left[\mathcal{M}_1(D_1) = O_1\right] \cdot \mathbb{P}\left[\mathcal{M}_2(D_1) = O_2\right] \\
&\leq e^{d^1_{\mathcal{D}}(D_1, D_2)} \cdot \mathbb{P}\left[\mathcal{M}_2(D_2) = O_1\right] \\
&\qquad \cdot e^{d^2_{\mathcal{D}}(D_1, D_2)} \cdot \mathbb{P}\left[\mathcal{M}_2(D_2) = O_2\right] \\
&\leq e^{d^{1+2}_{\mathcal{D}}(D_1, D_2)} \cdot \mathbb{P}\left[\mathcal{M}_1(D_2) = O_1 \text{ and } \mathcal{M}_2(D_2) = O_2\right].
\end{aligned}
$$

Each definition listed in Proposition 13 can also be combined with $d_{\mathcal{D}}$-Pr, and the composition proofs can be similarly adapted. □

Proposition 15 (Th. 4 in [36])**.** *If $\mathcal{M}_1$ is $(d_{\mathcal{D}}, \varepsilon_1, \delta_1)$-PsM DP and $\mathcal{M}_2$ is $(d_{\mathcal{D}}, \varepsilon_2, \delta_2)$-PsM DP then $\mathcal{M}_{1+2}$ is $(d_{\mathcal{D}}, \varepsilon_1 + \varepsilon_2, \delta_1 + \delta_2)$-PsM DP.*

Proposition 16 (Prop. 5.1.2 in [82])**.** *If $\mathcal{M}_1$ satisfies $(\pi, \gamma_1, \varepsilon_1)$-Ran DP and $\mathcal{M}_2$ satisfies $(\pi, \gamma_2, \varepsilon_2)$-Ran DP then $\mathcal{M}_{1+2}$ satisfies $(\pi, \gamma_1 + \gamma_2, \varepsilon_1 + \varepsilon_2)$-Ran DP.*

This can be also extended to variants in **Q**, see for example Lemma 2.5 in [10].

Proposition 17. *In general, definitions which assume limited background knowledge from the adversary do not compose.*

Proof. The proof of Proposition 14 cannot be adapted to a context in which the attacker has limited background knowledge: as the randomness partially comes from the data-generating distribution, the two probabilities are no longer independent. A typical example considers two mechanisms which answer e.g., queries "how many records satisfy property P" and "how many records satisfy property P and have an ID different from 4217": the randomness in the data might make each query private, but the combination of two queries trivially reveals something about a particular user. Variants of this proof can easily be obtained for all definitions with limited background knowledge. □

References

[1] Erfan Aghasian, Saurabh Garg, and James Montgomery. User's privacy in recommendation systems applying online social network data, a survey and taxonomy. *arXiv preprint arXiv:1806.07629*, 2018.

[2] Mohammad Alaggan, Sébastien Gambs, and Anne-Marie Kermarrec. Heterogeneous differential privacy. *arXiv preprint arXiv:1504.06998*, 2015.

[3] Joshua Allen, Bolin Ding, Janardhan Kulkarni, Harsha Nori, Olga Ohrimenko, and Sergey Yekhanin. An algorithmic framework for differentially private data analysis on trusted processors. *arXiv preprint arXiv:1807.00736*, 2018.

[4] Mário S Alvim, Konstantinos Chatzikokolakis, Catuscia Palamidessi, and Anna Pazii. Metric-based local differential privacy for statistical applications. *arXiv preprint arXiv:1805.01456*, 2018.

[5] Fredrik Andersson, John M Abowd, Matthew Graham, Jeremy Wu, and Lars Vilhuber. Formal privacy guarantees and analytical validity of onthemap public-use data.

https://ecommons.cornell.edu/handle/1813/47672, 2009.

[6] Miguel E Andrés, Nicolás E Bordenabe, Konstantinos Chatzikokolakis, and Catuscia Palamidessi. Geo-indistinguishability: Differential privacy for location-based systems. In *Proceedings of the 2013 ACM SIGSAC conference on Computer & communications security*. ACM, 2013.

[7] Hafiz Asif, Periklis A Papakonstantinou, and Jaideep Vaidya. How to accurately and privately identify anomalies. In *Proceedings of the 2019 ACM SIGSAC Conference on Computer and Communications Security*. ACM, 2019.

[8] Michael Backes, Aniket Kate, Sebastian Meiser, and Tim Ruffing. Differential indistinguishability for cryptography with (bounded) weak sources. *Grande Region Security and Reliability Day (GRSRD)*, 2014.

[9] Rina Foygel Barber and John C Duchi. Privacy and statistical risk: Formalisms and minimax bounds. *arXiv preprint arXiv:1412.4451*, 2014.

[10] Raef Bassily and Yoav Freund. Typicality-based stability and privacy. *arXiv preprint arXiv:1604.03336*, 2016.

