

# Differential Privacy

## Datenschutz mit technischen Mitteln

### Bachelorarbeit

Bachelor of Science in  
Information & Cyber Security

Ausgeführt von:

**Joshua Drexel**

18-274-126

Betreuungsperson: Prof. Dr. Esther Hänggi

Experte: Urs Rufer

Rotkreuz, 10. Juni 2022

## **Bachelorarbeit an der Hochschule Luzern – Informatik**

**Titel:** Differential Privacy – Datenschutz mit technischen Mitteln

**Student:** Joshua Drexel

**Studiengang:** BSc Information & Cyber Security

**Jahr:** 2022

**Betreuungsperson:** Prof. Dr. Esther Hänggi

**Experte:** Urs Rufer

**Auftraggeber:** Hochschule Luzern – Informatik

### **Codierung / Klassifizierung der Arbeit:**

- Öffentlich
- Vertraulich

**Eidesstattliche Erklärung** Ich erkläre hiermit, dass ich die vorliegende Arbeit selbständig und ohne unerlaubte fremde Hilfe angefertigt habe, alle verwendeten Quellen, Literatur und andere Hilfsmittel angegeben habe, wörtlich oder inhaltlich entnommene Stellen als solche kenntlich gemacht habe, das Vertraulichkeitsinteresse des Auftraggebers wahren und die Urheberrechtsbestimmungen der Hochschule Luzern respektieren werde.

Winterthur, 10. Juni 2022: \_\_\_\_\_

**Abgabe der Arbeit auf der Portfolio Datenbank:**

Bestätigungsvisum Studentin/Student

Ich bestätige, dass ich die Bachelorarbeit korrekt gemäss Merkblatt auf der Portfolio Datenbank abgelegt habe. Die Verantwortlichkeit sowie die Berechtigungen habe ich abgegeben, so dass ich keine Änderungen mehr vornehmen kann oder weitere Dateien hochladen kann.

Winterthur, 10. Juni 2022: \_\_\_\_\_

**Verdankung**

Ich möchte mich herzlich bei Prof. Dr. Esther Hänggi für die gute Zusammenarbeit und Betreuung der Arbeit bedanken.

Weiter gilt mein Dank Cyrill, Matteo, Michael, Ramona und Thomas für die Teilnahme an der Testdurchführung der Übung. Eure Feedbacks waren sehr wertvoll für diese Arbeit!

Abschliessend möchte ich mich bei meinen Eltern und meiner Freundin bedanken. Sie haben mich im gesamten Studium und auch in dieser Arbeit begleitet und unterstützt.

**Hinweis:** Geistiges Eigentum gemäss der Studienordnung für die Ausbildung an der Hochschule Luzern, FH Zentralschweiz

# Abstract

Diese Arbeit befasst sich mit der Ausarbeitung einer Übung, anhand welcher die Grundlagen der Differential Privacy erlernt und mithilfe von praktischen Aufgaben vertieft werden können.

Die für ein Grundverständnis notwendigen theoretischen Grundlagen der Differential Privacy werden beschrieben. Die Definitionen der reinen und der annähernden Differential Privacy werden erklärt und das Konzept des Privacy Budgets in der Tiefe behandelt. Zudem werden weitere Konzepte erläutert, welche für das Grundverständnis der Differential Privacy notwendig sind, wie beispielsweise die Bestimmung der Sensitivität von Funktionen. Es werden drei Mechanismen für die Implementierung von Differential Privacy erklärt und die verschiedenen Anwendungsfälle sowie die Grenzen der Differential Privacy aufgezeigt.

Basierend auf den erarbeiteten theoretischen Grundlagen wird der Lerninhalt der Übung bestimmt und die Lernziele formuliert. Die Lernziele bilden die Grundlage für die Übungsumsetzung. In der Übung wird die Theorie zum Verständnis der Differential Privacy vermittelt und durch praktische Aufgaben ergänzt. Zudem sind in der Übung Implementierungen des Laplace, des Gauss sowie des exponentiellen Mechanismus umgesetzt. Die Codebeispiele können ausgeführt und dadurch die Auswirkungen der verschiedenen Parameter beobachtet werden.

Durch eine Testdurchführung der Übung werden die Arbeitsergebnisse überprüft. Anhand der Rückmeldungen der Übungsteilnehmenden der Testdurchführung kann die Lernzielerreichung, das Verständnis der Inhalte sowie die Vollständigkeit der Übung überprüft werden.

Es ist angedacht, die erarbeitete Übung innerhalb der Hochschule Luzern zur Verfügung zu stellen. Anhand der Übung sollen sich Studierende, Forschende und Dozierende in die Welt der Differential Privacy einarbeiten können.

# Inhaltsverzeichnis

<b>1 Vision</b>	<b>1</b>
1.1 Klassischer Ansatz der Datenanonymisierung . . . . .	2
1.2 Die Bedeutung von Differential Privacy . . . . .	4
1.3 Ziele dieser Arbeit . . . . .	5
1.4 Nomenklatur . . . . .	5
<b>2 Stand der Praxis</b>	<b>6</b>
2.1 Die Grundidee von Differential Privacy . . . . .	7
2.2 Die formale Definition von Differential Privacy . . . . .	8
2.3 Das Privacy-Budget . . . . .	9
2.3.1 Quantifizierung des Angreiferwissens . . . . .	10
2.3.2 Komposition des Privacy-Budgets . . . . .	11
2.3.3 Das Privacy-Budget in der Praxis . . . . .	12
2.4 Die Sensitivität . . . . .	13
2.5 Implementierung von Differential Privacy . . . . .	15
2.5.1 Der Laplace Mechanismus . . . . .	15
2.5.2 Der Gauss Mechanismus . . . . .	16
2.5.3 Vergleich des Laplace und des Gauss Mechanismus . . . . .	17
2.5.4 Der exponentielle Mechanismus . . . . .	18
2.6 Anwendungsfälle von Differential Privacy . . . . .	20
2.7 Grenzen der Differential Privacy . . . . .	21
<b>3 Ideen und Konzepte</b>	<b>23</b>
3.1 Ideen für die Übungsart . . . . .	23
3.1.1 Freie Laborübung . . . . .	23
3.1.2 Geführte Laborübung . . . . .	23
3.1.3 Zeitbasierte Challenge . . . . .	24
3.1.4 Inhaltbasierte Challenge . . . . .	24
3.2 Ideen für die Umsetzung der Übungsumgebung . . . . .	24
<b>4 Methoden</b>	<b>25</b>
4.1 Projektmethodik . . . . .	25

4.2	Fachmethodik . . . . .	27
4.2.1	Methoden für die Übungsgestaltung . . . . .	27
4.2.2	Methoden für die Überprüfung der Zielerreichung . . . . .	28
<b>5</b>	<b>Realisierung</b>	<b>30</b>
5.1	Gestaltung der Übung . . . . .	30
5.1.1	Voraussetzungen klären . . . . .	30
5.1.2	Lerninhalt bestimmen . . . . .	31
5.1.3	Lernziele setzen . . . . .	31
5.1.4	Lehr-Lernform finden . . . . .	33
5.2	Ausarbeitung der Übung . . . . .	34
5.2.1	Notebook 0: Übersicht und Inhalt . . . . .	34
5.2.2	Notebook 1: Einführung in die Thematik . . . . .	34
5.2.3	Notebook 2: Die Definition der Differential Privacy . . . . .	35
5.2.4	Notebook 3: Der Laplace Mechanismus . . . . .	36
5.2.5	Notebook 4: Der Gauss Mechanismus . . . . .	36
5.2.6	Notebook 5: Der exponentielle Mechanismus . . . . .	37
5.2.7	Notebook 6: Anwendungsfälle und Grenzen der Differential Privacy . . . . .	38
5.2.8	Notebook 7: Musterlösungen der Übungen . . . . .	39
5.3	Bereitstellung der Übung . . . . .	39
5.4	Testdurchführung der Übung . . . . .	40
5.4.1	Zusammensetzung der Testgruppe . . . . .	40
5.4.2	Ausgestaltung der Testdurchführung . . . . .	41
5.4.3	Auswertung der Testdurchführung . . . . .	41
<b>6</b>	<b>Evaluation und Validation</b>	<b>43</b>
6.1	Beurteilung Ziel 1: Verständnis der Thematik . . . . .	43
6.2	Beurteilung Ziel 2: Vermittlung der Theorie . . . . .	44
6.3	Beurteilung Ziel 3: Praktische Übung . . . . .	44
<b>7</b>	<b>Ausblick</b>	<b>45</b>
7.1	Persönliche Reflexion . . . . .	45
7.2	Ideen für mögliche Ergänzungen . . . . .	47
7.3	Projektabschluss . . . . .	47
<b>8</b>	<b>Anhang</b>	<b>48</b>
8.1	Aufgabenstellung der Bachelorarbeit . . . . .	48
8.1.1	Ausgangslage und Problemstellung . . . . .	48
8.1.2	Ziel der Arbeit und erwartete Resultate . . . . .	50
8.1.3	Gewünschte Methoden, Vorgehen . . . . .	50
8.1.4	Kreativität, Varianten, Innovation . . . . .	51

Inhaltsverzeichnis

---

<b>9</b>	<b>Abbildungsverzeichnis</b>	<b>52</b>
<b>10</b>	<b>Tabellenverzeichnis</b>	<b>53</b>
<b>11</b>	<b>Literaturverzeichnis</b>	<b>54</b>

# 1 Vision

Bendig vom Frauenhofer-Verbund IUK-Technologie beschreibt Daten als unendlich wertvoll, viel wertvoller als Gold oder Öl [1]. Dies ist den vielseitigen Verwendungszwecken zu verdanken. So können Daten einmalig gesammelt und anschliessend mehrfach verkauft und verwertet werden. Bendig beschreibt weiter den Vorteil, wenn Daten geteilt und gemeinsam genutzt werden. Dieselben Daten müssen nicht mehrfach erhoben oder errechnet werden, was auch einen ökologischen Nutzen mit sich bringen kann. Dies zeigt den grossen Mehrwert der Sammlung, Verarbeitung und dem Teilen von Daten und substantiiert das Geschäftsmodell mit den Daten. [1]

Im internationalen Pakt über bürgerliche und politische Rechte der Vereinten Nationen (UNO-Pakt II), welchem die Schweiz am 18. Juni 1992 beigetreten ist, wird der Schutz der Privacy gesetzlich auf oberster Stufe verankert [2]. Mit dem Bundesgesetz über den Datenschutz (DSG) hat die Schweiz ein nationales Gesetz verabschiedet. So muss jede Verarbeitung von Personendaten dem DSG entsprechen. Gemäss Art. 22 Abs. 1 DSG «dürfen Bundesorgane Personendaten für nicht personenbezogene Zwecke, insbesondere für Forschung, Planung und Statistik bearbeiten, wenn:

- die Daten **anonymisiert** werden, sobald es der Zweck des Bearbeitens erlaubt;
- der Empfänger die Daten nur mit Zustimmung des Bundesorgans weitergibt; und
- die Ergebnisse so veröffentlicht werden, dass die betroffenen Personen **nicht bestimmbar** sind. » [3]

Um die gesetzlichen Grundlagen zum Schutz der Privacy einzuhalten, sind Organisationen daran interessiert, gesammelte Daten zu anonymisieren und so aufzubereiten, dass Individuen nicht mehr bestimmbar sind.

Sweeney zeigte in ihrer Arbeit an der Carnegie Mellon University, dass dies keine einfache Aufgabe ist [4]. Gemäss Sweeney lassen sich Menschen oftmals anhand von einfachen demografischen Daten eindeutig identifizieren [4]. Sweeney zeigt, dass 87% der US-amerikanischen Bürger\*innen alleine über die Postleitzahl, das Geschlecht und das Geburtsdatum eindeutig identifizierbar sind [4]. Weiter sind rund die Hälfte der US-amerikanischen Bürger\*innen über den Ort (z.B. Stadt oder Gemeinde), das Geschlecht und das Geburtsdatum identifizierbar [4]. Verlinken ist eine Technik, mit welcher mehrere Datensammlun-

gen so zusammengeführt werden, dass die eigentlich anonymisierte Datensammlung de-anonymisiert werden kann. Sweeney konnte eine anonymisierte Datensammlung über medizinische Informationen mit einer Wählerliste verlinken. Anhand der Postleitzahl, dem Geburtsdatum und dem Geschlecht konnten den anonymisierten medizinischen Daten zum grossen Teil wieder Namen zugeordnet werden. Die eigentlich anonymisierte Datensammlung wurde dadurch grösstenteils de-anonymisiert [4]. Daraus lässt sich schlussfolgern, dass das Entfernen von Identifikatoren wie der Vor- und Nachname keine ausreichende Massnahme zum Schutz der Privacy ist.

### 1.1 Klassischer Ansatz der Datenanonymisierung

Ein klassischer Ansatz für den Schutz der Privacy innerhalb einer Datensammlung ist die Anonymisierung der Daten. Hauf beschreibt, dass bezüglich der Anonymität in einer Datensammlung zwischen drei Typen von Attributen unterschieden wird [5]:

- **Identifikatoren:** Erlauben die direkte Identifizierung einer Person.
- **Quasi-Identifikatoren:** Eine Kombination dieser können zur direkten Identifizierung einer Person führen.
- **Sensitive Attribute:** Attribute welche besonders schützenswert sind und nicht mit einer Person in Verbindung gebracht werden sollten.

Als einen der klassischen Ansätze um Daten zu anonymisieren, nennt Hauf die  $k$ -Anonymität [5]. Bei der  $k$ -Anonymität gilt es sicherzustellen, dass eine Person in einer Gruppe kleiner oder gleich  $k$  nicht von den anderen Personen in dieser Gruppe unterscheidbar ist. Dadurch soll die Re-Identifizierung von Personen innerhalb der Datensammlung erschwert werden. Die Schwachstelle eines solchen Ansatzes liegt unter anderem darin, dass die Methode durch Hintergrundinformationen gebrochen werden kann. Das heisst, durch das Verlinken von Informationen aus anderweitigen Quellen, können Personen in einer  $k$ -anonymen Datensammlung re-identifiziert werden. [5]

Zur Verdeutlichung der Schwachstellen von Ansätzen wie der  $k$ -Anonymität wird ein fiktives Beispiel gezeigt. In Tabelle 1 ist eine erfundene Sammlung von Krankheitsdaten aus den zwei Gemeinden Grub AR (9035) und Grub SG (9036) ersichtlich.

Aus dieser Datensammlung wird eine 3-anonyme Datensammlung erstellt. Hierfür werden die Identifikatoren «Vorname» und «Nachname» entfernt. Weiter wird das «Geburtsdatum» generalisiert, sodass nur der Jahrgang ausgewiesen wird. In Tabelle 2 ist die daraus resultierende 3-anonyme Datensammlung aufgeführt. Diese ist vollständig anonymisiert und erfüllt die Vorgaben der  $k$ -Anonymität.

Vorname	Nachname	Geburtsdatum	Geschlecht	PLZ	Krankheit
Michael	Meier	19.11.1985	M	9035	Bluthochdruck
Julia	Künzler	05.02.1974	W	9036	Diabetes
Karl	Lutz	02.12.1985	M	9035	Bluthochdruck
Felix	Wagner	29.06.1985	M	9035	Bluthochdruck
Maria	Ikonic	15.10.1974	W	9036	Diabetes
Jolanda	Meile	30.07.1974	W	9036	Diabetes

Tabelle 1. Originale Datensammlung zu Krankheiten in Grub AR und Grub SG

Vorname	Nachname	Geburtsdatum	Geschlecht	PLZ	Krankheit
		1985	M	9035	Bluthochdruck
		1974	W	9036	Diabetes
		1985	M	9035	Bluthochdruck
		1985	M	9035	Bluthochdruck
		1974	W	9036	Diabetes
		1974	W	9036	Diabetes

Tabelle 2. Die 3-anonyme Datensammlung zu Krankheiten in Grub AR und Grub SG

Anhand von Hintergrundinformationen lässt sich die 3-anonyme Tabelle jedoch de-anonymisieren bzw. lässt sich das sensitive Attribut «Krankheit» Personen zuweisen. In Grub AR (9035) und Grub SG (9036) wohnen jeweils nur einige Hundert Personen. Es besteht die Möglichkeit von insgesamt nur 3 männlichen Einwohnern mit Jahrgang 1985 in Grub AR. In einem kleinen Dorf kennt man sich und die Wahrscheinlichkeit ist gross, dass mit diesem Wissen Michael Meier, Karl Lutz und Felix Wagner (siehe Tabelle 1) re-identifiziert werden könnten, bzw. könnte das sensitive Attribut «Krankheit» in 3 Fällen einer Person zugeordnet werden. So wäre nun bekannt, dass Michael Meier, Karl Lutz und Felix Wagner an Bluthochdruck leiden.

Die  $k$ -Anonymität stellt keine Anforderungen an die sensitiven Attribute in einer Datensammlung, weshalb die  $l$ -Diversität eingeführt wurde. Die  $l$ -Diversität ist ein Ansatz zum Schutz der sensitiven Attribute und setzt eine  $k$ -anonyme Datensammlung voraus. Die  $l$ -Diversität schreibt vor, dass pro  $k$ -Gruppe von Datensätzen jedes sensitive Attribut  $l$ -«ausreichend repräsentiert» sein muss. [5]

Beim Beispiel in Tabelle 2 hiesse dies, das Attribut «Krankheit» darf nicht für alle in derselben Gruppe enthaltenen Datensätze gleich sein. In Tabelle 3 findet sich ein weiteres Beispiel von erfundenen Krankheitsdaten aus zwei anderen Gemeinden. Diese Datensammlung erfüllt die Eigenschaften der 3-Anonymität sowie der 3-Diversität. Es ist ersichtlich, dass im Gegensatz zum Beispiel in Tabelle 2 verschiedene Krankheiten in jeder  $k$ -Gruppe enthalten sind.

Bei demselben Angriff aus dem ersten Beispiel wüsste ein Angreifer bei den Daten aus Tabelle 3 nicht, welche Krankheit zu welcher Person gehört. Jeder der drei Männer könnte entweder an Magersucht, Bluthochdruck oder Diabetes leiden.

Vorname	Nachname	Geburtsdatum	Geschlecht	PLZ	Krankheit
		1964	M	4535	Magersucht
		1972	W	3983	Magersucht
		1964	M	4535	Bluthochdruck
		1964	M	4535	Diabetes
		1972	W	3983	Diabetes
		1972	W	3983	Bluthochdruck

Tabelle 3. Die 3-diverse Datensammlung zu Krankheiten in Kammersrohr und Bister

Das Hintergrundwissen eines Angreifers ist auch bei der  $l$ -Diversität weiterhin von grosser Relevanz. Kann der Angreifer beispielsweise herausfinden, dass Johannes Sieber an Magersucht und Mathias Meister an Bluthochdruck leiden, so weiss er dann auch, dass Richard Inauen offensichtlich an Diabetes erkrankt sein muss.

Dies ist ein wesentlicher Schwachpunkt von Ansätzen die auf der Anonymisierung von Daten beruhen. Aus diesem Grund ist die Differential Privacy sehr interessant, da sie nicht auf der Anonymisierung von Daten basiert und auch funktioniert, wenn ein Angreifer unbegrenzte Hintergrundinformationen hat.

