DEPARTMENT OF INFORMATICS

TECHNISCHE UNIVERSITÄT MÜNCHEN

Master's Thesis in Information Systems

Supporting diverse stakeholders to make informed decisions about the use of differential privacy with a web-based e-learning application

Marcus Land

DEPARTMENT OF INFORMATICS

TECHNISCHE UNIVERSITÄT MÜNCHEN

Master's Thesis in Information Systems

Supporting diverse stakeholders to make informed decisions about the use of differential privacy with a web-based e-learning application

Unterstützung verschiedener Stakeholder im Treffen fundierter Entscheidungen zur Nutzung von Differential Privacy mittels einer web-basierten e-Learning Anwendung

Author:	Marcus Land
Supervisor:	Prof. Dr. Florian Matthes
Advisor:	Sascha Nägele, M. Sc.
Submission Date:	15.02.2021

I confirm that this master's thesis in information systems is my own work and I have documented all sources and material used.

Munich, 15.02.2021 Marcus Land

Acknowledgments

I would like to express my sincere gratitude to my advisor Sascha Nägele, for his great support throughout this thesis. I would also like to thank his colleagues Alexandra Klymenko and Gonzalo Munilla Garrido, who showed great interest in my work and always made me feel as part of the team.

I would also like to thank Prof. Dr. Florian Matthes for the opportunity to write this thesis at his chair and for his valuable feedback.

Finally, I would like to thank my friends and family for their ongoing support.

Abstract

Privacy has long been overshadowed by security but is becoming increasingly important as ever more data is collected, shared and analyzed. This has also been noticed by legislators, leading to new privacy laws like the GDPR. It has become evident that cybersecurity measures such as authentication and encryption are insufficient to ensure privacy. There is a need for dedicated privacy-preserving methods. However, recent research has shown that traditional approaches to privacy are deeply flawed, as they lack provable guarantees. This is addressed by differential privacy which emerged in 2006. Although it is regarded as a gold standard in privacy by research, a large part of the industry has not adapted it yet. The only exception being key players like Google, Microsoft or the US Census Bureau. Other and especially smaller companies seem to be hesitant in adopting it.

In this thesis, we want to investigate the reasons behind this imbalance. We evaluate the challenges and benefits and overall maturity of differential privacy to close the gap between research and practice. Even tough there are many tools available to facilitate the use of differential privacy, only few address the question whether differential privacy should be used in the first place. We want to guide practitioners in these earlier stages of their decision-making process. To do so, we look into the current state of differential privacy in the industry by studying the grey literature and conducting expert interviews. Based on these findings, we design learning material to support practitioners in making informed decisions about the use of differential privacy.

Kurzfassung

Privacy stand lange Zeit im Schatten der klassischen Security, wird aber zunehmend wichtiger, da immer mehr Daten gesammelt, geteilt und analysiert werden. Dies wurde auch von den Gesetzgebern bemerkt und führte zu neuen Datenschutzgesetzen wie der GDPR. Es hat sich gezeigt, dass Cybersecurity-Maßnahmen wie Authentifizierung und Verschlüsselung nicht ausreichen, um die Privatsphäre zu schützen. Es besteht ein Bedarf an spezifischen Methoden zur Wahrung der Privatsphäre. Jüngste Forschungen haben jedoch gezeigt, dass herkömmliche Ansätze zur Wahrung der Privatsphäre Mängel aufweisen, da sie keine nachweisbaren Garantien bieten. Diesem Problem widmet sich Differential Privacy, eine Definition, die im Jahr 2006 aufkam. Obwohl sie von der Forschung bereits als Goldstandard für Privacy angesehen wird, wird sie in der Industrie kaum angewandt. Während zwar große Organisationen wie Google, Microsoft oder das US Census Bureau Differential Privacy bereits nutzen, scheinen kleinere Unternehmen eher zögerlich zu sein.

Wir wollen bestehende Hindernisse identifizieren und Unternehmen einen besseren Zugang zu Differential Privacy ermöglichen. Hierbei untersuchen wir die Herausforderungen und Vorteile sowie den allgemeinen Reifegrad von Differential Privacy. Obwohl es bereits viele Tools gibt, die den Einsatz von Differential Privacy erleichtern, beschäftigen sich nur wenige mit der Frage, ob Differential Privacy überhaupt verwendet werden sollte. Wir wollen Praxisanwender in diesen frühen Phasen ihres Entscheidungsprozesses unterstützen. Dafür betrachten wir den aktuellen Stand von Differential Privacy in der Praxis, indem wir die graue Literatur studieren und Experteninterviews durchführen. Basierend auf diesen Ergebnissen erstellen wir Lernmaterial, um Praxisanwender bei ihren Entscheidungen bezüglich Differential Privacy zu unterstützen.

Contents

1 Introduction

This chapter motivates this thesis and states the research objectives and research approach of this thesis

1.1 Motivation

Privacy has long been overshadowed by security but is becoming increasingly important as ever more data is collected, shared, and analyzed. This has also been noticed by legislators, leading to new privacy laws like the GDPR. It has become evident that cybersecurity measures such as authentication and encryption are insufficient to ensure privacy. There is a need for dedicated privacy-preserving methods. However, recent research has shown that traditional approaches to privacy are deeply flawed. As soon as the supposedly anonymized data or even statistics about it are released, conclusions can be drawn about individuals in said data.

An example of this is the Netflix Prize. A competition announced by Netflix in 2006, offering one million dollars to those who managed to improve its current recommendation algorithm, "Cinematch," by 10%. As a training dataset, they released the data of nearly half a million of their users. Any personal information was removed. All that remained was which films the users had seen and when and how they had rated them. Nonetheless, researchers Narayanan and Shmatikov were able to re-identify users by comparing the data with that of the internet movie database IMDb. It turned out that already 99% of all users could be uniquely identified based on eight of their ratings from which two could even be completely wrong [1][2].
Similarly spectacular were Latanya Sweeney's findings, who managed to identify individuals in a medical dataset based on their date of birth, gender, and zip code. She further showed that for 87% of the US population, the combination of these three values was a unique identifier [3]. For aggregate statistics, the US Census Bureau demonstrated that based on the statistics it released in 2010, it could reconstruct the records of all 308 million Americans that took part in the US Census. The reconstructed records exactly matched those of 142 million Americans (46% of the population), 50 million of which could be re-identified by linking the reconstructed records to commercial datasets. [4][5]
Dinur and Nissim further proved that a shockingly small number of queries to a database are sufficient to reconstruct it almost completely, even if the answers to those queries are perturbed [6]. Their results raised the question of how much perturbation is required to prevent such disastrous results, which ultimately led to the definition of differential privacy.

1.1.1 Differential Privacy

Differential privacy was introduced in 2006 by Cynthia Dwork [7]. It is a definition that has to be satisfied by a mechanism releasing information about a dataset. It guarantees every individual within that dataset that nothing specific to them can be learned through such a data release. This guarantee can be mathematically proven and holds under any circumstances. So no matter what an attacker does to the output of a differentially private mechanism or which additional information he possesses, he cannot weaken the guarantee. Therefore, differential privacy protects against all the aforementioned attacks. Because of its strong guarantee and unique properties, differential privacy has already been declared the gold standard of privacy protection. But even though it is promoted by major organizations like Google, Microsoft, or the US Census Bureau, smaller companies seem to be hesitant in adopting it. There are already attempts to facilitate the use of differential privacy, offering tools to use differential privacy [8][9][10][11][12][13] or those that help in implementing it [14][15][16]. But only few tackle the question whether differential privacy should be used in the first place. We want to guide stakeholders in these earlier stages of their decision-making process.

1.2 Research Objectives

As the title of this thesis implies, our goal is threefold. First, we want to support practitioners in making informed decisions about the use of differential privacy. Second, this support should be tailored to the stakeholders, meaning their role and prior knowledge. Third, this is achieved with a web-based e-learning application.
The first aspect is broad and encompasses several questions practitioners could have regarding the use of differential privacy. The most important one being whether they should use differential privacy in the first place. Why should they use differential privacy? Is it even necessary in their use case? What are the advantages and disadvantages of using it? What are the possible alternatives? Moreover, they might want to know if they can even apply differential privacy to their use case and, if they do so, in which form? There are approximately 225 variations and extensions of differential privacy [17], the local and the global model, and different mechanisms to achieve differential privacy. All of these questions demand answers. And none of them have definitive answers. Thus, we want to provide practitioners with guidance and the necessary knowledge to answer them for themselves, i.e., make informed decisions. This leads us to our first research question.

> **Research question 1:** What are benefits and challenges of using differential privacy?

First, we wanted to investigate the question whether differential privacy should be used. What are the benefits of using it, and which challenges does it entail. Are there clear limitations when it can be used, and is it even necessary? These findings should give practitioners an understanding and will also be part of the learning nuggets. This question is answered in chapter 5.

The tricky part is that we cannot ask the target audience what they need because they do not know about differential privacy. Instead, we have to ask those that already acquired the necessary knowledge. We answer this research question by conducting a multivocal literature review. It differs from a regular literature review in that it also includes literature that has not formally been published. This is important to capture voices out of the industry by including their blogs and documentations. This is further motivated in the research approach. It also allows us to collect and compare existing learning materials on differential privacy, such as lecture slides. Additionally, we conduct interviews with practitioners to explore important themes.
After exploring whether differential privacy is worth investigating, we want to assess what are the most important topics practitioners should know about.

> **Research question 2:** What are the most important topics practitioners should know about?

This question is answered on the one hand through expert interviews and on the other hand by looking into existing learning material on differential privacy and assessing popular topics.

> **Research question 3:** How can these topics be conveyed?

For this last question, the findings of the previous two are consolidated. It mainly revolves around the design of the learning content. To answer it, learning nuggets for differential privacy are created and evaluated.

1.3 Research Approach

The research approach of this thesis is aligned with the design science research paradigm. It was introduced by Hevner et. al. in the field of information systems [18]. Peffers et. al. presented a methodology, which incorporates practices and procedures required to carry out such research [19]. Design science research is a pragmatic and outcome-based research approach. It revolves around the creation and evaluation of artifacts. Such a design research artifact can be any designed object in which a research contribution is embedded in the design. In the case of this thesis, the artifact is a set of learning units about core aspects of differential privacy.

1.4 Outline

In the next two chapters, we will explain the foundations of differential privacy and discuss related work. In chapter 4 we summarize the findings of the interviews we conducted, which will be the basis for chapter 5, in which we discuss the first research question, the benefits, and challenges of differential privacy. Chapter 6 gives a broad overview of the learning nuggets and the learning platform we conceptualized. Chapter 7 describes the evaluation we carried out and paves the way for the conclusion and future work discussed in chapter 8

2 Foundations

In this chapter, we will explain the basics of differential privacy required to understand the next chapters.

2.1 K-Anonymity

In 1998, one year after she famously re-identified the governor of Massachusetts, Latanya Sweeney introduced k-anonymity together with Pierangela Samarati [20]. It is nowadays among the most popular and well known privacy-preserving methods. Re-examining her successful attack, Sweeney found that removing direct identifiers like name, address, or social security number was not enough to prevent re-identification [3]. The combination of so-called quasi-identifiers such as gender, age, and zip code was also unique to most individuals and could therefore be used to identify them. K-anonymity addresses this vulnerability by ensuring that for no record, the combination of its quasi-identifiers is unique.

K-Anonymity A dataset is k-anonymous if every record in it is indistinguishable from at least $k-1$ other records with respect to its quasi-identifiers. Such a group of k indistinguishable records is called an equivalence class. To ensure that every equivalence class contains the minimum of k records, the dataset is usually transformed by generalization or suppression. When quasi-identifiers are generalized, they become less precise. So, for example, instead of exact age, only an age range will be displayed, and in the case of zip codes, the last digits are redacted. To prevent too broad generalization caused by outliers, they can be suppressed. This is the case for the record of Mrs. Wong in the example shown in Table 2.2. Including her would change the first equivalence class from $\{Female, 60\text{-}70, 0214*\}$ to $\{Female, 10\text{-}70, *\}$ and thus cause a huge loss of information. However, suppression can lead to bias and should therefore be used carefully.
For most datasets, there are multiple possible k-anonymous representations. Meyerson and Williams [21] showed that for any $k \geq 3$ finding, the optimal k-anonymity is NP-hard. However, there are plenty of approximation algorithms available [22][23].
Let us come back to the example from Table 2.2 and examine the protection it provides. Even if an attacker knows the gender, age, and zip code of his target Mr. Smith, he cannot deduce which record belongs to him. A larger k is considered to be more secure but has a larger impact on the utility. In the example of the medical dataset, a k of 3 or higher would have forced all entries into one equivalence class. If, aside from Wong, no other record is suppressed, this would result in $\{*, 40-70, 021**\}$.

Identifier	Quasi-Identifier			Sensitive Attribute
Name	Gender	Age	Zip code	Diagnosis
Garcia	Female	78	02141	Parkinson's disease
Smith	Male	47	02138	Cancer
Brown	Female	71	02142	Stroke
Wilson	Male	41	02139	Cancer
Wong	Female	19	19136	Flu

Table 2.1: A medical dataset

Identifier	Quasi-Identifier			Sensitive Attribute
Name	Gender	Age	Zip code	Diagnosis
*	Female	60-70	0214*	Parkinson's disease
*	Male	40-50	0213*	Cancer
*	Female	60-70	0214*	Stroke
*	Male	40-50	0213*	Cancer

Table 2.2: A 2-anonymous version of the dataset

Weaknesses of K-Anonymity

Although k-anonymity provides a parameter to trade privacy for utility k has no clear relation to the two. A k of 3 can provide reasonable protection and utility on some datasets and have devastating effects on others. Likewise, increasing k does not evenly decrease utility and increase privacy. Increasing a k up to 10 might be fine, but a k of 11 could render the entire dataset useless. This can depend on the number of rows and the number of quasi-identifiers, as k-anonymity is susceptible to "the curse of dimensionality" [24]. Moreover, it cannot guarantee privacy and is vulnerable to several attacks.

Homogeneity Attack If all records within an equivalence class have the same sensitive value, an attacker can still infer the sensitive value of his target even though he cannot tell which record belongs to him [25]. In Table 2.2 an attacker that knows one of the men in the second equivalence class can deduce that he has cancer.

Even if not all sensitive attributes in an equivalence class are equal, an attacker can still gain a lot of information. If he targets Mrs. Brown, he knows that she either had a stroke or has Parkinson's disease. This in itself could be considered a privacy breach, as it narrows down the possible disease she is suffering from. And the attacker might have additional information that helps him determine which of the two options is more realistic.

Background Knowledge Attack An attacker can have background information about his target or the general population. He might know that his neighbor Mrs. Brown was put into an ambulance, which would have been an uncommon measure if she had a long time

suffering like Parkinson's disease. He could also make inferences based on age, ethnicity, or other well-known facts.

Variants

To address these issues, Machanavajjhala, Kifer, Gehrke, and Venkitasubramaniam introduced ℓ-diversity in 2006 [25]. A dataset is ℓ-diverse if, in every equivalence class, there exist at least ℓ different values for the sensitive attribute. To breach any individual's privacy within the table, an adversary would need ℓ-1 pieces of background knowledge.

However, it reduces the utility even more and is still not sufficient to preserve privacy. These flaws were pointed out by Ninghui, Tiancheng, and Venkatasubramanian which in turn introduced t-closeness [26].

Composition Attack K-Anonymity and its variants all share the flaw that they do not compose securely [27]. Two k-anonymous datasets taken together do not provide k-anonymity. This means that two independently anonymized releases about an overlapping population can be used to breach privacy. This can be used when organizations independently release anonymized data about overlapping populations. For example a different hospital could also release k-anonymous versions of their patient data. Someone who knows the quasi-identifiers of Mrs. Garcia and that she visited both hospitals can infer that she has Parkinson's disease, as this is the only sensitive attribute present in both datasets (for these quasi-identifiers).

Identifier	Quasi-Identifier			Sensitive Attribute
Name	Gender	Age	Zip code	Diagnosis
*	Female	≥ 70	02141	Parkinson's disease
*	Female	≥ 70	02141	Arthritis
*	Female	≥ 70	02141	Heart Disease
*	Male	< 70	0213*	Flu
*	Male	< 70	0213*	Viral Infection
*	Male	< 70	0213*	Tuberculosis

Table 2.3: A 3-anonymous dataset of a different hospital

2.2 Syntactic and Semantic Notions of Privacy

All these definitions share a common flaw. They are only heuristic approaches, meaning that it cannot be mathematically proven how secure they are. So, whenever a new weakness is discovered, they have to be adapted. Nissim and Wood argue that, although these "heuristic approaches [...] have been defined using mathematical language, they are syntactic in nature (i.e., specifying properties of how an anonymized dataset should look) rather than semantic

(i.e., specifying restrictions on what an attacker may infer about the personal information that is the input for the anonymization process by observing its outcome)". [28]
In 1977 Dalenius articulated a desideratum for statistical databases: nothing about an individual should be learnable from the database that cannot be learned without access to the database. In 2006 Dwork showed that this is not possible. For example, if a study finds out that smoking causes cancer, a smoker is clearly harmed, as his insurance premiums might rise. However, this could even happen if he was not among the study participants. This caused Dwork to refine the definition from "nothing can be learned about an individual" to nothing "specific," meaning nothing that could not have been learned without their data [7].

2.3 Differential Privacy

2.3.1 The Intuition

Differential privacy is a definition. It is a mathematical guarantee that can be satisfied by an algorithm that releases statistical information about a dataset. Cynthia Dwork describes it as "a promise, made by a data curator to a data subject: you will not be affected, adversely or otherwise, by allowing your data to be used in any study, no matter what other studies, data sets, or information from other sources is available"[29]. Differential privacy keeps this promise by ensuring that the same conclusions will be reached, meaning that any output is "essentially" equally likely to occur, independent of the presence or absence of any individual in the dataset. Consequently, nothing specific about an individual is revealed because an adversary seeing the output of an analysis cannot even tell whether an individual's data was used to compute it, regardless of any auxiliary information he possesses. Therefore, every single individual has "essentially" the same level of protection they would have had if they did not contribute their data Likewise, even individuals that are not in the dataset have "essentially" the same protection as if they were. The term "essentially" is captured by a parameter ε; it bounds how much more likely one scenario is than the other. We want to promise every individual i, that for every possible output O:

$$e^{-\varepsilon} \leq \frac{\text{Probability to see output O under } \mathrm{D}_{\text{with individual i}}}{\text{Probability to see output O under } \mathrm{D}_{\text{without individual i}}}$$

So, no matter what the output looks like, none of the two cases should ever be more than e^{ε}-times more likely than the other one.

The Definition

More generally, this intuition can be expressed by the notion of neighboring databases D and D', i.e., pairs of databases that differ in at most one element. This brings us to the formal definition of differential privacy, as first introduced by Cynthia Dwork in 2006 [7].