[11] Raef Bassily, Adam Groce, Jonathan Katz, and Adam Smith. Coupled-worlds privacy: Exploiting adversarial uncertainty in statistical data privacy. In *Foundations of Computer Science (FOCS), 2013 IEEE 54th Annual Symposium on*. IEEE, 2013.

[12] Debabrota Basu, Christos Dimitrakakis, and Aristide Tossou. Differential privacy for multi-armed bandits: What is it and what is its cost? *arXiv preprint arXiv:1905.12298*, 2019.

[13] Johes Bater, Xi He, William Ehrich, Ashwin Machanavajjhala, and Jennie Rogers. Shrinkwrap: efficient sql query processing in differentially private data federations. *Proceedings of the VLDB Endowment*, 12(3):307–320, 2018.

[14] Raghav Bhaskar, Abhishek Bhowmick, Vipul Goyal, Srivatsan Laxman, and Abhradeep Thakurta. Noiseless database privacy. In *International Conference on the Theory and Application of Cryptology and Information Security*. Springer, 2011.

[15] Andrea Bittau, Úlfar Erlingsson, Petros Maniatis, Ilya Mironov, Ananth Raghunathan, David Lie, Mitch Rudominer, Ushasree Kode, Julien Tinnes, and Bernhard Seefeld. Prochlo: Strong privacy for analytics in the crowd. In *Proceedings of the 26th Symposium on Operating Systems Principles*. ACM, 2017.

[16] Daniel M Bittner, Anand D Sarwate, and Rebecca N Wright. Using noisy binary search for differentially private anomaly detection. In *International Symposium on Cyber Security Cryptography and Machine Learning*. Springer, 2018.

[17] Jeremiah Blocki, Avrim Blum, Anupam Datta, and Or Sheffet. Differentially private data analysis of social networks via restricted sensitivity. In *Proceedings of the 4th conference on Innovations in Theoretical Computer Science*, pages 87–96. ACM, 2013.

[18] Mark Bun, Cynthia Dwork, Guy N Rothblum, and Thomas Steinke. Composable and versatile privacy via truncated cdp. In *Proceedings of the 50th Annual ACM SIGACT Symposium on Theory of Computing*. ACM, 2018.

[19] Mark Bun and Thomas Steinke. Concentrated differential privacy: Simplifications, extensions, and lower bounds. In *Theory of Cryptography Conference*. Springer, 2016.

[20] Sébastien Canard and Baptiste Olivier. Differential privacy in distribution and instance-based noise mechanisms. *IACR Cryptology ePrint Archive*, 2015, 2015.

[21] TH Chan, Kai-Min Chung, Bruce M Maggs, and Elaine Shi. Foundations of differentially oblivious algorithms. In *Proceedings of the Thirtieth Annual ACM-SIAM Symposium on Discrete Algorithms*. Society for Industrial and Applied Mathematics, 2019.

[22] Konstantinos Chatzikokolakis, Miguel E Andrés, Nicolás Emilio Bordenabe, and Catuscia Palamidessi. Broadening the scope of differential privacy using metrics. In *International Symposium on Privacy Enhancing Technologies Symposium*. Springer, 2013.

[23] Kostantinos Chatzikokolakis, Ehab ElSalamouny, Catuscia Palamidessi, Pazii Anna, et al. Methods for location privacy: A comparative overview. *Foundations and Trends® in Privacy and Security*, 2017.

[24] Kamalika Chaudhuri, Jacob Imola, and Ashwin Machanavajjhala. Capacity bounded differential privacy. In *Advances in Neural Information Processing Systems*, 2019.

[25] Kamalika Chaudhuri and Nina Mishra. When random sampling preserves privacy. In *Annual International Cryptology Conference*. Springer, 2006.

[26] Rui Chen, Benjamin C Fung, Philip S Yu, and Bipin C Desai. Correlated network data publication via differential privacy. *The VLDB Journal—The International Journal on Very Large Data Bases*, 2014.

[27] Shixi Chen and Shuigeng Zhou. Recursive mechanism: towards node differential privacy and unrestricted joins. In *Proceedings of the 2013 ACM SIGMOD International Conference on Management of Data*. ACM, 2013.

[28] Zhili Chen, Xianyue Bao, Zuobin Ying, Ximeng Liu, and Hong Zhong. Differentially private location protection with continuous time stamps for vanets. In *International Conference on Algorithms and Architectures for Parallel Processing*. Springer, 2018.

[29] Chris Clifton and Tamir Tassa. On syntactic anonymity and differential privacy. In *2013 IEEE 29th International Conference on Data Engineering Workshops (ICDEW)*. IEEE, 2013.

[30] Léo Colisson. L3 internship report: Quantum analog of differential privacy in term of rényi divergence. *http://perso.ens-lyon.fr/omar.fawzi/docs/CollissonReport2016.pdf*, 2016.

[31] Paul Cuff and Lanqing Yu. Differential privacy as a mutual information constraint. In *Proceedings of the 2016 ACM SIGSAC Conference on Computer and Communications Security*. ACM, 2016.

