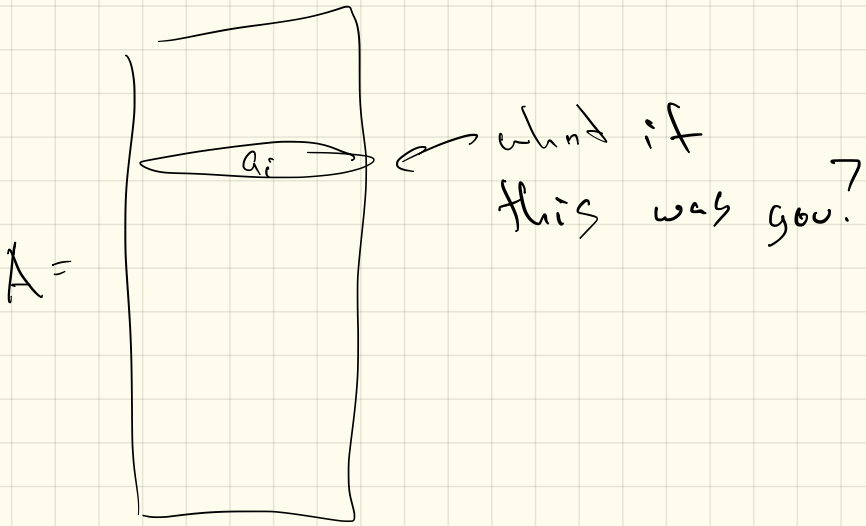
L21: Privacy

Jeff M. Phillips

April 6, 2020

Ethics = Empathy

What if you were the data point?



In early 2000s,

- lot new tech companies
- eagerness, interest scientists

Place data set online,

- anonymize ?
- state goal
- have competition to solve
  ↳ model, predictions.

Example: Heath Records

Story Time:

Example: Heath Records

Story Time:

- In 2000, Massachusetts released all State employee's medical records in an effort for researchers to be able to study them.

Example: Heath Records

Story Time:

- In 2000, Massachusetts released all State employee's medical records in an effort for researchers to be able to study them.
- They wiped all ids, but kept zip codes, birthday, gender. Was declared anonymized by the government.

Example: Heath Records

Story Time:

- In 2000, Massachusetts released all State employee's medical records in an effort for researchers to be able to study them.
- They wiped all ids, but kept zip codes, birthday, gender. Was declared anonymized by the government.
- In Massachusetts, it was possible to buy voter data for $20. It has names, zip codes, and gender of all voters.

Example: Heath Records

Story Time:

- In 2000, Massachusetts released all State employee's medical records in an effort for researchers to be able to study them.
- They wiped all ids, but kept zip codes, birthday, gender. Was declared anonymized by the government.
- In Massachusetts, it was possible to buy voter data for $20. It has names, zip codes, and gender of all voters.
- A grad student, Latanya Sweeney combined the two to identify the governor of Massachusetts. Story is, she mailed him his own health records!

Example: Heath Records

Story Time:

- In 2000, Massachusetts released all State employee's medical records in an effort for researchers to be able to study them.
- They wiped all ids, but kept zip codes, birthday, gender. Was declared anonymized by the government.
- In Massachusetts, it was possible to buy voter data for $20. It has names, zip codes, and gender of all voters.
- A grad student, Latanya Sweeney combined the two to identify the governor of Massachusetts. Story is, she mailed him his own health records!
- *Dr.* Sweeney now teaches at Harvard.

How to release data anonymously while preserving individual info?

---

k-anonymity: data set has **public traits** {age, zip code, gender}

private traits: (categorical) has cancer, has covid-19

enforce at least k each person has at least k-1 other people w/ same publically released traits.

l-diversity : k-anonymity : and each group had l-diverse traits.

e.g. some have cancer
some don't

---

issue either have cancer
or have diabetes

---

t-closeness : l-diversity and the distribution of traits is t-close to distribution of all people in data.

