Differential Privacy

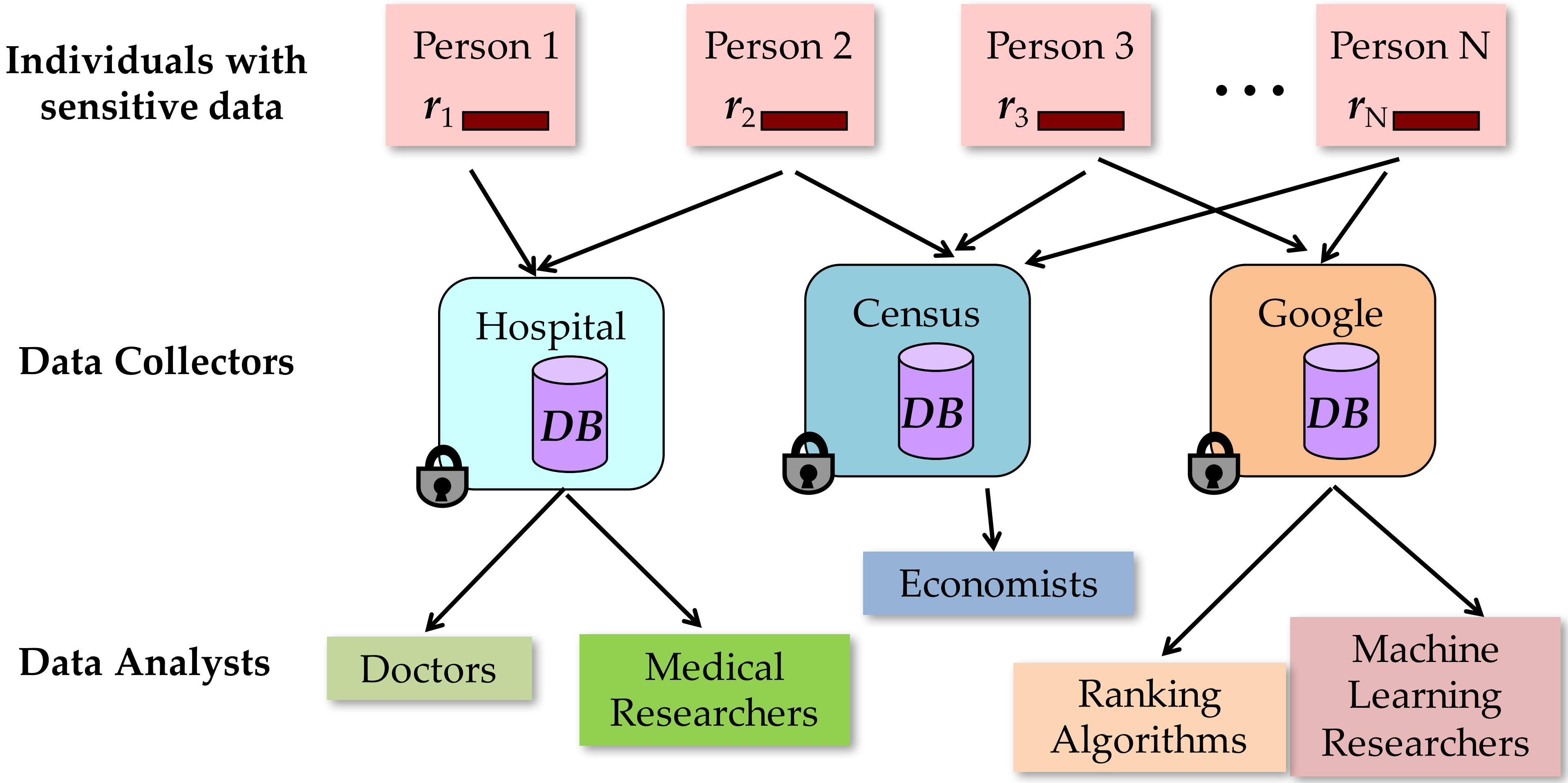
Privacy & Fairness in Data Science
CompSci 590.01 Fall 2018