Definition 2.3.1 (Differential Privacy)**.** A randomized mechanism M gives ε-differential privacy if for any two databases D and D' differing on at most one element, and any set of

possible outputs $S \subseteq Range(M)$

$$\Pr[M(D) \in S] \leq e^{\varepsilon} * \Pr[M(D') \in S] \tag{2.1}$$

Since D and D' can be arbitrarily exchanged, the same has to hold the other way around, resulting in a lower as well as an upper bound for the odds, that an observed output O was computed based on D instead of D':

$$e^{-\varepsilon} \leq \frac{\Pr[M(D) \in S]}{\Pr[M(D') \in S]} \leq e^{\varepsilon} \tag{2.2}$$

2.3.2 The Privacy Loss Parameter ε

The smaller we choose ε to be, the more private each individual is, as it becomes less likely that they are in the dataset, and thus every associated risk becomes less likely as well. Therefore, the parameter ε can quantify privacy risk. An ε of zero provides perfect privacy because the probability distributions of two neighboring databases D and D' are identical, i.e., for any set of possible outputs S.

$$\Pr[M(D) \in S] = \Pr[M(D') \in S]$$

However, their neighbors and the neighbors of their neighbors must also have this probability distribution. This propagates to all possible datasets, including the empty set $\emptyset$

$$\Pr[M(D) \in S] = \Pr[M(D') \in S] = \Pr[M(D'') \in S] = ... = \Pr[M(\emptyset) \in S]$$

Consequently, if the probability distribution is always the same, regardless of the underlying dataset, nothing about the data can be learned. An ε of infinity, on the other hand, will yield maximum utility but no meaningful privacy guarantee. Already an ε of 5 can hardly be considered privacy-preserving. One scenario being $e^5 = 150$ times more likely than the other, is far from being "essentially" equally likely, and there is little protection for an individual if an adversary can tell with 99% certainty whether they are in the dataset or not. Usually, an $\varepsilon \leq 1$ is recommended. As every output is possible under any dataset, it is always possible to get an output with less or even no utility. But the larger the dataset and are, the less likely this is. Hence, ε can be used to trade-off between utility and privacy.

2.3.3 Group Privacy

Since every output is possible under every dataset, the privacy guarantee also extends to groups. This can be important if some records share the same value (e.g., household members have the same address, (total) rent, living space etc.). As seen before, the inequalities propagate, i.e., the bounds extend to neighbors of neighboring datasets as well:

$$\Pr[M(D) \in S] \leq e^{\varepsilon} * \Pr[M(D') \in S] \leq e^{\varepsilon} * e^{\varepsilon} * \Pr[M(D'') \in S] = e^{\varepsilon + \varepsilon} * \Pr[M(D'') \in S] \tag{2.3}$$

From this observation follows, that two datasets that differ in k rows are bounded by $k * \varepsilon$. Hence, an ε-differentially private mechanism is $k\varepsilon$-differentially private for groups of size k. However, only small groups are protected, as, in the end, the purpose of analytics is to learn general trends about the population. As a rule of thumb, groups of size $1/\varepsilon$ can expect a reasonable level of privacy [30, p.276].

2.3.4 Robustness Under Composition

It is inevitable, that multiple analyses on data about the same individuals will lead to an increased privacy risk. For many privacy-preserving methods, this can even break privacy entirely [27]. Differential privacy on the other hand is robust under composition, which means that the combined result of two differentially private mechanisms is still differentially private. Similarly, as in group privacy, the guarantee extends in a degraded fashion to the combined results of multiple differentially private analyses. Intuitively, if an output O_1 is e^{ε_1}-times and the output O_2 is e^{ε_2}-times more likely to be observed under D than under D', then the probability to see both is $e^{\varepsilon_1} * e^{\varepsilon_2}$ times more likely under D than under D'.

$$\text{Probability to see } O_1 \text{ and } O_2 \text{ under } D \leq e^{\varepsilon_1} * e^{\varepsilon_2} * \text{Probability to see } O_1 \text{ and } O_2 \text{ under } D' \quad (2.4)$$

Formally, this can be proven by defining a mechanism, that is the paring of M_1 and M_1: $M_{1,2}(D) = (M_1(D), M_2(D))$
Repetitively applying this theorem shows that combining n differentially private mechanisms is $\sum_{i=1}^{n}$- differentially private. This form of composition is called sequential composition and even accounts for subsequent computations that use the result of previous analyses of the same dataset. Parallel composition allows for even better bounds: If the mechanisms operate on disjoint datasets, then combined they are $\max_i \varepsilon_i$- differentially private [11]. Intuitively, this makes sense, as every individual is only in one of k datasets and an analysis on one of the other $k-1$ datasets cannot disclose anything about them.
Sequential composition is essential, to allow multiple (adaptive) queries on the same dataset, whereas parallel composition is useful when computing histograms or when aggregating individual contributions in local differential privacy. Together they make it possible to create more complex structures out of basic building blocks and to define a privacy budget, that can be split up among several analyses and analysts.

2.3.5 Immunity to Post-Processing

As opposed to other privacy-preserving methods, differential privacy is immune to post-processing. This means that the guarantees of differential privacy hold, no matter how the output is manipulated, and which auxiliary information are used to do so. It ensures that the results of differentially private computations can be safely released. "A data analyst cannot

increase privacy loss — either under the formal definition or even in any intuitive sense — simply by sitting in a corner and thinking about the output of the algorithm" [31, p.18]. If there is a 75% chance that an individual answered truthfully, there is nothing that could be done to the answer that could in- or decrease this probability. Any operation an adversary might use to process the result of an analysis can be modeled as a function f; it suffices to show that:

$$\frac{Probability\ to\ see\ output\ O\ under\ D}{Probability\ to\ see\ output\ O\ under\ D'} = \frac{Probability\ to\ see\ f(output\ O)\ under\ D}{Probability\ to\ see\ f(output\ O)\ under\ D'} \tag{2.5}$$

Formally, the composition of a data-independent mapping f with an ε-differentially private algorithm M is also ε-differentially private. Let $M : \mathbb{N}^{|X|} \to R$ be a randomized algorithm that is ε-differentially private and let $f : R \to R'$ be be an arbitrary randomized mapping, then $f \cdot M : \mathbb{N}^{|X|} \to R'$ is ε-differentially private.

Proof. We prove the proposition for a deterministic function $f : R \to R'$. The result then follows because any randomized mapping can be decomposed into a convex combination of deterministic functions, and a convex combination of differentially private mechanisms is differentially private. Fix any pair of neighboring databases under D, D' and fix any event $S \subseteq R'$. Let $T = \{r \in R : f(r) \in S\}$. We then have:

$$\Pr[f(M(D)) \in S] = \Pr[M(D) \in T] \leq e^{\varepsilon} * \Pr[M(D') \in T] = e^{\varepsilon} * \Pr[f(M(D')) \in S] \tag{2.6}$$

□

2.4 Achieving Differential Privacy

The Setting

Before we explain how differential privacy can be achieved, let us clarify the terminology we will use and a few assumptions we make about the setting. We have a database or dataset D that is a collection of records, each containing the data of a different individual. The database is maintained by a trusted curator that has access to the raw data and either provides an interface through which analysts can query the database or he releases statistical information about it by himself. Since the latter can also be modeled as the result of a query, we will not make any further distinctions between these two cases. We will use the terms query, analysis, or study interchangeably. We assume that all of them can be modeled as a function f that takes as input a database D and produces some output $f(D)$. This output can be any possible outcome of a study or analysis: a query result, a statistic, a diagram, or an entire machine learning model. The curators' task is to preserve the privacy of everyone within the database while still allowing statistical analyses of the database as a whole.
The goal of differential privacy is to ensure that based on a query result, an adversary is not able to tell whether an individual was part of this analysis, even when he knows all records in the database except for that of this target. If he cannot even determine this, then he cannot infer anything else about his target.

2.4.1 Mechanisms

There are four major mechanisms to achieve differential privacy.

The Randomized Response Mechanism

Surveying participants on touchy subjects is often difficult. As truthful answers can be embarrassing or even incriminating, the respondents may lie or refuse to participate in the survey. To circumvent this problem, Warner introduced the randomized response technique in 1965 [32]. It provides privacy through randomness, which can be introduced with a coin toss, a spinner, or shuffled cards The process is simple; the interviewer asks a yes-no question, the respondents flip a coin:

- If they get heads, they answer truthfully
- If they get tails, they answer falsely, i.e., the opposite of their true answer

Obviously, the coin has to be biased, otherwise observing "Yes" or "No" is equally likely for every respondent, regardless of their true value. A common choice is a coin that comes up heads 75% of the time The respondents are provided with "plausible deniability" because, with a 25% chance, they did not answer truthfully. The underlying statistics can later be recovered, as the collected "Yes" answers will, in expectation, consist of $\frac{3}{4}$ of the real "Yes" answers and $\frac{1}{4}$ of the real "No" answers, i.e.

$$E[Yes_{observed}] = \frac{3}{4} * Yes_{real} + \frac{1}{4} * No_{real} = \frac{3}{4} * Yes_{real} + \frac{1}{4} * (1 - Yes_{real}) = \frac{1}{2} * Yes_{real} + \frac{1}{4}$$

This can be rearranged to get an estimate of the true distribution:

$$E\left[2 * \left(Yes_{obs} - \frac{1}{4}\right)\right] = Yes_{real}$$

Thus, meaningful conclusions about the population can be drawn without revealing the true answer of any respondent.

Accuracy. Unfortunately, it is unlikely that heads will come up exactly 75% of the time. Small deviations could already lead to greatly over-/underestimating trends in the population. In the worst case, 50% of the coin flips are tails, rendering the survey completely useless. Although these risks can never be completely ruled out, they become negligible for larger sample sizes. For 100 coin tosses, there is already an 80% chance to get between 70- and 80-times heads, which corresponds to an additive error of at most ±5%. We can analyze the accuracy of this algorithm more precisely by calculating the variance and standard deviation of our estimator:
Variance:

$$VAR\left[\frac{1}{n}\sum_{i=1}^{n} 2 * \left(Yes_{obs} - \frac{1}{4}\right)\right] = \frac{4}{n^2} * \sum_{i=1}^{n} VAR[Yes_{real}]^* = \frac{4}{n} * \frac{3}{4} * \frac{1}{4} = \frac{3}{4n} \quad (2.7)$$

Standard deviation:

$$\frac{\sqrt{3}}{2 * \sqrt{n}} \tag{2.8}$$

The latter already indicates that we have noise in the order of $O(\frac{1}{\sqrt{n}})$.
This conclusion can also be reached with the Chernoff bound. It gives bounds on how much the sum of independent random variables deviates from its expected value. It can be used to determine a threshold below which this difference stays with probability $1 - \beta$.

$$\Pr\left[\left|2\left(Yes_{obs} - \frac{1}{4}\right) - Yes_{real}\right| \geq 2\sqrt{\frac{\ln(2/\beta)}{2n}}\right] \leq \beta \tag{2.9}$$

Privacy-Utility-Trade-Off. We have seen that the result becomes more accurate the more people we survey. But a larger sample size requires considerably more effort. Alternatively, the accuracy could also be improved by increasing the bias of the coin. If head comes up more often, there are less random answers that we need to account for. The direct effect of such an adjustment becomes apparent when generalizing the formulas from before, with [0.5,1]* being the probability of getting heads in the coin flip:

Estimator:

$$E[Yes_{obs}] = \gamma * Yes_{real} + (1 - \gamma) * (1 - Yes_{real}) = (2\gamma - 1) * Yes_{real} + (1 - \gamma) \tag{2.10}$$

Variance:

$$VAR\left[\frac{1}{n}\sum_{i=1}^{n}\frac{(Yes_{obs} - (1 - \gamma))}{2\gamma - 1}\right] = \frac{1}{n^2 * (2\gamma - 1)^2} * \sum_{i=1}^{n} VAR[Yes_{obs}] = \frac{\gamma * (1 - \gamma)}{n * (2\gamma - 1)^2} \tag{2.11}$$

Standard deviation:

$$\frac{\sqrt{\gamma * (1 - \gamma)}}{\sqrt{n} * (2\gamma - 1)} \tag{2.12}$$

Chernoff bound:

$$\Pr\left[\left|\frac{Yes_{obs} - (1 - \gamma)}{2\gamma - 1)} - Yes_{real}\right| \geq \frac{\sqrt{\ln(2/\beta)}}{\sqrt{2n} * (s\gamma - 1}\right] \leq \beta \tag{2.13}$$

The parameter can be used to trade privacy with utility. With a of 99% the standard deviation is in the range of ±1%. But at the same time, there is only a 1% chance that a respondent did not answer truthfully, which cannot be considered privacy-preserving. Finding the right

balance in this trade-off is a challenge you will also encounter when dealing with differential privacy.

Randomized Response and Differential Privacy. As indicated in the introduction, the randomized response technique is differentially private. Differential privacy is a definition that has to be satisfied by an algorithm. It guarantees that the result of an analysis is "essentially" equally likely to occur, independent of the presence or absence of any individual in the dataset. This is achieved by enforcing for every individual that, based on the result of an analysis, it is at most e^{ε} -times more likely that they contributed their (real) data than that they did not. A simple proof suffices to show that the randomized response algorithm fulfills this requirement:

$$\frac{\Pr[Response = Yes|Truth = Yes]}{\Pr[Response = Yes|Truth = No]} = \frac{\gamma}{1-\gamma} = \frac{\Pr[Response = No|Truth = No]}{\Pr[Response = No|Truth = Yes]} = e^{\ln\left(\frac{\gamma}{1-\gamma}\right)}$$

So, with $\gamma = 0,75$ it is 3 times more likely that an individual responded truthfully, thus the algorithm would be ln(3)-differentially private. With $\gamma = \frac{e^{\varepsilon}}{e^{\varepsilon}+1}$ a truthful answer is e^{ε}-times more likely than the complementary event $\frac{1}{e^{\varepsilon}+1}$. The formulas can be adjusted accordingly:

Standard deviation:

$$\frac{\sqrt{\frac{e^{\varepsilon}}{e^{\varepsilon}+1} * \frac{1}{e^{\varepsilon}+1}}}{\sqrt{n} * \left(\frac{2 * e^{\varepsilon}}{e^{\varepsilon}+1} - 1\right)} = \frac{\sqrt{\frac{e^{\varepsilon}}{(e^{\varepsilon}+1)^2}}}{\sqrt{n} * \left(\frac{e^{\varepsilon}-1}{e^{\varepsilon}+1}\right)} = \frac{\sqrt{e^{\varepsilon}}}{\sqrt{n} * \left(\frac{e^{\varepsilon}-1}{e^{\varepsilon}+1}\right) * (e^{\varepsilon}+1)} = \frac{\sqrt{e^{\varepsilon}}}{\sqrt{n} * (e^{\varepsilon}-1)} \quad (2.14)$$

Chernoff bound:

$$\Pr\left[\left|\frac{1+e^{\varepsilon}}{e^{\varepsilon}-1}\left(r - \frac{1}{1+e^{\varepsilon}}\right) - q(D)\right| \geq \frac{1+e^{\varepsilon}}{e^{\varepsilon}-1}\sqrt{\frac{\log(s/\beta)}{2n}}\right] \leq \beta \quad (2.15)$$

The randomized response algorithm is a technique to preserve the privacy of respondents in a survey. It provides them with "plausible deniability" since, with a certain probability, they did not answer truthfully. This probability can be adjusted to either improve the accuracy of the analysis or the privacy of the individuals. The accuracy can be determined through probability bounds, i.e., by assessing how likely it is that the difference between the true and the estimated value is above a certain threshold. The randomized response algorithm is differentially private and introduces noise in the order of $O(\sqrt{n})$

Another approach to achieve differential privacy is the Laplace mechanism. It assumes a trusted data curator and adds noise to the output instead of the input of analysis, leading to significantly lower noise. However, the randomized response algorithm is locally differentially private, meaning it does not require a trusted data curator and provides privacy already in the data collection stage. An example where this comes in handy is Google RAPPOR, a real-world application of differential privacy based on the randomized response algorithm. It is used in the Google Chrome browser to collect user data.

The Laplace Mechanism

The Laplace mechanism is the best-known mechanism to achieve differential privacy. It was introduced in 2006 by Dwork, McSherry, Nissim, and Smith shortly before the term differential privacy was coined [33].
We assume that a query, analysis, or study* can be modeled as a function f that takes as input a database or dataset* D and returns a numeric value $f(D)$. Differential privacy requires that an adversary should not be able to tell apart two neighboring databases D and D' based on a query result. The Laplace mechanism achieves this by adding random values to the query result $f(D)$. Intuitively, we want to hide the presence (or absence) of any individual in the database by masking the maximum impact a single individual could have on the result of a query f, i.e., the maximum difference in the query result $f(D)$ between any two neighboring databases D and D'. This is captured by the sensitivity Δf.

Definition 2.4.1 (ℓ_1-sensitivity)**.** The sensitivity Δf of a query f is defined as:

$$\max_{\substack{D,D' \in \mathbb{N}^{|\mathcal{X}|} \\ \|D-D'\|_1=1}} \|f(D) - f(D')\|_1$$

According to differential privacy, for any output, it should be at most e^{ε}-times more likely that it was computed on the true database D than on a neighboring database D'. Hence, the random number added to the result should be drawn from a probability distribution that decreases by a factor of e^{ε} over intervals of length Δf, as this is the maximum distance between two neighboring databases. One distribution that has these properties is the Laplace distribution; it is a symmetric version of the exponential distribution.
The Laplace distribution:

$$Lap(x|b) = \frac{1}{2b} e^{-\frac{|x|}{b}}$$

As its name suggests, the Laplace mechanism draws this noise from the Laplace distribution. Formally, it is defined as follows:
Given a function $f : D \in \mathbb{N}^{|\mathcal{X}|} \to \mathbb{R}$ over a dataset D, mechanism M provides ε-differential privacy if it follows:

$$M(D) = f(D) + Lap\left(\frac{\Delta f}{\varepsilon}\right)$$

What immediately stands out is that the noise does not depend on the size of the dataset n. This means it has more impact on queries over smaller datasets. Intuitively this makes sense, as the presence or absence of a single individual has a bigger impact on smaller datasets, needing relatively more noise to cover it up.
The smaller we choose ε to be, the more do the probability distributions of neighboring databases overlap and the harder it will get for an adversary to tell them apart. But at the same time, it becomes more likely to get larger random values and consequently less accurate results.
This allows us to answer k queries at once. Assuming all of them are counting queries, i.e., a single individual can influence each of them by at most 1, we have a sensitivity of k. This means we must add $Lap(k/\varepsilon)$ to every coordinate of the output vector. Note that this will lead to the same amount of noise as if we had (made use of sequential composition and) posed the queries sequentially, splitting the privacy budget ε equally among them. Then we would have had to add Lap $\left(1/\frac{\varepsilon}{k}\right) = Lap\left(\frac{k}{\varepsilon}\right)$ to each of the k $\frac{k}{\varepsilon}$-differentially private queries. We can do better if a single individual can only affect one of the counts, like in a histogram with k disjunct bins. In such a case, it suffices to add Lap to the count of each bin (parallel composition).

