Differential Privacy

Seminar: Ist künstliche Intelligenz gefährlich?

Referentin: Mira Boehme 07. Juni 2017

- Ankündigung von Apple auf der WWDC 2016
- Mit iOS 10 mehr Nutzerdaten sammeln, aber trotzdem Privatsphäre schützen
- Kompromiss Datenverarbeitung vs. Datenschutz

→ Konzept nennt man Differential Privacy

- Verschlüsselung nicht immer zielführend

  → Operationen auf verschlüsselten Daten auszuführen ist teuer

- Bsp: Patientendaten eines Krankenhauses

→ **gesucht:** Verfahren zur Veröffentlichung von Daten ohne Datenschutz zu verletzen

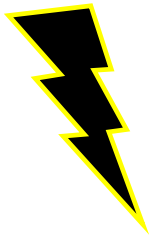
- Anonymisieren von Daten
- Unterscheidung in 3 Arten von Daten: (direkte) Identifikatoren, Quasi-Identifikatoren, sensible Werte
- **Aber:** Angreifer können Hintergrundwissen haben

→ Anonymisieren reicht nicht aus

NETFLIX

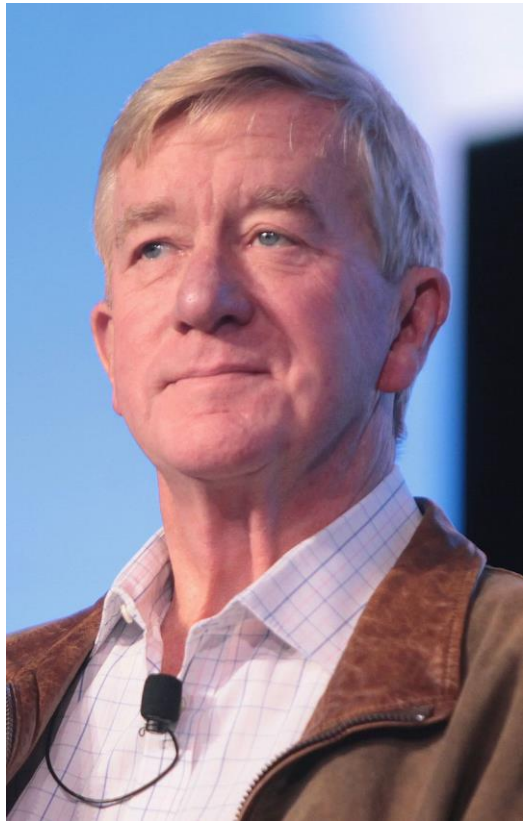
- Netflix Prize (2006)
- Wettbewerb für den besten Vorhersage-Algorithmus
- Anonymisierte Trainings-Datensätze:
  < user, movie, date of grade, grade >
- Abgleich mit IMDb-Ratings

→ Identifizierung von einzelnen Usern



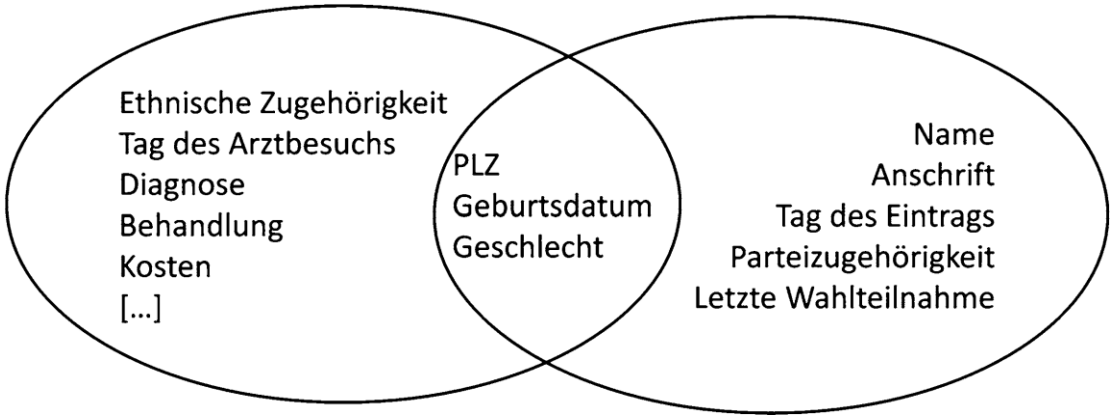
Massachusetts-Missgeschick

- GIC: Krankenversicherung von 135.000 Staatsbediensteten
- Weitergabe von (anonymisierten) Daten: Gesundheitsdaten, PLZ, Geschlecht etc.
- Abgleich mit (öffentlichem) Wählerverzeichnis