[32] Rachel Cummings and David Durfee. Individual sensitivity preprocessing for data privacy. *arXiv preprint arXiv:1804.08645*, 2018.

[33] Tore Dalenius. Towards a methodology for statistical disclosure control. *statistik Tidskrift*, 1977.

[34] Fatemeh Deldar and Mahdi Abadi. Pldp-td: Personalized-location differentially private data analysis on trajectory databases. *Pervasive and Mobile Computing*, 2018.

[35] Damien Desfontaines, Esfandiar Mohammadi, Elisabeth Krahmer, and David Basin. Differential privacy with partial knowledge. *arXiv preprint arXiv:1905.00650*, 2019.

[36] Christos Dimitrakakis, Blaine Nelson, Aikaterini Mitrokotsa, Benjamin Rubinstein, et al. Bayesian differential privacy through posterior sampling. *arXiv preprint arXiv:1306.1066*, 2013.

[37] Bolin Ding, Janardhan Kulkarni, and Sergey Yekhanin. Collecting telemetry data privately. In *Advances in Neural Information Processing Systems*, 2017.

[38] Xuan Ding, Wei Wang, Meng Wan, and Ming Gu. Seamless privacy: Privacy-preserving subgraph counting in interactive social network analysis. In *Cyber-Enabled Distributed Computing and Knowledge Discovery (CyberC), 2013 International Conference on*. IEEE, 2013.

[39] Irit Dinur and Kobbi Nissim. Revealing information while preserving privacy. In *Proceedings of the twenty-second ACM SIGMOD-SIGACT-SIGART symposium on Principles of database systems*. ACM, 2003.

[40] Jinshou Dong, Aaron Roth, and J. Weijie Su. Gaussian differential privacy. *arXiv preprint arXiv:1905.02383*, 2019.

[41] Kai Dong, Taolin Guo, Haibo Ye, Xuansong Li, and Zhen Ling. On the limitations of existing notions of location privacy. *Future Generation Computer Systems*, 2018.

[42] Stelios Doudalis, Ios Kotsogiannis, Samuel Haney, Ashwin Machanavajjhala, and Sharad Mehrotra. One-sided differential privacy. *arXiv preprint arXiv:1712.05888*, 2017.

[43] Flávio du Pin Calmon and Nadia Fawaz. Privacy against statistical inference. In *Communication, Control, and Computing (Allerton), 2012 50th Annual Allerton Conference on*. IEEE, 2012.

[44] Yitao Duan. Privacy without noise. In *Proceedings of the 18th ACM conference on Information and knowledge management*. ACM, 2009.

[45] John C Duchi, Michael I Jordan, and Martin J Wainwright. Local privacy and statistical minimax rates. In *Foundations of Computer Science (FOCS), 2013 IEEE 54th Annual Symposium on*. IEEE, 2013.

[46] John C Duchi and Feng Ruan. The right complexity measure in locally private estimation: It is not the fisher information. *arXiv preprint arXiv:1806.05756*, 2018.

[47] David Durfee and Ryan Rogers. Practical differentially private top-k selection with pay-what-you-get composition. *arXiv preprint arXiv:1905.04273*, 2019.

[48] Cynthia Dwork. Differential privacy. In *Proceedings of the 33rd international conference on Automata, Languages and Programming*. ACM, 2006.

[49] Cynthia Dwork. Differential privacy: A survey of results. In *International Conference on Theory and Applications of Models of Computation*. Springer, 2008.

[50] Cynthia Dwork. The differential privacy frontier. In *Theory of Cryptography Conference*. Springer, 2009.

[51] Cynthia Dwork. Differential privacy in new settings. In *Proceedings of the twenty-first annual ACM-SIAM symposium on Discrete Algorithms*. SIAM, 2010.

[52] Cynthia Dwork, Krishnaram Kenthapadi, Frank McSherry, Ilya Mironov, and Moni Naor. Our data, ourselves: Privacy via distributed noise generation. In *Eurocrypt*. Springer, 2006.

[53] Cynthia Dwork and Frank McSherry. Differential data privacy. *United States, US7698250B2*, 2005.

[54] Cynthia Dwork, Frank McSherry, Kobbi Nissim, and Adam Smith. Calibrating noise to sensitivity in private data analysis. In *Theory of Cryptography Conference*. Springer, 2006.

[55] Cynthia Dwork, Moni Naor, Toniann Pitassi, and Guy N Rothblum. Differential privacy under continual observation. In *Proceedings of the forty-second ACM symposium on Theory of computing*. ACM, 2010.

[56] Cynthia Dwork, Moni Naor, Toniann Pitassi, Guy N Rothblum, and Sergey Yekhanin. Pan-private streaming algorithms. In *ICS*, 2010.

[57] Cynthia Dwork, Aaron Roth, et al. The algorithmic foundations of differential privacy. *Foundations and Trends® in Theoretical Computer Science*, 2014.