The Exponential Mechanism

The Laplace mechanism can only be used for queries that are robust to perturbation and relatively insensitive to changes in the data of a single individual. It is not suitable for optimization problems or when dealing with non-numeric values.
The Laplace and the Gaussian mechanism are not suitable for such problems, as they achieve differential privacy by adding real-valued noise to the result of an analysis. They work best if these results are relatively insensitive to changes in the data of a single individual and if their usefulness is relatively unaffected by additive perturbations. So, for example, in counting queries, where an individual can only influence the count by one and adding noise to the result affects its accuracy just marginally Optimization problems, on the other hand, are not robust to additive perturbations and highly sensitive. Additionally, they may contain structural information that cannot be easily perturbed, like in bin packing, where a binary variable b_i states whether bin i is used or not.
The same applies to machine learning, where models and classifiers are produced or, in general for queries with non-numeric outputs, such as asking for the most common eye color of a population.
To address these issues Frank McSherry and Kunal Talwar developed the exponential mechanism It assigns a utility to every possible output and returns those with better utility with an exponentially higher probability The set of possible outputs R has to be determined in advance and independent of the underlying dataset. Otherwise, this could lead to outputs with non-zero probability in one dataset and zero probability in a neighboring set and thus violate differential privacy For the eye colors this could be $\{blue, brown, green, grey\}$ or $\{0, 1, 2, 3, 4, 5\}$

in the case of the optimal price.
Formally, the algorithm is defined with respect to a utility function $u : D \in \mathbb{N}^{|\mathcal{X}|} \times R \to \mathbb{R}$, which maps a database-output pair (D, r) to utility scores. The utility function $u(D, r)$ tells us how good the output r is for the dataset D. So, in our example, u would be the revenue, and a price r of 4 would be optimal. But for a different set of customers D', where no one is willing to pay more than 3$, it would be an unfavorable result The query result can be arbitrarily sensitive; we only consider the sensitivity of the utility function u, i.e., the maximum possible impact a single individual could have on the utility

$$\Delta u = \max_{r \in R} \max_{D,D'} |u(D, r) - u(D', r)|$$

The exponential mechanism on inputs D,R and u outputs some object $r \in R$, where the probability a particular r is selected is proportional to:

$$\exp\left(\frac{\varepsilon * u(D, r)}{2 * \Delta u}\right)$$

This means the probability to see an element r is:

$$\Pr[M(D) = r] = \frac{\exp\left(\varepsilon * \frac{u(D, r)}{2 * \Delta u}\right)}{\sum_{r' \in R} \exp\left(\frac{\varepsilon * u(D, r')}{2 * \Delta u}\right)}$$

Intuitively, the result r should be at most e^{ε}-times more likely on D, than on a neighboring dataset D', where the utility is (up to) Δu lower. However, a single individual cannot only affect the utility of one result but the normalization term as well. The exponential mechanism reserves half of the privacy budget to account for this, explaining the factor 2 in the exponent.

$$\exp\left(\frac{u(D, r)}{\frac{\Delta u}{\varepsilon/2}}\right) = \exp\left(\varepsilon * \frac{u(D, r)}{2 * \Delta u}\right)$$

The exponential mechanism M is ε-differentially private Fix D,D' as neighboring datasets, and some outcome $r \in R$. Then we express the ratio of the probability of r being output under D and D' as follows:

$$\frac{\Pr[M(D)=r]}{\Pr[M(D')=r]} = \frac{\dfrac{\exp\left(\varepsilon * \dfrac{u(D,r)}{2*\Delta u}\right)}{\sum\limits_{r'\in R}\exp\left(\dfrac{\varepsilon * u(D,r')}{2*\Delta u}\right)}}{\dfrac{\exp\left(\varepsilon * \dfrac{u(D',r)}{2*\Delta u}\right)}{\sum\limits_{r'\in R}\exp\left(\dfrac{\varepsilon * u(D',r')}{2*\Delta u}\right)}}$$

$$= \exp\left(\frac{\varepsilon * (u(D,r)-u(D,r'))}{2*\Delta u}\right) * \left(\frac{\sum\limits_{r'\in R}\exp\left(\dfrac{\varepsilon * u(D',r')}{2*\Delta u}\right)}{\sum\limits_{r'\in R}\exp\left(\dfrac{\varepsilon * u(D,r')}{2*\Delta u}\right)}\right) \leq \exp\left(\frac{\varepsilon}{2}\right) * \exp\left(\frac{\varepsilon}{2}\right) = e^{\varepsilon} \tag{2.16}$$

1. From the definition of the sensitivity Δu we know that $u(D,r) - u(D',r) \leq \Delta u$ and thus follows:

$$\exp\left(\frac{\varepsilon * (u(D,r)-u(D,r'))}{2*\Delta u}\right) \leq \exp\left(\frac{\varepsilon * \Delta u}{2*\Delta u}\right) = \exp\left(\frac{\varepsilon}{2}\right) \tag{2.17}$$

2. Likewise, we can conclude that $u(D',r') \leq u(D,r') + \Delta u$, plugged into the formula this shows that:

$$\frac{\sum\limits_{r'\in R}\exp\left(\dfrac{\varepsilon * u(D',r')}{2*\Delta u}\right)}{\sum\limits_{r'\in R}\exp\left(\dfrac{\varepsilon * u(D,r')}{2*\Delta u}\right)} \leq \frac{\sum\limits_{r'\in R}\exp\left(\dfrac{\varepsilon * u(D',r') + \Delta u}{2*\Delta u}\right)}{\sum\limits_{r'\in R}\exp\left(\dfrac{\varepsilon * u(D,r')}{2*\Delta u}\right)}$$

$$= \frac{\sum\limits_{r'\in R}\exp\left(\dfrac{\varepsilon * u(D',r') + \frac{\varepsilon}{2}}{2*\Delta u}\right)}{\sum\limits_{r'\in R}\exp\left(\dfrac{\varepsilon * u(D,r')}{2*\Delta u}\right)} = \frac{\exp\left(\frac{\varepsilon}{2}\right) * \sum\limits_{r'\in R}\exp\left(\dfrac{\varepsilon * u(D',r')}{2*\Delta u}\right)}{\sum\limits_{r'\in R}\exp\left(\dfrac{\varepsilon * u(D,r')}{2*\Delta u}\right)} = \exp\left(\frac{\varepsilon}{2}\right) \tag{2.18}$$

$$\Pr\left[u(M(D)) \leq \mathrm{OPT}(D) - \frac{2*\Delta u}{\varepsilon}\left(\ln\left(\frac{|R|}{|R_{OPT}|}\right) + t\right)\right] \leq e^{-t} \tag{2.19}$$

This can be rewritten to:

$$\Pr\left[u(M(D)) \geq OPT(D) - \frac{2 * \Delta u}{\varepsilon}(\ln(|R|) + t)\right] \geq 1 - e^{-t} \tag{2.20}$$

By setting $t = \frac{1}{1-\beta}$ we can determine the minimum amount of utility we will get with probability β

$$\Pr\left[u(M(D)) \geq OPT(D) - \frac{2 * \Delta u}{\varepsilon} * \ln\left(\frac{|R|}{1-\beta}\right)\right] \geq \beta \tag{2.21}$$

With f being a function that returns a real number, the Laplace mechanism can be modeled with the utility function

$$u(D, r) = -|f(D) - r| \tag{2.22}$$

Consequently, this results in the sensitivity

$$\Delta u = \max_{D,D'}(|f(D) - r| - |f(D') - r|) \leq \max_{D,D'}(|f(D) - f(D')|) = \Delta f \tag{2.23}$$

The exponential mechanism returns an output $r = f(D) + x$ with probability proportional to

$$\exp\left(\frac{\varepsilon * u(D, r)}{2 * \Delta u}\right) = \exp\left(\frac{-\varepsilon}{2 * \Delta f} * |x|\right) \tag{2.24}$$

which is Laplace noise with parameter $\frac{2\Delta f}{\varepsilon}$. Output the true value with probability

$$\Pr[M(0) = 0] = \Pr[M(1) = 1] = \frac{\exp\left(\frac{\varepsilon * u(D, r)}{2 * \Delta u}\right)}{\sum_{r' \in R} \exp\left(\frac{\varepsilon * u(D, r')}{2 * \Delta u}\right)} = \frac{e^{\frac{\varepsilon}{2}}}{e^{\frac{\varepsilon * 0}{2}} + e^{\frac{\varepsilon}{2}}} = \frac{e^{\frac{\varepsilon}{2}}}{1 + e^{\frac{\varepsilon}{2}}} \tag{2.25}$$

In both cases, the exponential mechanism is worse by a factor of two. This is due to half of the privacy budget being reserved for changes an individual could have on the normalization term. In the case of these two mechanisms, it can be shown that the normalization term is unaffected, e.g., for randomized response, it is $1 + e^{\varepsilon/2}$, no matter if D is 0 or 1, thus leading to better accuracy.

3 Related Work

In this chapter, we briefly summarize the core literature this thesis is based on and discuss existing efforts to teach differential privacy. They were also crucial in selecting important topics for our learning content.

3.1 Scientific Literature

To date, there are more than 3400 scientific papers available on differential privacy. However, most of them focus on highly specific, narrow topics.

The Algorithmic Foundations of Differential Privacy

After introducing differential privacy in 2006, Dwork subsequently published further work on it. She summarized first results [34], explored its use in the area of statistics [35][36], explained the guarantees differential privacy provides and clarified common misconceptions about it [37]. She then consolidated these findings in 2014 in "the Algorithmic Foundations of Differential privacy" “The algorithmic foundations of differential privacy” together with Aaron Roth, who previously wrote a thesis on “New Algorithms for Preserving Differential Privacy” [38]. To this date, "Algorithmic Foundations of Differential Privacy" is the most comprehensive work on differential privacy. It explains the definition and properties, the basic mechanisms, sensitivity, variants and alternate use cases. However, as the book is already seven years old, it lacks information about current advancements, tools and real-world applications.

The Complexity of Differential Privacy

Together with "the Algorithmic Foundations of Differential Privacy", this work constitutes the basis of most lectures on differential privacy [39][40][41][42]. They mostly overlap in the topics they cover.

Differential Privacy: A Primer for a Non-Technical Audience

Wood et al. wrote a simplified and informal introduction to differential privacy [30]. On 70 pages, they primarily describe the guarantee provided by differential privacy and explain it with illustrative examples. About the purpose of their paper, they write that: "It can help guide practitioners as they make decisions regarding whether to use differential privacy and, if so, what types of promises they should make to data subjects about the guarantees

differential privacy provides." This is closely related to our goal. However, it lacks the depth we strive for as it does not go beyond the definition and only explains in an abstract manner how differential privacy is achieved without mentioning any of the mechanisms. Based on this source alone, a practitioner will not be able to apply differential privacy. They also briefly touch on differential privacy tool, other privacy preserving methods, and discuss benefits of differential privacy [30, p.213].

Differential Privacy and Applications

In their work, Zhu et. al. discuss the application of differential privacy to different use cases, ranging from recommender systems to location privacy. They further provide a flowchart showcasing the various steps that can be followed when solving a privacy preservation problem for a particular application. They also consider possible challenges, most of which are related to utility, e.g. high sensitivity, large query sets, data sparsity and correlated data. They

3.2 Grey Literature

The grey literature is less involved with niche applications. Nevertheless, it also gained a lot of traction. Many tools and resources emerged in the last two years, and practitioners are beginning to make their first experiences with them. An especially interesting kind of grey literature for us are learning materials on differential privacy.

Programming Differential Privacy Joseph P. Near, is an assistant professor of the University of Vermont and also helped in developing Ubers differential privacy solutions [43][44]. In the fall of 2020 he held a course on data privacy. The textbook of the course is hosted online as a Jupyter notebook and openly accessible. He explains core themes of differential privacy with code examples and programming exercises. He covers de-identification, k-anonymity, differential privacy, and its properties, approximate differential privacy and other variants, (local) sensitivity, the core mechanisms, the sparse vector technique, local differential privacy, machine learning, and synthetic data. He also provides learning objectives aligned to Blooms taxonomy, exercises and homework.

Algorithms for Private Data Analysis Also in the fall of 2020, Prof. Gautam Kamath held a course on differential privacy. He provides detailed lecture nodes and recordings of the lectures [45].

4 Interviews

As outlined in the research approach, we wanted to include voices out of the industry to ensure the practical relevance of our work. This was accomplished in two ways, on the one hand, by studying the grey literature, and on the other hand by conducting semi-structured expert interviews.

The use of differential privacy by key players like Google [46], Microsoft [47][48], LinkedIn [49], Apple [50], Amazon [51][52], Uber [43][44], Snapchat[54], and the US Census Bureau [55][56][57] is already well-known and well-documented. Although the documentations are mostly about technical details, the only exception being the US Census Bureau which also wrote about the challenges it faced [58][59] and the lessons it learned from using differential privacy [60].

For most companies, however, little information was available about their experiences with Differential Privacy. It was difficult to assess how widely used or known it was in the industry. To address this gap, we conducted interviews with practitioners.

4.1 Procedure

Five interviews were conducted between September and November 2020. They were scheduled for one hour and held online using video conferencing tools. The respondents received the questionnaire in advance so that they could prepare for the interview or request changes. Four of them allowed us to record the interview to facilitate the transcription process. In total, 40 pages of transcripts were produced.

Identifying interview partners

We searched for companies that had patents or publications on differential privacy, offered it as a product, attended conferences about it, or stated on their website, blog, or privacy policy that they used it. We then reached out to them via email.

We also contacted those that did not meet these criteria but were involved in privacy and security related issues or had corresponding research departments. We hoped to find practitioners that deliberately decided against using differential privacy. Their reasoning behind this decision and the methods they were using instead would have been an interesting addition to our findings. However, we could not find anyone to whom this applied.

Questionnaire

The questionnaires were adapted to the respondents and consisted of up to 10 questions as well as some sample follow-up questions. The following six key questions were part of every questionnaire:

- What motivated you to use differential privacy?
- What are the benefits of using differential privacy?
- What challenges did you face when working with differential privacy, and how did you overcome them?
- What tools and sources did you rely on when working with differential privacy?
- How mature is differential privacy, in your opinion?
- What topics should we cover in our learning content?

We discuss the answers to these questions in the next section.

4.2 The Findings

In this section, we will discuss the insights we gained from the interviews. It is structured according to the key questions we asked. To preserve the privacy of our interview partners, we refer to them by a randomly assigned identifier E_i. The experts are quoted within the text. As not all interviews were recorded or conducted in English, the quotes may not reflect the exact wording. However, the interviews were transcribed and translated faithfully, and the interviewees had the opportunity to review them and request changes.

Background Information

All of our interview partners started working with differential privacy in 2017/2018. So about the time, the GDPR entered into force, and the first differential privacy libraries emerged. They all either had a Ph.D. or were working towards it. Most came from privacy-related fields, such as cryptography.

The use cases of our interviewees and their clients ranged from banking, financial services, and insurance to the automotive sector, healthcare, the public sector, telecommunication, location data, and trajectories as well as machine learning in general.

Motivation

We started the interview by asking the experts what motivated them to use differential privacy. According to E_4, the general interest in privacy-preserving methods stemmed from the desire to analyze sensitive data. IT-security measures ensured that the data was stored securely,

with no unauthorized person having access to it. But then the crucial part was to be able to use this data. Without privacy-preserving methods, the data could only be used in a very restricted fashion, always requiring the user's consent. E_1 argued similarly: "If [...] there is a dataset somewhere in your company that you cannot access because of restrictions such as the GDPR, it takes six to eight months of internal compliance processes just to have a look at it. And then you might realize that this does not work for you and you need additional information, then the whole process starts over. So, for us, the goal was to speed up that process to get a sneak peek of datasets within a couple of days, making everything more agile."

All experts were or quickly became aware of the shortcomings of traditional de-identification techniques. Their search for a more sophisticated privacy-preserving method eventually led them to differential privacy. For E_2 and E_5 this search started out of research interest. In the case of E_3, an existing system had to be upgraded with as little effort as possible. He explained: "It used k-anonymity. This may be sufficient for privacy-preserving data publishing, but our clients wanted to use it for data mining, i.e., to draw conclusions from the results and to use these conclusions to request the same data again. And because k-anonymity does not compose at all, i.e., multiple requests on the same datasets are no longer k-anonymous, even if the individual outputs are k-anonymous, they needed a better approach."

Alternatives

Since the topic was already brought up in the motivation, we continued by asking for possible alternatives to differential privacy. There was, however, little variety in the responses. All experts were aware of common practices such as de-identification or pseudonymization and agreed that they were outdated and insufficient. For privacy-preserving machine learning E_2 mentioned methods such as homomorphic encryption, secure multi-party computation, and federated learning. They all had their own application domain. But for a normal machine learning workflow, with data that did not need to be shared, and trust in the infrastructure on which the training was performed, differential privacy was the best choice.

K-Anonymity

The experts' opinions diverged when it came to k-anonymity. On the one hand, it got heavily criticized for being a purely syntactic notion of privacy, lacking composability and any real guarantee. E_3 illustrated how, in theory, even mechanisms where it was obvious that they were not privacy-preserving could satisfy k-anonymity. For example, if each row of the original data was replicated k times, then the resulting dataset was k-anonymous, although an attacker could simply reverse this process. E_2 also deemed it unsuitable for machine learning. Aggregating data points into classes would destroy individual correlations, which were exactly the dependencies that machine learning was about. Furthermore, E_2 argued that for some data types, too much information got lost through generalization. This was better with differential privacy, as here, the noise was added to the gradient instead of the dataset. Thus, any gradient-based method could be used, regardless of the data type.

On the other hand, E_4 also defended k-anonymity. For him, it was equivalent to differential privacy and would not be replaced by it. Both had their justification in the real world. Moreover, he said that it was not about k-anonymity or differential privacy, but about pseudonymizing and hoping for the best, or using anonymization methods, that offer privacy guarantees. According to him, this included differential privacy as well as k-anonymity, l-diversity, and the like. He further argued that k-anonymity was easier to understand. This was crucial when several stakeholders with different backgrounds were involved in deciding which method to use. Not only was the concept itself more intuitive, but also how it affected the utility. He explained this based on an example: "Imagine I would like to know how many 35-year-old women watch a certain TV show between 7 and 8 pm. [...]. Here I can argue that generalizing the age to 5-year steps is absolutely fine because I would have done it anyway. This loss of information is then not considered to be severe. I probably do not care whether someone is 36 or 37 because I have to aggregate them anyway for a sufficiently large population". The impact of such changes could be worse than that of those caused by differential privacy, but in return, they were more comprehensible. Nonetheless, he also acknowledged that differential privacy provided stronger protection, was more broadly applicable, and allowed for a better analysis of the utility of the results.

Benefits

The benefits our interviewees named coincided with those we found in the literature. Across the board, the strong protection and formal guarantee differential privacy provide were viewed as its greatest advantages. E_1 argued that while other methods only claimed to be privacy-preserving, this could actually be proven for differential privacy. According to E_4, this made it easier to justify the use of differential privacy to a data protection officer or supervisory authorities. E_3 stated that in general, they advised their clients to choose a solution that provides mathematically provable guarantees instead of one where it might not be known how to attack it today, but where this could change in the future. Ideally, one should use a parameterizable procedure, as this made it easier to react to changes in privacy laws and regulations. Having a parameter to trade off between privacy and utility and the resulting quantifiability of privacy risk were also regarded as major benefits by the other interviewees.

Versatility turned out to be another advantage of differential privacy. Initially, we asked for the limitations of differential privacy and if there were cases where it could not be used. But the common response was that it was universally applicable because it was such a general principle. E_3 even believed that whenever there was a solution to a problem, then there would be a differentially private one as well. He further acknowledged that there were use cases that were difficult to implement, for example, when the data was highly correlated. To ensure privacy, there would probably only be noise left. But he questioned whether this was a limitation of differential privacy or if it was mathematically simply not possible to learn something from this data without violating privacy. E_3 elaborated that there was often just

a misunderstanding of what privacy was about. "They want you to make the anonymity requirements magically disappear so that they can do everything they could have done on the raw data.

Another aspect was data accessibility and availability. E_1 explained how differential privacy could accelerate the approval processes he mentioned in the motivation. E_5 further added that differential privacy could open up new areas for data collection and processing.

Challenges

We proceeded to ask our interview partners which challenges they faced when using differential privacy and how they overcame them. The most frequent answer was the limited utility. E_5 noted that too much noise had to be inserted for real privacy guarantees. E_3 pointed out that with a fixed privacy budget, either the number of queries was limited or more noise had to be introduced with every subsequent query. The implications of the second approach were criticized by E_4: "It is very difficult to deal with information decay in a database environment. Why? Well, how would a data scientist behave when he sees data? He asks this, then that he wants to be exploratory. And if he accidentally chooses the wrong order, he has less accurate information than if he had asked the other way around". However, a limited number of queries was also problematic unless it was known in advance how often the dataset would be queried.