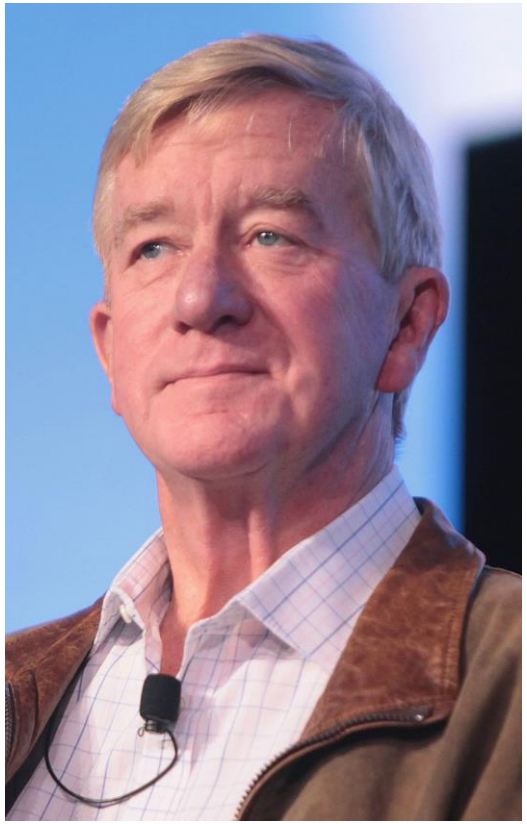
William Held

Massachusetts-Missgeschick



- Quasi-Identifikator:
  < PLZ, Geburtsdatum, Geschlecht >

  → Krankenakte des Gouverneurs



William Held

Imagine you have two otherwise identical databases, one with your information in it, and one without it. Differential Privacy ensures that the probability that a statistical query will produce a given result is (nearly) the same whether it's conducted on the first or second database.

- Matthew Green

- Maß für das Risiko einer einzelnen Person, an einer statistischen DB teilzunehmen
- **Grundidee:** Personenbezogene Daten einer Person dürfen zu keinem Unterschied des Ergebnis führen
- **DENN:** Wenn Ergebnis einer DB-Abfrage nicht von Daten einer einzelnen Person abhängig ist, dann ist Datenschutz dieser Person gewährleistet

- Funktion beantwortet DB-Anfragen und stellt sicher, dass Datenschutz nicht verletzt wird
- Daten mit Rauschen versehen
  → bspw. Generierung neuer Einträge
- DP ist eine Definition, kein Algorithmus!

Cynthia Dwork

- Geboren 1958
- Informatikerin an der Harvard University
- Forscht u.a. im Bereich Kryptographie
- Konzept der DP 2006 veröffentlicht
- Dijkstra-Preis (2007)

Cynthia Dwork

ε-Differential-Privacy

Eine randomisierte Funktion κ bietet ε-DP, wenn

- für alle Datensätze D_1 und D_2, die sich nur in höchstens einem Element unterscheiden
- für alle Teilmengen S des Wertebereichs W von κ gilt:

$$P[\kappa(D_1) \in S] \leq e^{\varepsilon} * P[\kappa(D_2) \in S]$$

ε-Differential-Privacy

Geburtsjahr	PLZ	Geschlecht	Diagnose
1982	33098	Männlich	Migräne
1982	33098	Männlich	Erkältung
1983	33098	Männlich	Rheuma
1983	33098	Männlich	Depression
1985	33100	Weiblich	Heuschnupfen
1985	33100	Weiblich	Hypochondrie
1983	33098	Weiblich	Migräne

ε-Differential-Privacy

Geburtsjahr	PLZ	Geschlecht	Diagnose
1982	33098	Männlich	Migräne
1982	33098	Männlich	Erkältung
1983	33098	Männlich	Rheuma
1983	33098	Männlich	Depression
1985	33100	Weiblich	Heuschnupfen
1985	33100	Weiblich	Hypochondrie
1983	33098	Weiblich	Migräne
1985	**33100**	**Männlich**	**Erkältung**