Height of Sylvester Stallone

- Information: Sly Stallone is height of the average NJ man.

- Independent survey: Average height of men in NJ 5'8"

Example: Netflix Prize

Story Time:

Example: Netflix Prize

Story Time:

- In 2006, Netflix (e.g., DVDs) released awesome data sets
  $D_1 = \{\langle \text{user-id, movie, date of grade, grade} \rangle\}$.
  And another set $D_2 = \{\langle \text{user-id, movie, date of grade} \rangle\}$. no grade
  Wants researchers to predict grade on D_2.
  (Had another similar private data D_3 to evaluate grades : cross validation.)

Example: Netflix Prize

Story Time:

- In 2006, Netflix (e.g., DVDs) released awesome data sets $D_1 = \{\langle \text{user-id, movie, date of grade, grade} \rangle\}$.
  And another set $D_2 = \{\langle \text{user-id, movie, date of grade} \rangle\}$.
  Wants researchers to predict grade on D_2.
  (Had another similar private data D_3 to evaluate grades : cross validation.)
- If certain improvement over Netflix's algorithm, get \$1 million!

Example: Netflix Prize

Story Time:

- In 2006, Netflix (e.g., DVDs) released awesome data sets $D_1 = \{\langle \text{user-id, movie, date of grade, grade} \rangle\}$.
  And another set $D_2 = \{\langle \text{user-id, movie, date of grade} \rangle\}$.
  Wants researchers to predict grade on D_2.
  (Had another similar private data D_3 to evaluate grades : cross validation.)
- If certain improvement over Netflix's algorithm, get \$1 million!
- Led to lots of cool research!

Example: Netflix Prize

Story Time:

- In 2006, Netflix (e.g., DVDs) released awesome data sets
  $D_1 = \{\langle$user-id, movie, date of grade, grade$\rangle\}$.
  And another set $D_2 = \{\langle$user-id, movie, date of grade$\rangle\}$.
  Wants researchers to predict grade on D_2.
  (Had another similar private data D_3 to evaluate grades : cross validation.)
- If certain improvement over Netflix's algorithm, get \$1 million!
- Led to lots of cool research!
- Raters of movies also rated on IMDB (w/ user id, time stamp)

Example: Netflix Prize

Story Time:

- In 2006, Netflix (e.g., DVDs) released awesome data sets $D_1 = \{\langle$user-id, movie, date of grade, grade$\rangle\}$.
  And another set $D_2 = \{\langle$user-id, movie, date of grade$\rangle\}$.
  Wants researchers to predict grade on D_2.
  (Had another similar private data D_3 to evaluate grades : cross validation.)
- If certain improvement over Netflix's algorithm, get \$1 million!
- Led to lots of cool research!
- Raters of movies also rated on IMDB (w/ user id, time stamp)
- Researchers showed that by linking who rated similar sets of movies, with similar scores and times, they could identify many people.

Example: Netflix Prize

Story Time:

- In 2006, Netflix (e.g., DVDs) released awesome data sets $D_1 = \{\langle$user-id, movie, date of grade, grade$\rangle\}$.
  And another set $D_2 = \{\langle$user-id, movie, date of grade$\rangle\}$.
  Wants researchers to predict grade on D_2.
  (Had another similar private data D_3 to evaluate grades : cross validation.)
- If certain improvement over Netflix's algorithm, get \$1 million!
- Led to lots of cool research!
- Raters of movies also rated on IMDB (w/ user id, time stamp)
- Researchers showed that by linking who rated similar sets of movies, with similar scores and times, they could identify many people.
- (maybe watched embarrassing films on Netflix, not listed on IMDB)

Example: Netflix Prize

Story Time:

- In 2006, Netflix (e.g., DVDs) released awesome data sets $D_1 = \{\langle$user-id, movie, date of grade, grade$\rangle\}$.
  And another set $D_2 = \{\langle$user-id, movie, date of grade$\rangle\}$.
  Wants researchers to predict grade on D_2.
  (Had another similar private data D_3 to evaluate grades : cross validation.)
- If certain improvement over Netflix's algorithm, get \$1 million!
- Led to lots of cool research!
- Raters of movies also rated on IMDB (w/ user id, time stamp)
- Researchers showed that by linking who rated similar sets of movies, with similar scores and times, they could identify many people.
- (maybe watched embarrassing films on Netflix, not listed on IMDB)
- Class action lawsuit filed (lated dropped) against Netflix.