According to E_2, low utility was less of an issue in machine learning. Good accuracy could be achieved through hyperparameter tuning in the differentially private stochastic gradient descent. The problem was rather the search for the right hyperparameters. This was a lengthy process since differentially private training took significantly longer than regular model training. Each data point was processed individually, clipped, and perturbed as opposed to regular machine learning, where the whole batch was processed at once, followed by a joint update. This was said to be a significant performance difference. Therefore, it was often not feasible to find the right parameters for high accuracy. E_3 also considered implementing differential privacy efficiently to be the biggest challenge. As an example, he brought up the exponential mechanism. It required to iterate over all possible outputs, which was not feasible in most cases. The universe of all possible outputs was just too huge to compute this in the lifetime of machines.
Of course, it was case-dependent whether the performance was considered to be a challenge. For simple applications, differential privacy only requires that noise is added to a query result. So it was neither a surprise nor a contradiction that other experts praised differential privacy for being computationally inexpensive. Especially when it was compared to other privacy-preserving methods like secure multi-party computation or homomorphic encryption.

However, all of our interviewees agreed that the complexity of differential privacy was a major challenge. They argued that differential privacy was hard to understand and to explain. E_2 said that if a company used it, its customers would not understand it and therefore could

not verify how well protected they were. This might even lead to a false sense of security. E_4 supported E_2 in that the parameter epsilon was abstract and hard to interpret. E_3 also agreed that k-anonymity was more intuitive in this regard.
E_1 mentioned that differential privacy was difficult to implement and to prove. There was an "unimaginable number of things that could go wrong." Reportedly, this was even more severe in the past, as there were no open-source projects or examples to get inspiration from. "You just had the paper and had to implement whatever was on the paper," E_1 explained. But according to him, this changed for the better. E_2 complained that existing differential privacy libraries were not very user-friendly. This further added to the complexity. "It is often unclear how parameters are set. You could improve a lot by suggesting values to the users or by explaining which values have which consequences."

How could these challenges be overcome

The discussion surrounding the complexity of differential privacy already touched on how this challenge could be addressed. The experts emphasized the need for good explanations of differential privacy. E_1 went on about the importance of developing differential privacy tools in an open-source environment. That way, anyone could test them and help to prevent grave mistakes. At his work, they also tried to be based on as many open-source projects as possible. When it came to interpreting the parameter epsilon, E_2 explained that as a proxy, they were simulating attacks, such as the membership inference attack. These were pre-implemented in libraries like Tensorflow Privacy. They then compared the differentially private machine learning model with a regular model and checked whether the risk had decreased. This, however, was only a temporary solution. In the long run, we needed a framework that quantified which effects which epsilon would have on the real world for different datasets and machine learning models. It should show how likely it was that a specific datapoint could be recovered or that certain attributes could be revealed.

Regarding the performance E_2 argued that this could become less of an issue in the future. Either through more efficient implementations of existing libraries and algorithms or because of the natural growth in computing power. This would simultaneously improve the utility in the aforementioned cases. Nonetheless, it would never be possible to completely eliminate inaccuracies due to the inherent trade-off between privacy and utility. When it came to dealing with the introduced noise, E_3 exemplified how this could be handled in practice: "You do not assume that all attackers in the world, who have ever seen a differentially private statistic will collude. Instead, you have agreements with data recipients, for example, that they will only use the data in a specific context and not simply put it online. That way, the privacy budget can be limited to specific use cases, customers, or time periods. You can classify the customers according to their trustworthiness. Those where I do not know what they will do with the data receive a fixed budget. With others, I have legal agreements and only have to ensure that an individual statistician [...] cannot learn anything about individuals in the dataset. But I am legally protected against the company joining forces with others and trying to reconstruct my raw data." To avoid the limited number of queries or the information decay,

E_4 suggested using local instead of global differential privacy.

Tools

As already indicated, E_2 used TensorFlow Privacy and its attack frameworks. The decision to choose it over alternatives like PySyft was rooted in TensorFlow Privacy being developed by Google. Since Google's scientists were contributing extensively to the field of differentially private machine learning, E_2 assumed that their implementations would always be incorporated into TensorFlow Privacy. This ensured that the library would be frequently updated and got new features faster than others.

The other experts did not use differential privacy tools, although they were aware of them. E_3 further elaborated that he mostly dealt with individual solutions and that it would be hard to find a broadly applicable library. Most of the existing libraries were from the area of machine learning, which he had little to do with or general implementations of algorithms, with an outcome that was differentially private but not usable for individual applications. He questioned if there ever would be the jack of all trades that everyone could use.

Sources

There was also no broad range of recommended sources. The RAPPOR paper from Google [46] was said to be an interesting read, and PrivBayes [61] was mentioned in the context of differentially private synthetic data generation. E_4 advised us to stick to "the algorithmic foundations of differential privacy." This was sufficient to cover the basics, and then, depending on the use case, one should look for a suitable follow-up paper.

Maturity

All experts reported that they saw a growing interest in privacy and received more inquiries about it. According to E_3 and E_4 supervisory authorities were already aware of differential privacy. It was further said to be a well-established definition and mature enough to be used in practice.

Naturally, this raised the question why this is then rarely the case. The experts came back to the challenges they mentioned earlier. The complexity, performance issues, and loss of utility were obstacles that slowed down the spread of differential privacy. E_2 argued: "I think many companies are starting to experiment with it, but the number of companies using differential privacy in their core products is still small. One of the most prominent players is Google, but they only use it for analytics in their chrome browser and not in their core product, the search algorithm. They would lose money if the predictions worsened, and they could no longer place the right ads. I think that is how many companies feel. They like to use it as a gimmick, where it does not hurt, but they do not dare to modify their core products yet." E_4 added: "Of course, every data scientist is afraid of changes to the data in any form. The whole point of

storing data is to be able to retrieve it as it was before. Every form of perturbation raises the question [...] whether my analysis is still reliable." But not only were companies discouraged from using differential privacy, some just did not see the need for it. E_5 explained: "In business, the question always is 'why do we have to change what we are doing right now?' If the legislation says that removing PII is enough, then they see no need for change." For E_3 this missing awareness was the biggest obstacle: "People in charge do not yet really see what differential privacy does and why they should use it. So it is the same problem that security has: 'Why do I need security? Until now, nothing has ever gone wrong?'" Furthermore, differential privacy required analysts to adapt their tools unless they were using synthetic data. The proposed solution to overcome these obstacles, which were mostly human factors, was once again education. E_5 deemed it also as helpful if differential privacy was legally required. E_2 explained that security was taken more seriously than privacy because legal regulations demanded it. So far, only the GDPR existed for privacy. No one knew what this meant for machine learning, which was why many did nothing in this regard. As often everything was developed in a hurry and privacy cost extra time and money, it was often left out.

In light of these challenges, we asked our interviewees how long they thought it would take for differential privacy to be widely used. We mentioned the estimate of the Gartner Hype Cycle as a reference. According to their analyst Van Baker [62], it would take five to ten years until differential privacy reached the plateau of productivity. Most of our interviewees shared this assessment. E_1 guessed that it would take more than five years to be commonly adopted. It was emerging, but nowhere near where it needed to be, in terms of wide adaptation. E_5 thought that until then, we would probably see developments in niche applications. E_3 was indecisive whether differential privacy was the definitive solution everyone would end up using, but he believed that it would gain in popularity. We were still in the early stages of implementing the GDPR, and analyzing personal data became increasingly important. For him, the estimate of five to ten years was realistic. But this was difficult to predict. He compared it to other methods and brought up an example from cryptography. Just a few years ago, TLS 1.3 had been standardized. Public-key cryptography was now 40 years old. So, pessimistically, one could say that it took almost 40 years to establish a reasonable standard. Likewise, E_5 argued that for multi-party computation it took 40 years to be where we were today and that there the only challenge was performance.

Promising Trends

To wrap up the predictions about the future of differential privacy, we asked our interviewees if there were any developments they looked forward to. They mostly answered with general statements, like improved utility or efficiency. Among the more specific answers were smooth sensitivity, local differential privacy, distributed learning, and differentially private synthetic data.

Which Topics Should Be Covered?

Finally, we asked which topics we should cover in our learning content. E_5 recommended us to explain the known mechanisms and when they were applied, the local and the global model, and the interaction between utility and noise. This was supported by E_3 who told us that everything he had seen so far got along quite well with the absolute basics. This included a good understanding of the definition, the composition theorem, and the Laplace mechanism. He further regarded it as beneficial to develop a broad understanding of the subject. So knowing the difference between semantic and syntactic notions of privacy, being aware of the risks, and understanding the limits of anonymity. E_1 suggested to structure the content into multiple layers of complexity. The first layer should be "super easy to understand and super intuitive," without being technically untrue. With every additional layer, this had to be backed up with more technical examples, definitions, and formulas. E_4 advised us to explain the concepts based on a concrete scenario. This was endorsed by E_5: "Examples would be good too. Datasets, even if they are mocked, where you can observe the effect of adding a certain level of privacy and see how this impacts the queries. What does it mean to add noise with epsilon 0.5 compared to noise with an epsilon of 1?" E_1 also proposed tools, where users could "play with a few parameters (and) see how the animation changes."

4.3 Conclusion

The interviews helped us to gain valuable insights into the use of differential privacy in practice. They laid the foundation for answering our research question about the potentials and limitations of differential privacy. Additionally, they contributed significantly to the selection of relevant topics for our learning materials as well as to their design and contents.

Although the benefits of differential privacy are evident, and it was deemed a promising solution to address privacy risks, it is still at an early stage and has not yet fully advanced from research into practice. It will take more than five years until we will see differential privacy widely used. At the moment, its complexity and the missing awareness for privacy-related issues are among the biggest obstacles hindering its spread. Both challenges can be addressed through education, which reinforces the importance of our efforts to provide learning materials on differential privacy.

5 Benefits and Challenges of Differential Privacy

In this chapter, we will discuss the benefits and challenges of differential privacy we discovered in the literature research and the interviews. It addresses the question whether differential privacy should be used.

5.1 The Benefits of Differential Privacy

The benefits of differential privacy can in large part be derived from the properties of its definition, namely the quantification of privacy risk, group privacy, robustness under composition, and the immunity to post-processing.

Strong Privacy Protection Every individual in the dataset has essentially the same protection as if they had not contributed their data. Therefore, nothing specific to them can be learned. This protects against arbitrary risks incurred from participating in a dataset that go beyond re-identification (e.g., rise in insurance premiums). Additionally, it can be mathematically proven whether a mechanism provides this guarantee. The guarantee also extends to groups, albeit to a lesser extent.

This formal guarantee holds, no matter what an adversary does with the output or which auxiliary information he possesses. This includes attacks that are currently unknown as well as future data sources. Thus, the guarantee is future proof. Naturally, the strength of the guarantee will degrade when combining the results of multiple differentially private analyses on the same dataset, but it will not break, as it is robust under composition.
Differential privacy is the only approach that provides such a formal, mathematically provable privacy guarantee [30, p.214, p.270][63, p.33] Other privacy-preserving methods are mostly syntactic in nature and, therefore, susceptible to linkage attacks, might break under composition, or only protect against certain types of attacks or risks [28].

Versatility Differential privacy is not a method that is applied but a definition that must be fulfilled, and there are numerous ways to do so. Thus, differential privacy is quite versatile and can be adapted to various use cases, ranging from simple statistics to machine learning [64]. This is further facilitated by its robustness under composition, as it allows to construct complex algorithms from smaller building blocks.

Since all the properties of differential privacy are part of the definition and independent of the implementation, there is a high degree of freedom in how it can be realized. Differential privacy can also handle various data types, such as numeric, non-numeric, graph, or location data.
Unlike other approaches, differential privacy protects against arbitrary risks and attacks Moreover, it can even be used for purposes apart from privacy. For example, to prevent strategic lies in mechanism design, to ensure generalizability in adaptive data analysis, or to prevent overfitting in machine learning.
Not only can the privacy loss parameter ε be used to adjust the level of protection, but there are also different variations and relaxations of differential privacy that adapt it to settings with a computationally bounded adversary. They modify various parts of the definition, e.g., how neighboring databases or the sensitivity are defined or which properties are protected. Moreover, differential privacy can be combined with other privacy-enhancing methods like federated learning or secure multiparty computation to rule out additional attack vectors, like a data leakage at the curator or during transmission.

Availability If the outcome for users is the same, no matter if they are in the dataset or not, it will be easier to incentivize them to contribute their data. More data reduces the sampling error and other biases in the data and therefore outweighs the loss in utility caused by the noise introduced to preserve privacy.
Due to the hard worst-case assumption of differential privacy and its immunity against post-processing, there is no need for attack modeling, i.e., assessing what an adversary might know or how he might attack Hence, data can be shared without hesitation. This can significantly reduce the time it takes to get access to sensitive data.
Collect, process, and share data, where this was not possible before, due to privacy concerns. Sensitive data could become more accessible, enabling more people to research it for the greater good

A Framework to Reason About Privacy Risk The privacy loss parameter ε quantifies privacy risk and how it accumulates across multiple differentially private analyses. This makes it possible to compare how private different algorithms are and to reason about the appropriate level of privacy. "Traditional de-identification techniques often require concealment of the extent to which the data has been transformed, thereby leaving users with uncertainty regarding the accuracy of analyses on the data" 2,p.271 When using differential privacy on the other hand, it isn't necessary to keep implementation details like parameters secret Thus, users can precisely determine the accuracy of their analysis, making it easier to compare differentially private algorithms in terms of their accuracy and level of privacy.

5.2 The Challenges of Differential Privacy

Limited Utility The most frequently mentioned argument against differential privacy is the limited utility. A fixed privacy budget for a database results either in a finite number of

queries, after which the database may never be queried again, or in increasingly higher noise with each query, that quickly renders query results useless. This is even worse for everything beyond counting queries, such as analyses of highly skewed or correlated data And when counting the number of people suffering from a rare disease, even a small amount of noise could lead to false conclusions. This raises the question whether it can be afforded to worsen the results of (medical) research just to prevent attacks that are theoretically possible under a hard worst-case assumption.

This is not an issue of differential privacy, but privacy itself. Too many too accurate queries will eventually break privacy. Every analysis with non-zero utility leaks information about the underlying dataset, this in inevitable. Differential privacy just makes this issue visible. It quantifies the privacy loss and states the amount of noise necessary to bound it The worst-case assumption of an omnipotent adversary might be too strong, but this definition can be relaxed. Noise can also be decreased using local sensitivity or advanced composition theorems. In general, it is advisable to use robust statistics, which are less sensitive to outliers. For example, using the median instead of the mean. Moreover, "the claim that differential privacy is inconsistent with a specific accuracy goal is hard to prove. The poor utility of a specific differentially private algorithm should not be confused with a failure of differential privacy per se" [37] although better algorithms might be (computationally) more complex, adding efficiency to the trade-off between privacy and utility. When weighing up the benefits of more accurate study results against privacy, it should be considered that the data might not be available otherwise. Ultimately, the data subjects should decide for themselves whether they want to contribute their data and which level of privacy they demand in return. Variants of differential privacy such as personalized privacy could make this possible in the future [65].

Complexity Differential privacy is complex. It is hard to understand, explain and implement. There are many misconceptions regarding its guarantee and implications, even among scholars. Communicating differential privacy to clients and colleagues can be cumbersome. Customers might either not acknowledge it or have a false sense of security if they hear that differential privacy is used. Like in the case of Apple, that permits an overall daily privacy loss of 16 [50], which is far from private. This is something a user can hardly check, especially since Apple kept their implementation details secret. Implementing a differentially private algorithm has many pitfalls and requires mathematical proof that it satisfies the guarantee. Even simple post-processing operations, like bounding the results to positive values, can lead to significantly less utility. Choosing the right parameters is not a trivial task either. Neither are there clear guidelines on how to choose , nor is it easy to determine the sensitivity of a query, i.e., the biggest impact a single individual could have on the result of the query. When it comes to an understanding and explaining differential privacy, there are already plenty of resources available. The same goes for differential privacy tools that help in implementing it, either by providing basic building blocks or by verifying whether an algorithm is differentially private. If the sensitivity is hard to determine, clamping the values or falling back on the

sample and aggregate model are possible solutions.

Requires Change Analysts have to change the way they work with data. They cannot access raw data directly anymore and need new approaches to interpret the results. They "are accustomed to the data looking a certain way, and to interpreting those data as the 'ground truth.' As such, they are unaccustomed to seeing population counts with fractional or negative values". Moreover, the data can have logical inconsistencies, for example, that the sum of subpopulations does not equal the total population. Also, they have to resist the temptation to repeat analysis if the result is far off. Besides this cultural change, the tools they work with have to be adjusted as well. Differential privacy is future proof, but to ensure this, the remaining privacy budget of every user has to be tracked and once the overall privacy budget is depleted, the database may never be queried again.
There are vendors that offer differentially private synthetic data. It mimics the original data and thus allows to work with it as one would have done with the raw data. However, it doesn't capture all use cases as a single row doesn't correspond to an existing individual anymore. Additionally, several differential privacy tools are already designed to be similar to their non-private counterpart [8][9][15], thus facilitating the transition.

6 Learning Nuggets

6.1 Categories

The learning nuggets are intended to support practitioners in their decision whether they should use differential privacy. We identified four major questions they might have when making this decision and aligned the learning nuggets accordingly.

1. Should I use differential privacy?
   The learning nuggets from the motivation category address the question whether it is necessary to use differential privacy and if there are any alternatives available. The discussion category then provides arguments for and against the use of differential privacy by listing the benefits and challenges of using differential privacy.

2. Can I use differential privacy? Another question practitioners might have, is whether they can use differential privacy in their particular use case. To which settings can it be applied, and which problems it solves is addressed in the learning nuggets from the definition category.

3. How should I use it? There are different forms of differential privacy (variants and extensions, the local and the global model, global or smooth sensitivity etc.) and different ways to achieve it. Learning nuggets from the application and outlook category address these topics.

4. Who or what can help me to use it? The Outlook category presents vendors offering differential privacy as a product, real-world applications to get inspiration from, and tools that help in using differential privacy.

6.2 Structure

All of the learning nuggets were created in PowerPoint. This makes it easier to edit and maintain them. To ensure that the learners use them as intended, the presentation is set to kiosk mode and saved as a .ppsx file (PowerPoint Show). That way, users cannot edit the slides and can only view them in the presentation mode. Moreover, they cannot use the arrow keys, click or scroll to change the slide they are currently on. Instead, they have to rely on the navigation options we provide them with.

Navigation Elements

Users can navigate within and across learning nuggets. The navigation bar depicted in figure 6.1 is located in the upper right corner of every slide. Additionally, a user can click on words highlighted in blue to go to the corresponding learning nugget or glossary entry. They can then use the curved arrow from the navigation bar to go back to the learning nugget they were coming from.