[58] Cynthia Dwork and Guy N Rothblum. Concentrated differential privacy. *arXiv preprint arXiv:1603.01887*, 2016.

[59] Hamid Ebadi, David Sands, and Gerardo Schneider. Differential privacy: Now it's getting personal. In *Acm Sigplan Notices*. ACM, 2015.

[60] Ehab ElSalamouny and Sébastien Gambs. Differential privacy models for location-based services. *Transactions on Data Privacy*, 2016.

[61] Úlfar Erlingsson, Vasyl Pihur, and Aleksandra Korolova. RAPPOR: Randomized aggregatable privacy-preserving ordinal response. In *Proceedings of the 2014 ACM SIGSAC conference on computer and communications security*. ACM, 2014.

[62] Alexandre Evfimievski, Johannes Gehrke, and Ramakrishnan Srikant. Limiting privacy breaches in privacy preserving data mining. In *Proceedings of the twenty-second ACM SIGMOD-SIGACT-SIGART symposium on Principles of database systems*. ACM, 2003.

[63] Chengfang Fang and Ee-Chien Chang. Differential privacy with delta-neighbourhood for spatial and dynamic datasets. In *Proceedings of the 9th ACM symposium on Information, computer and communications security*. ACM, 2014.

[64] Farhad Farokhi. Discounted differential privacy: Privacy of evolving datasets over an infinite horizon. *arXiv preprint arXiv:1908.03995*, 2019.

[65] Farhad Farokhi. Noiseless privacy. *arXiv preprint arXiv:1910.13027*, 2019.

[66] Vitaly Feldman, Ilya Mironov, Kunal Talwar, and Abhradeep Thakurta. Privacy amplification by iteration. In *2018 IEEE 59th Annual Symposium on Foundations of Computer Science (FOCS)*. IEEE, 2018.

[67] Vitaly Feldman and Thomas Steinke. Calibrating noise to variance in adaptive data analysis. *arXiv preprint arXiv:1712.07196*, 2017.

[68] Natasha Fernandes, Mark Dras, and Annabelle McIver. Generalised differential privacy for text document processing. In *International Conference on Principles of Security and Trust*. Springer, 2019.

[69] Srivatsava Ranjit Ganta, Shiva Prasad Kasiviswanathan, and Adam Smith. Composition attacks and auxiliary information in data privacy. In *Proceedings of the 14th ACM SIGKDD international conference on Knowledge discovery and data mining*. ACM, 2008.

[70] Simson L Garfinkel, John M Abowd, and Sarah Powazek. Issues encountered deploying differential privacy. In *Proceedings of the 2018 Workshop on Privacy in the Electronic Society*, pages 133–137. ACM, 2018.

[71] Johannes Gehrke, Michael Hay, Edward Lui, and Rafael Pass. Crowd-blending privacy. In *Advances in Cryptology–CRYPTO 2012*. Springer, 2012.

[72] Johannes Gehrke, Edward Lui, and Rafael Pass. Towards privacy for social networks: A zero-knowledge based definition of privacy. In *Theory of Cryptography Conference*. Springer, 2011.

[73] Joseph Geumlek and Kamalika Chaudhuri. Profile-based privacy for locally private computations. In *Proceedings of the 2019 IEEE International Symposium on Information Theory*. IEEE, 2019.

[74] Joseph Geumlek, Shuang Song, and Kamalika Chaudhuri. Renyi differential privacy mechanisms for posterior sampling. In *Advances in Neural Information Processing Systems*, 2017.

[75] Arpita Ghosh and Robert Kleinberg. Inferential privacy guarantees for differentially private mechanisms. *arXiv preprint arXiv:1603.01508*, 2016.

[76] Arpita Ghosh and Aaron Roth. Selling privacy at auction. *Games and Economic Behavior*, 2015.

[77] Shafi Goldwasser and Silvio Micali. Probabilistic encryption. *Journal of computer and system sciences*, 1984.

[78] Shafi Goldwasser, Silvio Micali, and Charles Rackoff. The knowledge complexity of interactive proof systems. *SIAM Journal on computing*, 18(1):186–208, 1989.

[79] Adam Groce, Jonathan Katz, and Arkady Yerukhimovich. Limits of computational differential privacy in the client/server setting. In *Theory of Cryptography Conference*, pages 417–431. Springer, 2011.

[80] Rachid Guerraoui, Anne-Marie Kermarrec, Rhicheek Patra, and Mahsa Taziki. D 2 p: distance-based differential privacy in recommenders. *Proceedings of the VLDB Endowment*, 2015.

[81] Mehmet Emre Gursoy, Acar Tamersoy, Stacey Truex, Wenqi Wei, and Ling Liu. Secure and utility-aware data collection with condensed local differential privacy. *arXiv preprint arXiv:1905.06361*, 2019.

[82] Rob Hall et al. *New Statistical Applications for Differential Privacy*. PhD thesis, PhD thesis, Carnegie Mellon, 2012.