ε-Differential-Privacy

- Je größer ε, desto schwächer Garantie für Datenschutz
- In der Praxis sowohl sehr kleine als auch sehr große ε
- gewisse Einschränkungen bezüglich der Anwendbarkeit

  → Erweiterung: (ε,δ)-Differential Privacy

(ε,δ)-Differential-Privacy

- Erweiterung um (additiven) Parameter δ
- erlaubt, dass Voraussetzungen bis zu einem gewissen Grad δ unerfüllt bleiben
- δ sollte möglichst klein sein

(ε,**δ**)-Differential-Privacy

Eine randomisierte Funktion κ bietet (ε,δ)-DP, wenn

- für alle Datensätze D_1 und D_2, die sich nur in höchstens einem Element unterscheiden
- für alle Teilmengen S des Wertebereichs W von κ gilt:

$$P[\kappa(D_1) \in S] \leq e^{\varepsilon} * P[\kappa(D_2) \in S] + \boldsymbol{\delta}$$

Definition „Sensibilität“

Sensibilität einer Funktion f

$\Delta f = \max_{D1,D2} |f(D_1) - f(D_2)|$ für „benachbarte“ Datensätze D_1, D_2

→ Dies wollen wir Verbergen mit DP

→ Maß, wie sehr eine Person das Ergebnis beeinflussen kann

→ Sensibilität einer Zählabfrage ist 1

Geburtsjahr	PLZ	Geschlecht	Diagnose
1982	33098	Männlich	Migräne
1982	33098	Männlich	Erkältung
1983	33098	Männlich	Rheuma
1983	33098	Männlich	Depression
1985	33100	Weiblich	Heuschnupfen
1985	33100	Weiblich	Hypochondrie
1983	33098	Weiblich	Migräne
1985	**33100**	**Männlich**	**Erkältung**

Laplace-Mechanismus

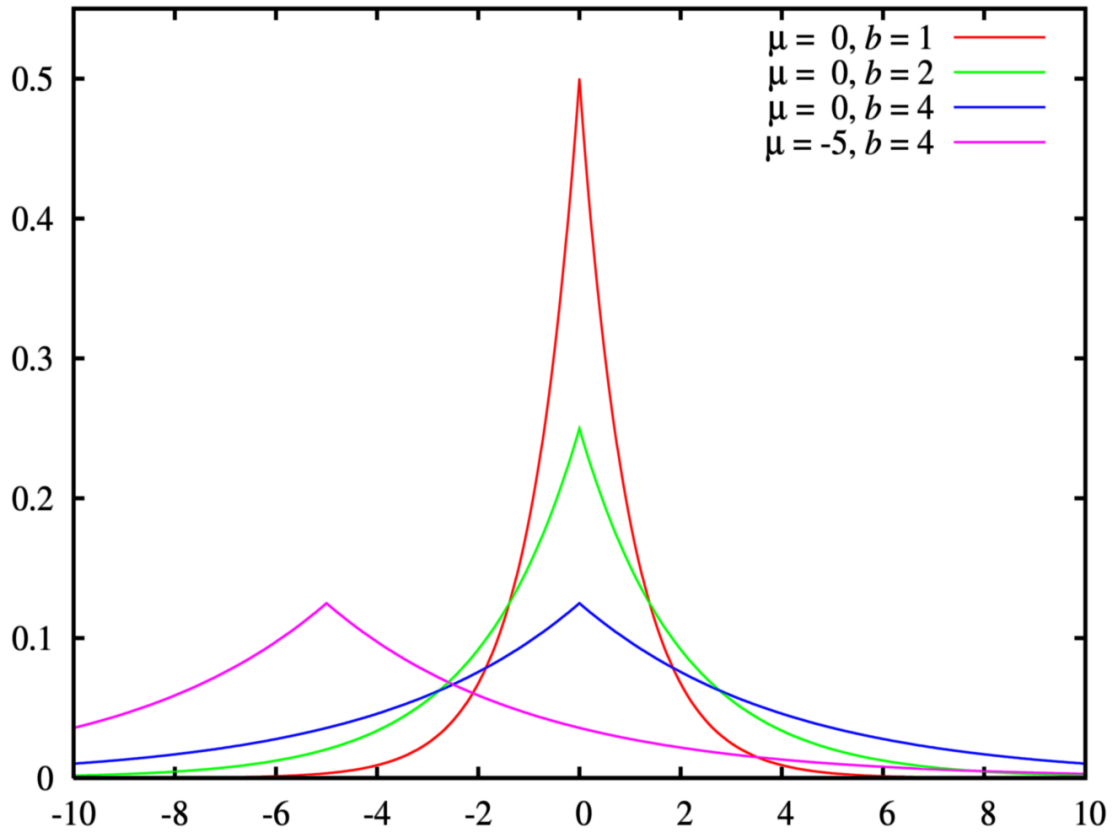
- Hinzufügen von kontrolliertem Rauschen
- Laplace-Rauschen: Laplace-Verteilung