## 1.2 Die Bedeutung von Differential Privacy

Entgegen klassischer Ansätze wie der  $k$ -Anonymität oder  $l$ -Diversität bietet die Differential Privacy ein Modell, mit dem eine beweisbare Garantie für die Privacy von Individuen in einer Datensammlung erreicht werden kann. Weiter erlaubt die Differential Privacy durch das Festlegen eines Privacy-Budgets den potenziellen Wissenszuwachs eines Angreifers zu quantifizieren und zu limitieren. Dies ermöglicht es, einen Kompromiss zwischen Datengenauigkeit und Schutz der Privacy zu finden und beliebig dem Schutzbedarf der Datensammlung anzupassen.

Die Differential Privacy wird bereits von verschiedenen Organisationen eingesetzt. So setzt laut Abowd *et al.* das United States Census Bureau (dt. Volkszählungsbüro der Vereinigten Staaten) für die Veröffentlichung deren Statistiken seit 2020 auf Differential Privacy [6], ebenso Technologiekonzerne wie Apple. Nach deren Angaben nutzen sie diese beispielsweise für die sichere Auswertung der Häufigkeit der genutzten Emojis oder der besuchten Domains. Laut Apple ist Differential Privacy eine der stärksten Methoden für die Wahrung der Privacy. [7]

### 1.3 Ziele dieser Arbeit

Gemäss Aufgabenstellung (vgl. Abschnitt 8.1) soll im Rahmen dieser Arbeit das Konzept der Differential Privacy verstanden und so aufbereitet werden, um es einem breiten Publikum verständlich machen zu können. Es soll aufgezeigt werden, wie Differential Privacy erreicht und angewendet werden kann. Weiter gilt es darzulegen, welche Arten von Differential Privacy existieren und für welche Anwendungsfälle diese eingesetzt werden kann. Insbesondere sollen die möglichen Stellen für das Hinzufügen von Rauschen, sowie verschiedene Verteilungen von Rauschen beschrieben werden. Zudem soll das Privacy-Budget als zentrales Konzept der Differential Privacy detailliert erklärt werden.

Basierend auf der erarbeiteten Theorie soll eine Übung ausgearbeitet werden, in welcher Differential Privacy implementiert und ausprobiert werden kann.

### 1.4 Nomenklatur

Der englische Begriff «Privacy» kann nicht eindeutig ins Deutsche übersetzt werden und ist abhängig vom Zusammenhang. Privacy kann Datenschutz bedeuten und sich mit dem Schutz vor unbefugter Erhebung, Speicherung und Weitergabe von Personendaten befassen. Privacy kann auch Privatsphäre bedeuten und damit den Schutz der privaten Atmosphäre beschreiben. Um eine Verwechslung mit den Begriffen «Datenschutz» und «Privatsphäre» zu vermeiden, wird in dieser Arbeit auf eine Übersetzung des Begriffs Privacy verzichtet. Unter **Privacy** wird in dieser Arbeit der Schutz von persönlichen Informationen verstanden.

Eine Funktion, welche die Eigenschaften der Differential Privacy erfüllt, wird als **Mechanismus** bezeichnet. Die Institution, welche die Daten sammelt oder publiziert, wird als **Aggregator** bezeichnet. Das Resultat eines Mechanismus ist vom Anwendungsfall abhängig und kann stark variieren. Dies kann das Ergebnis einer Datenbankabfrage, ein Histogramm, die Schnittstelle zu einem anderen Mechanismus u.v.m. sein. Das Resultat des Mechanismus wird in dieser Arbeit unabhängig vom Anwendungsfall als **Ausgabe** bezeichnet.

## 2 Stand der Praxis

Die Differential Privacy ist eine Definition, welche Anforderungen an einen Mechanismus stellt. Entsprechend können die Eigenschaften der Differential Privacy nur für Mechanismen erfüllt werden, nicht für eine Datensammlung [8]. Dies ist ein essentieller Unterschied zu anderen Ansätzen, wie beispielsweise der  $k$ -Anonymität.

Bei der Differential Privacy sammelt oder publiziert der Aggregator die Daten nicht direkt. Die Daten werden anhand eines Mechanismus gesammelt oder zur Verfügung gestellt. Nach Desfontaines wird zwischen dem zentralen und dem lokalen Modell von Differential Privacy unterschieden [9].

Abbildung 1 zeigt das zentrale Modell. In diesem Modell hat der Aggregator Zugriff auf die ursprünglichen Daten. Also jene Daten, die nicht verändert wurden und dem originalen Informationsgehalt entsprechen. Ein Aggregator kann zum Beispiel ein Dienstleister oder Service sein, welcher Daten erhebt und zur Verfügung stellt oder weiterverarbeitet. Der Aggregator stellt die Daten über einen Mechanismus zur Verfügung und setzt dadurch die Eigenschaften der Differential Privacy um.

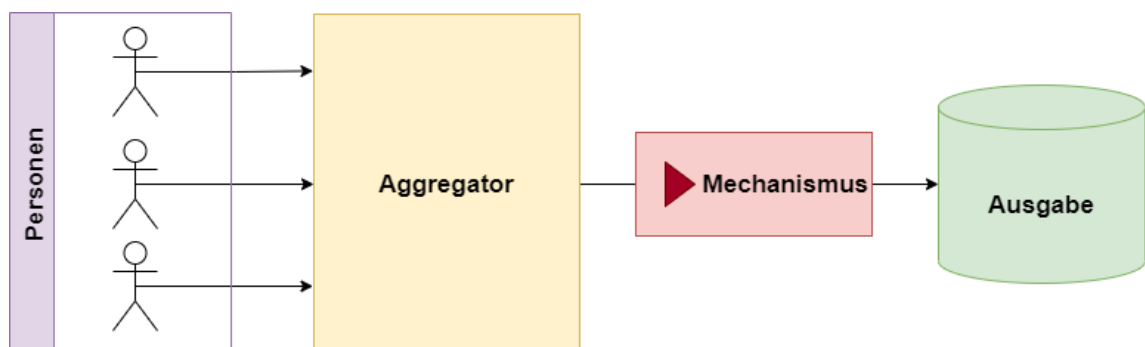


Abb. 1. Zentrales Modell (angelehnt an Abbildung von Desfontaines [9])

In Abbildung 2 wird das lokale Modell von Differential Privacy gezeigt. Bei diesem Modell wird der Mechanismus beim Sammeln von Daten ausgeführt. Der Aggregator kann nur über den Mechanismus auf die Daten zugreifen und hat somit keinen direkten Zugriff auf die originalen Daten. [9]

Desfontaines erwähnt, dass auch ein hybrides Modell implementiert werden kann, welches das zentrale und das lokale Modell vereint [9]. Im Rahmen dieser Arbeit sollen die grundle-

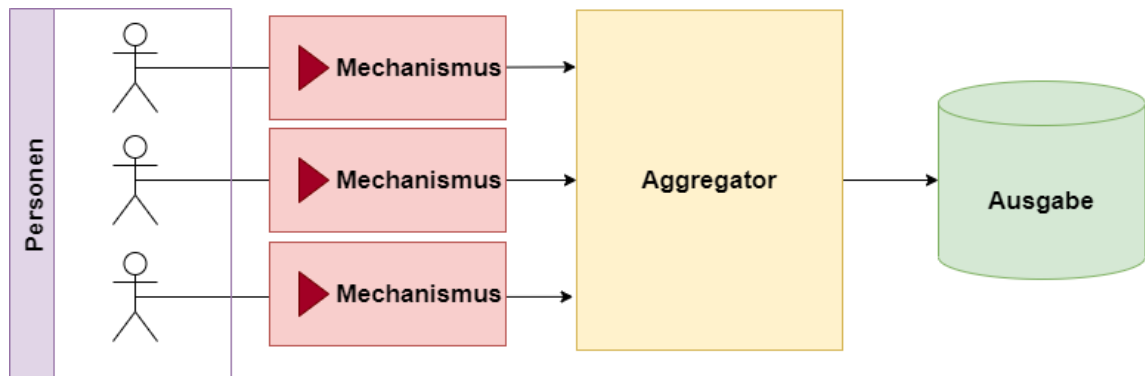


Abb. 2. Lokales Modell (angelehnt an Abbildung von Desfontaines [9])

genden Eigenschaften der Differential Privacy behandelt werden, weshalb bewusst der Fokus auf das lokale und zentrale Modell gelegt wird. Das hybride Modell wird nicht weiter berücksichtigt.

### 2.1 Die Grundidee von Differential Privacy

Zhu *et al.* beschreiben als Grundidee der Differential Privacy, dass ein Angreifer nicht herausfinden kann, ob eine bestimmte Person in einer Datensammlung enthalten ist oder nicht [8]. Diese Eigenschaft lässt sich wie folgt beschreiben (vgl. Abbildung 3):

- Es ist eine Datensammlung  $D_1$  mit Anzahl  $n$  Datensätzen gegeben.
- Es existiert eine Kopie dieser Datensammlung, welche sich in genau einem Datensatz unterscheidet. Dies ist eine sogenannte benachbarte Datensammlung  $D_2$  mit  $n - 1$  Datensätzen.
- In  $D_1$  ist die gesuchte Person enthalten und in  $D_2$  ist diese nicht enthalten.
- Ein Angreifer macht eine Abfrage und erhält die Ausgabe entweder basierend auf  $D_1$  (gesuchte Person ist enthalten) oder  $D_2$  (gesuchte Person ist nicht enthalten). Der Angreifer sieht nicht aus welcher der beiden Datensammlungen die Ausgabe stammt.
- Die Differential Privacy verlangt, dass ein Angreifer bei der erhaltenen Ausgabe mit sehr hoher Wahrscheinlichkeit nicht unterscheiden kann, ob es sich bei der Quelle um  $D_1$  oder  $D_2$  handelt. Der Angreifer soll also basierend auf der erhaltenen Ausgabe nicht bestimmen können, ob die gesuchte Person enthalten ist oder nicht.
- Sind diese Eigenschaften für alle Datensätze gegeben, kann ein Angreifer für keine der Personen herausfinden, ob sie enthalten sind oder nicht. Dadurch ist der Schutz jedes Individuums gewährleistet.

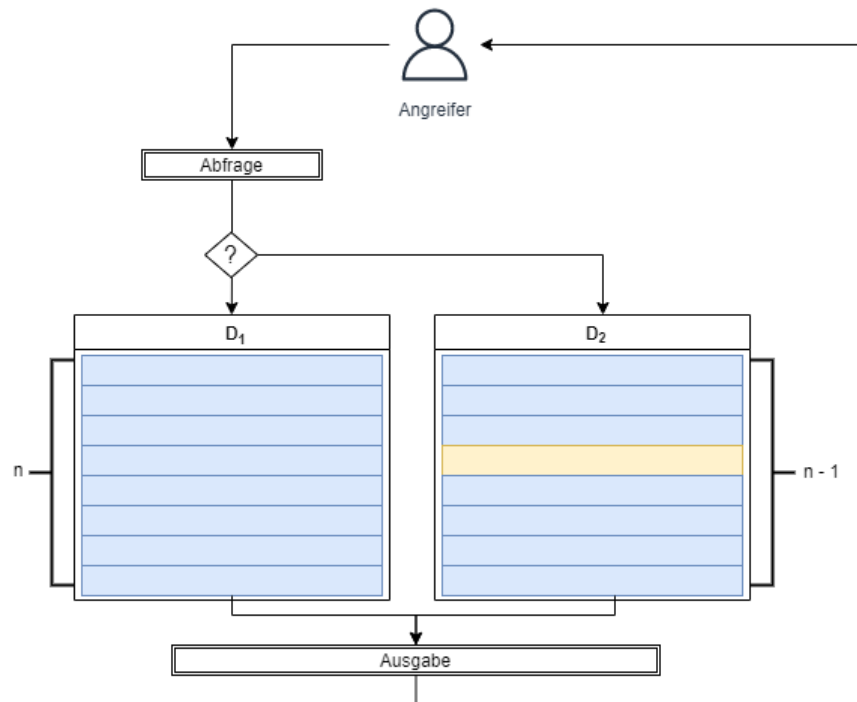


Abb. 3. Grundidee von Differential Privacy

Für eine bessere Verständlichkeit wurde in diesem Beispiel in der benachbarten Datensammlung  $D_2$  eine Zeile vollständig entfernt. Dies muss nicht zwingend der Fall sein. Es könnte auch ein Datensatz entsprechend verändert worden sein, sodass die gesuchte Person nicht mehr enthalten ist (z.B. ersetzt durch eine andere Person). Wichtig ist, dass sich die beiden Datensammlungen in genau einem Datensatz unterscheiden.

## 2.2 Die formale Definition von Differential Privacy

Die Grundidee von Differential Privacy lässt sich in folgende formale Definition übersetzen [10]:

**Definition 1 ( $\epsilon$ -Differential Privacy)** Ein Mechanismus  $M$  erfüllt dann die Eigenschaften von  $\epsilon$ -Differential Privacy, wenn für alle Datensammlungen  $D_1$  und  $D_2$ , welche sich in genau einem Datensatz unterscheiden, für alle möglichen Ausgaben  $A$  gilt:

$$\mathbb{P}[M(D_1) = A] \leq e^\epsilon * \mathbb{P}[M(D_2) = A]$$

$\mathbb{P}[M(D_1) = A]$  beschreibt die Wahrscheinlichkeit, dass wenn der Mechanismus  $M$  auf der Datensammlung  $D_1$  ausgeführt wird, die Ausgabe  $A$  resultiert. Der Mechanismus  $M$  ist aufgrund des hinzugefügten Rauschens probabilistisch. Wird somit der Mechanismus  $M$  mehr-

mals nacheinander ausgeführt, können sich die Ausgaben unterscheiden. [10]

$\mathbb{P}[M(D_2) = A]$  beschreibt die Wahrscheinlichkeit, dass wenn der Mechanismus  $M$  auf der Datensammlung  $D_2$  ausgeführt wird, die Ausgabe  $A$  resultiert. Diese Wahrscheinlichkeit wird mit  $e^\epsilon$  multipliziert. Das  $\epsilon$  spielt eine essentielle Rolle, weshalb in dieser Definition auch von  $\epsilon$ -Differential Privacy gesprochen wird. [10]

Die Definition ist symmetrisch. Das heisst, dass  $D_1$  und  $D_2$  vertauscht werden können, was formal wie folgt dargestellt werden kann [10]:

$$e^{-\epsilon} * \mathbb{P}[M(D_2) = A] \leq \mathbb{P}[M(D_1) = A] \leq e^\epsilon * \mathbb{P}[M(D_2) = A]$$

Nach Zhu *et al.* wird die  $\epsilon$ -Differential Privacy auch als reine Differential Privacy bezeichnet. Es ist nicht in jedem Fall möglich die reine Differential Privacy zu implementieren. In diesen Fällen kann die annähernde Differential Privacy umgesetzt werden. In der Definition der annähernden Differential Privacy kommt ein  $\delta$  dazu, welches eine kontrollierte Verletzung der reinen Differential Privacy erlaubt. [8]

**Definition 2 (( $\epsilon, \delta$ )-Differential Privacy)** *Ein Mechanismus  $M$  erfüllt dann die Eigenschaften von ( $\epsilon, \delta$ )-Differential Privacy, wenn für alle Datensammlungen  $D_1$  und  $D_2$ , welche sich in genau einem Datensatz unterscheiden, für alle möglichen Ausgaben  $A$  gilt:*

$$\mathbb{P}[M(D_1) = A] \leq e^\epsilon * \mathbb{P}[M(D_2) = A] + \delta$$

Die  $\epsilon$ -Differential Privacy kann auch als ( $\epsilon, \delta$ )-Differential Privacy mit einem Wert von  $\delta = 0$  definiert werden. Das  $\delta$  bildet somit den Unterschied von der reinen zur annähernden Differential Privacy und bedeutet, dass der Mechanismus mit einer Wahrscheinlichkeit von  $1 - \delta$  die Eigenschaften von  $\epsilon$ -Differential Privacy erfüllt. Oder umgekehrt kann das  $\delta$  als die Wahrscheinlichkeit betrachtet werden, dass die  $\epsilon$ -Differential Privacy verletzt wird. Somit kann es bei der annähernden Differential Privacy zu Ereignissen kommen, in welchen die  $\epsilon$ -Differential Privacy nicht erfüllt wird. Je mehr sich das  $\delta$  an 0 annähert, desto näher kommt der Mechanismus an die reine und striktere Differential Privacy. [11]

### 2.3 Das Privacy-Budget

Das  $\epsilon$  beschreibt den möglichen Wissensgewinn eines Angreifers. Das heisst, das  $\epsilon$  definiert, wie viel Wissen ein Angreifer maximal zusätzlich erhalten könnte. Das  $\epsilon$  erlaubt somit den Privacy-Verlust genau zu quantifizieren und wird deshalb auch als «Privacy-Budget» bezeichnet. [8]

Das  $\varepsilon$  muss per Definition  $\geq 0$  sein. Je kleiner das  $\varepsilon$  wird, umso mehr nähert sich die Exponentialfunktion  $e^\varepsilon$  an 1 an. Die Ähnlichkeit der Ergebnisse hängt somit von  $\varepsilon$  ab. Je kleiner das  $\varepsilon$  ist, desto ähnlicher sind die Ausgaben und somit wird auch der maximale Wissenszuwachs eines Angreifers kleiner. [10]

Zusammengefasst heisst das: Je kleiner das  $\varepsilon$ , umso höher der Schutz der Privacy.

### 2.3.1 Quantifizierung des Angreiferwissens

Die Bedeutung des Privacy-Budgets wird nachfolgend anhand eines Beispiels verdeutlicht. Hierbei wird gezeigt, wie das Wissen des Angreifers quantifiziert wird und welche Auswirkungen die Wahl des  $\varepsilon$ -Wertes hat. Dieses Beispiel wurde dem Differential Privacy-Blog von Desfontaines [10] entnommen.

Es seien folgende Rahmenbedingungen gegeben:

- Es sei ein Mechanismus  $M$  gegeben, welcher die Eigenschaften von  $\varepsilon$ -Differential Privacy erfüllt.
- Dieser Mechanismus wird auf der Datensammlung  $D$  ausgeführt, woraus die Ausgabe  $M(D) = A$  resultiert.
- Ein Angreifer versucht herauszufinden, ob die Zielperson in  $D$  enthalten ist oder nicht. Für den Angreifer existieren somit zwei Szenarien: Die Zielperson ist in  $D$  enthalten oder sie ist nicht in  $D$  enthalten. Dies kann abstrahiert werden, indem zwischen zwei Datensammlungen unterschieden wird.  $D_1$  für die Datensammlung, in welcher die Zielperson enthalten ist und  $D_2$  für die Datensammlung, in welcher die Zielperson fehlt. Der Angreifer versucht anhand der Ausgabe  $A$  herauszufinden, ob diese basierend auf  $D_1$  oder  $D_2$  zustande kam.
- Es wird angenommen, dass der Angreifer maximales Hintergrundwissen hat und in der Datensammlung somit nur die Zielperson unbekannt ist.