[83] Rob Hall, Alessandro Rinaldo, and Larry Wasserman. Random differential privacy. *arXiv preprint arXiv:1112.2680*, 2011.

[84] Samuel Haney, Ashwin Machanavajjhala, and Bolin Ding. Design of policy-aware differentially private algorithms. *Proceedings of the VLDB Endowment*, 2015.

[85] Michael Hay, Chao Li, Gerome Miklau, and David Jensen. Accurate estimation of the degree distribution of private networks. In *Data Mining, 2009. ICDM'09. Ninth IEEE International Conference on*. IEEE, 2009.

[86] Xi He, Ashwin Machanavajjhala, and Bolin Ding. Blowfish privacy: Tuning privacy-utility trade-offs using policies. In *Proceedings of the 2014 ACM SIGMOD international conference on Management of data*. ACM, 2014.

[87] Xi He, Ashwin Machanavajjhala, Cheryl Flynn, and Divesh Srivastava. Composing differential privacy and secure computation: A case study on scaling private record linkage. In *Proceedings of the 2017 ACM SIGSAC Conference on Computer and Communications Security*. ACM, 2017.

[88] Johannes Heurix, Peter Zimmermann, Thomas Neubauer, and Stefan Fenz. A taxonomy for privacy enhancing technologies. *Computers & Security*, 2015.

[89] Naoise Holohan, Spiros Antonatos, Stefano Braghin, and Pól Mac Aonghusa. (k,e)-anonymity: k-anonymity with e-differential privacy. *arXiv preprint arXiv:1710.01615*, 2017.

[90] Márk Jelasity and Kenneth P Birman. Distributional differential privacy for large-scale smart metering. In *Proceedings of the 2nd ACM workshop on Information hiding and multimedia security*. ACM, 2014.

[91] Bo Jiang, Ming Li, and Ravi Tandon. Context-aware data aggregation with localized information privacy. In *2018 IEEE Conference on Communications and Network Security (CNS)*. IEEE, 2018.

[92] Noah Johnson, Joseph P Near, and Dawn Song. Towards practical differential privacy for sql queries. *Proceedings of the VLDB Endowment*, 2018.

[93] Austin Jones, Kevin Leahy, and Matthew Hale. Towards differential privacy for symbolic systems. In *2019 American Control Conference (ACC)*. IEEE, 2019.

[94] Zach Jorgensen, Ting Yu, and Graham Cormode. Conservative or liberal? personalized differential privacy. In *Data Engineering (ICDE), 2015 IEEE 31st International Conference on*. IEEE, 2015.

[95] Peter Kairouz, Sewoong Oh, and Pramod Viswanath. The composition theorem for differential privacy. *IEEE Transactions on Information Theory*, 2017.

[96] Shiva P Kasiviswanathan and Adam Smith. On the 'semantics' of differential privacy: A bayesian formulation. *Journal of Privacy and Confidentiality*, 6(1), 2014.

[97] Yusuke Kawamoto and Takao Murakami. Local distribution obfuscation via probability coupling. In *2019 57th Annual Allerton Conference on Communication, Control, and Computing (Allerton)*. IEEE, 2019.

[98] Yusuke Kawamoto and Takao Murakami. Local obfuscation mechanisms for hiding probability distributions. In *European Symposium on Research in Computer Security*. Springer, 2019.

[99] Michael Kearns, Mallesh Pai, Aaron Roth, and Jonathan Ullman. Mechanism design in large games: Incentives and privacy. In *Proceedings of the 5th conference on Innovations in theoretical computer science*. ACM, 2014.

[100] Michael Kearns, Aaron Roth, Zhiwei Steven Wu, and Grigory Yaroslavtsev. Private algorithms for the protected in social network search. *Proceedings of the National Academy of Sciences*, 2016.

[101] Georgios Kellaris, George Kollios, Kobbi Nissim, and Adam O'Neill. Accessing data while preserving privacy. *arXiv preprint arXiv:1706.01552*, 2017.

[102] Georgios Kellaris, Stavros Papadopoulos, Xiaokui Xiao, and Dimitris Papadias. Differentially private event sequences over infinite streams. *Proceedings of the VLDB Endowment*, 2014.

[103] Daniel Kifer and Bing-Rong Lin. Towards an axiomatization of statistical privacy and utility. In *Proceedings of the twenty-ninth ACM SIGMOD-SIGACT-SIGART sym-*

posium on Principles of database systems. ACM, 2010.

[104] Daniel Kifer and Bing-Rong Lin. An axiomatic view of statistical privacy and utility. *Journal of Privacy and Confidentiality*, 2012.

[105] Daniel Kifer and Ashwin Machanavajjhala. No free lunch in data privacy. In *Proceedings of the 2011 ACM SIGMOD International Conference on Management of data*. ACM, 2011.

[106] Daniel Kifer and Ashwin Machanavajjhala. A rigorous and customizable framework for privacy. In *Proceedings of the 31st ACM SIGMOD-SIGACT-SIGAI symposium on Principles of Database Systems*. ACM, 2012.