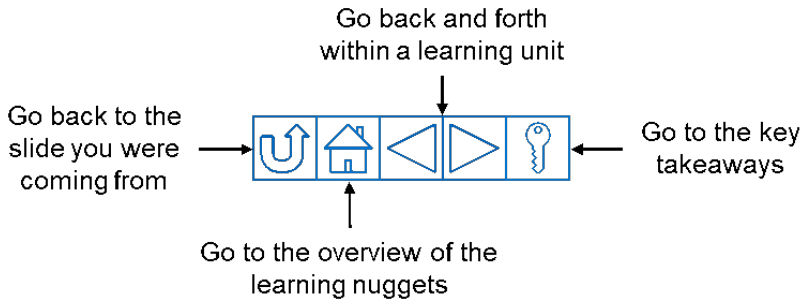


Figure 6.1: Navigation elements

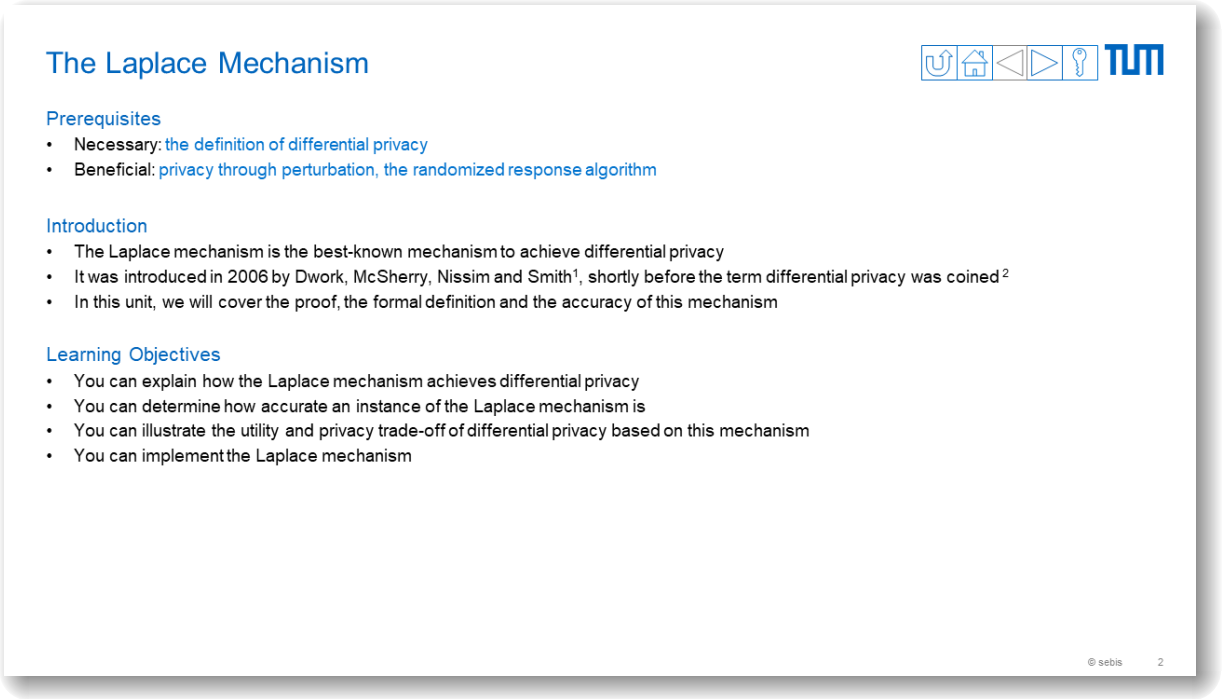
The learning nuggets themselves all follow the same structure. They begin with an introduction, followed by the content, which is then summarized in the Key Takeaways. Afterward, the outlook and the list of sources provide the user with suggestions on where to go next.

Introduction

The first slide states which learning nuggets the user should have read in order to understand the contents of the following slides. These prerequisites are divided into necessary and beneficial prior knowledge. Afterward, the topic of the learning nugget is motivated. A few introductory sentences briefly summarize the most important aspects of the previous units and outline how the current one will tie in with them. Lastly, the learning objectives are listed. They are structured according to Bloom's taxonomy and represent the order of the content.

Content

The content spans over several slides. To avoid overwhelming learners with large blocks of text, it is split up into a bulleted list of sentences that no slide has more than ten of. Graphs and images are used to further loosen the content up. Additionally, examples and mathematical proofs are contained in extra boxes or slides since they can be regarded as optional content, and some users might want to skip them altogether.
Some slides, such as the one in figure 6.3 also have interactive elements. By double-clicking on the table on the right, the user can access an excel sheet. Here he can change parameters and see how they affect the other values in the table.



The Laplace Mechanism

Prerequisites

- Necessary: the definition of differential privacy
- Beneficial: privacy through perturbation, the randomized response algorithm

Introduction

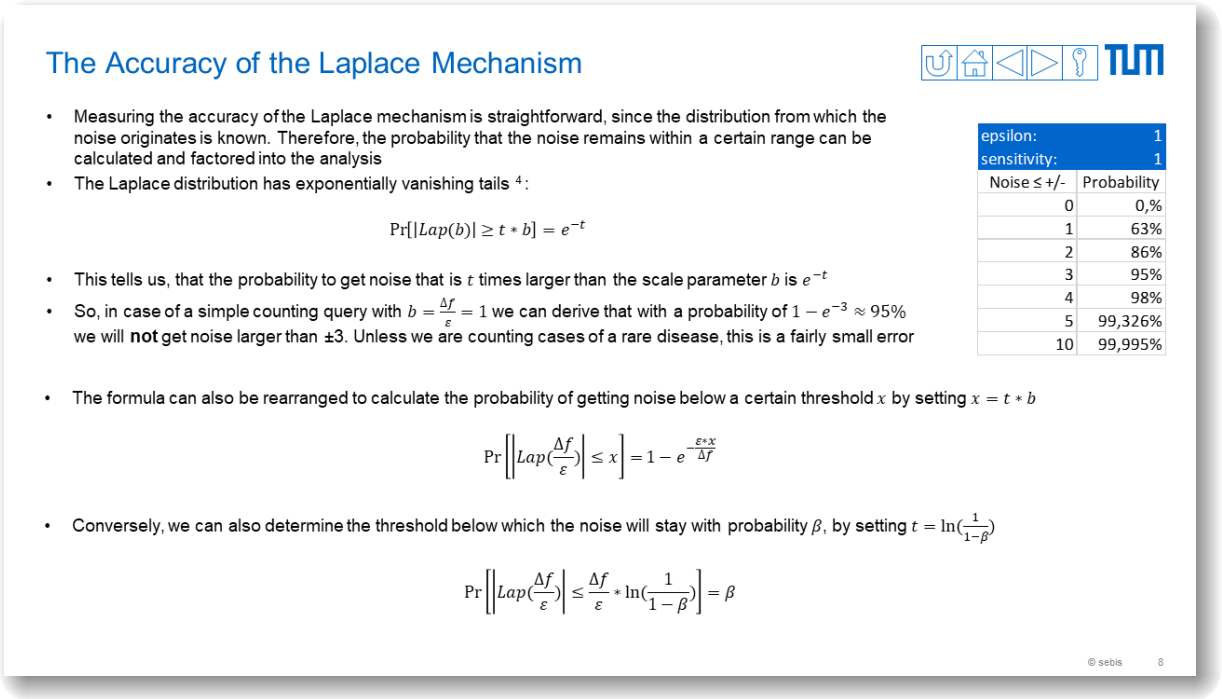
- The Laplace mechanism is the best-known mechanism to achieve differential privacy
- It was introduced in 2006 by Dwork, McSherry, Nissim and Smith[1], shortly before the term differential privacy was coined[2]
- In this unit, we will cover the proof, the formal definition and the accuracy of this mechanism

Learning Objectives

- You can explain how the Laplace mechanism achieves differential privacy
- You can determine how accurate an instance of the Laplace mechanism is
- You can illustrate the utility and privacy trade-off of differential privacy based on this mechanism
- You can implement the Laplace mechanism

© sebis 2

Figure 6.2: Introduction slide



The Accuracy of the Laplace Mechanism

- Measuring the accuracy of the Laplace mechanism is straightforward, since the distribution from which the noise originates is known. Therefore, the probability that the noise remains within a certain range can be calculated and factored into the analysis
- The Laplace distribution has exponentially vanishing tails [4]:

$$\Pr[|Lap(b)| \geq t * b] = e^{-t}$$

- This tells us, that the probability to get noise that is t times larger than the scale parameter b is e^{-t}
- So, in case of a simple counting query with $b = \frac{\Delta f}{\varepsilon} = 1$ we can derive that with a probability of $1 - e^{-3} \approx 95\%$ we will **not** get noise larger than ±3. Unless we are counting cases of a rare disease, this is a fairly small error

epsilon:	1
sensitivity:	1
Noise ≤ +/-	Probability
0	0,%
1	63%
2	86%
3	95%
4	98%
5	99,326%
10	99,995%

- The formula can also be rearranged to calculate the probability of getting noise below a certain threshold x by setting $x = t * b$

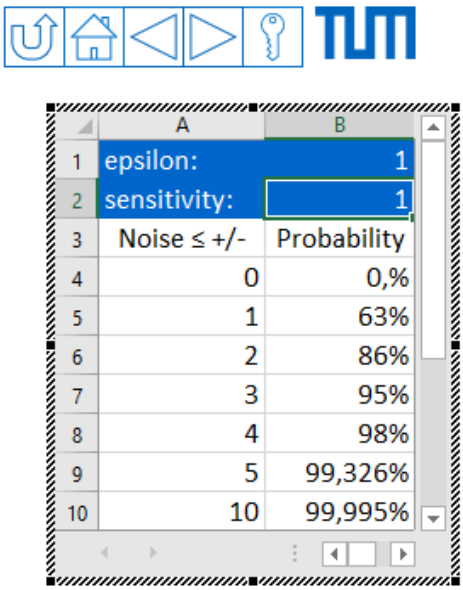
$$\Pr\left[\left|Lap(\frac{\Delta f}{\varepsilon})\right| \leq x\right] = 1 - e^{-\frac{\varepsilon * x}{\Delta f}}$$

- Conversely, we can also determine the threshold below which the noise will stay with probability β, by setting $t = \ln(\frac{1}{1-\beta})$

$$\Pr\left[\left|Lap(\frac{\Delta f}{\varepsilon})\right| \leq \frac{\Delta f}{\varepsilon} * \ln(\frac{1}{1-\beta})\right] = \beta$$

© sebis 8

Figure 6.3: A slide with an interactive element



	A	B
1	epsilon:	1
2	sensitivity:	1
3	Noise ≤ +/-	Probability
4	0	0,%
5	1	63%
6	2	86%
7	3	95%
8	4	98%
9	5	99,326%
10	10	99,995%

Figure 6.4: An interactive element

Exercises

Some learning nuggets also contain exercises. They allow users to apply the theory they just read about. This also helps them to check if they really understood it. Other learning nuggets rather contain comprehension questions. They usually ask the learner to apply formulas and algorithms from the learning nugget. They also transfer knowledge, extend the content to things that would be out of scope. Sometimes they require programming or bring in knowledge from other learning nuggets The solution to the exercises are shown on the following slides; the users have to check for themselves whether they were correct.

Key Takeaways

At the end of the learning nugget, the most important points are summarized. The learner can use this opportunity to check if he got them all while reading the learning nuggets. Others might want to skip the content altogether and just read the key takeaways to move on to other learning nuggets. The outlook gives a preview of related nuggets The last slide lists the sources referenced throughout the learning nuggets as well as those that might be interesting for further reading.

6.3 The List of Learning Nuggets

In total, 22 learning nuggets were designed. They can be grouped into five categories.

Motivation

Solution – Implementing the Laplace Mechanism – 2/2

1. Show what happens if the same query is repeated multiple times. You do not necessarily have to program to solve this task

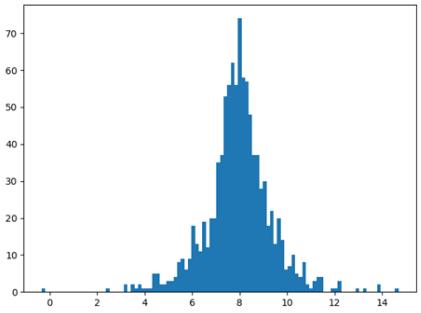
Since the Laplace mechanism is differentially private it composes. This means that k ε-differentially private queries will still be differentially private when combined, but the level of protection degrades to $k\varepsilon$-differential privacy.
We can show this by plotting a histogram of the outputs we get from repeating the same query 1000 times

```
import matplotlib.pyplot as plt

plt.hist([laplace_mechanism(8,1,1) for x in range(1000)], 100)
plt.show()
```

As you can see, the histogram takes on the form of the Laplace distribution centered at 8. With every query it becomes more likely that this is the true result.

This illustrates the inevitability that too many too accurate query results will eventually break privacy. Differential privacy attempts to postpone this inevitability but cannot change the fact that either the number of queries we can ask, or their usefulness is limited.



 13

Figure 6.5: Exercise solution slide

Key Takeaways

- The Laplace mechanism is the best-known mechanism to achieve ε-differential privacy when dealing with numeric data
- It adds random noise drawn from the Laplace distribution, scaled to $\frac{sensitivity\ \Delta f}{\varepsilon}$, to the result of a query
- The introduced noise is independent of n, this means it gets relatively smaller for larger datasets
- With probability $1 - e^{-\frac{\varepsilon * x}{\Delta f}}$, the noise is below the threshold x

Outlook

- The Laplace mechanism adds noise to the output as opposed to the randomized response algorithm, which adds noise to the input of a query. This leads to significantly less noise. Visit local differential privacy to further explore these differences
- The Laplace mechanism is limited to numeric queries, whereas the exponential mechanism can deal with non-numeric data and with queries where adding noise directly to the result does not make sense (e.g. optimization problems)
- Learn more about the sensitivity, to get the most out of the Laplace mechanism

 14

Figure 6.6: Key takeaways

1. Attacks on Privacy
2. Privacy-preserving Methods
3. K-Anonymity

Definition

1. Privacy Through Perturbation
2. The Definition of Differential Privacy
3. The Adversary's Perspective

Discussion

1. The Benefits of Differential Privacy
2. The Challenges of Differential Privacy
3. Putting Noise Into Perspective

Application

1. The Randomized Response Algorithm
2. The Laplace Mechanism
3. The Exponential Mechanism
4. The Gaussian Mechanism
5. Comparing the Mechanisms
6. The Sensitivity
7. Local Differential Privacy

Outlook

1. The Maturity of Differential Privacy
2. Differential Privacy Tools
3. Real-world Applications of Differential Privacy
4. Differential Privacy Vendors
5. Differential Privacy in Machine Learning

6. The Differential Privacy Family

We will discuss them in the next section. In doing so, we will address the choices made in designing them, the primary sources they are based on, and the key learnings they should convey.

6.4 Learning Nuggets Overview

6.4.1 Motivation

The goal of this set of learning nuggets is to raise awareness for privacy issues. It emphasizes the need to take them seriously and the shortcomings of traditional privacy-preserving methods. None of the learning nuggets from this category are necessary to understand differential privacy. They do, however, highlight desiderata for privacy-preserving methods, properties such methods should have, the attacks they should protect against and criteria that are met by differential privacy. Of course, someone who is already aware of these issues or wants to get straight to the point can proceed directly to the definition of differential privacy.

Attacks on Privacy

This learning nugget illustrates prominent attacks on privacy and further serves as an introduction of core terminology and concepts like re-identificiation, de-identification, quasi-identifiers, auxiliary information, linkage and reconstruction attacks.
It starts by presenting three infamous cases of re-identification: the AOL scandal, the re-identification of governor William Welds and the Netflix Prize. Based on the examples, the following insights are discussed: the flaws of de-identification, the power of linkage attacks and the fact that even innocuous information can be used to breach privacy and might entail even greater harm. This shows that one never knows which auxiliary information is available and what an attacker might learn.

Privacy-Preserving Methods

This learning nugget discusses various approaches to preserve privacy and their shortcomings. This encompasses traditional de-identification and pseudonymization techniques, query restriction, query auditing, output, and data perturbation. It also touches on security measures (e.g., encryption, access control, homomorphic encryption, secure multiparty computation) and highlights how they are also essential and address different privacy risks. Learning: We need methods that provide provable guarantees, and thus there is no real alternative to differential privacy.

Exercises - Sketch how methods could be combined - Unknown method: Show how it could be broken

k-Anonymity

We revisit the re-identification of Governor William Welds and how this led to the definition of k-anonymity. We explain the property, how generalization and suppression work and then proceed to explain the weaknesses. That it is susceptible to homogeneity and background knowledge attacks and that it is not robust under composition. This leads to the final conclusion that the issue lies in them being syntactic instead of semantic notions of privacy.

Exercises

1. The learners have to make a dataset k-anonymous
2. A k-anonymous dataset is presented, and the learners have to break it

6.4.2 Definition

This category consists of the basic learning nuggets necessary to understand what differential privacy is. To understand the guarantees it provides and what it protects. Based on these learnings, practitioners should be able to decide if it addresses their privacy concerns.

Privacy Through Perturbation

This learning nugget was initially not intended but created when the first Think-aloud sessions showed that learners had some open questions left. It explores the basic intuition behind output perturbation as a means of preserving privacy. It shows that infinite noise is required, that it is no good idea to rely on privacy through hiding information Based on the example of a simple output perturbation technique, the users learn about the need for infinite noise and that security through obscurity is not advice-able according to Kerckhoff's principle.

The Definition of Differential Privacy

This learning nugget covers the guarantee differential privacy provides and describes how this translates into the definition. It further explains the properties of differential privacy, namely group privacy, robustness under composition, and immunity to post-processing.

Exercises To strengthen the intuition about the protection differential privacy provides, learners are presented several cases of privacy breaches and have to decide which of them could have been prevented through the use of differential privacy. Additionally learners have to argue why any non-trivial mechanism satisfying differential privacy, or any reasonable notion of privacy in general, has to be non-deterministic.

The Adversary's Perspective

This learning nugget facilitates to interpret the privacy loss parameter ε and also showcases how it can be used to quantify privacy risk. It introduces the Bayesian interpretation of differential privacy and how the level of protection provided can be determined in the form of the maximum information gain of an adversary in the worst case. This possibility is showcased on the example of Gertrude, presented in the differential privacy primer [30].

6.4.3 Discussion

The learning nuggets in this category provide the learner with arguments to consider when deciding about using differential privacy.

The Benefits of Differential Privacy

We list the benefits of differential privacy according to the structure already presented in chapter 5. We start off with the strong protection, followed by its versatility, how it increases the availability and accessibility of data, and end with the quantifiability of privacy risk and the transparency it enables and how this makes it a perfect framework to reason about risk. The benefits were those stated in literature and expert interviews.

The Challenges of Differential Privacy

The learning nuggets present the limited utility and the complexity of differential privacy and that the way we handle data requires re-orientation. We discuss how these challenges can be overcome. The challenges were obtained from the literature and the expert interviews.

Putting Noise Into Perspective

Practitioners tend to worry about the inaccuracy caused by differential privacy. We put the introduced noise into perspective by comparing it to the sampling error and re-stating the database reconstruction theorem. We further argue that the noise is an opportunity because it rewards the information minimization principle.

Exercises Learners have to decide which one out of two queries is less private than the other, and thus requires more noise.

6.4.4 Application

This category revolves around achieving differential privacy. It introduces the basic mechanism and targets everyone interested in implementing differential privacy.

The Randomized Response Algorithm

The randomized response algorithm is an ideal introduction to the topic of differential privacy as it originated from the social sciences, is simple and intuitive but in fact, also differentially private. This allows us to explore important aspects of differential privacy, such as the utility privacy trade-off. The randomized response algorithm provides local differential privacy, but it introduces more noise than the other mechanisms.

Exercises Learners have to design an own variant of the randomized response algorithm that allows three instead of two answers. Further, they can explore the properties of differential privacy based on the randomized response algorithm by showing how the algorithm behaves under composition and by investigating the level of privacy it provides for groups.

Local Differential Privacy

This learning nugget teaches the difference between the global and the local model of differential privacy. They are compared in terms of security they provide and the amount of noise introduced. The comparison enables practitioners to decide which fits better to their case. Also, they learn that most companies using differential privacy apply the local model because it does not require trust in the data curator. We further touch on the shuffle model and using secure multiparty computation to simulate the global model.

The Sensitivity

This learning nugget teaches how to get the most out of the previously presented mechanisms as most of them introduce noise scales to the sensitivity. We discuss the sensitivity of different statistics. Further, local sensitivity (AI) is introduced, which has to be determined independent of the database to be analyzed. We talk about ways to use this local sensitivity like the propose-test-release-approach or smooth sensitivity. We will show practices to reduce sensitivity by clipping, which bounds the sensitivity. Finally, a subsample-and-aggregate-framework is illustrated for cases in which it is difficult to determine the sensitivity.

Exercises Learners are provided with examples and have to determine the sensitivity.