Der Angreifer hat einen initialen Verdacht, ob die Zielperson enthalten ist oder nicht. Dieser Verdacht wird durch die Wahrscheinlichkeit  $\mathbb{P}[D = D_1]$  repräsentiert, welche zwischen 0 (Angreifer weiss, dass Zielperson nicht enthalten ist) und 1 (Angreifer weiss, dass Zielperson enthalten ist) liegt. Hat der Angreifer keinen Verdacht, ob die Zielperson enthalten ist oder nicht, liegt die Wahrscheinlichkeit  $\mathbb{P}[D = D_1]$  bei 0.5.

Der Mechanismus  $M$  wird ausgeführt und resultiert in der Ausgabe  $A$ . Nun stellt sich die Frage, um wie viel der Verdacht des Angreifers geändert hat, nachdem dieser die Ausgabe  $A$  erhalten hat. Der neue Verdacht wird durch  $\mathbb{P}[D = D_1 | M(D) = A]$  repräsentiert. Wird die Wahrscheinlichkeit des initialen Verdachts  $\mathbb{P}[D = D_1]$  mit der Wahrscheinlichkeit des neuen

Verdachts  $\mathbb{P}[D = D_1 | M(D) = A]$  verglichen, erhält man einen genauen Wert für den Wissensgewinn des Angreifers. In Abbildung 4 ist zu sehen, wie Desfontaines die unteren und oberen Grenzen des Wissensgewinns für unterschiedliche  $\epsilon$ -Werte grafisch dargestellt hat. Die weiße Linie von  $(0.0|0.0)$  zu  $(1.0|1.0)$  stellt den Fall dar, wenn der Angreifer kein zusätzliches Wissen gewonnen hat und der initiale Verdacht («initial suspicion») dem neuen Verdacht («updated suspicion») entspricht. Es ist ersichtlich, dass bei steigendem  $\epsilon$ -Wert der Wissensgewinn des Angreifers wächst. So würde ein initialer Verdacht von 0.1 bei einem  $\epsilon$ -Wert von 5 zu einem neuen Verdacht von 0.94 führen. Der Angreifer wäre sich also nach einer einzigen Abfrage zu 94% sicher, dass die Zielperson in der Datensammlung enthalten ist, obwohl er sich zu Beginn nur zu 10% sicher war. [10]

Dies zeigt deutlich, dass die richtige Wahl des Privacy-Budgets essentiell für den Schutz der Privacy ist. Ein zu hoch gewähltes Privacy-Budget kann zu unzureichendem Schutz der Privacy führen.

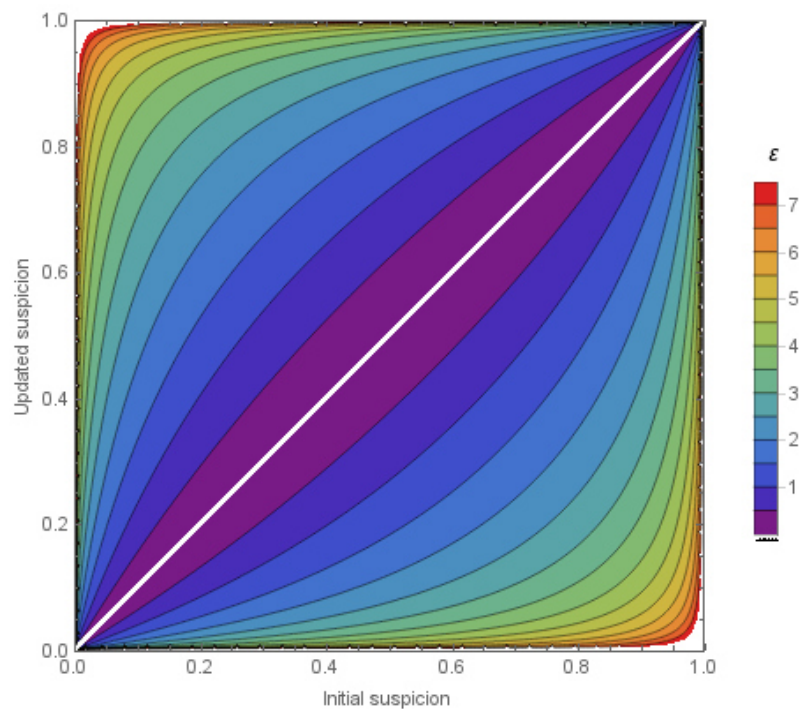


Abb. 4. Auswirkung von  $\epsilon$  auf Wissensgewinn nach Desfontaines [10]

### 2.3.2 Komposition des Privacy-Budgets

Zhu *et al.* beschreiben als weiteren grossen Vorteil von Differential Privacy, dass das Privacy-Budget auch für Kompositionen von mehreren Mechanismen bestimmt werden kann. Ganz im Kontrast zu beispielsweise der  $k$ -Anonymität. Bei der Kombination von zwei  $k$ -anonymen Datensammlungen kann keine Aussage über den  $k$ -Wert der resultierten Datensammlung

gemacht werden. [8]

Nach Zhu *et al.* wird bei Differential Privacy zwischen paralleler und sequentieller Komposition unterschieden. Bei der parallelen Komposition (vgl. Abbildung 5) werden mehrere Mechanismen  $M = \{M_1, \dots, M_n\}$  ausgeführt, wobei jeder Mechanismus  $M_i$  das Privacy-Budget von  $\varepsilon_i$  für eine disjunkte Teilmenge der gesamten Datensammlung garantiert. Die Komposition dieser Mechanismen ( $M$ ) garantiert in diesem Fall das Privacy-Budget von  $\max(\varepsilon_1, \dots, \varepsilon_n)$ , was dem grössten  $\varepsilon$ -Wert aller Mechanismen innerhalb der Komposition entspricht. [8]

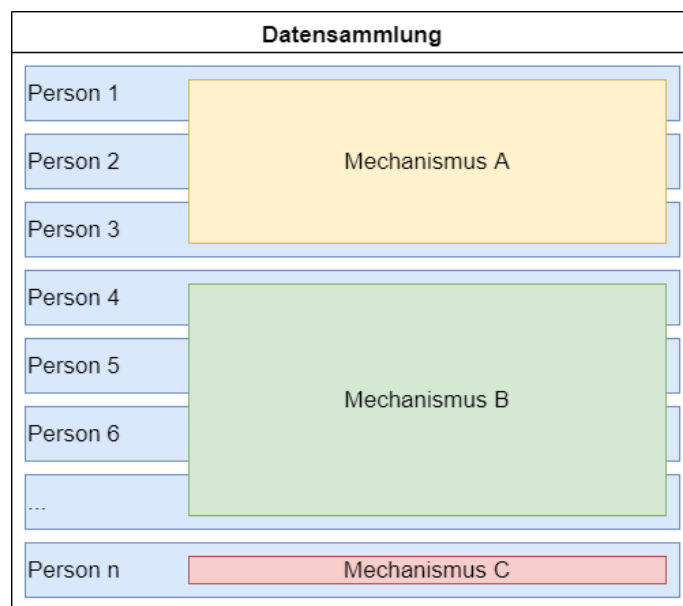


Abb. 5. Parallele Komposition

Bei der sequentiellen Komposition (vgl. Abbildung 6) werden die Mechanismen  $M = \{M_1, \dots, M_n\}$  sequentiell und somit nacheinander auf derselben Datensammlung ausgeführt, wobei jeder Mechanismus  $M_i$  das Privacy-Budget von  $\varepsilon_i$  garantiert. Die Komposition dieser Mechanismen ( $M$ ) garantiert in diesem Fall das Privacy-Budget von  $\sum_{i=1}^n \varepsilon_i$ , was der Summe der  $\varepsilon$ -Werte aller Mechanismen innerhalb der Komposition entspricht. [8]

### 2.3.3 Das Privacy-Budget in der Praxis

Gemäss Zhu *et al.* wird ein  $\varepsilon$ -Wert von kleiner als 1 empfohlen. Die Werte 0.1 und  $\ln(2)$  werden als typische Beispiele genannt. [8] Desfontaines hat in seinem Blog über Differential Privacy diverse Beispiele aus der Praxis zusammengefasst. Apple soll für den Mechanismus, welcher die Emoji-Empfehlungen schützt, einen  $\varepsilon$ -Wert von 4 verwenden. Für die QuickType Vorschläge, um von den Benutzern eingegebene, bisher unbekannte Wörter zu lernen, soll Apple einen Wert von 16 für das  $\varepsilon$  festgelegt haben. Facebook nutzt gemäss Desfontaines einen  $\varepsilon$ -Wert von 0.45 für den Schutz der Benutzerdaten über Benutzerinteraktionen mit URLs innerhalb

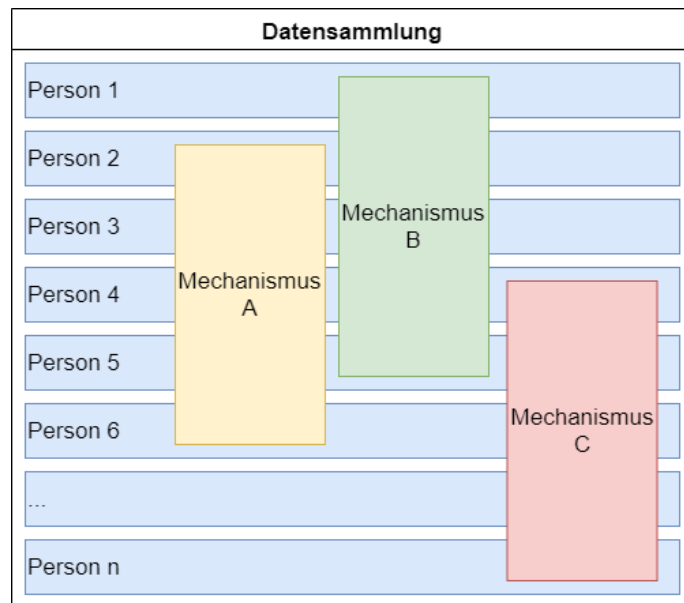


Abb. 6. Sequentielle Komposition

von Facebook. [12]

Diese Beispiele aus der Praxis zeigen, dass das Privacy-Budget sehr individuell und dem jeweiligen Schutzbedarf der Datensammlung angepasst gewählt werden muss. Eine allgemein gültige Empfehlung für den Wert von  $\epsilon$  scheint nicht ableitbar zu sein, ausser dass das  $\epsilon$  grundsätzlich so klein wie möglich gewählt werden sollte.

### 2.4 Die Sensitivität

Für die praktische Implementierung der Differential Privacy ist es wichtig das Konzept der Sensitivität von Funktionen zu verstehen. Gemäss Near und Abuah entspricht die Sensitivität einer Funktion dem Betrag, um den sich die Ausgabe der Funktion ändert, wenn sich ihre Eingabe ändert [13]. Zur Veranschaulichung drei Beispiele für die Sensitivitätswerte von einfachen Funktionen:

- Die Sensitivität von  $f(x) = x$  entspricht 1, denn: ändert sich  $x$  um 1, so ändert sich auch  $f(x)$  um 1.
- Die Sensitivität von  $f(x) = x + x$  entspricht 2, denn: ändert sich  $x$  um 1, so ändert sich  $f(x)$  um 2.
- Die Sensitivität von  $f(x) = 5 * x$  entspricht 5, denn: ändert sich  $x$  um 1, so ändert sich  $f(x)$  um 5.

Für die Zählfunktion (`count`) kann nach Near und Abuaq stets die Sensitivität 1 gewählt werden. Dies ist darauf zu begründen, dass wenn ein Datensatz hinzugefügt oder entfernt wird, sich die Zählfunktion um maximal 1 ändern kann. [13]

Im Gegensatz zur Zählfunktion hängt bei der Summenfunktion (`sum`) die Ausgabe vom Inhalt der einzelnen Datensätze ab. Zum Beispiel sind in einer Datensammlung die Anzahl Arztbesuche in einem Jahr erfasst. Die Summenfunktion zählt sämtliche Arztbesuche zusammen. Wird nun ein Datensatz hinzugefügt in welchem 23 Arztbesuche erfasst sind, ändert sich die Ausgabe der Summenfunktion um 23, obwohl nur 1 Datensatz hinzugefügt wurde. In der Annahme, niemand würde mehr als 100 Arztbesuche pro Jahr wahrnehmen, könnte beispielsweise eine Sensitivität für die Summenfunktion der Arztbesuche von 100 definiert werden. Wird nun jedoch ein Datensatz mit 101 Arztbesuchen hinzugefügt, wird die Sensitivität verletzt.

Aus diesem Grund wird zwischen der unbegrenzten Sensitivität und der begrenzten Sensitivität unterschieden. Die unbegrenzte Sensitivität beschreibt den oben genannten Fall, wenn keine Unter- und Obergrenze des Werts gegeben ist. Um eine begrenzte Sensitivität handelt es sich, wenn eine Unter- und Obergrenze definiert ist. Dies könnte beispielsweise eine Funktion sein, welche die Anzahl der verkauften Sitzplätze eines Kinos zusammenzählt. Hier ist die untere Grenze 0, wenn in keinem Saal ein Sitzplatz verkauft wurde. Die obere Grenze ist dann erreicht, wenn sämtliche Säle ausgebucht und somit alle vorhandenen Sitzplätze verkauft wurden. Die Sensitivität berechnet sich aus der Differenz der unteren und oberen Grenze. Im Beispiel des Kinos würde die Sensitivität der Summenfunktion somit der Anzahl Sitzplätze aller Säle entsprechen.

Funktionen mit unbegrenzter Sensitivität können nach Near und Abuaq in solche mit begrenzter Sensitivität transformiert werden, um eine bestimmte Sensitivität garantieren zu können. Dafür muss bei der Implementierung der Funktion eine Unter- und Obergrenze erzwungen werden. Hierzu werden zu niedrige oder zu hohe Werte abgeschnitten. Wird beispielsweise eine Funktion implementiert, welche die Alter aller Personen in der Datensammlung zusammenzählt, so könnte eine untere Grenze von 0 und obere Grenze von 125 erzwungen werden. Für den sehr seltenen Fall, dass eine Person mit einem Alter von über 125 erfasst würde, zählte die Funktion für diese Person nur ein Wert von 125 dazu. Dank der nun klar gesetzten Grenzen, kann die Sensitivität begrenzt und in jedem Fall eingehalten werden. [13]

In der Literatur wird zwischen lokaler und globaler Sensitivität unterschieden. Diese Unterscheidung würde die fachliche Tiefe dieser Arbeit übersteigen, weshalb auf eine Einführung der lokalen Sensitivität verzichtet wird. Innerhalb dieser Arbeit wird mit «Sensitivität» stets die globale Sensitivität bezeichnet.

## 2.5 Implementierung von Differential Privacy

Zhu *et al.* beschreiben, dass in der Praxis die folgenden drei Mechanismen häufig Anwendung finden: Der Laplace, der Gauss sowie der exponentielle Mechanismus [8].

Während die Laplace und Gauss Mechanismen ausschliesslich für numerische Funktionen (wie z.B. «Wie viele Personen haben den Jahrgang 1985?») verwendet werden, wird der exponentielle Mechanismus in der Regel für nicht-numerische Anfragen (wie z.B. «Welches ist die häufigste Krankheit?») eingesetzt [8].

### 2.5.1 Der Laplace Mechanismus

Beim Laplace Mechanismus wird zum Resultat der Abfrage Rauschen nach der Laplace Verteilung hinzugefügt. Durch das Hinzufügen von Rauschen wird die Ausgabe des Mechanismus probabilistisch. Das heisst, die Ausgabe wird zu einem bestimmten Mass zufällig. Abbildung 7 zeigt die Funktionsweise des Laplace Mechanismus.

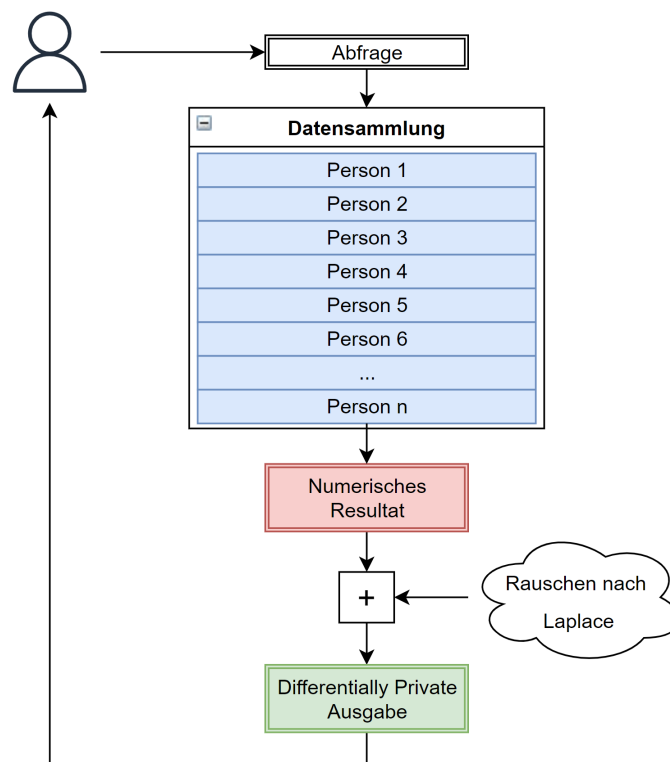


Abb. 7. Funktionsweise des Laplace Mechanismus

Nach Zhu *et al.* lautet die Definition des Laplace Mechanismus wie folgt [8].

**Definition 3 (Laplace Mechanismus)** Die Funktion  $f$  wird über die Datensammlung  $D$  ausge-

führt und liefert als Resultat eine Zahl zurück. Zu dieser Zahl wird Rauschen nach Laplace hinzugefügt ( $Lap(\frac{s}{\epsilon})$ ). Durch das Hinzufügen des Laplace Rauschens, erfüllt der Mechanismus  $M(D)$  die Eigenschaften der  $\epsilon$ -Differential Privacy. Wobei  $s$  die Sensitivität der Funktion  $f$  ist.

$$M(D) = f(D) + Lap(\frac{s}{\epsilon})$$

Near und Abuah weisen bei der Implementierung des Laplace Mechanismus darauf hin, dass der Bruch  $\frac{s}{\epsilon}$  dem Skalenparameter der Laplace Wahrscheinlichkeitsdichtefunktion entspricht. Wird beispielsweise die Funktion `random.laplace()` der Python Bibliothek «Numpy» verwendet, müssen ein Lageparameter `loc` und ein Skalenparameter `scale` mitgegeben werden. [13]

Ein Funktionsaufruf für das Generieren des Rauschens nach Laplace könnte dann wie folgt aussehen: `random.laplace(loc=0, scale=sensitivity/epsilon)`.

### 2.5.2 Der Gauss Mechanismus

Im Gegensatz zum Laplace Mechanismus wird beim Gauss Mechanismus Rauschen nach der Gauss Verteilung (Normalverteilung) hinzugefügt (vgl. Abbildung 8).