[107] Sara Krehbiel. Choosing epsilon for privacy as a service. *Proceedings on Privacy Enhancing Technologies*, 2019.

[108] Vishaal Krishnan and Sonia Martínez. A distributional framework for moving-horizon estimation: Stability and privacy guarantees. *arXiv preprint arXiv:1812.09672*, 2018.

[109] Peeter Laud, Alisa Pankova, and Pettai Martin. Achieving differential privacy using methods from calculus. *arXiv preprint arXiv:1811.06343*, 2018.

[110] Jaewoo Lee and Chris Clifton. Differential identifiability. In *Proceedings of the 18th ACM SIGKDD international conference on Knowledge discovery and data mining*. ACM, 2012.

[111] Jaewoo Lee and Daniel Kifer. Concentrated differentially private gradient descent with adaptive per-iteration privacy budget. In *Proceedings of the 24th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining*. ACM, 2018.

[112] Samantha Leung and Edward Lui. Bayesian mechanism design with efficiency, privacy, and approximate truthfulness. In *International Workshop on Internet and Network Economics*. Springer, 2012.

[113] Ninghui Li, Tiancheng Li, and Suresh Venkatasubramanian. t-closeness: Privacy beyond k-anonymity and l-diversity. In *Data Engineering, 2007. ICDE 2007. IEEE 23rd International Conference on*. IEEE, 2007.

[114] Ninghui Li, Wahbeh Qardaji, Dong Su, Yi Wu, and Weining Yang. Membership privacy: a unifying framework for privacy definitions. In *Proceedings of the 2013 ACM SIGSAC conference on Computer & communications security*. ACM, 2013.

[115] Ninghui Li, Wahbeh H Qardaji, and Dong Su. Provably private data anonymization: Or, k-anonymity meets differential privacy. *CoRR, abs/1101.2604*, 49:55, 2011.

[116] Changchang Liu, Supriyo Chakraborty, and Prateek Mittal. Dependence makes you vulnberable: Differential privacy under dependent tuples. In *NDSS*, 2016.

[117] Jinfei Liu, Li Xiong, and Jun Luo. Semantic security: Privacy definitions revisited. *Trans. Data Privacy*, 2013.

[118] Ziqi Liu, Yu-Xiang Wang, and Alexander Smola. Fast differentially private matrix factorization. In *Proceedings of the 9th ACM Conference on Recommender Systems*. ACM, 2015.

[119] Yunhui Long, Vincent Bindschaedler, and Carl A Gunter. Towards measuring membership privacy. *arXiv preprint arXiv:1712.09136*, 2017.

[120] Edward Lui and Rafael Pass. Outlier privacy. In *Theory of Cryptography Conference*. Springer, 2015.

[121] Ashwin Machanavajjhala, Johannes Gehrke, and Michaela Götz. Data publishing against realistic adversaries. *Proceedings of the VLDB Endowment*, 2009.

[122] Ashwin Machanavajjhala, Johannes Gehrke, Daniel Kifer, and Muthuramakrishnan Venkitasubramaniam. l-diversity: Privacy beyond k-anonymity. In *Data Engineering, 2006. ICDE'06. Proceedings of the 22nd International Conference on*. IEEE, 2006.

[123] Ashwin Machanavajjhala and Xi He. Analyzing your location data with provable privacy guarantees. In *Handbook of Mobile Data Privacy*. Springer, 2018.

[124] Ashwin Machanavajjhala, Daniel Kifer, John Abowd, Johannes Gehrke, and Lars Vilhuber. Privacy: Theory meets practice on the map. In *Proceedings of the 2008 IEEE 24th International Conference on Data Engineering*. IEEE Computer Society, 2008.

[125] David R McClure. *Relaxations of differential privacy and risk/utility evaluations of synthetic data and fidelity measures*. PhD thesis, Duke University, 2015.

[126] H Brendan McMahan, Daniel Ramage, Kunal Talwar, and Li Zhang. Learning differentially private recurrent language models. *arXiv preprint arXiv:1710.06963*, 2017.

[127] Sebastian Meiser. Approximate and probabilistic differential privacy definitions. *Cryptology ePrint Archive, Report 2018/277*, 2018.

[128] Ilya Mironov. Renyi differential privacy. In *Computer Security Foundations Symposium (CSF), 2017 IEEE 30th*. IEEE, 2017.

[129] Ilya Mironov, Omkant Pandey, Omer Reingold, and Salil Vadhan. Computational differential privacy. In *Advances in Cryptology-CRYPTO 2009*. Springer, 2009.

[130] Takao Murakami and Yusuke Kawamoto. Utility-optimized local differential privacy mechanisms for distribution estimation. In *28th USENIX Security Symposium)*, pages 1877–1894, 2019.

[131] Boel Nelson and Jenni Reuben. Chasing accuracy and privacy, and catching both: A literature survey on differentially private histogram publication. *arXiv*, 2019.