The Laplace Mechanism

The Laplace mechanism is the best-known mechanism for differential privacy. In this unit, we will cover the proof, the formal definition, and the accuracy of this mechanism. It is taught how to apply the Laplace mechanism. We start to explain it for single-valued queries and end with vector-valued queries. Learnings: After studying this learning nugget, learners are able to explain how the Laplace mechanism achieves differential privacy and how accurate an instance of the Laplace mechanism is. Further, learners can illustrate the utility and privacy trade-off of differential privacy based on this mechanism.

Exercises In the respective exercise, learners have to implement the Laplace mechanism and plot a histogram with the results of repeated queries. This visualizes the composition because the histogram will take on the shape of the Laplace distribution.

The Exponential Mechanism

The Laplace and the Gaussian mechanism can only be used for queries that are robust to perturbation and relatively insensitive to changes in the data of a single individual. It is not suitable for optimization problems or when dealing with non-numeric values. This learning nugget showcases how differential privacy can be achieved in these settings using the exponential mechanism. We will cover the motivation, the formal definition, the proof, and the accuracy of this mechanism. Moreover, we show that it is a general framework that captures any mechanism that gives differential privacy.

Exercises The exercise is to apply the exponential mechanism to a simple counting case to return the most popular item of a set. Learners, therefore, implement the exponential mechanism.

The Gaussian Mechanism

We explore the Gaussian mechanism and the benefits we obtain from sampling from the Normal distribution. We prove why this only achieves ε, δ-differential privacy, and we explain the ℓ_2-sensitivity and why it can be applied in this case, and the advanced composition theorem.

Comparing the Mechanisms

We compare the different mechanisms to achieve differential privacy. We give guidelines to assist practitioners in their decision for one of the mechanism. The randomized response algorithm as a means to achieve local differential privacy, the Laplace mechanism to achieve pure ε-differential privacy for numeric valued queries, the Gaussian mechanism for better composition properties / i.e. more queries with less noise, as long as it is acceptable to just achieve ε, δ-differential privacy, and lastly the exponential mechanism for dealing with non-numeric values or optimization problems. We also visualize and compare their performances (error terms, amount of noise introduced etc.).

Exercises Now that the learners know about all the basic mechanisms to achieve differential privacy, they are tasked to design a differentially private version of the k-means algorithm by choosing the right (combinations of) mechanisms.

6.4.5 Outlook

The learning nuggets within the outlook category will give the practitioners guidance on where to proceed (after the completion of the previous learning nuggets). So if a practitioner

decides to use differential privacy, he might be interested in tools and vendors that can help him use differential privacy and real-world applications from which he can get inspiration. It also touches on the maturity of differential privacy, the vast field of differentially private machine learning, and variations and extensions of differential privacy (the DP family).

The Maturity of Differential Privacy

In this learning nugget, we share the insights we gained in our interviews and give the learner an impression of the current state of differential privacy by providing them with a timeline listing the most important milestones.

Real-World Applications of Differential Privacy

We will give an overview of the areas in which Google, Apple, Uber, the US Census Bureau, Microsoft, LinkedIn, Snapchat, and Salesforce use differential privacy. We take a more in-depth look into Google RAPPOR and Apple's differential privacy solution.

Differential Privacy Vendors

A major obstacle of using differential privacy is its complexity. Although there are already differential privacy tools practitioners can use, they might not want to implement it themselves and rather refer to a commercial solution. This learning nugget presents companies that assist practitioners interested in commercial differential privacy solutions: The SAP HANA Data Anonymization platform and differentially private synthetic datasets as they are provided by Statice, Hazy, or LeapYear.

Differential Privacy Tools

We list the most popular differential privacy tools and discuss their feature set.

Differential Privacy in Machine Learning

This learning unit gives an overview of differential privacy in machine learning. For a long time machine learning models were thought to be private simply because of their complex nature. But as illustrated in attacks on privacy, research has already proven otherwise. Since covering the whole field and various models of machine learning would be too extensive, only neural networks are discussed in more depth. Nevertheless, a list of machine learning models for which a differentially private version already exists is provided, to give an impression how broadly applicable it already is in the field of machine learning, and if applicable the library that implements them.

The Differential Privacy Family

There are approximately 225 variations and extensions of differential privacy. As practitioners are probably overwhelmed by this amount of all those different notions, we want to guide

them based on the taxonomy developed by Desfontaines and Pejó[17]. We introduce the seven dimensions of the taxonomy and provide an example for each of them. Users will be able to use them to find the right variant of differential privacy for their use case.

6.5 Learning Platform

In this section, we will outline the concept of the e-learning platform that was conceptualized within this thesis and later implemented by a group of students in a practical course. They further refined these initial ideas. Their solution is covered in the outlook of this thesis.

6.5.1 Motivation

Our goal was to make differential privacy more accessible. We wanted to offer learners more than just text and visualizations. They should be able to learn about differential privacy in an exploratory manner by using interactive widgets. However, existing e-learning platforms and learning management systems are restrictive in the way content can be created and displayed. Their goal is to keep the creation process simple to protect the content creator from himself. Therefore, they are often limited to text, pictures, and videos. As we wanted to provide various interactive elements alongside our learning content, we decided to develop an own dedicated e-learning platform for differential privacy.

6.5.2 Views

The application is divided into multiple views. The landing page provides an overview of all learning nuggets. Each of them is represented by a tile, displaying the title and a short description of the learning nugget. They are ordered by relevance. Those that were deemed helpful by other users with the same role as the current one are ranked higher. A learner can also use search and filter functions to access specific content that might interest him. Clicking on a learning unit will forward the user to its detail view.
In the detail view, the whole content of the unit is displayed. The content is mainly text-based but may contain tables, pictures, videos, and interactive elements. These range from simulations to graphs that are plotted based on the parameters the users have chosen.

6.5.3 Prototype

A small prototype (Figure 6.8) was developed as a proof of concept. The tool allows a user to upload a .csv file, choose an ε value and either compute a differentially private linear regression or histogram. An unmodified result is provided as a reference. The differentially private results are computed by diffprivlib, the differential privacy library of IBM [15]. The unmodified results are provided by sci-kit-learn. The prototype was later integrated into the learning platform.

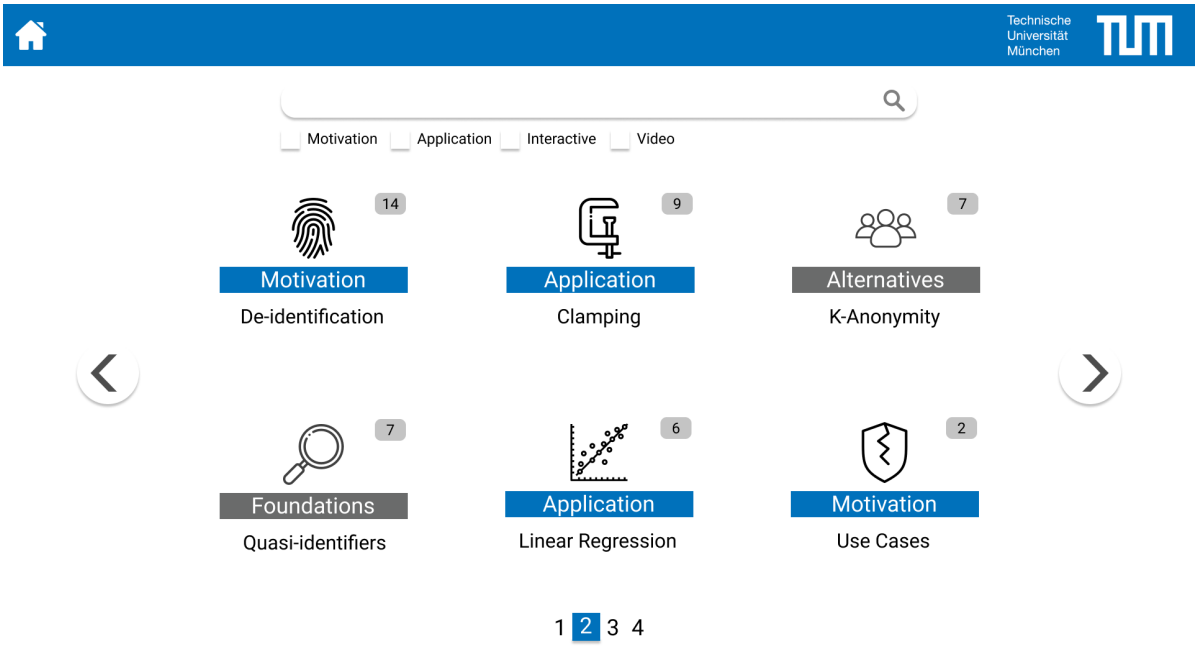


Figure 6.7: MockUp of the Landing Page

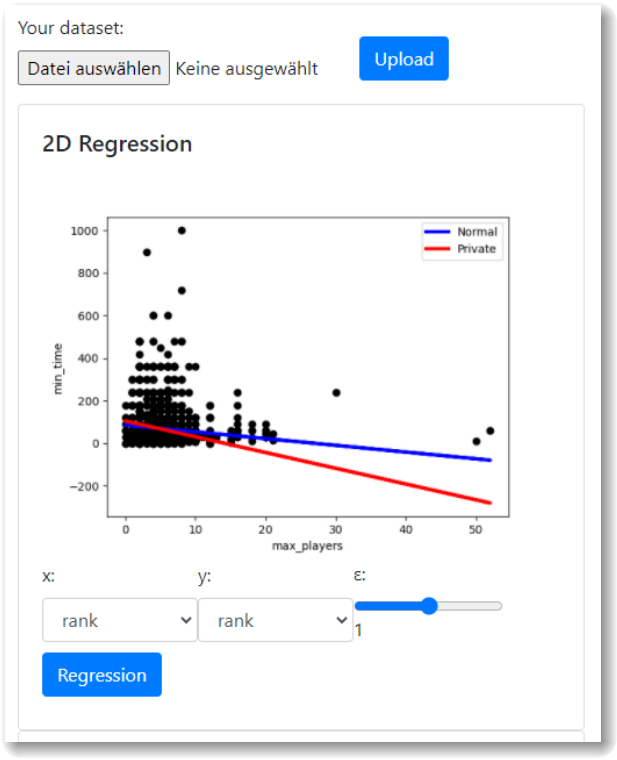


Figure 6.8: Prototype

7 Evaluation

The learning nuggets went through multiple evaluation steps. In an early stage, we had informal Think-Aloud tests with five students. We observed how they approached the learning nuggets, how long they stayed on each slide, and discussed the issues they had with the content. This first input primarily revolved around the complexity of the topic and led to the inclusion of more examples and visual elements. Second, the learning nuggets were reviewed by other researchers working on differential privacy to ensure their overall quality and correctness. Lastly, a formal evaluation with 17 participants was conducted. Due to time reasons they were only provided with a sample of seven learning nuggets:

1. Privacy Through Perturbation
2. The Randomized Response Algorithm
3. The Definition of Differential Privacy
4. The Laplace Mechanism
5. The Exponential Mechanism
6. The Benefits of Differential Privacy
7. The Challenges of Differential Privacy

The evaluation form consisted of 27 items, most of which were Likert scale questions with five response options ranging from strongly disagree to strongly agree. The questions were grouped into five categories:

1. General Information
2. Usability
3. Content
4. Acquired Competence
5. Personal Opinion

The questions about the usability and content were based on the "Learning Object Review Instrument (LORI)" proposed by Leacock and Nesbit [66]. Those about the acquired competence relied on self-assessment and were aligned to the six levels of the revised edition of Bloom's taxonomy [67]. In the following sections, we will discuss the results of each category.

General Information

The first set of questions assessed the profession or field of study of the participants, their prior knowledge, which learning nuggets they completed, how long it took them, and how they approached the learning nuggets, i.e., whether they skimmed them or read everything thoroughly. These questions helped us interpret individual responses. Clearly, someone who completed all learning nuggets and took 45 minutes for each of them is going to report a vastly different learning effect than someone who only briefly skimmed two or three of them. The profession and prior knowledge of the respondents might further influence their judgment.

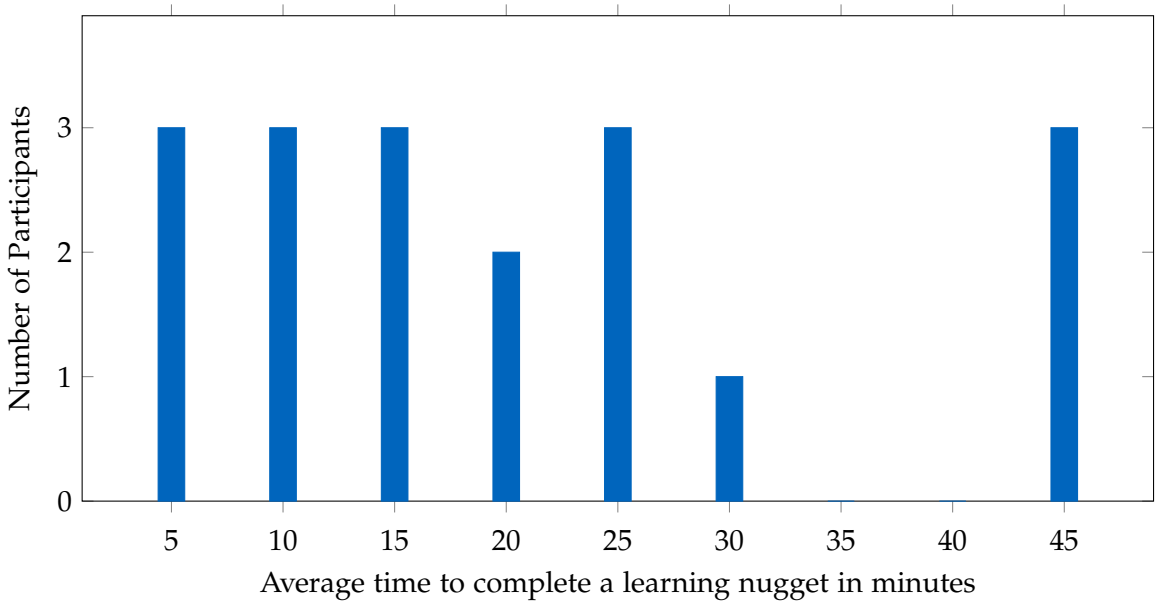


Figure 7.1: Evaluation Results - Average Time

The average time it took respondents to complete a learning nugget ranged from 5 to 45 minutes. Multiple factors can explain this significant difference. First, those who took less than 10 minutes reported that they have skipped complex formulas, proofs, and the exercises or only skimmed the text. Naturally, someone who does the exercises takes significantly longer. Second, it depended on the learning nuggets they completed. Whereas the benefits and challenges of differential privacy are easy to digest, the mechanisms are far more complex. Lastly, participants with prior knowledge of differential privacy have far less trouble understanding it.
We consider it an advantage that the learning nuggets can be completed in less than 10 minutes, but that someone who wants to understand the topic in depth can spend up to 45 minutes on them. That way, different audiences are addressed, which ties neatly into the philosophy behind our learning nuggets. By offering shortcuts like the key takeaways and optional content like examples and proofs, we encourage users to learn at their own paste and to study the topics they are most interested in, in greater detail. Nevertheless, we consider an

average completion time of 21 minutes to be too high. In future refinements existing learning nuggets should be split up and reduced in content.

Usability

Since it was safe to assume that everyone knew how to use a PowerPoint slide deck and there were only a few interactive elements embedded, the usability section primarily revolved around the structure and design of the slides, as most usability aspects become important once the learning platform is deployed [68].

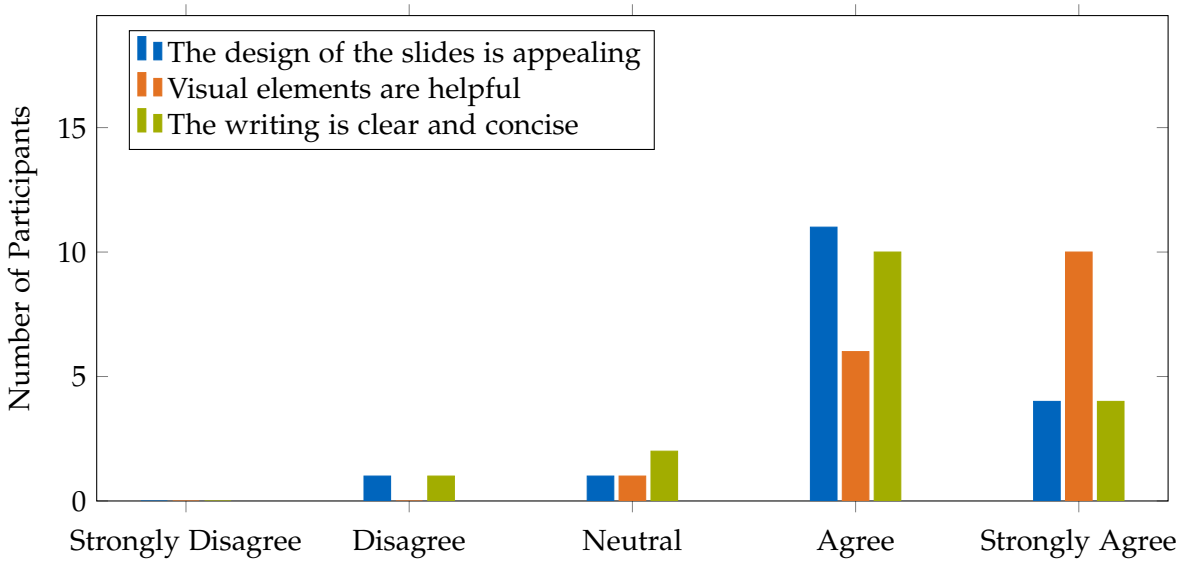


Figure 7.2: Evaluation Results - Usability

The design of the slides was predominantly perceived as appealing. Visual elements such as pictures, diagrams, and tables were widely appreciated and seen as useful aids. The writing was mostly seen as clear and concise, but still seems to deserve improvement.
We ended the section with an open question on how the design and usability could be improved. An overwhelming number of respondents suggested reducing the amount of content per slide and trimming the text. This must be considered in further refinements.

Content

Next up, we evaluated the content of the learning nuggets. First, we wanted to assess if the learning nuggets' scope was perceived as too broad or even too narrow and whether the content had an adequate level of detail. Thirteen participants reported that the scope of the learning nuggets was neither too broad nor too narrow. The remaining four reported that it was a bit too broad. This showed us that we are on a good track, but as the answers

to previous questions already indicated, it would be beneficial to reduce the content of the learning nuggets. The level of detail was well-received. Sixteen respondents agreed or strongly agreed that key points and significant ideas were emphasized with an appropriate level of detail.