Während der Laplace Mechanismus die Eigenschaften der reinen  $\epsilon$ -Differential Privacy erfüllt, kann der Gauss Mechanismus nur die Eigenschaften der annähernden  $(\epsilon, \delta)$ -Differential Privacy erfüllen. Nach Near und Abuah wird der Gauss Mechanismus wie folgt definiert. [13]

**Definition 4 (Gauss Mechanismus)** Zur Ausgabe der Funktion  $f(D)$ , welche eine Zahl zurück liefert, wird Rauschen nach der Gauss Verteilung  $\mathcal{N}(0, \sigma^2)$  hinzugefügt. Dadurch erfüllt der Mechanismus  $M(D)$  die Eigenschaften von  $(\epsilon, \delta)$ -Differential Privacy.

$$M(D) = f(D) + \mathcal{N}(0, \sigma^2)$$

Wobei:

$$\sigma^2 = \frac{2 * s^2 * \ln(1.25/\delta)}{\epsilon^2}$$

Das  $s$  entspricht der Sensitivität der Funktion  $f(D)$ .  $\mathcal{N}(0, \sigma^2)$  bezeichnet das Gauss Rauschen, zentriert bei 0 und mit der Varianz  $\sigma^2$ .

Near und Abuah weisen bei der Implementierung des Gauss Mechanismus darauf hin, dass das Sigma  $\sigma$  dem Skalenparameter der Gauss Wahrscheinlichkeitsdichtefunktion entspricht. Wird beispielsweise die Funktion `random.normal()` der Python Bibliothek «Numpy» verwendet, müssen ein Lageparameter `loc` und ein Skalenparameter `scale` mitgegeben werden. [13]

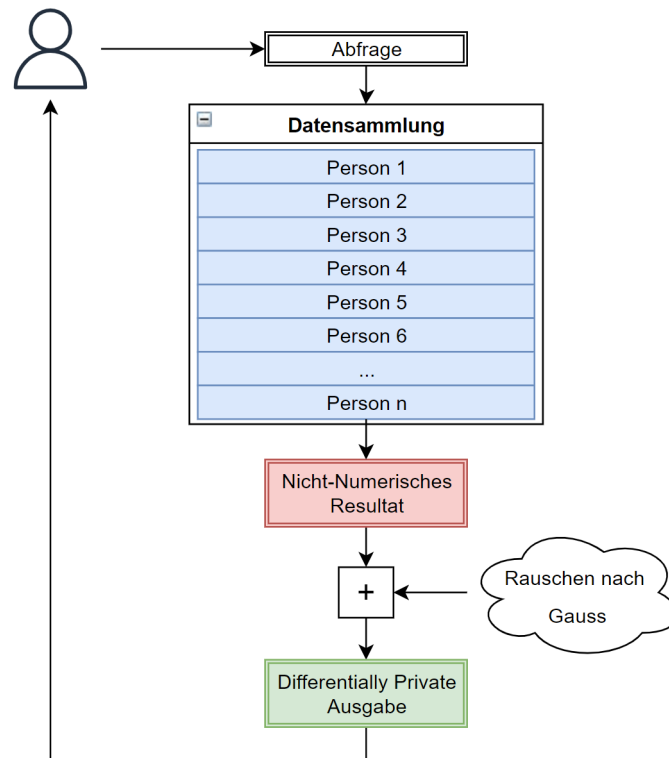


Abb. 8. Funktionsweise des Gauss Mechanismus

Ein Funktionsaufruf für das Generieren des Rauschens nach Gauss könnte dann wie folgt aussehen: `random.normal(loc=0, scale=sigma)`.

### 2.5.3 Vergleich des Laplace und des Gauss Mechanismus

Ein grosser Vorteil des Laplace Mechanismus ist, dass die Eigenschaften der  $\epsilon$ -Differential Privacy in jedem Fall erfüllt werden. Dies im Gegensatz zum Gauss Mechanismus, bei welchem die Definition der Differential Privacy gelockert werden muss, indem ein  $\delta$  hinzugefügt wird.

Nach Desfontaines liegt ein wesentlicher Vorteil des Gauss Mechanismus in der Verteilung des Rauschens. Die Normalverteilung ist sehr verbreitet, intuitiv und hat viele gute statistische Eigenschaften. [14]

Ein weiterer Unterschied liegt gemäss Desfontaines in der Menge des hinzugefügten Rauschens. Hat eine Person in der Datensammlung nur Einfluss auf wenige Statistiken, muss mit dem Laplace Mechanismus weniger Rauschen hinzugefügt werden. Hat eine Person Einfluss auf viele Statistiken, so benötigt der Gauss Mechanismus weniger Rauschen. In Abbildung 9 hat Desfontaines die beiden Mechanismen bezüglich der Menge des Rauschens gegenübergestellt.

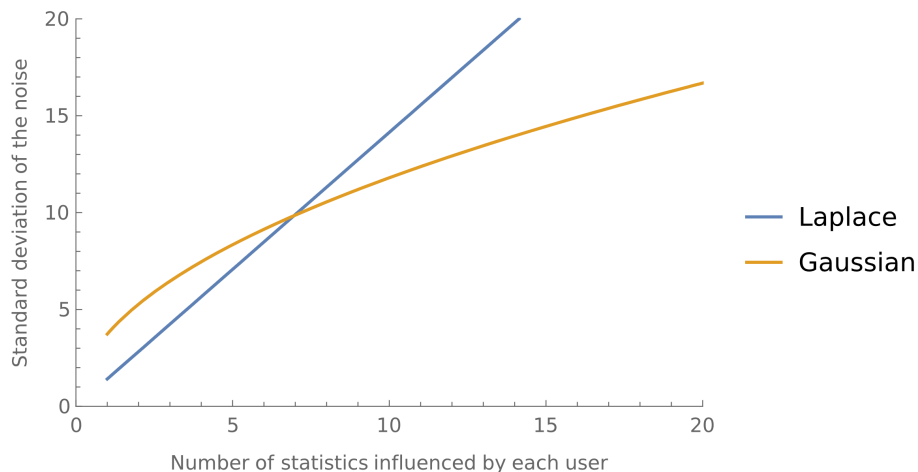


Abb. 9. Vergleich des Laplace und Gauss Rauschens nach Desfontaines [14]

Desfontaines hat die Menge des Rauschens des Laplace und des Gauss Mechanismus bei einem  $\varepsilon = 1.0$  und einem  $\delta = 10^{-5}$  abgebildet. Es ist ersichtlich, dass der Laplace Mechanismus weniger Rauschen benötigt, wenn eine Person in der Datensammlung Einfluss auf 1 bis ca. 7 Statistiken hat. Ab 7 Statistiken benötigt in diesem Beispiel der Gauss Mechanismus weniger Rauschen.

«Einfluss auf mehrere Statistiken» heisst, eine Person kann Einfluss auf mehrere Werte haben, welche auf Basis derselben Datensammlung errechnet werden.

Zur Verdeutlichung ein Beispiel: Die 50 meist geschauten Kinofilme sollen veröffentlicht werden. Dazu soll die Anzahl der Kinobesuche pro Film ausgegeben werden. Hat nun eine Person genau einen Kinofilm geschaut, so hat diese auch nur Einfluss auf genau einen Wert. Handelt es sich jedoch um eine Person, welche viele Kinofilme geschaut hat, so beeinflusst diese Person die Anzahl mehrerer Filme und dadurch auch mehrerer Statistiken. Bei vielen Statistiken wird das Rauschen beim Gauss Mechanismus geringer als beim Laplace Mechanismus.

Ob der Laplace oder der Gauss Mechanismus implementiert wird, hängt somit vom Anwendungsfall ab und muss vor der Implementierung gut überlegt werden.

### 2.5.4 Der exponentielle Mechanismus

In Abbildung 10 wird die Funktionsweise des exponentiellen Mechanismus gezeigt. Im Gegensatz zum Laplace oder Gauss Mechanismus wird beim exponentiellen Mechanismus kein Rauschen zur Ausgabe selbst hinzugefügt. Es kommt eine Bewertungsfunktion zum Einsatz, welche die möglichen Ausgaben bewertet. Basierend auf diesen Bewertungen wird für jede mögliche Ausgabe eine Wahrscheinlichkeit berechnet, mit welcher diese Ausgabe vom Mechanismus ausgegeben werden soll. Dadurch werden die Eigenschaften der  $\varepsilon$ -Differential Pri-

vacy eingehalten. Zhu *et al.* weisen darauf hin, dass die Art der Bewertungsfunktion abhängig vom Anwendungsfall ist und entsprechend spezifisch auf den Anwendungsfall implementiert werden muss [8].

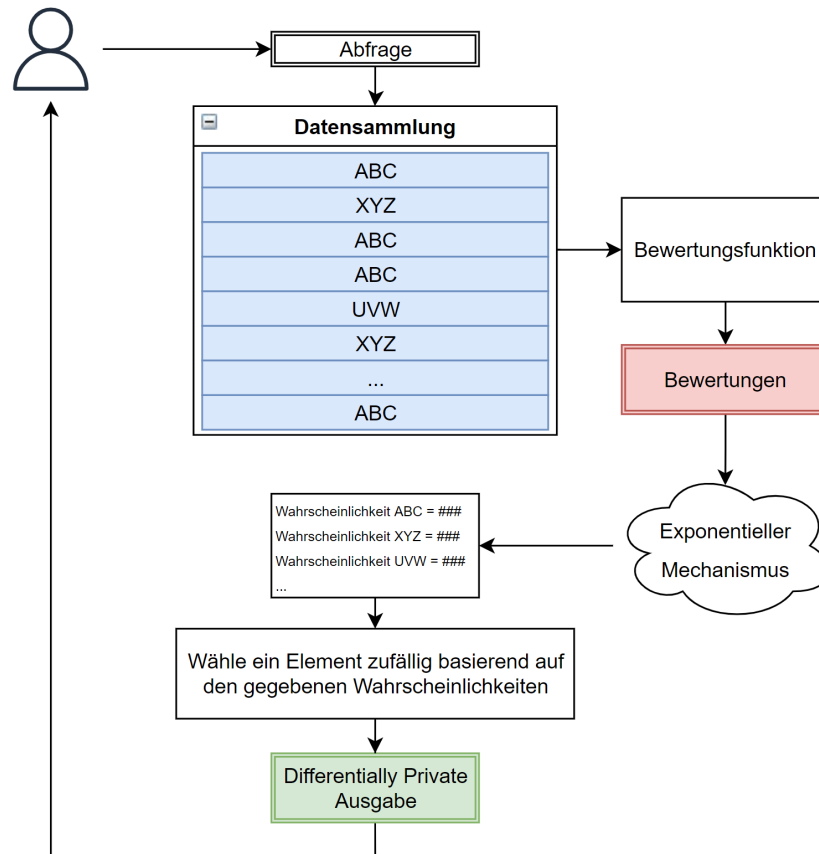


Abb. 10. Funktionsweise des exponentiellen Mechanismus

Nach Zhu *et al.* wird der exponentielle Mechanismus wie folgt definiert. [8]

**Definition 5 (Exponentieller Mechanismus)** Es ist die Bewertungsfunktion  $q(D, \phi)$  gegeben. Der Mechanismus  $M(D)$  erfüllt die Eigenschaften von  $\epsilon$ -Differential Privacy, wenn diese die Ausgabe  $\phi \in \Phi$  mit der Wahrscheinlichkeit proportional zu

$$\exp\left(\frac{\epsilon * q(D, \phi)}{2 * s}\right)$$

ausgibt. Wobei  $s$  die Sensitivität der Bewertungsfunktion und  $\exp()$  die Eulersche Exponentialfunktion ist.

Zur Veranschaulichung des exponentiellen Mechanismus machen Zhu *et al.* ein Beispiel. In Tabelle 4 sind Krankheiten und die Anzahl erkrankter Personen aufgeführt. Die Bewertungsfunktion wird in diesem Beispiel so definiert, dass diese Funktion die Anzahl Personen zählt. Die

Sensitivität der Bewertungsfunktion ist 1, da es sich um eine Zählfunktion handelt. Basierend auf der Sensitivität und dem  $\epsilon$  wird nun für jede Krankheit eine Wahrscheinlichkeit berechnet. In Tabelle 4 sind die Wahrscheinlichkeiten für unterschiedliche Epsilon-Werte aufgeführt. Die Ausgabe des Mechanismus wird basierend auf diesen Wahrscheinlichkeiten gemacht.

Bei einem  $\epsilon$  von 0 wird jede Krankheit mit derselben Wahrscheinlichkeit ausgegeben. Dies hat zwar den höchsten Schutz der Privacy zur Folge, reduziert aber sehr stark den Nutzen der Daten. Bei einem  $\epsilon$ -Wert von 0.1 wird zu 40% als Ausgabe «Grippe» ausgegeben, aber zu 32% kann die Ausgabe auch «Diabetes» sein. Bei einem  $\epsilon$ -Wert von 1 unterscheiden sich die Wahrscheinlichkeiten der Krankheiten signifikant. Zu 88% wird das Resultat «Grippe» sein. Dadurch wird zwar ein grosser Nutzen der Daten erreicht, jedoch auf Kosten der Privacy.

Krankheit	Anzahl Personen	$\epsilon = 0$	$\epsilon = 0.1$	$\epsilon = 1.0$
Diabetes	24	0.25	0.32	0.12
Hepatitis	8	0.25	0.15	$4 * 10^{-5}$
Grippe	28	0.25	0.40	0.88
HIV	5	0.25	0.13	$8.9 * 10^{-6}$

Tabelle 4. Exponentieller Mechanismus mit Krankheitsdaten nach Zhu *et al.* [8]

## 2.6 Anwendungsfälle von Differential Privacy

Die Anwendungsfälle von Differential Privacy sind sehr vielfältig. Grundsätzlich kann die Definition von Differential Privacy für sämtliche Funktionen angewendet werden, welche auf Datensammlungen zugreifen, Daten sammeln oder veröffentlichen.

Die folgenden Anwendungsfälle werden als besonders interessant eingeschätzt:

- **Statistische Analysen:** Es wird eine Schnittstelle angeboten, welche es erlaubt Abfragen zu tätigen. Die Ausgaben werden von einem Mechanismus zurückgegeben, welcher die Eigenschaften von Differential Privacy erfüllt. [8]
- **Veröffentlichung von Daten / Statistiken:** Die Daten / Statistiken werden vor der Veröffentlichung durch einen Mechanismus aufbereitet, welcher die Eigenschaften der Differential Privacy erfüllt. [8]
- **Generieren von synthetischen Daten:** Synthetische Daten sind künstlich erzeugte Daten, welche aber im Optimalfall die Eigenschaften von realen Daten aufweisen. Synthetische Daten können beispielsweise für Softwaretests genutzt werden, um in der Testumgebung nicht mit echten Daten arbeiten zu müssen. Es kann ein Mechanismus implementiert werden, welcher aus der originalen Datensammlung eine synthetische Datensammlung generiert. Dieser Mechanismus wird so implementiert, dass dieser die

Eigenschaften von Differential Privacy erfüllt. So kann die Privacy der in der originalen Datensammlung enthaltenen Individuen gewahrt werden. [8]

- **Machine Learning:** Machine Learning Algorithmen funktionieren, indem sie eine grosse Menge von Daten untersuchen und versuchen daraus Erkenntnisse zu ziehen bzw. Modelle zu errechnen. Das Ziel von Machine Learning ist es, dass allgemeingültige Muster erkannt werden und nicht die Erkenntnisse über einzelne Individuen in das Modell einfließen. Gemäss dem National Institute of Standards and Technology (NIST) kann die Differential Privacy hierfür einen grossen Mehrwert erbringen. Hierbei wird über einen Mechanismus auf die Trainingsdaten zugegriffen, welcher die Eigenschaften der Differential Privacy erfüllt und verhindert, dass Erkenntnisse über einzelne Individuen ins Modell einfließen [15].
- **Data Mining / Recommender System:** Beim Data Mining wird eine grosse Datenmenge statistisch untersucht, um neue Verbindungen und Trends zu erkennen. Ein Recommender System ist ein System, welches basierend auf historischen Benutzerdaten versucht die Interessen und potentiellen Aktionen eines Benutzenden vorauszusagen bzw. Empfehlungen abgeben zu können. Um zu verhindern, dass aus diesen Daten Erkenntnisse über Individuen gewonnen werden können, wird dem Data Miner der Zugriff auf die Daten nur über einen Mechanismus ermöglicht, welcher die Eigenschaften der Differential Privacy erfüllt. [8]

Die beschriebenen Anwendungsfälle zeigen, wie die Differential Privacy in vielen verschiedenen Themenfeldern Anwendung finden kann.

### 2.7 Grenzen der Differential Privacy

Die Möglichkeiten und Einsatzgebiete der Differential Privacy scheinen unbegrenzt und es könnte geschlussfolgert werden, die Differential Privacy könnte sämtliche bisherigen Konzepte zum Schutz der Privacy ersetzen. Aus diesem Grund ist es wichtig die Grenzen der Differential Privacy zu betrachten und deren Anwendungen kritisch zu hinterfragen.

Domingo-Ferrer *et al.* äussern in ihrer Arbeit über die Grenzen der Differential Privacy deutliche Kritik, dass die Differential Privacy gerade von den grossen Technologie-Konzernen oftmals missbraucht und als Tarnung für «gute Privacy» verwendet wird [16]. In Abschnitt 2.3.3 wurden Beispiele für Privacy-Budgets aus der Praxis gelistet. Apple soll nach Desfontaines Privacy-Budgets von  $\epsilon = 4$  und  $\epsilon = 16$  verwenden [12]. Domingo-Ferrer *et al.* greifen diese Beispiele ebenfalls auf und schätzen solch hohe Privacy-Budgets als sinnlos ein. Die Privacy werde dadurch nicht genügend geschützt. [16]

Weiter werden nach Domingo-Ferrer *et al.* für Machine Learning Anwendungen sehr hohe

Privacy-Budgets benötigt, um sinnvolle Ergebnisse zu erzielen. Deshalb schätzen sie die Verwendung von Differential Privacy für Machine Learning als anspruchsvoll und nicht in jedem Fall sinnvoll ein. [16]

Weiter führen Domingo-Ferrer *et al.* aus, dass teilweise bewusst nur die annähernde Differential Privacy implementiert wird, um zu hohe Privacy-Budgets zu verhindern. Dafür sollen dann aber  $\delta$ -Werte verwendet werden, welche die Definition so stark lockern, dass die Privacy nicht mehr genügend geschützt wird. Dies werde oftmals gemacht, um schlechte Presse aufgrund zu hoher Privacy-Budgets zu vermeiden. [16]

Ein weiterer wichtiger Punkt den es zu beachten gilt, ist die Wahl des Privacy-Budgets bei Kompositionen von mehreren Mechanismen. Dies ist besonders beim kontinuierlichen Sammeln von Daten ein Knackpunkt. Werden beispielsweise laufend die Benutzerdaten über die Verwendung der Emojis (wie Apple dies macht [12]) gesammelt, müsste für jede einzelne Übermittlung ein eigenes Privacy-Budget alloziert werden. Da diese Daten zu demselben Individuum gehören, kommt die sequentielle Komposition zum Zug, bei welcher die einzelnen Privacy-Budgets aufsummiert werden. Die Definition von Differential Privacy erlaubt aus diesem Grund kein kontinuierliches Sammeln von Daten, da zu Beginn das Gesamt-Privacy-Budget bestimmt werden müsste. Nach Domingo-Ferrer *et al.* hat Apple deshalb die Definition soweit vereinfacht, dass nur für Daten innerhalb desselben Tages die sequentielle Komposition gilt, was eine weitere Schwächung des Privacy-Schutzes darstellt. [16]

Es ist essentiell, dass das Privacy-Budget sinnvoll und durchdacht gewählt und strikt eingehalten wird. Nur bei einem entsprechend tiefen Privacy-Budget wird auch ein guter Schutz der Privacy garantiert. Es ist wichtig für jeden Anwendungsfall kritisch zu hinterfragen, ob das Konzept der Differential Privacy sinnvoll umgesetzt werden kann.