[132] Yiwen Nie, Wei Yang, Liusheng Huang, Xike Xie, Zhenhua Zhao, and Shaowei Wang. A utility-optimized framework for personalized private histogram estimation. *IEEE Transactions on Knowledge and Data Engineering*, 2018.

[133] Nadia Niknami, Mahdi Abadi, and Fatemeh Deldar. Spatialpdp: A personalized differentially private mechanism for range counting queries over spatial databases. In *Computer and Knowledge Engineering (ICCKE), 2014 4th International eConference on*. IEEE, 2014.

[134] Kobbi Nissim, Sofya Raskhodnikova, and Adam Smith. Smooth sensitivity and sampling in private data analysis. In *Proceedings of the thirty-ninth annual ACM symposium on Theory of computing*, pages 75–84. ACM, 2007.

[135] Sarvar Patel, Giuseppe Persiano, and Kevin Yeo. What storage access privacy is achievable with small overhead? *arXiv preprint arXiv:1904.05452*, 2019.

[136] Rafael Pinot. Minimum spanning tree release under differential privacy constraints. *arXiv preprint arXiv:1801.06423*, 2018.

[137] Rafael Pinot, Florian Yger, Cédric Gouy-Pailler, and Jamal Atif. A unified view on differential privacy and robustness to adversarial examples. *arXiv preprint arXiv:1906.07982*,

2019.

[138] Davide Proserpio, Sharon Goldberg, and Frank McSherry. Calibrating data to sensitivity in private data analysis: a platform for differentially-private analysis of weighted datasets. *Proceedings of the VLDB Endowment*, 2014.

[139] Vibhor Rastogi, Michael Hay, Gerome Miklau, and Dan Suciu. Relationship privacy: output perturbation for queries with joins. In *Proceedings of the twenty-eighth ACM SIGMOD-SIGACT-SIGART symposium on Principles of database systems*. ACM, 2009.

[140] Jenni Reuben. Towards a differential privacy theory for edge-labeled directed graphs. *SICHERHEIT 2018*, 2018.

[141] Aaron Roth. New algorithms for preserving differential privacy. *Microsoft Research*, 2010.

[142] Benjamin IP Rubinstein and Francesco Aldà. Pain-free random differential privacy with sensitivity sampling. In *Proceedings of the 34th International Conference on Machine Learning-Volume 70*. JMLR. org, 2017.

[143] Pierangela Samarati. Protecting respondents identities in microdata release. *IEEE transactions on Knowledge and Data Engineering*, 2001.

[144] Adam Sealfon. Shortest paths and distances with differential privacy. In *Proceedings of the 35th ACM SIGMOD-SIGACT-SIGAI Symposium on Principles of Database Systems*. ACM, 2016.

[145] Goldwasser Shafi and Silvio Micali. Probabilistic encryption. *Journal of computer and system sciences*, 1984.

[146] Elaine Shi, HTH Chan, Eleanor Rieffel, Richard Chow, and Dawn Song. Privacy-preserving aggregation of time-series data. In *Annual Network & Distributed System Security Symposium (NDSS)*. Internet Society., 2011.

[147] Sean Simmons, Cenk Sahinalp, and Bonnie Berger. Enabling privacy-preserving gwass in heterogeneous human populations. *Cell systems*, 2016.

[148] David M Sommer, Sebastian Meiser, and Esfandiar Mohammadi. Privacy loss classes: The central limit theorem in differential privacy. *Proceedings on Privacy Enhancing Technologies*, 2019.

[149] Jordi Soria-Comas, Josep Domingo-Ferrer, David Sánchez, and David Megías. Individual differential privacy: A utility-preserving formulation of differential privacy guarantees. *IEEE Transactions on Information Forensics and Security*, 2017.

[150] Klara Stokes and Vicenç Torra. n-confusion: a generalization of k-anonymity. In *Proceedings of the 2012 Joint EDBT/ICDT Workshops*. ACM, 2012.

[151] Haipei Sun, Xiaokui Xiao, Issa Khalil, Yin Yang, Zhan Qin, Hui Wendy Wang, and Ting Yu. Analyzing subgraph statistics from extended local views with decentralized differential privacy. In *Proceedings of the 2019 ACM SIGSAC Conference on Computer and Communications Security*. ACM, 2019.

[152] Latanya Sweeney. k-anonymity: A model for protecting privacy. *International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems*, 2002.

[153] Christine Task and Chris Clifton. A guide to differential privacy theory in social network analysis. In *Proceedings of the 2012 International Conference on Advances in Social Networks Analysis and Mining (ASONAM 2012)*. IEEE Computer Society, 2012.

[154] Differential Privacy Team. Learning with privacy at scale, 2016.

[155] Raphael R Toledo, George Danezis, and Ian Goldberg. Lower-cost e-private information retrieval. *Proceedings on Privacy Enhancing Technologies*, 2016.