DUKE
COMPUTER SCIENCE

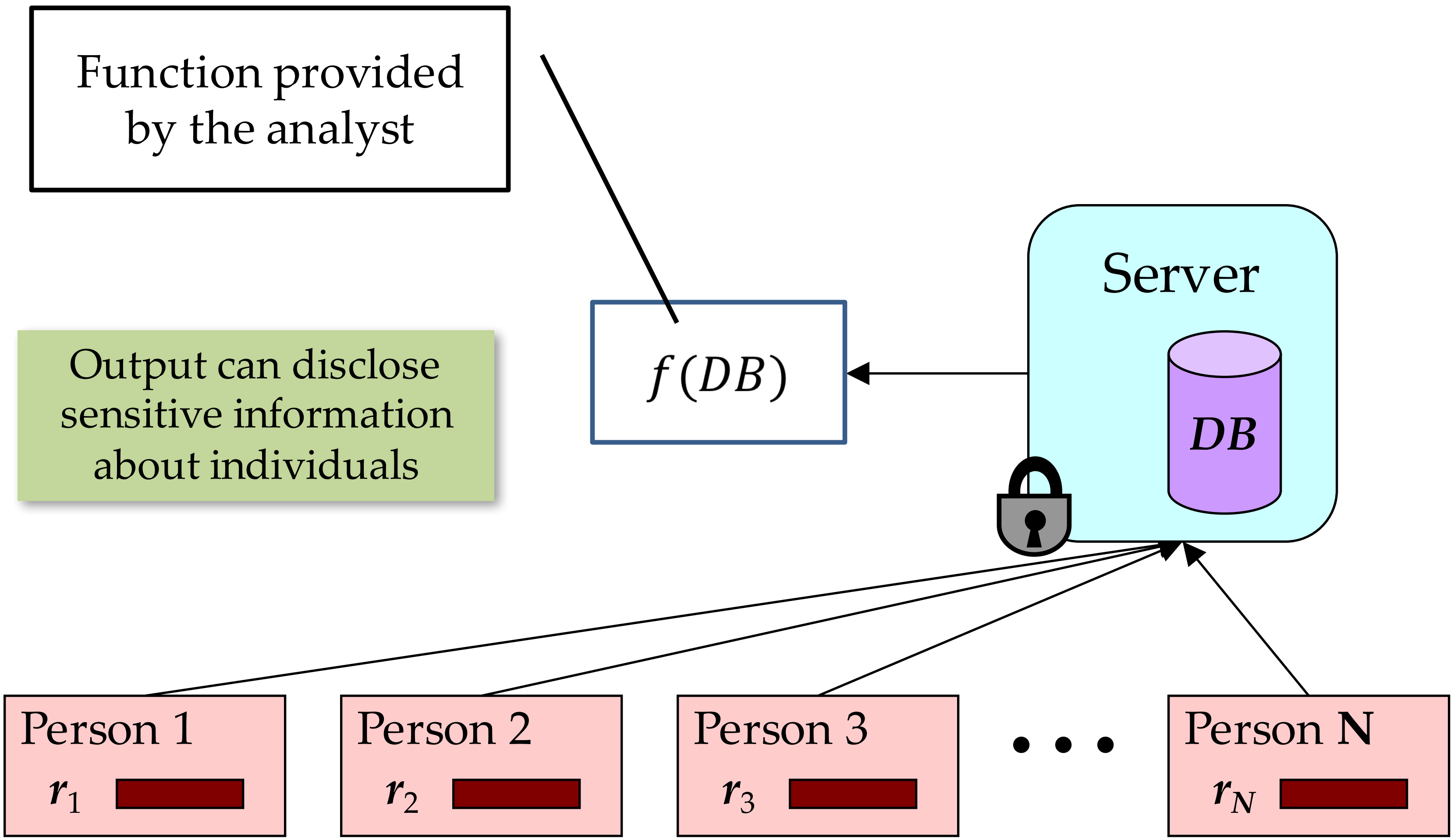
Outline

- Problem
- Differential Privacy
- Algorithms

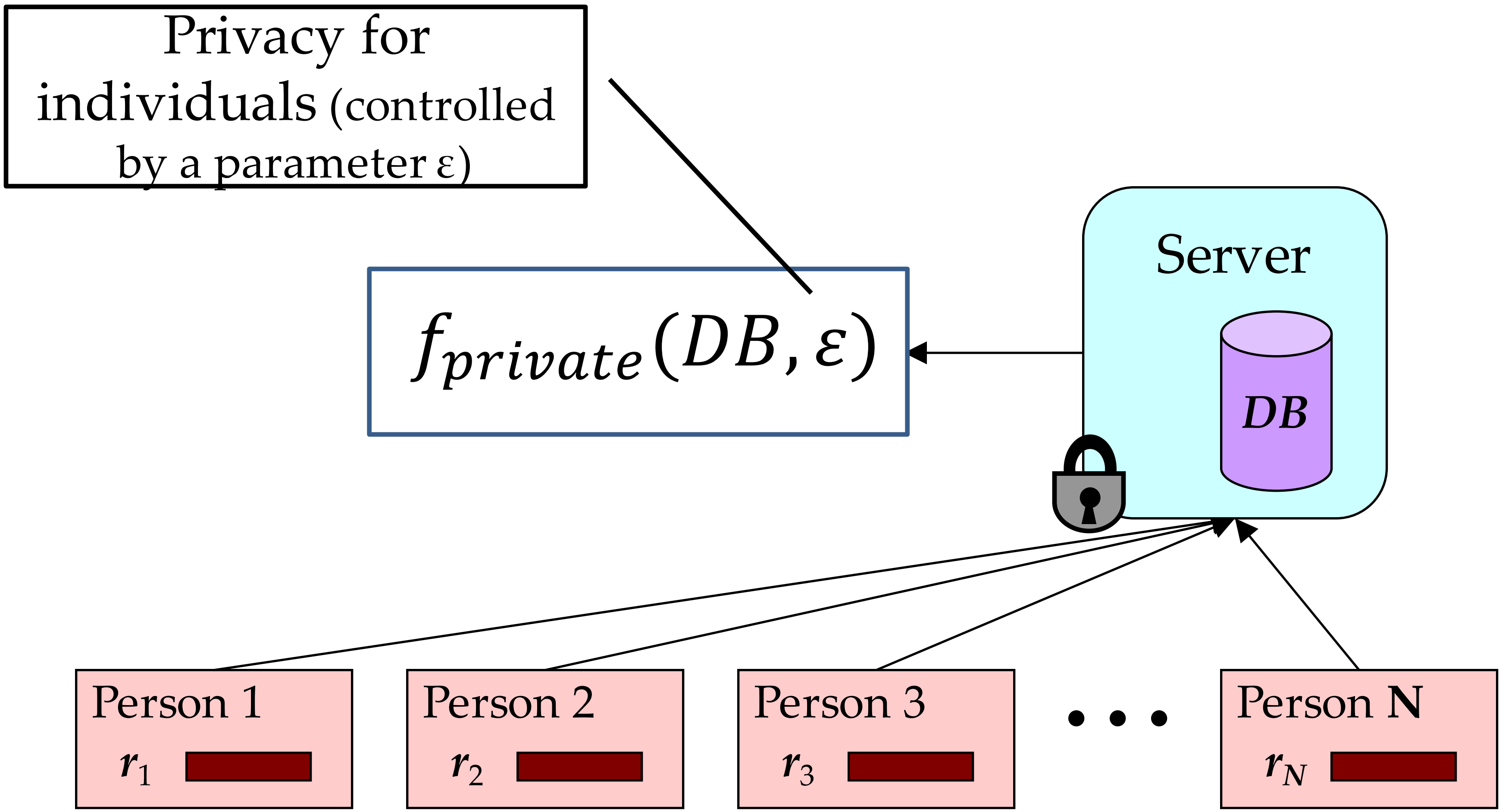
Statistical Databases



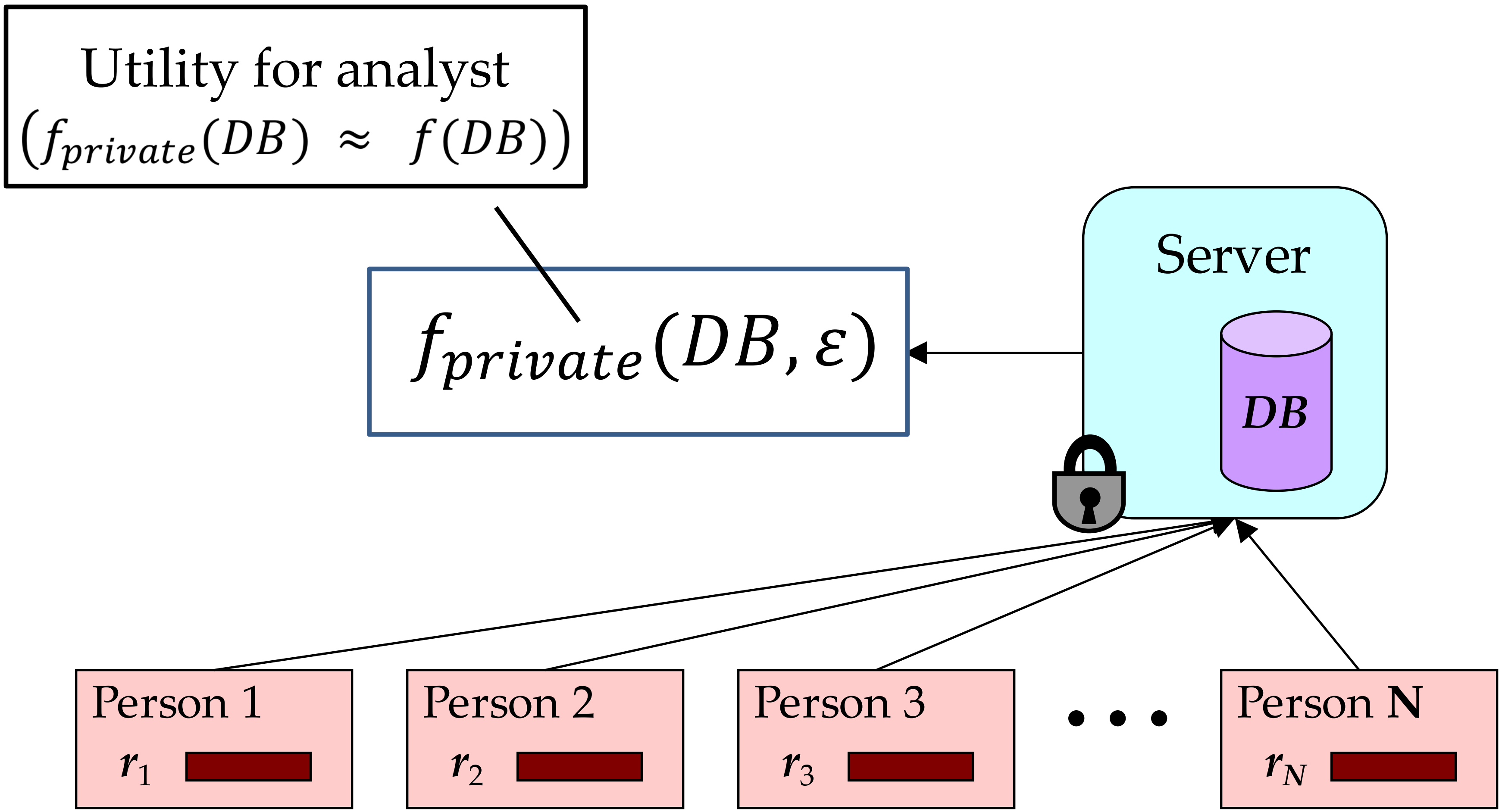
Statistical Database Privacy



Statistical Database Privacy



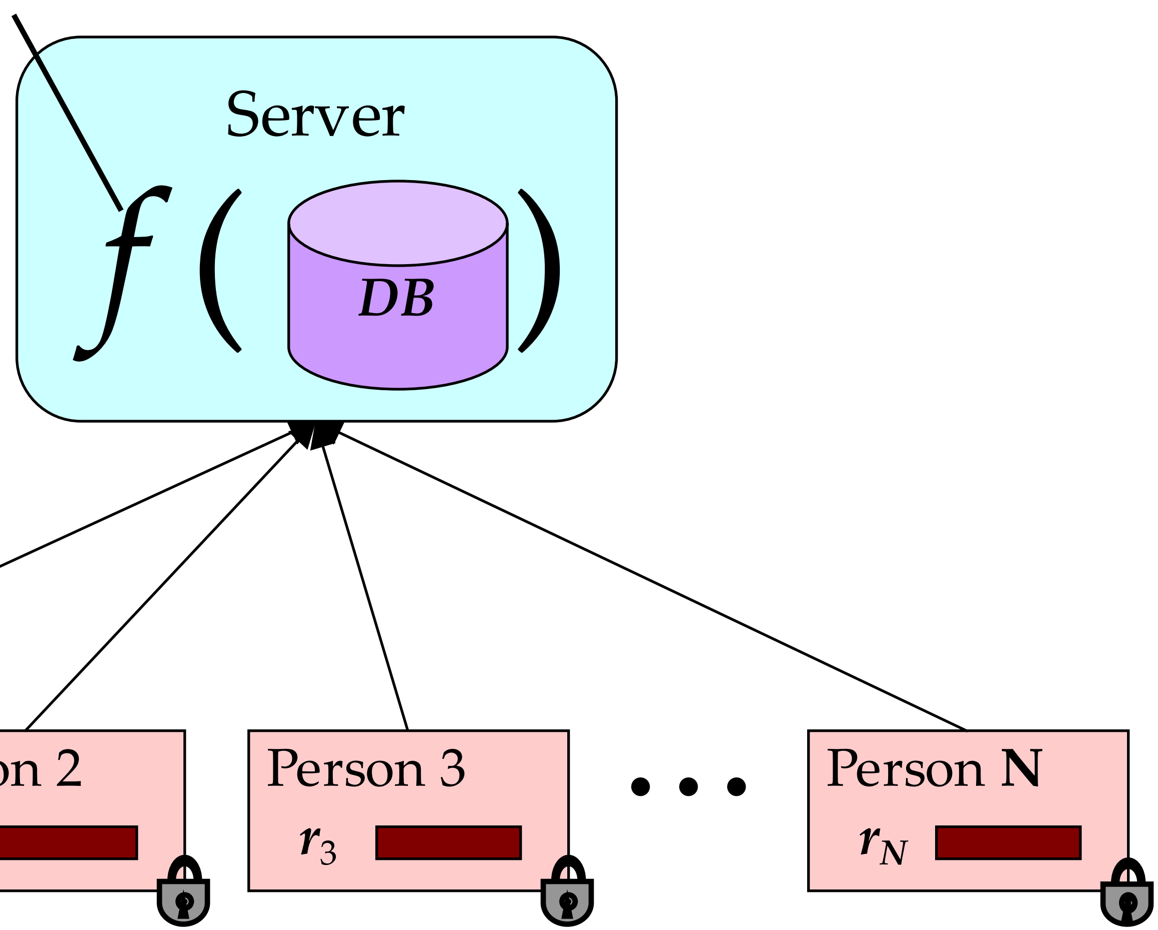
Statistical Database Privacy



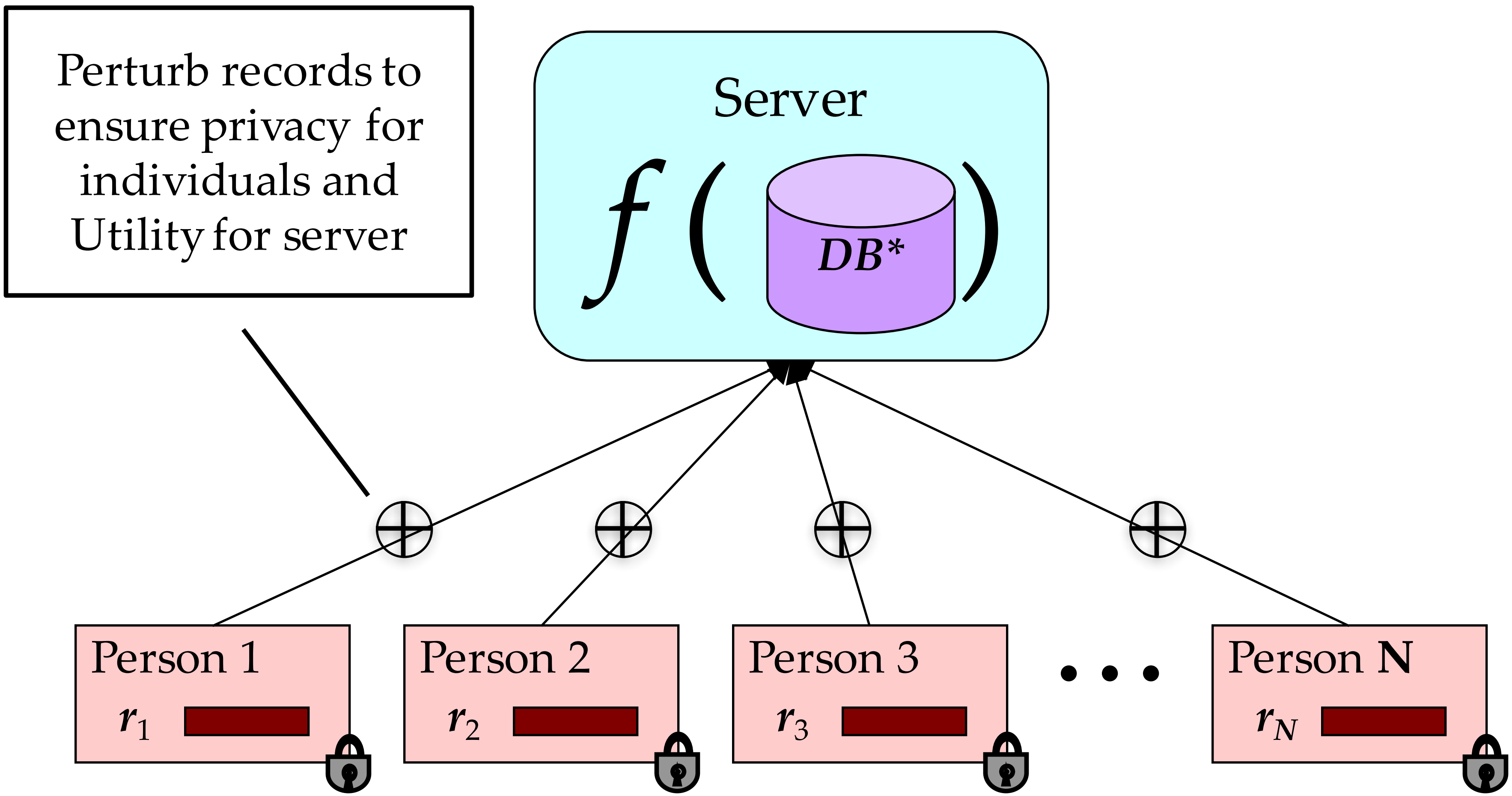
Statistical Database Privacy (untrusted collector)

Server wants to compute f