## 3 Ideen und Konzepte

In der Aufgabenstellung werden die Inhalte der Übung in groben Zügen vorgegeben. Bezüglich der Übungsgestaltung wird in der Aufgabenstellung darauf hingewiesen, dass das Verwenden des Tools «ARX» (<https://arx.deidentifier.org>) vorgesehen ist. In Absprache mit der Betreuungsperson wurde entschieden, dass ARX keine zwingende Vorgabe ist. Es soll ein Tool verwendet werden, welches sich optimal für die Gestaltung der Übung und die Vermittlung der Inhalte eignet.

Nachfolgend werden Ideen für mögliche Übungsarten und Übungsumgebungen aufgeführt. Diese Ideen bilden die Grundlage für die schlussendliche Ausgestaltung der Übung.

### 3.1 Ideen für die Übungsart

Um die Grundlagen und die Funktionsweise von Differential Privacy einem breiten Publikum näher bringen zu können, soll eine Übung ausgearbeitet werden. Die Übung muss einfach verständlich sein und selbständig bearbeitet werden können. Für die Art der Übungsdurchführung werden nachfolgend einige Ideen beschrieben.

#### 3.1.1 Freie Laborübung

In der freien Laborübung wird durch eine initiale Aufgabenstellung das Ziel der Übung vorgegeben. Die Teilnehmenden bearbeiten die Aufgabe selbständig und ohne weitere Unterstützung. Das theoretische Wissen wird vor der Übung vermittelt, in der Regel durch einführende Aufgaben oder ein Selbststudium vor dem Übungsbeginn.

#### 3.1.2 Geführte Laborübung

Bei der geführten Laborübung werden die Teilnehmenden schrittweise durch die Übungsbearbeitung geführt. Dies erlaubt das Aufteilen der Übung in mehrere Teilübungen. Weiter erlaubt diese Aufteilung, die Theorie ebenfalls in mehrere Teil-Blöcke zu gliedern. So könnte die Theorie abwechselnd zu den Teil-Übungen vermittelt werden.

### 3.1.3 Zeitbasierte Challenge

Bei der zeitbasierten Challenge treten die Übungsteilnehmenden gegeneinander an. Dies erzeugt ein kompetitives Umfeld und regt zur schnellen Übungsbearbeitung an. Die Übung muss so ausgestaltet sein, dass die Bearbeitungszeit der Teilnehmenden gemessen und verglichen werden kann. Die zeitbasierte Challenge setzt voraus, dass mehrere Teilnehmende die Übung bearbeiten. Die Theorie wird vor der Challenge vermittelt, sodass alle Teilnehmenden auf demselben Kenntnisstand sind.

### 3.1.4 Inhaltbasierte Challenge

Bei der inhaltbasierten Challenge entscheidet der erfolgreiche Abschluss der Übung über den Erfolg der Challenge. Die Übung muss so ausgestaltet sein, dass die Ergebnisse der Teilnehmenden vergleichbar sind. Die Challenge könnte in mehrere Teil-Challenges aufgeteilt werden. So liessen sich die Teilnehmenden anhand der bearbeiteten Teil-Challenges vergleichen. Die Theorie wird entweder vor der Challenge oder jeweils vor den einzelnen Teil-Challenges vermittelt, insofern die Challenge aufgeteilt wird.

## 3.2 Ideen für die Umsetzung der Übungsumgebung

Auf Grundlage der Ideen für die Übungsarten bieten sich folgende Übungsumgebungen an:

- **Virtuelle Maschine**, innerhalb welcher die Übung durchgeführt wird. In der virtuellen Maschine sind die Übungsdaten und Programmdateien enthalten. Alle Übungsteilnehmenden erhalten dieselbe virtuelle Maschine und betreiben diese auf dem persönlichen Gerät.
- **Web-Applikation**, welche die Funktionalitäten und Daten anbietet, um die Übung durchzuführen. Die Web-Applikation wird entweder auf einem Web-Server angeboten oder alle Übungsteilnehmenden starten die Web-Applikation lokal auf dem persönlichen Gerät.
- **Lokale Installation** der notwendigen Werkzeuge auf den Geräten der Übungsteilnehmenden. Es werden nur die dafür notwendigen Dateien und Anleitungen zur Verfügung gestellt.
- **Jupyter Notebook** welches die Theorie und Übungen dynamisch miteinander verbindet. Der Code kann innerhalb des Jupyter Notebooks ausgeführt werden und die Benutzenden erhalten ein direktes Feedback. Weiter erlaubt es Plots und Graphen zu generieren, was für statistische Anwendungen sehr nützlich ist.

## 4 Methoden

Während der Projektumsetzung werden unterschiedliche Methoden eingesetzt. Diese werden in den nachfolgenden Abschnitten beschrieben und begründet. Es wird zwischen der Projektmethodik und der Fachmethodik unterschieden. Die Projektmethodik beschreibt die Methoden der Projektplanung und -durchführung. Die für die Erarbeitung des Inhalts und die Umsetzung der Arbeitsziele benötigten Methoden werden als Fachmethodik zusammengefasst.

### 4.1 Projektmethodik

Diese Projektarbeit wird nach dem klassischen Vorgehensmodell geplant und durchgeführt. Ein wasserfallartiges und phasenweises Vorgehen bietet sich an, da die Konzeptbildung und Realisierung zeitlich getrennt sind. Weiter spricht für das klassische Vorgehensmodell, dass sich das Projekt in eine Initialisierungsphase, Konzeptionsphase, Realisierungsphase und Einführungsphase unterteilen lässt.

In Abbildung 11 und 12 ist der SOLL-Projektplan ersichtlich. Um einen kontinuierlichen Abgleich mit der Betreuungsperson sicherstellen zu können, wurden wöchentliche Termine eingeplant. Diese Termine wurden für eine bessere Übersichtlichkeit in einem Arbeitsvorgang (Nr. 001) zusammengefasst.

Der Projektplan hält zudem die geplante Dauer, sowie die effektive Dauer pro Arbeitsvorgang fest. Dies ermöglicht eine stetige SOLL-IST-Analyse des Projektfortschritts und erlaubt eine zeitnahe Reaktion auf potentielle Projektverzögerungen.

Durch die fest vorgegebene Zeitdauer der Projektarbeit ist eine Verlängerung des Projekts ausgeschlossen. Aus diesem Grund muss bei einer Projektverzögerung der Umfang des Projekts reduziert werden. Während der Projektinitialisierung hat sich ergeben, dass zusätzlich zu der Erstellung der Übung auch eine Testdurchführung der Übung sinnvoll ist. Diese Testdurchführung ist nicht vorgegeben und nicht verpflichtend. Bei der Projektplanung wurde auf das Einplanen eines Zeitpuffers bewusst verzichtet. Sollte es zu Projektverzögerungen kommen, wird der Umfang der Testdurchführung reduziert oder die Testdurchführung entfällt ganz.

## 4 Methoden

Nr.	Aufgabe	Status	SW		Geplante Dauer	Effektive Dauer
			Start	Ende		
001	Wöchentliche Abgleichsmeetings mit Esther Hänggi	Offen	01	16	360.0 h	0.0 h
<b>INITIALISIERUNGSPHASE</b>						
002	Kick-off Meeting mit Esther Hänggi	Offen	01	01	2.0 h	
003	Vorbereitung Arbeitsumgebung, Dokumentenablage, Vorlagen	Offen	01	01	8.0 h	
004	Projektplanung erstellen	Offen	01	01	6.0 h	
005	Einlesen in die Grundlagen von "Differential Privacy"	Offen	01	01	8.0 h	
<b>M1</b>	<b>PROJEKT FERTIG GEPLANT</b>	<b>Meilenstein</b>	-	-	-	-
006	Vision verfassen, Aufgabenstellung paraphrasieren, Nutzen aufzeigen	Offen	02	02	16.0 h	
007	Didaktische Methoden, Varianten für die Stoffvermittlung	Offen	03	03	16.0 h	
<b>KONZEPTIONSPHASE</b>						
008	Varianten und Entscheid für Übungsgestaltung	Offen	03	04	16.0 h	
<b>M2</b>	<b>GROBKONZEPT DER ÜBUNGSGESTALTUNG FERTIG</b>	<b>Meilenstein</b>	-	-	-	-
009	Aufarbeitung und Strukturierung der Theorie	Offen	04	05	20.0 h	
<b>REALISIERUNGSPHASE</b>						
010	Erstellen der Übungen und Unterlagen	Offen	05	08	80.0 h	
<b>M3</b>	<b>ÜBUNGEN UND UNTERLAGEN FERTIG</b>	<b>Meilenstein</b>	-	-	-	-
011	Zwischenpräsentation (Voraussichtlich in SW09)	Offen	09	09	16.0 h	
<b>EINFÜHRUNGSPHASE</b>						
012	Testdurchführung der Übung planen / Vorbereiten	Offen	09	10	14.0 h	
013	1. Zwischenkorrektur und Überarbeitung	Offen	10	10	14.0 h	
014	Testdurchführung der Übung	Offen	11	11	8.0 h	
015	Auswerten der Testdurchführung	Offen	11	11	12.0 h	
<b>M4</b>	<b>TESTDURCHFÜHRUNG ABGESCHLOSSEN</b>	<b>Meilenstein</b>	-	-	-	-
016	2. Zwischenkorrektur und Überarbeitung	Offen	11	13	14.0 h	
017	Mögliche Verbesserungsmöglichkeiten / Änderungen festhalten	Offen	13	14	16.0 h	
018	Erstellen Pitching-Video	Offen	14	14	16.0 h	
019	Abstract schreiben	Offen	15	15	12.0 h	
020	Webabstract schreiben	Offen	15	15	12.0 h	
021	Abschliessende Korrektur der Arbeit	Offen	15	15	16.0 h	
<b>M5</b>	<b>SCHRIFTLICHE ARBEIT ABGESCHLOSSEN</b>	<b>Meilenstein</b>	-	-	-	-
022	Vorbereitung der Abgabe (Erfassen auf Portfolio-DB, etc.)	Offen	16	16	8.0 h	
023	Abgabe Bachelorarbeit (am 10.06.2022)	Offen	16	16	2.0 h	
024	Vorbereitung Abschlusspräsentation / MEP	Offen	16	19	12.0 h	
025	Abschlusspräsentation / MEP	Offen	18	19	2.0 h	

Verbleibende Dauer: 0.00 360.00

Abb. 11. Arbeitspakete und Meilensteine des SOLL-Projektplans

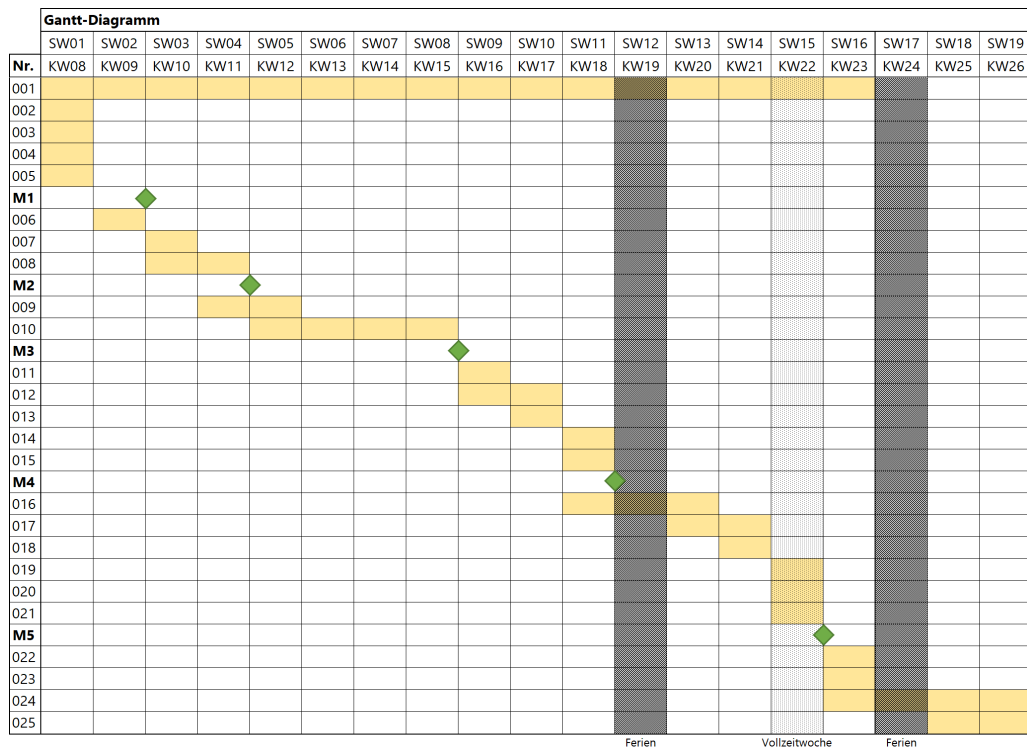


Abb. 12. Gantt-Diagramm des SOLL-Projektplans

## 4.2 Fachmethodik

Es werden die angewandten Methoden für die Übungsgestaltung sowie die Überprüfung der Zielerreichung erläutert. Weiter werden die für die Gestaltung der Übung relevanten Inhalte aus der Didaktik aufgeführt.

### 4.2.1 Methoden für die Übungsgestaltung

Die Übung ist so zu gestalten, dass die für die Durchführung der Übung benötigten theoretischen Kenntnisse vorab vermittelt werden. Die Übung soll so ausgelegt sein, dass möglichst geringe Vorkenntnisse notwendig sind.

Für die Planung einer Lehrveranstaltung sind gemäss Leitfaden der Hochschuldidaktik Universität Zürich folgende Planungsschritte sinnvoll [17]:

1. **Voraussetzungen klären:** Feststellen der Vorkenntnisse der Teilnehmenden, Anzahl der Teilnehmenden bestimmen, Ort der Durchführung festlegen.
2. **Lerninhalt bestimmen:** Lerninhalte identifizieren und abgrenzen.
3. **Lernziele setzen:** Formulieren der zu vermittelnden Kompetenzen.
4. **Lernzeiten unterscheiden:** Bestimmung der Verteilung von Präsenzzeit sowie der Vor- und Nachbereitungszeit (Selbststudium).
5. **Kompetenznachweis planen:** Festlegen der Form des Nachweises über die erlangten Kompetenzen gemäss den Lernzielen.
6. **Lehr-Lernformen finden:** Art und Weise der Wissensvermittlung definieren.

Diese Planungsschritte werden als Orientierung für die Übungsgestaltung verwendet. In einem ersten Schritt werden die Voraussetzungen geklärt. Danach wird der Lerninhalt bestimmt und Lernziele formuliert. Abschliessend soll eine passende Lehr-Lernform gefunden werden, welche die Art und Weise der Übungsumsetzung definiert.

Die Schritte «Lernzeiten unterscheiden» und «Kompetenznachweis planen» sind für diese Arbeit nicht relevant und werden deshalb weggelassen.

Nach Bloom *et al.* sind Lernziele nach den folgenden Taxonomiestufen zu gliedern [18]:

- K1. **Wissen:** Gelerntes Wissen abrufen und in gleichartiger Situation wiedergeben können.
- K2. **Verstehen:** Gelerntes Wissen in eigenen Worten wiedergeben und sinngemäss abbilden können.

- K3. **Anwenden:** Erlernte Fähigkeit in verschiedenen Situationen anwenden können.
- K4. **Analysieren:** Eine komplexe Situation systematisch untersuchen, Strukturen und Sachverhalte erkennen und gliedern können.
- K5. **Synthese:** Einzelne Elemente eines Sachverhalts zu einem Ganzen zusammenführen können.
- K6. **Bewerten:** Einen komplexen Sachverhalt untersuchen und nach eigenen Kriterien beurteilen können.

Die Hochschuldidaktik der Justus-Liebig-Universität Giessen veröffentlichte aktivierende Methoden für Seminare und Übungen. Dabei gliedert sich der Aufbau einer Übung in «Warm Up / Einstieg», «Work Out / Arbeitsphase» und «Cool Down / Ausstieg» [19]. Für die Übungsgestaltung werden folgende aktivierende Methoden pro Übungsphase als sinnvoll beurteilt:

- Warm Up / Einstieg: **Anekdote zu Beginn**

Durch eine Anekdote oder Geschichte zu Beginn, wird der Einstieg ins Thema interessant und praxisbezogen gestaltet. Den Übungsteilnehmenden soll daraus die Relevanz der folgenden Übung bewusst werden. Dies vereinfacht den Einstieg in die Thematik.

- Work Out / Arbeitsphase: **Simulation**

Vorgänge und Systeme werden anhand von Modellen gezeigt. Die Übungsteilnehmenden können durch das Verändern von Parametern selbständig die Auswirkung auf das Modell ausprobieren. Die Simulation wird dadurch interaktiv und die Thematik verständlicher. Es erlaubt das Durchspielen von verschiedenen Szenarien und stellt diese in Bezug zur Realität.

- Cool Down / Ausstieg: **Klausurfragen**

Zum Abschluss der Übung werden einige Fragen gestellt, welche die Übungsteilnehmenden beantworten. Besteht die Übung aus mehreren Teilen, können die Fragen am Ende jedes Teils gestellt und zu Beginn des nächsten Übungsteils beantwortet werden.

### 4.2.2 Methoden für die Überprüfung der Zielerreichung

Das Hauptziel der Arbeit ist es, das Thema «Differential Privacy» anhand einer praktischen Übung einem breiten Publikum verständlich zu machen. Die Erreichung dieses Ziels wird anhand einer Testdurchführung der Übung überprüft. Durch die Testdurchführung soll geprüft werden, ob die Lernziele den Test-Teilnehmenden auf den definierten Taxonomiestufen ver-

mittelt werden konnten.

Der Leitfaden der Hochschuldidaktik Universität Zürich weist darauf hin, dass das Einholen von Rückmeldungen der Teilnehmenden ein effektives Mittel für die Qualitätssicherung und Lektionsauswertung ist [17]. Deshalb wird nach der Testdurchführung der Übung umfangreiches Feedback der Teilnehmenden eingeholt. Anhand dieses Feedbacks soll die Übung verbessert und vervollständigt werden können.