Example: Netflix Prize

Story Time:

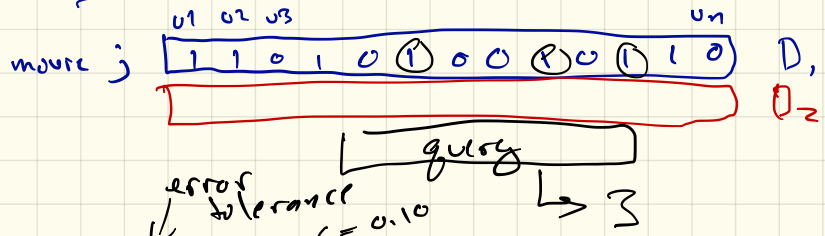
- In 2006, Netflix (e.g., DVDs) released awesome data sets $D_1 = \{\langle$user-id, movie, date of grade, grade$\rangle\}$.
  And another set $D_2 = \{\langle$user-id, movie, date of grade$\rangle\}$.
  Wants researchers to predict grade on D_2.
  (Had another similar private data D_3 to evaluate grades : cross validation.)
- If certain improvement over Netflix's algorithm, get \$1 million!
- Led to lots of cool research!
- Raters of movies also rated on IMDB (w/ user id, time stamp)
- Researchers showed that by linking who rated similar sets of movies, with similar scores and times, they could identify many people.
- (maybe watched embarrassing films on Netflix, not listed on IMDB)
- Class action lawsuit filed (lated dropped) against Netflix.
- Netflix Prize had proposed sequel, dropped in 2010 for more privacy concerns.

Differential Privacy

- Two similar data sets D_1 & D_2 (true / released)
  - ↳ global analysis on D_1 similar to D_2
  - ↳ for no particular data point in D_1 can I know its value.

Global analysis

Databases



Was $q(D_1) \geq 4$

for any $q \in Q$

$$\frac{Pr[q(D_1) \in R]}{Pr[q(D_2) \in R]} \leq \exp(\varepsilon) \approx 1 + \varepsilon$$

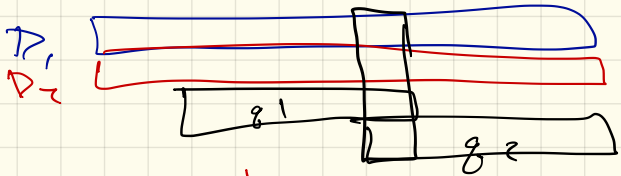
Two-Versions

- **Interactive Version**

  I control D_1, D_2

  I limit queries

  ↳ I return answer w/ noise



Change D_2 as I go.

- **Non-Interactive**

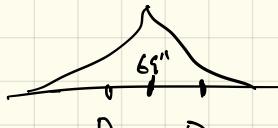
  I perturb $D_1 \rightarrow D_2$

  I release D_2.

  (typically done

  D_2 = Diff Laplace Noise.

Height of Sly Stallone



$D_2 \rightarrow 5'9'' \approx 69'' = D_1 + \text{Lap}(\alpha)$

example $D_1 = 68''$

e^{Lap}

$$\frac{P_r[D_1 \geq 70]}{P_r[D_1 \geq 70]} \approx \frac{e^{-2\alpha}}{e^{-\alpha}} = e^{-2\alpha} \cdot e^{\alpha} = e^{-2\alpha+\alpha}$$

$$= e^{-\alpha}$$

$$\frac{P_r[D_2 \geq 70]}{P_r[D_1 \geq 70]} = e^{\alpha} \approx 1 + \alpha$$