$\Delta f = \max_{D1,D2} |f(D1) - f(D2)|$

→ Hinzufügen von Rauschen Lap(b)

mit $b = \Delta f / \varepsilon$

Laplace-Verteilung



ε-Differential-Privacy

Geburtsjahr	PLZ	Geschlecht	Diagnose	Raucher
1982	33098	Männlich	Migräne	Nein
1982	33098	Männlich	Erkältung	Ja
1983	33098	Männlich	Rheuma	Nein
1983	33098	Männlich	Depression	Nein
1985	33100	Weiblich	Heuschnupfen	Ja
1985	33100	Weiblich	Hypochondrie	Nein
1983	33098	Weiblich	Migräne	Nein
1985	**33100**	**Männlich**	**Erkältung**	**Nein**

Laplace-Mechanismus

Bei einfachen Zählabfrage:

Wie viele Leute in der DB erfüllen Attribut A?

- Sensibilität = 1
- Laplace-Rauschen: Lap(1/ε)

Laplace-Mechanismus

- Bei mehreren Abfragen: Verknüpfen

  → ε_1-DP Mechanismus gefolgt von einem ε_2-DP Mechanismus, ist ein $(\varepsilon_1 + \varepsilon_2)$-DP Mechanismus

- Funktioniert auch für (δ, ε)-DP

„Randomized Response“-Mechanismus

- Konzept wird in sozialwissenschaftlichen Studien verwendet, um Rückschlüsse auf einzelne Personen zu verhindern
- Befragter kann ehrlich antworten ohne Angst vor Konsequenzen

Vorgehen

→ Befragter wirft Münze:

Bei Zahl antwortet er wahrheitsgemäß

Bei Kopf wirft er die Münze nochmal

→ Diesmal:

Bei Zahl antwortet er „Nein“

Bei Kopf antwortet er „Ja“

Stärken

- Maß für Datenschutzgarantie
- nicht abhängig von der Computer-Power der Gegner

Schwächen

- Nur nützlich bei interaktiven Mechanismen
- **DENN**: Anfragen können so lange mit hoher Genauigkeit beantwortet werden, bis weitere Ausgaben die Definition von DP verletzen würden
- DP stellt hohe Anforderungen an Mechanismen
  → Ergebnisse können stark an Nutzen verlieren

Fazit

- Mehr (große) Unternehmen sollten DP umsetzen
- **Aber:** es gibt nicht *das* Verfahren zum Datenschutz
- Bsp. Patientenakten: Mechanismen sehr kompliziert
- Diskussion: feel free ☺
- Fragen?

Quellen

[1] Petrlic, Ronald und Sorge, *Christoph. Datenschutz: Einführung in technischen Datenschutz, Datenschutzrecht und angewandte Kryptographie.* Springer Vieweg, 2017.

[2] Ji, Zhanglong, Lipton, Zachary C. und Elkan, Charles. *Differential Privacy and Machine Learning: a Survey and Review.* 2014.

[3] https://www.heise.de/mac-and-i/artikel/Besserer-Datenschutz-Wie-Apples-Differential-Privacy-funktioniert-Update-3678489.html

[4] https://www.seas.harvard.edu/directory/dwork

[5] https://www.wirelessweek.com/news/2016/06/behind-differential-privacy-apples-way-see-your-data-without-seeing-you

[6] https://www.youtube.com/watch?v=ekIL65D0R3o
Tutorial on Differential Privacy by Katrina Ligett, 2013.