## 5 Realisierung

Die Theorie aus Kapitel 2 wird aufbereitet und durch entsprechende praktische Aufgaben ergänzt, um das Thema «Differential Privacy» einem breiten Publikum verständlich zu machen. Die Eigenschaften, Varianten und Möglichkeiten von Differential Privacy sollen möglichst interaktiv und abwechslungsreich ausprobiert werden können.

### 5.1 Gestaltung der Übung

Die Schritte für die Übungsgestaltung richten sich an die in Abschnitt 4.2.1 beschriebenen Schritte für die Planung einer Lehrveranstaltung nach dem Leitfaden der Hochschuldidaktik Universität Zürich.

#### 5.1.1 Voraussetzungen klären

Die Übung richtet sich an ein breites Publikum. Dies können Studierende, Forschende, Lehrende oder anderweitig am Thema «Differential Privacy» interessierte Personen sein. Es ist wichtig die vorausgesetzten Kenntnisse der Übung klar zu definieren und abzugrenzen.

Um einfache Code-Beispiele für die praktische Implementierung von «Differential Privacy» verstehen zu können, sind Kenntnisse der Programmierung notwendig. Die Code-Beispiele sollen einfach gehalten werden, um die Komplexität der Übung zu reduzieren. Aus diesem Grund sind nur grundlegende Kenntnisse in der Programmierung notwendig, wie die Definition und das Aufrufen von einfachen Funktionen, ein Verständnis der primitiven Datentypen, Arrays und Schleifen, sowie die Ausgabe von Werten. Weiter sind für das Verständnis der theoretischen Inhalte Kenntnisse aus der Mathematik notwendig. Dazu gehört das Verstehen von Variablen und Gleichungen, sowie die Grundlagen der Funktionslehre, Wahrscheinlichkeitsrechnung und Statistik (Verteilungen und Varianz).

Die Übung wird so gestaltet, dass diese selbständig und in Einzelarbeit bearbeitet werden kann. Die Theorie wird abwechslungsweise zu den praktischen Aufgaben vermittelt. So können die Übungsteilnehmenden schrittweise das theoretische Wissen aneignen und direkt anhand von praktischen Aufgaben festigen.

Um ein breites Publikum erreichen zu können, soll der Durchführungsort flexibel sein. Es erscheint sinnvoll, die Übung online anzubieten. Die Installation von Software durch die Übungsteilnehmenden soll vermieden werden, da dies die Fehleranfälligkeit erhöht. Weiter wird die Aushändigung der Übung deutlich vereinfacht, wenn diese online angeboten wird.

An der Hochschule Luzern wird das Modul «Privacy» unterrichtet. Eine Integration dieser Übung in das Privacy-Modul ist nach Angaben des Modulverantwortlichen wünschenswert. Aus diesem Grund wird die Übung so gestaltet, dass diese in einen Unterrichtsblock integriert werden kann und einer Gesamtdauer für die Übungsdurchführung von ungefähr 2 Stunden entspricht.

### **5.1.2 Lerninhalt bestimmen**

Um die Notwendigkeit für Ansätze wie die Differential Privacy zu erkennen, wird das Bewusstsein für die Schwächen der klassischen Ansätze als wichtig betrachtet. Das Aufzeigen dieser Schwächen soll deshalb Bestandteil der Übung sein. Weiter sind die Definitionen und Funktionsweisen von Differential Privacy essentiell für das Verständnis des Themas. Diese werden entsprechend ebenfalls in die Übung aufgenommen. Das Privacy-Budget ist ein essentieller Aspekt der Differential Privacy und soll in der Übung ausführlich behandelt werden. Für ein praktisches Verständnis der Differential Privacy sollen der Laplace, Gauss sowie exponentielle Mechanismus implementiert werden. Dies erlaubt es den Übungsteilnehmenden die theoretischen Inhalte anhand von Code-Beispielen zu festigen. Abschliessend sollen die möglichen Anwendungsfälle sowie die Grenzen der Differential Privacy aufgezeigt werden. Dies wird als besonders relevant betrachtet, da den Übungsteilnehmenden einerseits gezeigt werden soll, wie vielfältig die Differential Privacy eingesetzt werden kann und andererseits, dass sie auch Grenzen hat und nicht für jeden Anwendungsfall geeignet ist.

Der Fokus der Übung soll auf der Vermittlung der genannten Inhalte liegen. Weitere Inhalte sollen erwähnt werden, liegen aber nicht im Fokus der Übung. Die fachliche Tiefe der Inhalte wird durch die Taxonomiestufen der Lernziele definiert. Je höher die Taxonomiestufe eines Lernziels, desto höher der Anspruch an die fachliche Tiefe der jeweiligen Inhalte.

### **5.1.3 Lernziele setzen**

Die Lernziele dienen einerseits der stufengerechten Umsetzung der Übung, andererseits sollen sie auch als Orientierungshilfe für die Übungsteilnehmenden dienen. In Tabelle 5 sind sämtliche Lernziele mit dazugehöriger Taxonomiestufe aufgeführt. Die Lernziele wurden mit der Betreuungsperson abgesprachen und als sinnvoll bestätigt.

ID	Taxonomiestufe	Lernziel
1	K5	Die Teilnehmenden sind in der Lage, die klassischen Ansätze der Datenanonymisierung kritisch zu beurteilen.
2	K3	Die Teilnehmenden sind in der Lage, das Prinzip der $k$ -Anonymität anzuwenden.
3	K3	Die Teilnehmenden sind in der Lage, die Eigenschaften der Differential Privacy für eine einfache Funktion mittels des Laplace Mechanismus zu implementieren.
4	K3	Die Teilnehmenden sind in der Lage, die Eigenschaften der Differential Privacy für eine einfache Funktion mittels des Gauss Mechanismus zu implementieren.
5	K3	Die Teilnehmenden sind in der Lage, die Eigenschaften der Differential Privacy für eine einfache Funktion mittels des exponentiellen Mechanismus zu implementieren.
6	K2	Die Teilnehmenden sind in der Lage, die Definition der $\epsilon$ -Differential Privacy zu erklären.
7	K2	Die Teilnehmenden sind in der Lage, das Konzept des Privacy-Budgets zu erklären.
8	K2	Die Teilnehmenden sind in der Lage, die beiden Modelle (lokal und zentral) der Differential Privacy zu unterscheiden.
9	K2	Die Teilnehmenden sind in der Lage, die Auswirkung des Privacy-Budgets auf den Wissensgewinn des Angreifers abzuschätzen.
10	K2	Die Teilnehmenden sind in der Lage, die Sensitivität einer einfachen Funktion abzuschätzen.
11	K2	Die Teilnehmenden sind in der Lage, zwischen der unbegrenzten und begrenzten Sensitivität zu unterscheiden.
12	K1	Die Teilnehmenden sind in der Lage, die Definition der $(\epsilon, \delta)$ -Differential Privacy abzurufen.
13	K1	Die Teilnehmenden sind in der Lage, die Vor- und Nachteile des Laplace und Gauss Mechanismus zu benennen.
14	K1	Die Teilnehmenden sind in der Lage, die verschiedenen Anwendungsfälle der Differential Privacy zu benennen.
15	K1	Die Teilnehmenden sind in der Lage, das Konzept der Komposition des Privacy-Budgets zu bezeichnen.
16	K1	Die Teilnehmenden sind in der Lage, die Grenzen der Differential Privacy zu benennen.

Tabelle 5. Lernziele absteigend nach Taxonomiestufe

### 5.1.4 Lehr-Lernform finden

Als Lehr-Lernform wird die Art und Weise bezeichnet, wie und in welchem Rahmen die Inhalte vermittelt werden. In diesem Fall ist die Form aus der Aufgabenstellung bereits als praktische Übung vorgegeben. Die Art und Weise, wie die Übung umgesetzt wird, ist nicht vorgegeben. In Kapitel 3 wurden folgende Ideen für mögliche Übungsarten eruiert: Freie Laborübung, geführte Laborübung, zeitbasierte Challenge oder inhaltbasierte Challenge.

Bei der freien Laborübung müssten die notwendigen Theorieinhalte vor der Übungsdurchführung vollständig vermittelt werden, sodass die Teilnehmenden anschliessend die Übung selbständig bearbeiten können. Aufgrund der hohen Komplexität des Themas bietet sich eher eine geführte Laborübung an. So werden die Übungsteilnehmenden schrittweise durch die Bearbeitung der Übung geführt. Weiter ermöglicht dieser Ansatz, die Theorie abwechselungsweise zu den praktischen Aufgaben zu vermitteln. Dies erlaubt es den Übungsteilnehmenden das Gelernte laufend anhand von kleineren praktischen Aufgaben zu festigen.

Eine Challenge (zeit- oder inhaltbasiert) wird für diese Übung ausgeschlossen, denn analog zur freien Laborübung müssten die Inhalte vor der Übungsdurchführung vollständig vermittelt werden.

Für die freie Laborübung muss eine geeignete Übungsumgebung gewählt werden. In Kapitel 3 wurden die folgenden möglichen Übungsumgebungen beschrieben: Virtuelle Maschine, Web-Applikation, lokale Installation oder Jupyter Notebook.

Da die Übung einem möglichst breiten Publikum zur Verfügung gestellt werden soll, sollte die Übungsumgebung möglichst einfach gestaltet sein. Die Varianten «virtuelle Maschine» und «lokale Installation» verlangen, dass die Übungsteilnehmenden entweder eine virtuelle Umgebung zur Verfügung stellen oder Software lokal installieren. Dies stellt viele Anforderungen und Bedingungen an die von den Übungsteilnehmenden verwendete Infrastruktur (Betriebssystem, Konfiguration der Software, usw.). Aus diesem Grund werden diese beiden Optionen für die Übungsumsetzung ausgeschlossen.

Jupyter Notebook bietet die optimalen Voraussetzungen, um eine geführte Laborübung umzusetzen. Die Theorieinhalte können fließend mit Code, Statistiken und anderen interaktiven Funktionen vereint werden. Weiter kann das Jupyter Notebook den Übungsteilnehmenden online zur Verfügung gestellt werden. Die Entwicklung einer eigenen Web-Applikation würde diese Aspekte ebenfalls erfüllen, der Aufwand für deren Umsetzung wäre aber im Vergleich zum Jupyter Notebook erheblich grösser.

Aus den genannten Gründen wird eine **geführte Laborübung** anhand von **Jupyter Notebook** realisiert.

## 5.2 Ausarbeitung der Übung

Die Übung wird in mehrere einzelne Jupyter Notebooks aufgeteilt, um die Inhalte übersichtlich zu strukturieren. Die Übungsteilnehmenden haben dadurch beim Abschluss jedes Jupyter Notebooks ein kleines Erfolgserlebnis, was sich motivierend für die weitere Übungsbearbeitung auswirken soll. Zudem erlaubt diese Struktur, dass einzelne Inhalte innerhalb eines Jupyter Notebooks abgeschlossen und mittels Repetitionsfragen gefestigt werden können. Die Musterlösungen zu den Repetitionsfragen sind jeweils im darauffolgenden Jupyter Notebook gegeben. So kann laufend überprüft werden, ob die Inhalte korrekt verstanden wurden.

### 5.2.1 Notebook 0: Übersicht und Inhalt

Dieses Jupyter Notebook gibt eine Übersicht über die Übung. Die weiteren Jupyter Notebooks sind verlinkt. Weiter ist beschrieben, an wen sich die Übung richtet und welche Vorkenntnisse empfohlen werden. Es wird eine Übersicht aller Lernziele gegeben, welche im Rahmen der Übung erreicht werden sollen.

### 5.2.2 Notebook 1: Einführung in die Thematik

In diesem Notebook werden die folgenden Lernziele erreicht:

- Die Teilnehmenden sind in der Lage, das Prinzip der  $k$ -Anonymität anzuwenden.
- Die Teilnehmenden sind in der Lage, die klassischen Ansätze der Datenanonymisierung kritisch zu beurteilen.

Zur Einführung in die Thematik wird anhand eines Beispiels zur  $k$ -Anonymität und  $l$ -Diversität die Funktionsweise klassischer Ansätze zum Schutz der Privacy gezeigt. Anhand einer praktischen Aufgabe soll eine gegebene Tabelle in eine 4-anonyme Version umgewandelt werden. Die Übungsteilnehmenden vertiefen dadurch einerseits die Funktionsweise der  $k$ -Anonymität, sollen aber andererseits auch feststellen, dass die Erreichung der  $k$ -Anonymität für eine Datensammlung keine einfache Aufgabe ist.

Nach der praktischen Aufgabe wird aufgezeigt, wie die  $k$ -anonyme und  $l$ -diverse Tabelle anhand von Hintergrundinformationen de-anonymisiert bzw. die darin enthaltenen Personen re-identifiziert werden können. Dadurch wird die Relevanz der Differential Privacy aufgezeigt und der Übergang zum nächsten Notebook gemacht.

### 5.2.3 Notebook 2: Die Definition der Differential Privacy

In diesem Notebook werden die folgenden Lernziele erreicht:

- Die Teilnehmenden sind in der Lage, die beiden Modelle (lokal und zentral) der Differential Privacy zu unterscheiden.
- Die Teilnehmenden sind in der Lage, die Definition der  $\epsilon$ -Differential Privacy zu erklären.
- Die Teilnehmenden sind in der Lage, das Konzept des Privacy-Budgets zu erklären.
- Die Teilnehmenden sind in der Lage, die Auswirkung des Privacy-Budgets auf den Wissensgewinn des Angreifers abzuschätzen.
- Die Teilnehmenden sind in der Lage, das Konzept der Komposition des Privacy-Budgets zu bezeichnen.

Zu Beginn dieses Notebooks werden das lokale und das zentrale Modell von Differential Privacy kurz erklärt, jedoch nicht im Detail behandelt. Wichtig ist, den Übungsteilnehmenden aufzuzeigen, dass es zwei Modelle von Differential Privacy gibt und wie sich diese unterscheiden.

Danach wird die Grundidee von Differential Privacy erklärt. Es werden Begriffe wie beispielsweise «benachbarte Datensammlung» eingeführt, welche anschliessend für das Verständnis der Definition von Differential Privacy benötigt werden. Es wird deutlich zwischen der Grundidee von Differential Privacy, der Definition von Differential Privacy und der Implementierung von Differential Privacy unterschieden. Weiter wird in der Erklärung der Grundidee ein einfaches Beispiel für das Hinzufügen von Rauschen zu einer Datenbankabfrage gegeben. Es wird gezeigt, wie das Hinzufügen von Rauschen die Privacy schützen kann.

Um die Unterscheidung von Hinweisen, Definitionen, Übungen und Erkenntnissen zu verdeutlichen, werden diese in einer unterschiedlich farbigen Textbox abgebildet. Hinweise werden blau hinterlegt, Definitionen grün, Übungen rot und Erkenntnisse gelb. Die Idee hinter den Erkenntnis-Textboxen ist es, die essentiell wichtigen Erkenntnisse nochmals zu wiederholen und hervorzuheben.

Von der formalen Definition der  $\epsilon$ -Differential Privacy wird zur Quantifizierung des Angreiferwissens und damit zur Bedeutung des Privacy-Budgets übergeleitet. Der Vorteil der Differential Privacy, den Wissensgewinn eines Angreifers genau bestimmen zu können, soll aufgezeigt werden.

Die Definition der annähernden  $(\epsilon, \delta)$ -Differential Privacy wird bewusst noch nicht eingeführt, um die Teilnehmenden nicht zu überfordern. Die Teilnehmenden sollen zuerst ein besseres Verständnis für die Differential Privacy entwickeln. Die Definition der annähernden Differential

Privacy soll deshalb erst bei der Implementierung des Gauss Mechanismus erläutert werden.

Das Jupyter Notebook wird mit der Komposition des Privacy-Budgets abgeschlossen. Es wird die parallele Komposition und die sequentielle Komposition anhand von je einem Beispiel erklärt. Zum besseren Verständnis der beiden Kompositionsarten werden diese anhand einer praktischen Aufgabe repetiert.

### 5.2.4 Notebook 3: Der Laplace Mechanismus

In diesem Notebook werden die folgenden Lernziele erreicht:

- Die Teilnehmenden sind in der Lage, die Sensitivität einer einfachen Funktion abzuschätzen.
- Die Teilnehmenden sind in der Lage, zwischen der unbegrenzten und begrenzten Sensitivität zu unterscheiden.
- Die Teilnehmenden sind in der Lage, die Eigenschaften der Differential Privacy für eine einfache Funktion mittels des Laplace Mechanismus zu implementieren.

Zum Beginn dieses Jupyter Notebooks wird die begrenzte und unbegrenzte Sensitivität eingeführt. Darauf folgt die Erklärung der Definition des Laplace Mechanismus.

Anschliessend wird der Laplace Mechanismus anhand eines einfachen Beispiels in Python implementiert. Da in Jupyter Notebook beliebig zwischen Code- und Text-Blöcken abgewechselt werden kann, können die jeweiligen Code-Blöcke laufend erklärt werden. Es wird eine einfache Zählfunktion implementiert, welche in einer Datensammlung bestimmte Datensätze zählt und ausgibt. In einer praktischen Aufgabe sollen unterschiedliche Privacy-Budgets gewählt und die Auswirkungen beobachtet werden.

Das erste Beispiel wird bewusst sehr einfach gehalten. Die Datensammlung umfasst nur 5 Datensätze. So kann das Prinzip des Laplace Mechanismus besser nachvollzogen werden. Anschliessend wird ein umfangreicheres Beispiel mit mehreren Tausend Datensätzen gezeigt. Die Teilnehmenden können erkennen, dass die Ausgaben bei grossen Datensammlungen genauer werden.

### 5.2.5 Notebook 4: Der Gauss Mechanismus

In diesem Notebook werden die folgenden Lernziele erreicht:

- Die Teilnehmenden sind in der Lage, die Definition der  $(\epsilon, \delta)$ -Differential Privacy abzurufen.

- Die Teilnehmenden sind in der Lage, die Eigenschaften der Differential Privacy für eine einfache Funktion mittels des Gauss Mechanismus zu implementieren.
- Die Teilnehmenden sind in der Lage, die Vor- und Nachteile des Laplace und Gauss Mechanismus zu benennen.

Während der Laplace Mechanismus die reine Differential Privacy erfüllt, kann der Gauss Mechanismus nur die annähernde Differential Privacy implementieren. Deshalb wird dieses Jupyter Notebook mit der Erklärung und Definition der annähernden Differential Privacy begonnen. Auf eine detaillierte Erklärung wird bewusst verzichtet, da dies die Komplexität der Übung deutlich erhöhen würde.

Im Anschluss zur Definition der annähernden Differential Privacy wird in die Definition des Gauss Mechanismus eingeleitet. Es wird dasselbe Beispiel wie beim Laplace Mechanismus praktisch implementiert und der Code laufend erklärt. Als praktische Aufgabe soll eigener Code implementiert werden, welcher den Mechanismus 1'000 Mal ausführt und den Durchschnitt der Ergebnisse berechnet. Es sollen verschiedene Werte für das Privacy-Budget und das Delta verwendet und die Auswirkungen beobachtet werden. Es wird ersichtlich, dass die Ergebnisse je nach Privacy-Budget und Delta stärker oder weniger stark vom tatsächlichen Wert abweichen. Dadurch sollen die Übungsteilnehmenden ein Gespür entwickeln, was diese beiden Parameter für eine Bedeutung haben.