[156] Aristide CY Tossou and Christos Dimitrakakis. Algorithms for differentially private multi-armed bandits. In *Thirtieth AAAI Conference on Artificial Intelligence*, 2016.

[157] Michael Carl Tschantz, Shayak Sen, and Anupam Datta. Differential privacy as a causal property. *arXiv preprint arXiv:1710.05899*, 2017.

[158] Sameer Wagh, Paul Cuff, and Prateek Mittal. Differentially private oblivious ram. *Proceedings on Privacy Enhancing Technologies*, 2018.

[159] Isabel Wagner and David Eckhoff. Technical privacy metrics: a systematic survey. *ACM Computing Surveys (CSUR)*, 2018.

[160] Weina Wang, Lei Ying, and Junshan Zhang. On the tradeoff between privacy and distortion in differential privacy. *CoRR, vol. abs/1402.3757*, 2014.

[161] Weina Wang, Lei Ying, and Junshan Zhang. On the relation between identifiability, differential privacy, and mutual-information privacy. *IEEE Transactions on Information Theory*, 2016.

[162] Yu-Xiang Wang. Per-instance differential privacy and the adaptivity of posterior sampling in linear and ridge regression. *arXiv preprint arXiv:1707.07708*, 2017.

[163] Yu-Xiang Wang, Borja Balle, and Shiva Kasiviswanathan. Subsampled rényi differential privacy and analytical moments accountant. *arXiv preprint arXiv:1808.00087*, 2018.

[164] Yu-Xiang Wang, Jing Lei, and Stephen E Fienberg. On-average kl-privacy and its equivalence to generalization for max-entropy mechanisms. In *International Conference on Privacy in Statistical Databases*. Springer, 2016.

[165] Stanley L Warner. Randomized response: A survey technique for eliminating evasive answer bias. *Journal of the American Statistical Association*, 1965.

[166] Larry Wasserman and Shuheng Zhou. A statistical framework for differential privacy. *Journal of the American Statistical Association*, 2010.

[167] Genqiang Wu, Yeping He, Jingzheng Wu, and Xianyao Xia. Inherit differential privacy in distributed setting: Multiparty randomized function computation. In *Trustcom/BigDataSE/I SPA, 2016 IEEE*. IEEE, 2016.

[168] Genqiang Wu, Xianyao Xia, and Yeping He. Extending differential privacy for treating dependent records via information theory. *arXiv preprint arXiv:1703.07474*, 2017.

[169] Xiaotong Wu, Wanchun Dou, and Qiang Ni. Game theory based privacy preserving analysis in correlated data publication. In *Proceedings of the Australasian Computer Science Week Multiconference*. ACM, 2017.

[170] Xiaotong Wu, Taotao Wu, Maqbool Khan, Qiang Ni, and Wanchun Dou. Game theory based correlated privacy preserving analysis in big data. *IEEE Transactions on Big Data*, 2017.

[171] Yonghui Xiao and Li Xiong. Protecting locations with differential privacy under temporal correlations. In *Proceedings of the 22nd ACM SIGSAC Conference on Computer and Communications Security*. ACM, 2015.

[172] Ziqi Yan, Jiqiang Liu, Gang Li, Zhen Han, and Shuo Qiu. Privmin: Differentially private minhash for jaccard similarity computation. *arXiv preprint arXiv:1705.07258*, 2017.

[173] Bin Yang, Issei Sato, and Hiroshi Nakagawa. Bayesian differential privacy on correlated data. In *Proceedings of the 2015 ACM SIGMOD international conference on Management of Data*. ACM, 2015.

[174] Xiaowei Ying, Xintao Wu, and Yue Wang. On linear refinement of differential privacy-preserving query answering. In *Pacific-Asia Conference on Knowledge Discovery and Data Mining*. Springer, 2013.

[175] Jinxue Zhang, Jingchao Sun, Rui Zhang, Yanchao Zhang, and Xia Hu. Privacy-preserving social media data outsourcing. In *IEEE INFOCOM 2018-IEEE Conference on Computer Communications*. IEEE, 2018.

[176] Zijian Zhang, Zhan Qin, Liehuang Zhu, Wei Jiang, Chen Xu, and Kui Ren. Toward practical differential privacy in smart grid with capacity-limited rechargeable batteries. *arXiv preprint arXiv:1507.03000*, 2015.

[177] Shuheng Zhou, Katrina Ligett, and Larry Wasserman. Differential privacy with compression. In *Information Theory, 2009. ISIT 2009. IEEE International Symposium on*. IEEE, 2009.

[178] Tianqing Zhu, Gang Li, Yongli Ren, Wanlei Zhou, and Ping Xiong. Differential privacy for neighborhood-based collaborative filtering. In *Proceedings of the 2013 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining*, pages 752–759. ACM, 2013.

[179] Tianqing Zhu, Ping Xiong, Gang Li, and Wanlei Zhou. Correlated differential privacy: hiding information in non-iid data set. *IEEE Transactions on Information Forensics and Security*, 10(2):229–242, 2015.