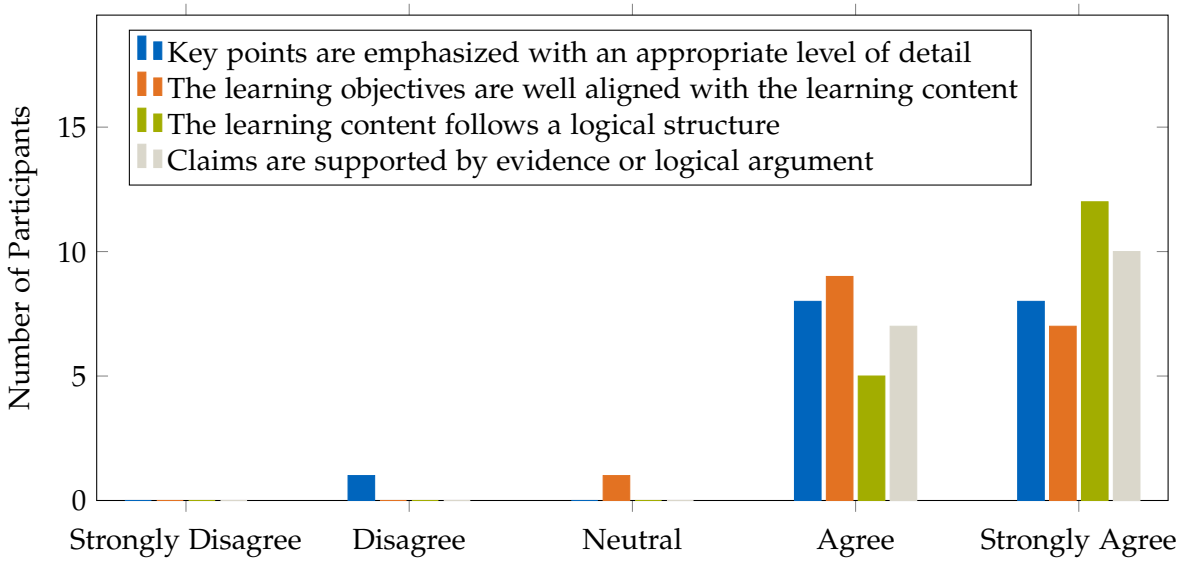


Figure 7.3: Evaluation Results - Content

Secondly, we wanted to assess whether the content was too complex and difficult to understand. However, we had to make sure that it was not the difficulty of the topic that was evaluated but whether the explanations about it were overly complicated. So instead of asking directly for the complexity of the content, we asked how it compared to other sources, i.e., if it was as simple as a blog post (1) or as complex as a scientific paper (5). Eight of the respondents placed it in between (3), and seven answered that it was almost as complex as an academic paper (4). We strive to be more formal than an ordinary blog, and, in the best case, a learner does not have to look into the scientific literature after completing our learning nugget. However, we do not intend to rival the scientific literature in terms of complexity. In future iterations learning nuggets should be kept shorter and simpler.

Thirdly, we focused on the overall quality of the content. The questions were tightly aligned with the characteristics of good content quality presented in LORI. All respondents agreed that the learning content followed a logical structure and that claims were supported by evidence or logical argument. The learning objectives were also considered to be well aligned with the learning content. Here we could strive for even tighter integration. Instead of revisiting them in the key takeaways, we could already highlight them on the respective slides.

Again, we ended with an open question about suggestions for improvement. Two respondents

emphasized that the content should be kept clear and simple. At the same time, one demanded more examples, subsections, and sub-key-takeaways, and another one that certain aspects are explained in more detail.

Acquired Competence

To capture the learning effect, we had to rely on self-assessment. The questions were aligned to the six levels of Bloom's Taxonomy.

Most of the respondents achieved the first two levels of Bloom's Taxonomy, remembering, and understanding. The participants were less confident that they could apply the methods and formulas covered in the learning nuggets. This can be partly explained by how they approached the learning nuggets. Nine reported that they ignored complex formulas, and six that they skipped the exercises. Nonetheless, this is another argument in favor of more examples.

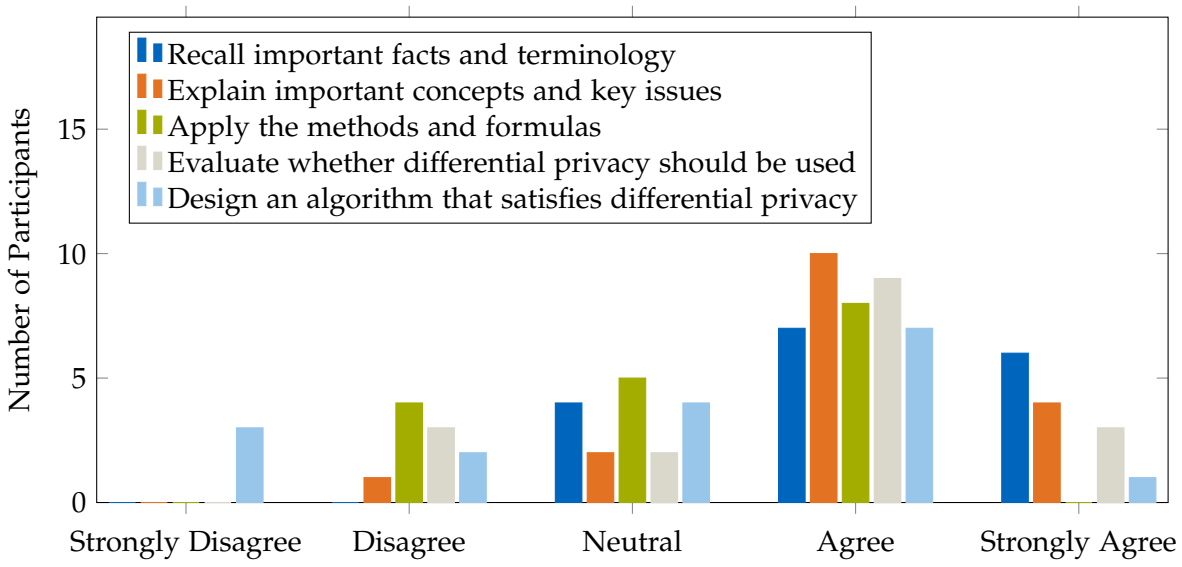


Figure 7.4: Evaluation Results - Acquired Competence 1/2

The answers diverged when it came to the higher levels of Bloom's Taxonomy. Only a few were skeptical that they could evaluate whether differential privacy should be used in a specific case. Considerably more doubted that they could design an algorithm that satisfies differential privacy. They might have underestimated themselves, as an algorithm that ignores the underlying dataset and returns random numbers would already suffice to satisfy differential privacy. However, we should address this gap with more exercises that require the learner to develop differentially private versions of known mechanisms.

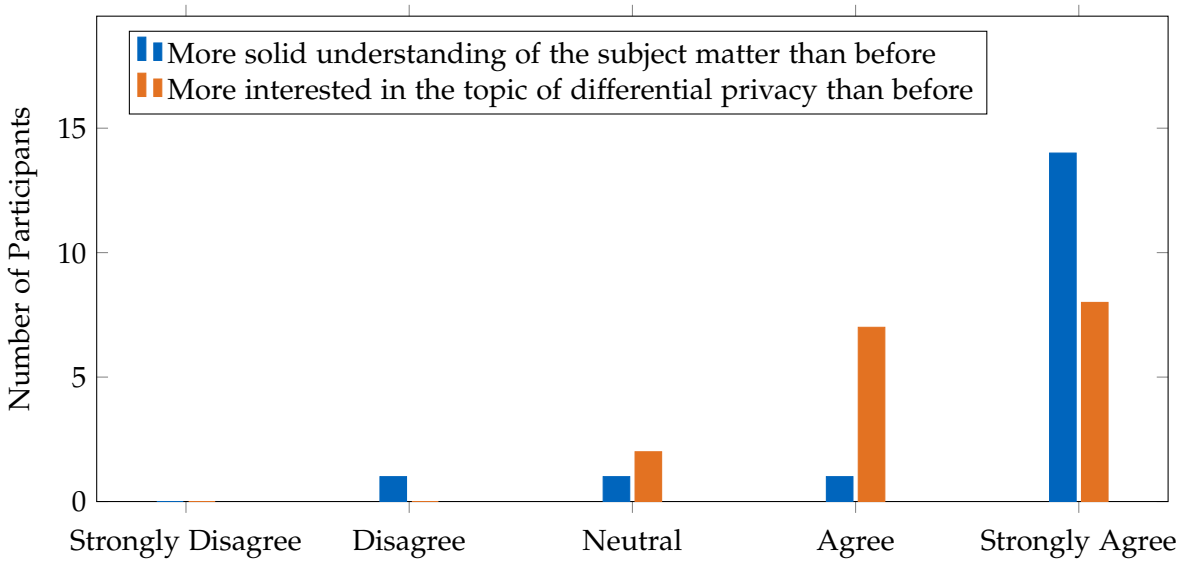


Figure 7.5: Evaluation Results - Acquired Competence 2/2

Overall, the respondents almost unanimously agreed that they now had a more solid understanding of differential privacy than before and were more interested in the subject. Outliers were those that already had prior knowledge about differential privacy.

Personal Opinion

In the end, we asked the participants what they liked about the learning nuggets and what should be improved. The structure of the slides, the examples and the exercises were praised. Suggested improvements were to reduce the text per slide and to add more visual elements.

Discussion

All in all, the learning nuggets were deemed helpful, led to a learning effect, and fulfilled common quality criteria. This confirms their overall usefulness and that they will be a good basis for future iterations. The evaluation also showed that they should be shortened and use even more examples, exercises, and visualizations. These are requirements the learning platform directly addresses. It will enrich the content with interactive elements that invite the users to try out the algorithms they read about. Next to interactive examples and visualizations, the exercises can also be designed to be more engaging, e.g., in the form of quizzes. The learning platform can also address the issue of overcrowded slides. Buttons that allow users to show or hide formulas, proofs, and examples prevent them from being overwhelmed by large amounts of content. The use of tooltips can further help to get rid of remarks, footnotes, and glossary slides.

Limitations

The evaluation had a limited scope, both in terms of the number of learning nuggets and the number of participants. There were too few learning nuggets to make out popular learning paths. Due to the limited number and diversity of participants, we could not conclude which topics different roles are most interested in. It was challenging to find enough participants, as studying the learning nuggets was a significant time investment. Additionally, we could not find respondents from our target audience, i.e., practitioners planning to use differential privacy. Moreover, the learning nuggets need to be evaluated together with the learning platform to see how they work in the environment they are supposed to be in. This will also enable us to monitor user behavior and to get more precise statistics about the time it takes to complete certain learning nuggets or about the most popular learning nuggets among different stakeholders.

8 Conclusion and Future Work

In this chapter, we will conclude the thesis, discuss the limitations of this work and the challenges we encountered. Further, an outlook on future work is given.

8.1 Conclusion

The goal of this thesis was to support diverse stakeholders to make informed decisions about the use of differential privacy. The core question being whether they should use differential privacy led to the first research question, assessing the benefits and challenges of differential privacy. With the second research question, "What are the most important topics practitioners should know about" we narrowed down the complexity of differential privacy to its key concepts. Finally, we researched how these topics could be conveyed in a convenient way to teach differential privacy to practitioners.

> **Research question 1:** What are benefits and challenges of using differential privacy?

We addressed this first research question by studying the grey literature and conducting expert interviews. We found that differential privacy can be rightfully considered to be a promising technique but is still in an early stage. Most benefits can be directly derived from the definition of differential privacy and thus come by default, independent of the use case or its implementation.

> **Research question 2:** What are the most important topics practitioners should know about?

The interviews also influenced our selection of important topics. The selection was further refined by looking into lecture slides of other universities and analyzing the most-taught topics. These topics encompass the definition of differential privacy, its properties, the core mechanisms, the different models, variations and extensions, and real-world applications and tools available for differential privacy.

> **Research question 3:** How can these topics be conveyed?

Among the biggest challenge was the complexity of differential privacy. To convey the most important topics on differential privacy in a less complex manner, we conceptualized a learning platform and designed several learning nuggets. They address different questions

practitioners might have and were grouped in five categories.

In the final evaluation, the learning nuggets were deemed helpful and led to a significant learning effect.

8.2 Limitations and Challenges

We already stated limitations and challenges in the respective chapters. We will briefly recap them and state overall challenges and limitations, paving the way for the outlook and future work.

Novelty of the Topic One of the biggest challenges was the novelty of differential privacy and the high research interest in it. When we started working on it, there were about 2800 scientific papers listed in Scopus; by the end, there were more than 3400. This increase of 600 papers is still a conservative estimate, as it only takes into account one literature database. The amount of grey literature grew even faster due to its shorter publication processes. Also, many of the university courses we evaluated started in Fall 2020, and new tools emerged during our work on differential privacy, such as SmartNoise, the joint project between Microsoft and Harvard, and Opacus from Facebook. This underscores the need for our learning content to be easily extensible, which is especially important for the learning nuggets from the outlook category.

The Maturity of Differential Privacy Differential privacy is still in an early stage, and many practitioners are not aware of it. This made it harder to find suitable interview partners and respondents for the evaluation. And even practitioners who already worked with differential privacy did so just for a short amount of time and might not have developed guidelines. As most tools were also rather new, there was little information that could be gained about them. There might even be challenges we are not aware of.
In the future, when more companies start looking into differential privacy, it would be interesting to conduct case studies and document their experiences with differential privacy.

Topic selection We assessed important topics by checking existing literature and courses on differential privacy for the topics they selected. More work is required on the needs of companies working with differential privacy. How to deduce the questions most important to them has to be evaluated in the future.

Legal Aspects We did not cover legal aspects. Of course, they have to be considered as well and might be among the most important arguments to decide for proper privacy-preserving methods.

8.3 Future Work

8.3.1 Learning content

In the future, the learning content should be extended to different levels of complexity. Many of the existing learning nuggets gave an overview of multiple methods or tools. However, a more in-depth analysis of the different tools available and learning nuggets grouped by use cases or areas of application, like recommender systems, location privacy, machine learning, would be beneficial. With more than 700 papers and tools in this area, those are really interesting for many practitioners. Further learning nuggets could provide an overview of the most important institutions and their work on differential privacy and short summaries of important literature to get a better overview of the field. As the basics of differential privacy are already covered, future iterations could focus stronger on application, including advanced algorithms. Finally, the content would also profit from including the legal aspects of privacy.

8.3.2 The Learning Platform

The learning platform conceptualized within this thesis was later refined and implemented by four students in a practical course. We discussed their progress in weekly sprint meetings and assisted them with differential privacy-related topics. Pictures of there current state can be found on the following pages. On the landing page (Fig. 8.1) users can filter the learning nuggets based on the categories they belong to, the time it takes to complete them or by searching for keywords. The "create a learning nugget"-view (Fig. 8.2) offers users a rich text editor with the functionality to add text, tables, pictures and formulas. It even allows users to create their on interactive graphs or tables. The prototype that was developed within this thesis also made its way into the application (Fig. 8.3).

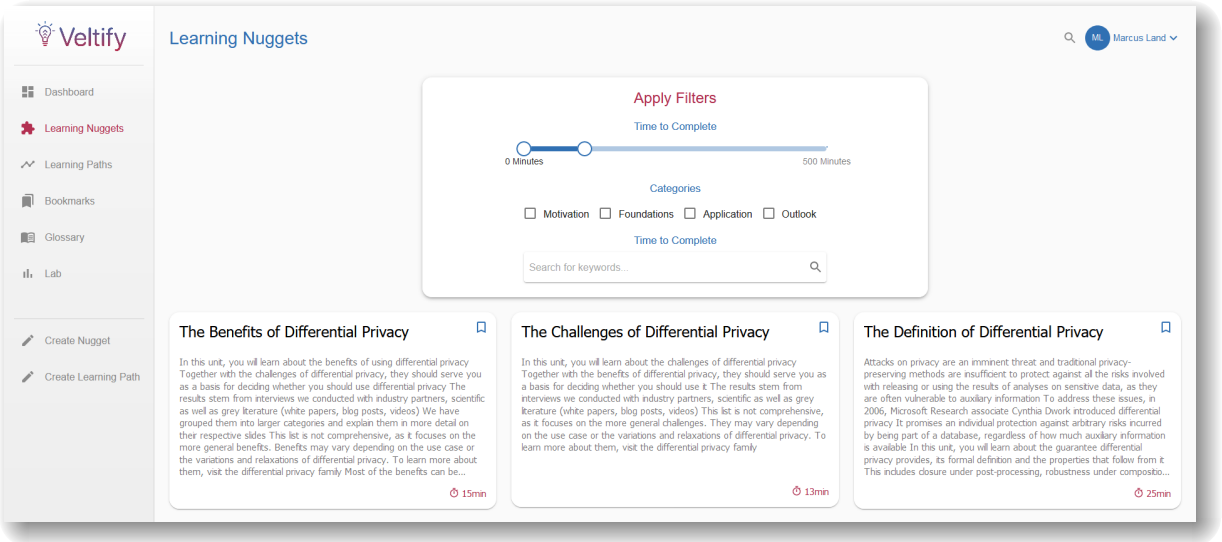


Figure 8.1: Landing Page

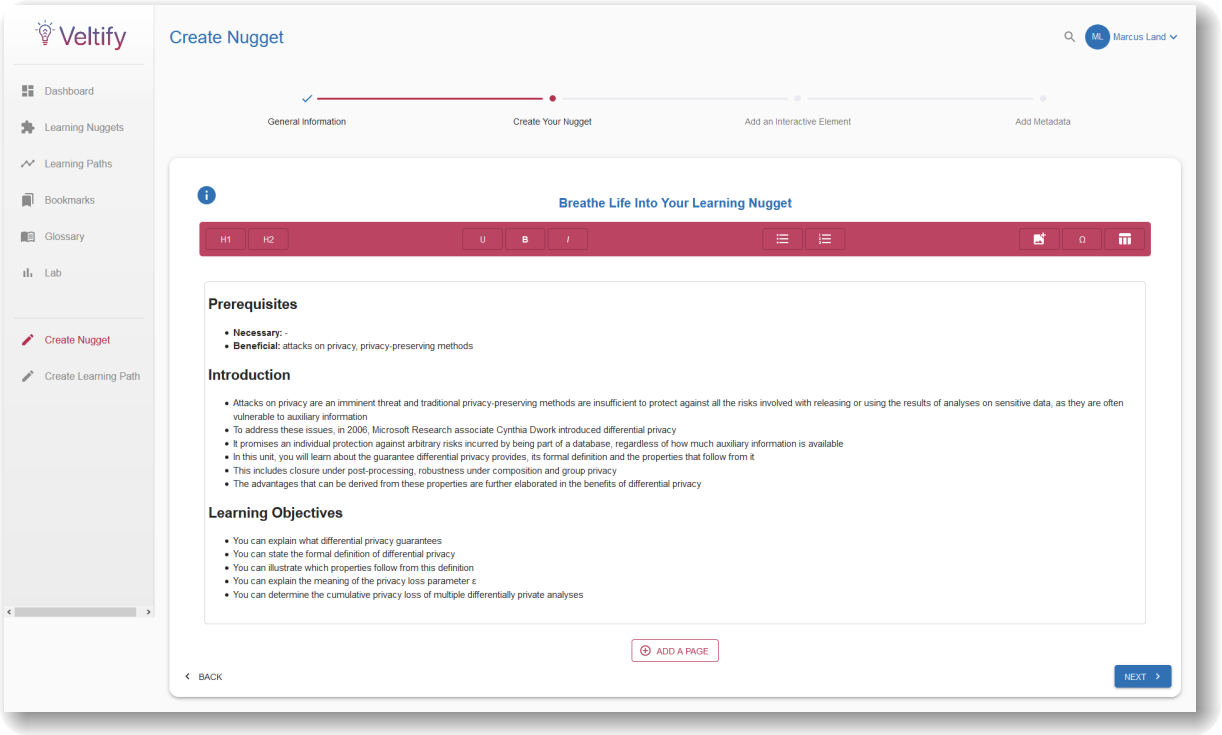


Figure 8.2: Create Learning Nugget

E-Learning Features

Another important aspect will be to further make use of the benefits e-learning offers and provide quizzes. As the backend of the learning platform is written in Python and the differential privacy library of IBM is already embedded, programming exercises directly in the learning platform are possible.

Recommendation Algorithm

A key element that is still missing in the learning platform is a recommendation algorithm that provides learners with learning nuggets that fit to their role, prior knowledge, and learning goals. The four students adapted our prototypical implementation.

Collaborative Tools

The exchange among students is crucial in learning. Communication among learners would be great. And as basic CRUD functionalities and an editor are already implemented, this should not be too hard. Especially in the discussion category, under learning nuggets like: "the challenges of differential privacy," this could lead to great discussions and new input. It could become a place to share experiences they made with differential privacy.