Anschliessend wird auch für den Gauss Mechanismus ein realitätsnahes Beispiel mit einer grossen Datensammlung implementiert.

Da es sich beim Laplace und beim Gauss Mechanismus um zwei ähnliche Mechanismen handelt, wird dieses Jupyter Notebook mit einem Vergleich dieser beiden Mechanismen abgeschlossen.

### **5.2.6 Notebook 5: Der exponentielle Mechanismus**

In diesem Notebook werden die folgenden Lernziele erreicht:

- Die Teilnehmenden sind in der Lage, die Eigenschaften der Differential Privacy für eine einfache Funktion mittels des exponentiellen Mechanismus zu implementieren.

Der exponentielle Mechanismus kann für nicht-numerische Funktionen benutzt werden, weshalb sich dieser deutlich vom Laplace und vom Gauss Mechanismus unterscheidet. Dieses Jupyter Notebook wird mit der Definition und einem einfachen Beispiel des exponentiellen Mechanismus eingeleitet. Dieses Beispiel wird anschliessend praktisch implementiert.

Als praktische Aufgaben sollen die Übungsteilnehmenden für die gegebene Implementierung des exponentiellen Mechanismus verschiedene Privacy-Budgets wählen und die Auswirkungen

gen beobachten. Bei dieser Aufgabe wird zum einen nochmals verdeutlicht, dass zur Ausgabe der Funktion zwar kein Rauschen hinzugefügt wird, die Funktion aber dennoch probabilistisch ist. Es sollen bewusst auch die Privacy-Budgets 0 und 1 gewählt und beobachtet werden. Die Auswirkungen auf den Schutz der Privacy werden dadurch sehr deutlich und die Übungsteilnehmenden sollen erkennen, dass ein zu hoch gewähltes Privacy-Budget jeglichen Schutz aufhebt.

Abschliessend wird auch für den exponentiellen Mechanismus ein umfangreicheres Beispiel implementiert, um ein realitätsnahes Beispiel geben zu können.

### 5.2.7 Notebook 6: Anwendungsfälle und Grenzen der Differential Privacy

In diesem Notebook werden die folgenden Lernziele erreicht:

- Die Teilnehmenden sind in der Lage, die verschiedenen Anwendungsfälle der Differential Privacy zu benennen.
- Die Teilnehmenden sind in der Lage, die Grenzen der Differential Privacy zu benennen.

Dieses Jupyter Notebook soll die Übung abrunden und abschliessen. Zuerst werden die vielfältigen Anwendungsfälle der Differential Privacy aufgezeigt. Damit sollen den Übungsteilnehmenden die vielseitigen Möglichkeiten der Differential Privacy gezeigt und deutlich gemacht werden, dass die Differential Privacy grundsätzlich für sämtliche Funktionen angewendet werden kann, welche Daten sammeln, veröffentlichen oder auf Datensammlungen zugreifen.

Von den vielfältigen Einsatzgebieten der Differential Privacy wird zu den Grenzen der Differential Privacy übergeleitet. Es wird kritisch hinterfragt, in welchen Anwendungsfällen der Einsatz von Differential Privacy tatsächlich Sinn macht. Weiter werden die in der Praxis von den grossen Technologie-Konzernen verwendeten Privacy-Budgets hinterfragt. Wie in Abschnitt 2.3.3 beschrieben, soll Apple nach Desfontaines Privacy-Budgets von  $\epsilon = 4$  und  $\epsilon = 16$  verwenden [12]. Den Übungsteilnehmenden soll bewusst werden, dass diese Privacy-Budgets sehr hoch gewählt sind. Auch wenn für die abschliessende Beurteilung des Privacy-Budgets detaillierte Kenntnisse zu dessen Implementierung notwendig sind, sollen die Übungsteilnehmenden dazu bewegt werden, die Wahl der Privacy-Budgets kritisch zu hinterfragen.

Dieses Jupyter Notebook wird durch ein kurzes, abschliessendes Kapitel abgerundet. Es werden nochmals die wichtigsten Erkenntnisse aufgelistet, welche aus der Übung mitgenommen werden sollen und damit die Übung abgeschlossen.

### 5.2.8 Notebook 7: Musterlösungen der Übungen

In diesem Jupyter Notebook sind die Musterlösungen der praktischen Aufgaben enthalten und detailliert erklärt. Die Musterlösungen wurden bewusst in ein separates Jupyter Notebook ausgelagert, um die Übersichtlichkeit zu verbessern.

Die Jupyter Notebooks 1-5 werden mit Kontrollfragen abgeschlossen. Die Lösungen dieser Kontrollfragen wurden bewusst in das jeweils nächste Jupyter Notebook und nicht in das Musterlösungs-Notebook integriert. Dies erlaubt es den Übungsteilnehmenden ihr Verständnis laufend zu Beginn des nächsten Notebooks zu kontrollieren.

## 5.3 Bereitstellung der Übung

Die Jupyter Notebooks werden im GitLab-Projekt «BAA\_FS22\_Differential-Privacy»<sup>1</sup> des EnterpriseLab der Hochschule Luzern zur Verfügung gestellt.

Das GitLab-Projekt ist wie folgt strukturiert:

- `exercise`: Dieser Ordner enthält die Jupyter Notebooks und die Dateien, welche in die Jupyter Notebooks eingebettet wurden.
- `.gitignore`: In dieser Datei werden die Dateien aufgeführt, welche von Git nicht synchronisiert werden sollen.
- `README.md`: Beim Öffnen des Root-Verzeichnisses des GitLab-Projekts wird standardmässig der Inhalt dieser Datei angezeigt. In dieser Datei sind der Zweck der Übung sowie Instruktionen für die Übungsteilnehmenden beschrieben.
- `environment.yml`: Diese Datei führt die notwendigen Bibliotheken auf, welche von MyBinder bei der Installation der Umgebung automatisch installiert werden sollen.

Für das Ausführen der Jupyter Notebooks sind zwei Varianten vorgesehen. Zum einen können die Jupyter Notebooks aus dem GitLab-Projekt heruntergeladen und lokal auf dem eigenen Rechner ausgeführt werden. Dazu muss von den Übungsteilnehmenden die dafür notwendige Software installiert werden. Eine Anleitung für die Installation ist in der Datei `README.md` enthalten.

Alternativ können die Jupyter Notebooks über MyBinder ausgeführt werden. MyBinder ist ein Cloud-Service und erlaubt es Jupyter Notebooks direkt aus einem Git-Repository heraus in einer Online-Umgebung auszuführen. Dadurch müssen die Übungsteilnehmenden keine Software installieren und können die Jupyter Notebooks ohne zusätzliche Schritte ausführen. Ab-

---

<sup>1</sup>[https://gitlab.enterprise-lab.ch/jdrexel/baa\\_fs22\\_differential-privacy](https://gitlab.enterprise-lab.ch/jdrexel/baa_fs22_differential-privacy)

## 5 Realisierung

bildung 13 zeigt die Arbeitsoberfläche von MyBinder während der Ausführung eines Jupyter Notebooks.

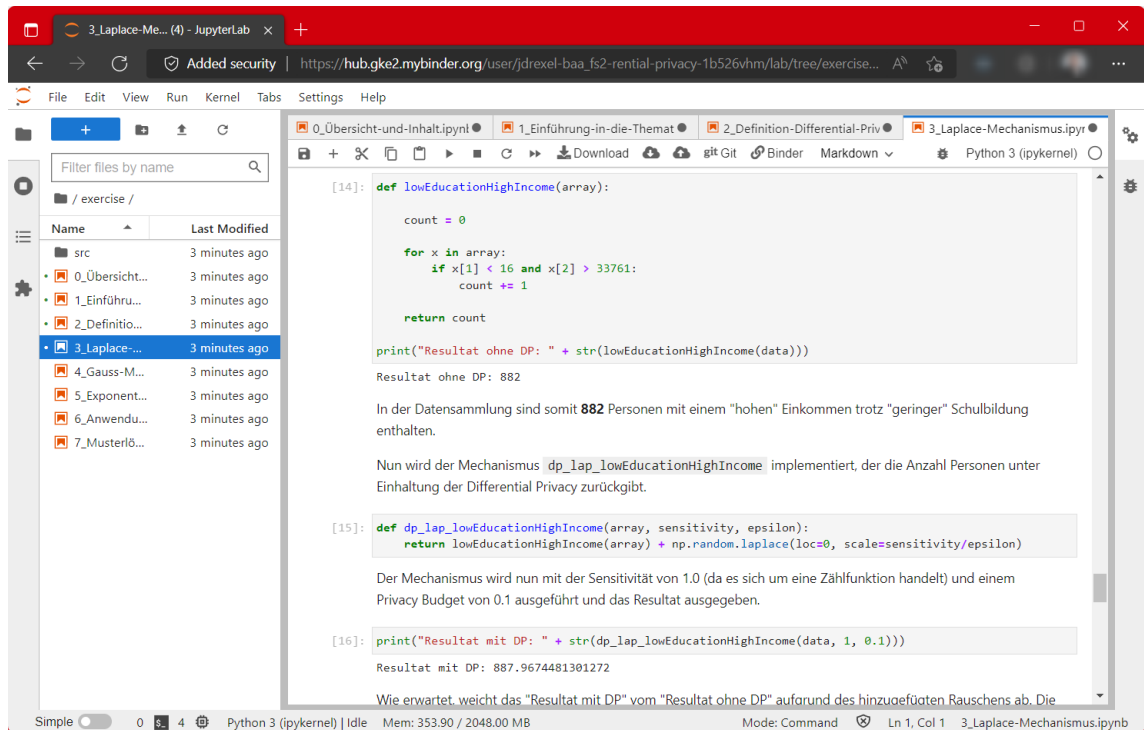


Abb. 13. Arbeitsoberfläche von MyBinder

## 5.4 Testdurchführung der Übung

Für die Testdurchführung wurde eine Testgruppe von 5 Personen zusammengestellt, in welcher Personen mit unterschiedlichen Vorkenntnissen teilnehmen. Die Personen werden kurz beschrieben, um die jeweiligen Vorkenntnisse festzuhalten und die Diversität der gewählten Testgruppe aufzuzeigen.

### 5.4.1 Zusammensetzung der Testgruppe

Person 1 verfügt als höchste abgeschlossene Ausbildung über eine Informatik-Lehre EFZ mit technischer Berufsmaturität. Person 1 hatte zum Zeitpunkt der Testdurchführung keine Ausbildung auf Tertiärstufe begonnen oder abgeschlossen. Person 1 hatte nach eigenen Angaben keine Vorkenntnisse und konnte vor der Testdurchführung keines der in Abschnitt 5.1.3 beschriebenen Lernziele erfüllen.

Person 2 verfügt als höchste abgeschlossene Ausbildung über eine Informatik-Lehre EFZ, sowie über die gymnasiale Maturität. Person 2 befand sich zum Zeitpunkt der Testdurchführung

im Bachelorstudium in Information & Cyber Security. Person 2 hatte nach eigenen Angaben wenige Vorkenntnisse und konnte vor der Testdurchführung das Lernziel Nr. 1 (vgl. Tabelle 5) bereits erfüllen.

Person 3 verfügt als höchste abgeschlossene Ausbildung über das HF-Studium «Techniker HF Informatik» und hatte zum Zeitpunkt der Testdurchführung keine andere Ausbildung begonnen. Person 3 hatte nach eigenen Angaben einige Vorkenntnisse und konnte vor der Testdurchführung die Lernziele Nr. 1, 2, 10, 11, 14 und 16 bereits erfüllen.

Person 4 verfügt als höchste abgeschlossene Ausbildung über den Bachelor of Science in Informatik. Person 4 befand sich zum Zeitpunkt der Testdurchführung im Masterstudium in Digitaler Forensik. Person 4 hatte nach eigenen Angaben gute Vorkenntnisse und konnte vor der Testdurchführung die Lernziele Nr. 1, 2, 8, 9, 13 und 16 bereits erfüllen.

Person 5 verfügt als höchste abgeschlossene Ausbildung über ein Masterstudium in Information Security und hatte zum Zeitpunkt der Testdurchführung keine andere Ausbildung begonnen. Person 5 hatte nach eigenen Angaben wenige Vorkenntnisse und konnte vor der Testdurchführung die Lernziele Nr. 1 und 2 bereits erfüllen.

### **5.4.2 Ausgestaltung der Testdurchführung**

Im Fokus der Testdurchführung stand die Überprüfung der Verständlichkeit der Inhalte, sowie der Lernzielerreichung. Die Teilnehmenden wurden angewiesen, die Zeit für die Bearbeitung der Übung zu messen und zu dokumentieren. Weiter wurde die Testgruppe instruiert, allfällige Unklarheiten festzuhalten und die Verständlichkeit der Inhalte laufend zu überprüfen.

Sämtliche Instruktionen für die Übungsbearbeitung waren entweder in der Datei `README.md` (siehe Abschnitt 5.3) oder in den Übungsunterlagen enthalten. Die Testgruppe hat keine zusätzlichen Instruktionen erhalten. Weiter haben alle Teilnehmenden die Übung selbständig und in Einzelarbeit durchgeführt.

Die Rückmeldungen der Teilnehmenden wurden schriftlich eingeholt. Die Einschätzung der Lernzielerreichung erfolgte basierend auf einer Selbstbeurteilung der Teilnehmenden.

### **5.4.3 Auswertung der Testdurchführung**

Alle Teilnehmenden konnten die gesamte Übung ohne weitere Instruktionen erfolgreich bearbeiten. Die Gesamtdauer für die Bearbeitung der Übung lag zwischen 80 und 120 Minuten. Für das Einlesen in die theoretischen Inhalte wurden zwischen 55 und 100 Minuten benötigt. Für die praktischen Aufgaben wurden von den Teilnehmenden zwischen 20 und 35 Minuten aufgewendet. Das Verhältnis zwischen den theoretischen und den praktischen Teilen wurde

von sämtlichen Teilnehmenden als sinnvoll und gut beurteilt.

Die Übung wurde von allen Teilnehmenden als sehr verständlich eingeschätzt. Der Umfang des Inhaltes und die Struktur der Übung wurde als sinnvoll beschrieben. Aufgrund der Vorkenntnisse wurden von Person 4 gewisse Inhalte als zu einfach eingestuft. Die anderen Teilnehmenden schätzten diese Inhalte als angemessen ein.

Die Jupyter Notebooks 1 - 5 werden mit Kontrollfragen abgeschlossen. Obwohl die Kontrollfragen von einigen Teilnehmenden als einfach eingeschätzt worden sind, wurden diese von allen Teilnehmenden sehr geschätzt und als sinnvolles Mittel zur Selbstkontrolle betrachtet.

Einige Teilnehmende wünschten sich zusätzliche Beispiele zu den Implementierungen. Basierend auf diesem Feedback wurden zusätzliche Beispiele von Mechanismen in die Übung integriert. Weiter konnten kleinere Fehler in der Darstellung einiger Inhalte nachgebessert werden.

Alle Teilnehmenden konnten, gemäss ihrer Selbsteinschätzung, trotz der stark unterschiedlichen Vorkenntnisse sämtliche Lernziele erreichen.

## 6 Evaluation und Validation

Aus der Aufgabenstellung (vgl. Abschnitt 8.1) können zusammenfassend die folgenden Ziele für diese Arbeit abgeleitet werden:

Ziel 1: Das Konzept der Differential Privacy verstehen.

Ziel 2: Die Theorie für das Verständnis der Differential Privacy so aufbereiten, dass diese einem breiten Publikum verständlich gemacht werden kann.

Ziel 3: Eine praktische Übung ausarbeiten, anhand welcher das Verständnis für die Differential Privacy gefestigt und praktisch ausprobiert werden kann.

Die Erreichung dieser Ziele wird nachfolgend überprüft und begründet.

### 6.1 Beurteilung Ziel 1: Verständnis der Thematik

Im Kapitel 2 werden die als relevant eingeschätzten Definitionen und Konzepte beschrieben. Dies sind insbesondere die Definitionen der  $\epsilon$ -Differential Privacy, der  $(\epsilon, \delta)$ -Differential Privacy sowie des Laplace, Gauss und exponentiellen Mechanismus. Des Weiteren werden die notwendigen Konzepte erklärt, welche für eine Implementierung von Differential Privacy verstanden werden müssen. Dies sind unter anderem die Sensitivität von Funktionen, sowie die Komposition von Mechanismen.

Ein gutes Verständnis des Privacy-Budgets ist für beide Definitionen der Differential Privacy von zentraler Bedeutung. Wird das Privacy-Budget falsch gewählt, kann dies jeglichen Schutz aufheben und zu folgenschweren Fehlimplementierungen führen. Aus diesen Gründen wurde das Privacy-Budget ausführlich behandelt.

Weiter werden die wichtigen praktischen Aspekte behandelt. Es wird gezeigt, welche beiden Modelle von Differential Privacy existieren und an welchen Stellen Rauschen hinzugefügt werden kann. Weiter werden zwei Verteilungen von Rauschen (Laplace und Gauss) beschrieben.

Abschliessend wird eine Übersicht der möglichen Anwendungsfälle, sowie die Grenzen der Differential Privacy aufgezeigt.

Aus diesen Gründen wird dieses Ziel als **vollständig erfüllt** beurteilt.

### 6.2 Beurteilung Ziel 2: Vermittlung der Theorie

Die theoretischen Inhalte wurden basierend auf den in Abschnitt 5.1.3 beschriebenen Lernzielen ausgearbeitet. Die Lernziele wurden von der Betreuungsperson überprüft und als sinnvoll beurteilt. Durch die Testdurchführung der Übung konnte anhand der 5 Testpersonen die Erreichung dieser Lernziele validiert werden. Sämtliche Testpersonen haben bestätigt, dass alle Lernziele vollumfänglich anhand der Theorieinhalte erfüllt werden konnten. Die Testpersonen hatten unterschiedliche Ausbildungen und Vorkenntnisse. Aus diesem Grund kann die Testgruppe als breites Publikum eingestuft werden.

Eine Integration der Übung in den Unterricht des Privacy-Moduls der Hochschule Luzern wäre nach Angaben des Modulverantwortlichen wünschenswert. Eine zusätzliche Testdurchführung der Übung im Rahmen der Privacy-Vorlesung war geplant und vorbereitet. Jedoch konnten aus der aktuellen Durchführung der Privacy-Vorlesung keine Freiwilligen gefunden werden, weshalb diese Testdurchführung nicht stattfinden konnte.

Bei der Testdurchführung lag die Gesamtdauer für die Bearbeitung der Übung zwischen 80 und 120 Minuten. Deshalb sollte eine Integration der Übung in das Privacy-Modul problemlos möglich sein.

Aus diesen Gründen wird dieses Ziel als **vollständig erfüllt** beurteilt.

### 6.3 Beurteilung Ziel 3: Praktische Übung

Aus der Aufgabenstellung geht hervor, dass im Rahmen der Übung zwei Anwendungsfälle umgesetzt werden sollten. Mögliche Anwendungsfälle werden in Abschnitt 2.6 beschrieben. Im Rahmen der Übungsausarbeitung wurde in Absprache mit der Betreuungsperson entschieden, kein Tool für die Umsetzung der Differential Privacy zu verwenden. Die Implementierung der Differential Privacy wurde in eigenem Code gemacht und bewusst einfach gehalten. Der Einsatz von Tools sowie die Implementierung mehrerer Anwendungsfälle hätte den sinnvollen Umfang der Übung überstiegen. Das Verständnis der Differential Privacy und der grundlegenden Mechanismen wurde bewusst in den Fokus gerückt.

Aus diesen Gründen wird dieses Ziel als **vollständig erfüllt** beurteilt.

## 7 Ausblick

Die persönliche Leistung wird reflektiert und beurteilt. Zudem werden mögliche Ergänzungen oder Erweiterungen der Übung beschrieben und einige abschliessende Kommentare zur Arbeit gegeben.

### 7.1 Persönliche Reflexion

Zu Beginn der Arbeit waren keine Vorkenntnisse zur Differential Privacy vorhanden. Die Einarbeitung in die Thematik war intensiv und anspruchsvoll. Das Projekt zu planen, ohne das Thema im Detail verstanden zu haben, war eine Herausforderung. Deshalb wurde der detaillierten Planung und dem stetigen Nachführen des Projektplans eine grosse Bedeutung zugeschrieben. Dadurch konnte der Projektstand jederzeit beurteilt und potentielle Projektverzögerungen frühzeitig erkannt werden. Zudem ermöglichten die wöchentlichen Abgleichsmeetings mit der Betreuungsperson eine regelmässige und unkomplizierte Absprache. Dadurch konnten Fragen schnell beantwortet und Entscheidungen zeitnah getroffen werden.

Die Abbildungen 14 und 15 zeigen den IST-Projektplan. Im Vergleich zur initialen Planung (vgl. Abbildungen 11 und 12) sind einige Unterschiede ersichtlich. Die Aufgabe Nr. 009 «Aufarbeitung und Strukturierung der Theorie» nahm deutlich mehr Zeit in Anspruch als initial geplant. Die Komplexität der theoretischen Grundlagen der Differential Privacy wurden zu Beginn des Projekts unterschätzt. Die Erkrankung an COVID-19 führte zu einem Arbeitsunterbruch in den Projektwochen 4 und 5. Der gewünschte Projektumfang war deshalb in deutlich weniger Zeit umzusetzen. Dies erforderte eine sehr strukturierte und effiziente Arbeitsweise und ein hohes Mass an Disziplin. Es mussten mehr Stunden an Abenden und Wochenenden geleistet werden, als dies zu Beginn des Projekts geplant war. Durch die intensive Einarbeitung in Aufgabe Nr. 009, wurde für die Aufgabe Nr. 010 «Erstellen der Übungen und Unterlagen» weniger Zeit benötigt. Gesamtheitlich betrachtet war die initiale Projektplanung sehr zutreffend und hilfreich.

Die Belastung durch die Berufstätigkeit während der Bachelorarbeit wurde unterschätzt. Es war geplant, vormittags an der Bachelorarbeit und nachmittags für den Arbeitgeber zu arbeiten. Diese Trennung war oftmals nicht möglich, beispielsweise wenn bestimmte Termine des Arbeitgebers nur vormittags geplant werden konnten. Vor allem mit der personellen und fachli-



chen Verantwortung als Teamleiter war es nicht in jedem Fall möglich nur halbtags zu arbeiten. In einem künftigen Projekt dieser Art sollte mehr Wert auf eine klare Trennung der Arbeitszeiten gelegt werden. Eine Aufteilung auf Basis von halben Tagen stellte sich als schwierig heraus.

### 7.2 Ideen für mögliche Ergänzungen

Die Differential Privacy ist ein umfangreiches Themengebiet. Die erarbeitete Übung erlaubt es sich grundlegend in die Thematik einzuarbeiten und erste, einfache Implementierungen von Mechanismen auszuprobieren. Für eine produktive Implementierung von Differential Privacy bedarf es aber einem tieferen Verständnis. Besonders die richtige Allokation des Privacy-Budgets ist kritisch für die korrekte und damit sichere Umsetzung von Differential Privacy.

Die Zusammenhänge und Auswirkungen des Privacy-Budgets werden in der Übung zwar gezeigt, jedoch wird in der Übung nicht behandelt, wie das korrekte Privacy-Budget für einen konkreten Anwendungsfall bestimmt werden kann. Dies wäre ein interessantes Thema für eine Erweiterung oder Fortsetzung dieser Übung. Auch weitere Parameter wie das Delta bei der annähernden Differential Privacy oder die globale und lokale Sensitivität wären interessante Aspekte für die weitere Vertiefung in die Thematik.

Es gibt Software-Bibliotheken, welche die Implementierung von Differential Privacy vereinfachen sollen. Beispielsweise das Projekt OpenDP der Harvard Universität. Die Anwendung solcher Bibliotheken wäre eine weitere interessante Erweiterung der bestehenden Übung.

### 7.3 Projektabschluss

Es ist angedacht, dass die erarbeitete Übung innerhalb der Hochschule Luzern zur Verfügung gestellt wird. Anhand der Übung sollen sich Studierende, Forschende und Dozierende in die Welt der Differential Privacy einarbeiten können. Weiter wurde bereits eine mögliche Integration der Übung in die Privacy-Vorlesung besprochen. Die Arbeitsergebnisse dieses Projekts werden nach Projektabschluss dem Modulverantwortlichen des Privacy-Moduls zur Verfügung gestellt.

# 8 Anhang

## 8.1 Aufgabenstellung der Bachelorarbeit

Die Aufgabenstellung wurde durch Prof. Dr. Esther Hänggi verfasst. Die Aufgabenstellung wurde ohne Änderungen übernommen und ist vollständig als Fremdleistung zu verstehen.

### 8.1.1 Ausgangslage und Problemstellung

Die technischen Möglichkeiten der Verarbeitung grosser Datenmengen sind in den letzten Jahren stark gewachsen. Gemeinsam mit der immer grösseren Verfügbarkeit und Vernetzung von Datensätzen eröffnet dies ganz neue Möglichkeiten mittels Data Science oder maschinellem Lernen wichtige Informationen aus diesen Daten zu erhalten. Mögliche Anwendungen befinden sich in fast allen Lebensbereichen, von der (personalisierten) Werbung, über Spracherkennung und Textverarbeitung bis zu besseren medizinischen Behandlungen; um nur ein paar wenige konkrete Beispiele zu nennen.

Bei allen Chancen beinhaltet die Datenanalyse aber auch Risiken. Insbesondere stellt sich die Frage, welche aggregierten Erkenntnisse aus Datensätzen veröffentlicht werden können ohne die berechtigten Interessen für Datenschutz der beteiligten Personen oder Organisationen zu verletzen.

Zwei konkrete Beispiele:

- Während der Corona-Pandemie wurde in der Schweiz diskutiert, ob die Anzahl Fälle nach Postleitzahl aufgeschlüsselt gemeldet werden sollen. Während dies bei Gemeinden mit grossen Einwohnerzahlen im Allgemeinen kein Problem darstellt, könnte dies bei kleinen Einwohnerzahlen Rückschlüsse auf den Teststatus von Individuen ermöglichen.
- Ein zweites Beispiel ist die Verbesserung von Smartphone-Anwendungen aufgrund von Benutzerdaten, z.B. Wortvorschläge bei der Texteingabe. Während alle Benutzer von einer Verbesserung profitieren, sollen konkrete Vorschläge keine Rückschlüsse auf die Texteingaben anderer Benutzer zulassen. Die eigenen Texteingaben sollen auch nicht unverändert an eine zentrale Datenbank (z.B. von Apple) hochgeladen werden.

Konkrete Angriffe haben gezeigt, dass die intuitive Annahme wann Daten anonymisiert sind, oft nicht genügend ist. So hat das US Census Bureau seine eigenen publizierten Daten aus der Volkszählung von 2010 analysiert und festgestellt, dass ein Grossteil der Datensätze rekonstruiert werden konnte. Zusammen mit öffentlich einsehbaren Informationen wie einem Telefonbuch oder einem Stimmregister konnten sogar ca. die Hälfte der Personen namentlich identifiziert werden. (<https://www.census.gov/data/academy/webinars/2021/disclosure-avoidance-series/simulated-reconstruction-abetted-re-identification-attack-on-the-2010-census.html>)

In den letzten Jahren ist deshalb das Studium von technischen Möglichkeiten für «Privacy» zu einem wichtigen Forschungszweig der Kryptographie geworden. Neu an der Forschung in diesem Gebiet ist, dass die Datenschutzerfordernungen und erreichten Datenschutzeigenschaften genau quantifiziert werden. Insbesondere das Konzept der «Differential Privacy» erfreut sich grosser Beliebtheit: die Idee hinter diesem Konzept ist, dass es für publizierte Daten keinen Unterschied machen sollte, ob eine einzelne Person im Datensatz enthalten ist oder nicht. Damit wird die Privatsphäre von Individuen gewahrt. Die erreichte «Privacy» kann dabei nicht nur genau quantifiziert werden, sondern auch mittels einem «Privacy Budget» je nach Sensibilität der Daten alloziert werden. Der erreichte Schutz der Daten ist unabhängig davon, welche weiteren Daten (Telefonbuch etc.) der Angreifer noch zur Verfügung hat.

Das Ziel «differential privacy» wird erreicht, indem persönliche Daten durch das Hinzufügen von Rauschen verborgen werden. Gemittelt auf grössere Datensätze sind die Auswirkungen des Rauschens aber klein oder heben sich sogar gegenseitig auf, sodass weiterhin gute statistische Resultate durch Datenanalyse möglich sind.

Erste Anwender aus der Industrie sind grosse Techfirmen und das US Census Bureau:

- Bei Apple wird das Konzept auf Smartphone-Benutzerdaten angewandt: nur modifizierte Daten werden vom Smartphone an die zentrale Datenbank hochgeladen und zur Verbesserung der Anwendungen verwendet. ([https://www.apple.com/privacy/docs/Differential\\_Privacy\\_Overview.pdf](https://www.apple.com/privacy/docs/Differential_Privacy_Overview.pdf)).
- Das US Census Bureau wendet es für die Daten der Volkszählung von 2020 an: nur leicht verrauschte Statistiken werden publiziert (<https://www.census.gov/programs-surveys/decennial-census/decade/2020/planning-management/process/disclosure-avoidance.html>)

Die notwendigen theoretischen Grundlagen zum Verständnis der Funktionsweise dieser «Tools» sind allerdings nicht einfach zu verstehen.

Ziel dieses Projekts ist es deshalb, «Differential Privacy» verständlich zu machen. Dafür soll insbesondere eine Übung entwickelt werden, mit der Studierende das Konzept selbst anwenden können.

### 8.1.2 Ziel der Arbeit und erwartete Resultate

Ziel der Arbeit ist es, das Konzept «Differential Privacy» zu verstehen. Insbesondere folgende Punkte:

- An welchen Stellen kann Rauschen zu Daten hinzugefügt werden? (z.B. beim Einlesen, Abändern der Datenbank oder Erstellen der Statistik)
- Welche Arten von Verteilung von Rauschen gibt es und wie unterscheiden sie sich?
- Welche Anwendungsfälle gibt es? (z.B. Erstellen von synthetischen Daten, publizieren von Statistiken, Machine Learning, interaktive Queries)
- Was bedeutet das «Privacy Budget» und wie kann es alloziert werden?
- Was muss bei den verschiedenen Use Cases beachtet werden für die Security/Privacy?

Diese Themen sollen so aufgearbeitet werden, dass sie einem breiteren Publikum verständlich gemacht werden können (z.B. Studierende).

Dann soll dazu eine Übung kreiert werden, bei der z.B. ein Datenset in eine «differentially private» Version umgewandelt wird oder statistische Daten auf «private» Art und Weise evaluiert. Die Übung soll mind. 2 Anwendungsfälle abdecken. Als Tool für diese Übung soll ARX verwendet werden, ein Standard-Tool im Bereich Datenschutz/Privacy (<https://arx.deidentifier.org/>). Die Verwendung von anderen Libraries o.ä. ist möglich, falls sich dies als sinnvoll erweist.

### 8.1.3 Gewünschte Methoden, Vorgehen

- Verstehen der Theorie
- Verstehen von Angriffen, welche die Notwendigkeit von Differential Privacy zeigen
- Aufarbeiten von praktischen Anwendungsfällen in der Industrie
- Aufbereiten des Materials, sodass es einem breiteren Publikum verständlich ist
- Erstellen von zwei Übungen, z.B.
  - Ein Datenset in eine «differentially private» Version zu überführen, wobei die statistischen Eigenschaften erhalten bleiben
  - Statistische Daten über ein Datenset auf eine «differentially private» Art berechnen

#### **8.1.4 Kreativität, Varianten, Innovation**

Bei der Umsetzung, wie das Thema «Differential Privacy» einem breiten Publikum verständlich gemacht werden soll sind verschiedenste Möglichkeiten denkbar. Ebenso kann der genaue Inhalt der Übung (in Absprache mit der Betreuungsperson) gewählt werden. Schliesslich sind auch bei den technischen Mitteln, e.g. ARX oder Library, Varianten möglich.

## 9 Abbildungsverzeichnis

1	Zentrales Modell (angelehnt an Abbildung von Desfontaines [9]) . . . . .	6
2	Lokales Modell (angelehnt an Abbildung von Desfontaines [9]) . . . . .	7
3	Grundidee von Differential Privacy . . . . .	8
4	Auswirkung von $\epsilon$ auf Wissensgewinn nach Desfontaines [10] . . . . .	11
5	Parallele Komposition . . . . .	12
6	Sequentielle Komposition . . . . .	13
7	Funktionsweise des Laplace Mechanismus . . . . .	15
8	Funktionsweise des Gauss Mechanismus . . . . .	17
9	Vergleich des Laplace und Gauss Rauschens nach Desfontaines [14] . . . . .	18
10	Funktionsweise des exponentiellen Mechanismus . . . . .	19
11	Arbeitspakete und Meilensteine des SOLL-Projektplans . . . . .	26
12	Gantt-Diagramm des SOLL-Projektplans . . . . .	26
13	Arbeitsoberfläche von MyBinder . . . . .	40
14	Arbeitspakete und Meilensteine des IST-Projektplans . . . . .	46
15	Gantt-Diagramm des IST-Projektplans . . . . .	46

## 10 Tabellenverzeichnis

1	Originale Datensammlung zu Krankheiten in Grub AR und Grub SG . . . . .	3
2	Die 3-anonyme Datensammlung zu Krankheiten in Grub AR und Grub SG . . . . .	3
3	Die 3-diverse Datensammlung zu Krankheiten in Kammersrohr und Bister . . . . .	4
4	Exponentieller Mechanismus mit Krankheitsdaten nach Zhu <i>et al.</i> [8] . . . . .	20
5	Lernziele absteigend nach Taxonomiestufe . . . . .	32

# 11 Literaturverzeichnis

- [1] T. Bendig. "Big Data - Viel wertvoller als Öl und Gold." (2020), [Online]. Available: <https://www.fraunhofer-innovisions.de/big-data/lebendige-zukunft/> (visited on 4.3.2022).
- [2] Vereinigte Nationen. "Internationaler Pakt über bürgerliche und politische Rechte." (1966), [Online]. Available: [https://www.fedlex.admin.ch/eli/cc/1993/750\\_750\\_750/de](https://www.fedlex.admin.ch/eli/cc/1993/750_750_750/de) (visited on 4.3.2022).
- [3] Schweizerische Eidgenossenschaft. "Bundesgesetz über den Datenschutz." (1992), [Online]. Available: [https://www.fedlex.admin.ch/eli/cc/1993/1945\\_1945\\_1945/de](https://www.fedlex.admin.ch/eli/cc/1993/1945_1945_1945/de) (visited on 4.3.2022).
- [4] L. Sweeney, "Simple Demographics Often Identify People Uniquely," Carnegie Mellon University, Working Paper, 2000.
- [5] D. Hauf, "Allgemeine Konzepte - K-Anonymity, I-Diversity and T-Closeness," IPD Uni-Karlsruhe, Working Paper.
- [6] J. M. Abowd *et al.*, "The modernization of statistical disclosure limitation at the U.S. Census Bureau," U.S. Census Bureau, Working Paper, 2020.
- [7] Apple Differential Privacy Team. "Learning with Privacy at Scale." (2017), [Online]. Available: <https://machinelearning.apple.com/research/learning-with-privacy-at-scale> (visited on 5.3.2022).
- [8] T. Zhu, G. Li, W. Zhou, and P. S. Yu, *Differential Privacy and Applications*. Switzerland: Springer International Publishing AG, 2017, ISBN: 978-3-319-62004-6.
- [9] D. Desfontaines. "Local vs. central differential privacy." (2019), [Online]. Available: <https://desfontain.es/privacy/local-global-differential-privacy.html#central> (visited on 25.3.2022).
- [10] D. Desfontaines. "Differential privacy in (a bit) more detail." (2018), [Online]. Available: <https://desfontain.es/privacy/differential-privacy-in-more-detail.html> (visited on 29.3.2022).
- [11] D. Desfontaines. "Almost differential privacy." (2019), [Online]. Available: <https://desfontain.es/privacy/almost-differential-privacy.html> (visited on 2.4.2022).

- [12] D. Desfontaines. "A list of real-world uses of differential privacy." (2021), [Online]. Available: <https://desfontain.es/privacy/real-world-differential-privacy.html> (visited on 2.4.2022).
- [13] J. P. Near and C. Abueh, *Programming Differential Privacy*. 2021, vol. 1. [Online]. Available: <https://uvm-plaid.github.io/programming-dp/>.
- [14] D. Desfontaines. "The magic of Gaussian noise." (2020), [Online]. Available: <https://desfontain.es/privacy/gaussian-noise.html> (visited on 8.4.2022).
- [15] NIST. "How to deploy machine learning with differential privacy." (Dec. 21, 2021), [Online]. Available: <https://www.nist.gov/blogs/cybersecurity-insights/how-deploy-machine-learning-differential-privacy> (visited on 1.6.2022).
- [16] J. Domingo-Ferrer, D. Sánchez, and A. Blanco-Justicia, "The Limits of Differential Privacy (and its Misuse in Data Release and Machine Learning)," Universitat Rovira i Virgili, Working Paper, 2020.
- [17] Hochschuldidaktik UZH, "Einstieg in die Hochschullehre," Universität Zürich, White Paper, 2013.
- [18] B. S. Bloom, M. D. Engelhart, E. J. Furst, W. H. Hill, and D. R. Krathwohl, *Taxonomy of Educational Objectives*. New York: David McKay Company Inc., 1956.
- [19] G. Dübbelde. "Aktivierende Methoden für Seminare und Übungen." (2017), [Online]. Available: <https://www.uni-giessen.de/fbz/zentren/zfbk/didaktik/informationen/downloads/lehreinsteiger-1/methodenkoffer-seminare> (visited on 9.3.2022).