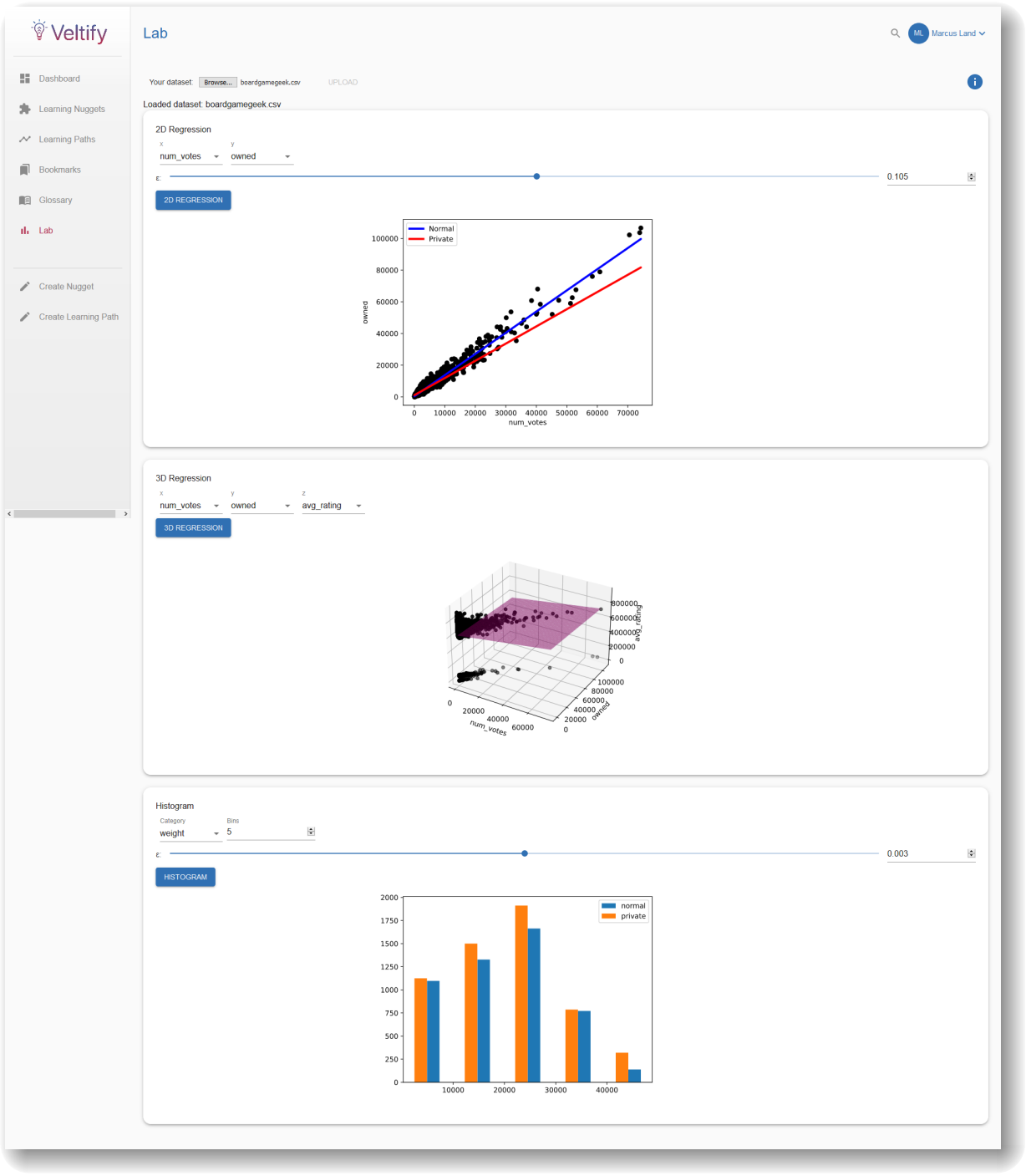


Figure 8.3: Differential Privacy Lab

List of Figures

List of Tables

Bibliography

[1] A. Narayanan and V. Shmatikov. "How To Break Anonymity of the Netflix Prize Dataset". In: December 1999 (2006), pp. 1–10. arXiv: 0610105 [cs]. URL: http://arxiv.org/abs/cs/0610105.

[2] A. Narayanan and V. Shmatikov. "Robust de-anonymization of large sparse datasets". In: *Proceedings - IEEE Symposium on Security and Privacy*. 2008. ISBN: 9780769531687. DOI: 10.1109/SP.2008.33.

[3] L. Sweeney. "Simple demographics often identify people uniquely". In: *Carnegie Mellon University, Data Privacy Working Paper 3. Pittsburgh 2000* (2000).

[4] J. M. Abowd. *Tweetorial: Reconstruction-abetted re-identification attacks and other traditional vulnerabilities*. 2019. URL: http://blogs.cornell.edu/abowd/special-materials/245-2/ (visited on 01/31/2021).

[5] S. Garfinkel, J. M. Abowd, and C. Martindale. "Understanding database reconstruction attacks on public data". In: *Communications of the ACM* (2019). ISSN: 15577317. DOI: 10.1145/3287287.

[6] I. Dinur and K. Nissim. "Revealing Information while Preserving Privacy". In: *Proceedings of the ACM SIGACT-SIGMOD-SIGART Symposium on Principles of Database Systems*. 2003. DOI: 10.1145/773153.773173.

[7] C. Dwork. "Differential privacy". In: *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*. 2006. ISBN: 3540359079. DOI: 10.1007/11787006_1.

[8] C. Radebaugh and U. Erlingsson. "Introducing TensorFlow Privacy: Learning with Differential Privacy for Training Data". In: *Medium* (2019).

[9] T. Ryffel, A. Trask, M. Dahl, B. Wagner, J. Mancuso, D. Rueckert, and J. Passerat-Palmbach. *A generic framework for privacy preserving deep learning*. 2018. arXiv: 1811.04017.

[10] D. Testuggine and I. Mironov. *Introducing Opacus: A high-speed library for training PyTorch models with differential privacy*. 2020. URL: https://ai.facebook.com/blog/introducing-opacus-a-high-speed-library-for-training-pytorch-models-with-differential-privacy/ (visited on 01/30/2021).

[11] F. McSherry. "Privacy integrated queries". In: *Communications of the ACM* (2010). ISSN: 0001-0782. DOI: 10.1145/1810891.1810916.

[12] S. Kessler, J. Hoff, and J. C. Freytag. "SAP HANA goes private - From privacy research to privacy aware enterprise analytics". In: *Proceedings of the VLDB Endowment*. 2018. DOI: 10.14778/3352063.3352119.

[13] J. Eicher, R. Bild, H. Spengler, K. A. Kuhn, and F. Prasser. "A comprehensive tool for creating and evaluating privacy-preserving biomedical prediction models". In: *BMC Medical Informatics and Decision Making* (2020). ISSN: 14726947. DOI: 10.1186/s12911-020-1041-3.

[14] *SmartNoise*. 2020. URL: https://smartnoise.org/ (visited on 02/14/2021).

[15] N. Holohan, P. M. Aonghusa, S. Braghin, and K. Levacher. *Diffprivlib: The IBM differential privacy library*. 2019. arXiv: 1907.02444.

[16] M. Hay, A. Machanavajjhala, G. Miklau, Y. Chen, and D. Zhang. "Principled evaluation of differentially private algorithms using DPBENCH". In: *Proceedings of the ACM SIGMOD International Conference on Management of Data*. 2016. ISBN: 9781450335317. DOI: 10.1145/2882903.2882931. arXiv: 1512.04817.

[17] D. Desfontaines and B. Pejó. "SoK: Differential privacies". In: *Proceedings on Privacy Enhancing Technologies* (2020). DOI: 10.2478/popets-2020-0028. arXiv: 1906.01337.

[18] A. R. Hevner, S. T. March, J. Park, and S. Ram. "Design science in information systems research". In: *MIS Quarterly: Management Information Systems* (2004). ISSN: 02767783. DOI: 10.2307/25148625.

[19] K. Peffers, T. Tuunanen, M. A. Rothenberger, and S. Chatterjee. "A design science research methodology for information systems research". In: *Journal of Management Information Systems* (2007). ISSN: 07421222. DOI: 10.2753/MIS0742-1222240302.

[20] P. Samarati and L. Sweeney. "Protecting Privacy when Disclosing Information: k-Anonymity and its Enforcement Through Generalization and Suppresion." In: *Proc of the IEEE Symposium on Research in Security and Privacy* (1998).

[21] A. Meyerson and R. Williams. "On the complexity of optimal k-anonymity". In: *Proceedings of the ACM SIGACT-SIGMOD-SIGART Symposium on Principles of Database Systems*. 2004. DOI: 10.1145/1055558.1055591.

[22] V. Ayala-Rivera, P. McDonagh, T. Cerqueus, and L. Murphy. "A Systematic comparison and evaluation of k-Anonymization algorithms for practitioners". In: *Transactions on Data Privacy* (2014). ISSN: 20131631.

[23] A. Gkoulalas-Divanis, G. Loukides, and J. Sun. *Publishing data from electronic health records while preserving privacy: A survey of algorithms*. 2014. DOI: 10.1016/j.jbi.2014.06.002.

[24] C. C. Aggarwal. "On k-anonymity and the curse of dimensionality". In: *VLDB 2005 - Proceedings of 31st International Conference on Very Large Data Bases*. 2005. ISBN: 1595931546.

[25] A. Machanavajjhala, D. Kifer, J. Gehrke, and M. Venkitasubramaniam. "L - diversity". In: *ACM Transactions on Knowledge Discovery from Data* (2007). ISSN: 15564681. DOI: 10.1145/1217299.1217302.

[26] L. Ninghui, L. Tiancheng, and S. Venkatasubramanian. "t-Closeness: Privacy beyond k-anonymity and -diversity". In: *Proceedings - International Conference on Data Engineering*. 2007. ISBN: 1424408032. DOI: 10.1109/ICDE.2007.367856.

[27] S. R. Ganta, S. P. Kasiviswanathan, and A. Smith. "Composition attacks and auxiliary information in data privacy". In: *Proceedings of the ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*. 2008. ISBN: 9781605581934. DOI: 10.1145/1401890.1401926. arXiv: 0803.0032.

[28] K. Nissim and A. Wood. "Is privacy privacy?" In: *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences* 376.2128 (2018), pp. 1–19. ISSN: 1364503X. DOI: 10.1098/rsta.2017.0358.

[29] C. Dwork. "The promise of differential privacy: A tutorial on algorithmic techniques". In: *Proceedings - Annual IEEE Symposium on Foundations of Computer Science, FOCS*. 2011. ISBN: 9780769545714. DOI: 10.1109/FOCS.2011.88.

[30] A. Wood, M. Altman, A. Bembenek, M. Bun, M. Gaboardi, J. Honaker, K. Nissim, D. O'Brien, T. Steinke, and S. Vadhan. "Differential Privacy: A Primer for a Non-Technical Audience". In: *SSRN Electronic Journal* (2019). ISSN: 1556-5068. DOI: 10.2139/ssrn.3338027.

[31] C. Dwork and A. Roth. "The algorithmic foundations of differential privacy". In: *Foundations and Trends in Theoretical Computer Science* (2013), p. 7. ISSN: 15513068. DOI: 10.1561/0400000042.

[32] S. L. Warner. "Randomized Response: A Survey Technique for Eliminating Evasive Answer Bias". In: *Journal of the American Statistical Association* (1965). ISSN: 1537274X. DOI: 10.1080/01621459.1965.10480775.

[33] C. Dwork, F. McSherry, K. Nissim, and A. Smith. "Calibrating noise to sensitivity in private data analysis". In: *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*. 2006. ISBN: 3540327312. DOI: 10.1007/11681878_14.

[34] C. Dwork. "Differential Privacy: A Survey of Results". In: *Theory and Applications of Models of Computation*. 2008. DOI: 10.1007/978-3-540-79228-4_1.

[35] C. Dwork and J. Lei. "Differential privacy and robust statistics". In: *Proceedings of the Annual ACM Symposium on Theory of Computing*. 2009. ISBN: 9781605585062. DOI: 10.1145/1536414.1536466.

[36] C. Dwork and A. Smith. "Differential Privacy for Statistics: What we Know and What we Want to Learn". In: *Journal of Privacy and Confidentiality* (2010). DOI: 10.29012/jpc.v1i2.570.

[37] C. Dwork, F. Mcsherry, K. Nissim, and A. Smith. "Differential Privacy – A Primer for the Perplexed". In: *Joint UNECE/Eurostat work session on statistical data confidentiality* (2011).

[38] A. Roth. "New Algorithms for Preserving Differential Privacy". In: (2010).

[39] M. Gaboardi. *Topics in Differential Privacy*. 2017. URL: http://cs-people.bu.edu/gaboardi/teaching/CSE660-fall17.html (visited on 02/14/2021).

[40] G. Kamath. *Algorithms for Private Data Analysis*. 2020. URL: http://www.gautamkamath.com/CS860-fa2020.html (visited on 02/14/2021).

[41] G. Rothblum. *Foundations of Privacy in Data Analysis, Fall 2017*. 2017. URL: https://guyrothblum.wordpress.com/foundations-of-privacy-in-data-analysis-fall-2017/ (visited on 02/14/2021).

[42] H. Kaplan. *Seminar on differential privacy*. 2019. URL: https://www.cs.tau.ac.il/%7B~%7Dhaimk/privacy-seminar/main-page.html.

[43] N. Johnson, J. P. Near, and D. Song. "Towards practical differential privacy for SQL queries". In: *Proceedings of the VLDB Endowment*. 2018. DOI: 10.1145/3177732.3177733. arXiv: 1706.09479.

[44] N. Johnson, J. P. Near, J. M. Hellerstein, and D. Song. "Chorus: A Programming Framework for Building Scalable Differential Privacy Mechanisms". In: *Proceedings - 5th IEEE European Symposium on Security and Privacy, Euro S and P 2020*. 2020. ISBN: 9781728150871. DOI: 10.1109/EuroSP48549.2020.00041.

[45] G. Kamath. *A Course In Differential Privacy [Video]*. 2020. URL: https://www.youtube.com/watch?v=FJMjNOcIqkc%7B%5C&%7Dlist=PLmd%7B%5C_%7DzeMNzSvRRNpoEWkVo6QY%7B%5C_%7D6rR3SHjp (visited on 02/14/2021).

[46] Ú. Erlingsson, V. Pihur, and A. Korolova. "RAPPOR: Randomized aggregatable privacy-preserving ordinal response". In: *Proceedings of the ACM Conference on Computer and Communications Security*. 2014. ISBN: 9781450329576. DOI: 10.1145/2660267.2660348. arXiv: 1407.6981.

[47] S. Bird. *Putting differential privacy into practice to use data responsibly*. 2020. URL: https://blogs.microsoft.com/ai-for-business/differential-privacy/ (visited on 01/31/2021).

[48] B. Ding, J. Kulkarni, and S. Yekhanin. "Collecting telemetry data privately". In: *Advances in Neural Information Processing Systems*. 2017. arXiv: 1712.01524.

[49] R. Rogers, S. Subramaniam, S. Peng, D. Durfee, S. Lee, S. K. Kancha, S. Sahay, and P. Ahammad. *LinkedIn's audience engagements API: A privacy preserving data analytics system at scale*. 2020. arXiv: 2002.05839.

[50] *Apple Differential Privacy Technical Overview*. URL: https://www.apple.com/privacy/docs/Differential%7B%5C_%7DPrivacy%7B%5C_%7DOverview.pdf.

[51] T. Diethe, O. Feyisetan, B. Balle, and T. Drake. "Preserving privacy in analyses of textual data". In: *CEUR Workshop Proceedings* 2573 (2020), pp. 1–3. ISSN: 16130073.

[52] *Protecting data privacy*. 2018. URL: https://www.aboutamazon.com/news/amazon-ai/protecting-data-privacy (visited on 01/31/2021).

[53] B. Balle and Y. X. Wang. "Improving the Gaussian mechanism for differential privacy: Analytical calibration and optimal denoising". In: *35th International Conference on Machine Learning, ICML 2018*. 2018. ISBN: 9781510867963. arXiv: 1805.06530.

[54] V. Pihur, A. Korolova, F. Liu, S. Sankuratripati, M. Yung, D. Huang, and R. Zeng. *Differentially-private "draw and discard" machine learning*. 2018. arXiv: 1807.04369.

[55] A. Machanavajjhala, D. Kifer, J. Abowd, J. Gehrke, and L. Vilhuber. "Privacy: Theory meets practice on the map". In: *Proceedings - International Conference on Data Engineering*. 2008. ISBN: 9781424418374. DOI: 10.1109/ICDE.2008.4497436.

[56] A. D. Foote, A. Machanavajjhala, and K. McKinney. "Releasing Earnings Distributions using Differential Privacy". In: *Journal of Privacy and Confidentiality* (2019). DOI: 10.29012/jpc.722.

[57] G. Long. *Formal Privacy Methods for the 2020 Census*. Tech. rep. 2020. URL: https://www2.census.gov/programs-surveys/decennial/2020/program-management/planning-docs/privacy-methods-2020-census.pdf.

[58] S. L. Garfinkel, J. M. Abowd, and S. Powazek. "Issues encountered deploying differential privacy". In: *Proceedings of the ACM Conference on Computer and Communications Security*. 2018. ISBN: 9781450359894. DOI: 10.1145/3267323.3268949. arXiv: 1809.02201.

[59] S. L. Garfinkel and P. Leclerc. "Randomness Concerns when Deploying Differential Privacy". In: *WPES 2020 - Proceedings of the 19th Workshop on Privacy in the Electronic Society*. 2020. ISBN: 9781450380867. DOI: 10.1145/3411497.3420211. arXiv: 2009.03777.

[60] Michael B. Hawes. "Implementing Differential Privacy: Seven Lessons From the 2020 United States Census". In: *Harvard Data Science Review* (2020). DOI: 10.1162/99608f92.353c6f99.

[61] J. Zhang, G. Cormode, C. M. Procopiuc, D. Srivastava, and X. Xiao. "PrivBayes: Private data release via Bayesian networks". In: *Proceedings of the ACM SIGMOD International Conference on Management of Data*. 2014. ISBN: 9781450323765. DOI: 10.1145/2588555.2588573.

[62] B. Woo, B. Willemsen, and V. Baker. *Hype Cycle for Privacy, 2020*. Tech. rep. 2020, pp. 13–14. URL: https://www.gartner.com/en/documents/3987903/hype-cycle-for-privacy-2020.

[63] I. Wagner and D. Eckhoff. *Technical privacy metrics: A systematic survey*. 2018. DOI: 10.1145/3168389.

[64] Z. Ji, Z. C. Lipton, and C. Elkan. "Differential Privacy and Machine Learning: a Survey and Review". In: (2014), pp. 1–30. arXiv: 1412.7584. URL: http://arxiv.org/abs/1412.7584.

[65] Z. Jorgensen, T. Yu, and G. Cormode. "Conservative or liberal? Personalized differential privacy". In: *Proceedings - International Conference on Data Engineering*. 2015. ISBN: 9781479979639. DOI: 10.1109/ICDE.2015.7113353.

[66] T. L. Leacock and J. C. Nesbit. *A framework for evaluating the quality of multimedia learning resources*. 2007.

[67] D. R. Krathwohl. "A Revision Of Bloom's Taxonomy Of Educational Objectives". In: *Theory into Practice* (2002). ISSN: 00405841. arXiv: arXiv:1202.2745v1.

[68] T. C. Reeves, L. Benson, D. Elliott, M. Grant, D. Holschuh, B. Kim, H. Kim, E. Lauber, and C. S. Loh. "Usability and Instructional Design Heuristics for E-Learning Evaluation." In: *World Conference on Educational Multimedia, Hypermedia and Telecommunications* (2002).