Individuals do not want server to infer their records



Statistical Database Privacy (untrusted collector)



Statistical Databases in real-world applications

Application	Data Collector	Private Information	Analyst	Function (utility)
Medical	Hospital	Disease	Epidemiologist	Correlation between disease and geography
Genome analysis	Hospital	Genome	Statistician/ Researcher	Correlation between genome and disease
Advertising	Google/FB	Clicks/Browsing	Advertiser	Number of clicks on an ad by age/region/gender …
Social Recommen-dations	Facebook	Friend links / profile	Another user	Recommend other users or ads to users based on social network

Statistical Databases in real-world applications

- Settings where data collector may not be trusted (or may not want the liability …)

Application	Data Collector	Private Information	Function (utility)
Location Services	Verizon/AT&T	Location	Traffic prediction
Recommen-dations	Amazon/Google	Purchase history	Recommendation model
Traffic Shaping	Internet Service Provider	Browsing history	Traffic pattern of groups of users

Privacy is *not* …

Statistical Database Privacy is not …

- Encryption:

Statistical Database Privacy is not …

- Encryption:
  Alice sends a message to Bob such that Trudy (attacker) does not learn the message. Bob should get the correct message …

- Statistical Database Privacy:
  Bob (attacker) can access a database
  - Bob must learn aggregate statistics, but
  - Bob must not learn new information about individuals in database.

Statistical Database Privacy is not …

- Computation on Encrypted Data:

Statistical Database Privacy is not …

- Computation on Encrypted Data:
  - Alice stores encrypted data on a server controlled by Bob (attacker).
  - Server returns correct query answers to Alice, without Bob learning *anything* about the data.

- Statistical Database Privacy:
  - Bob is allowed to learn aggregate properties of the database.

Statistical Database Privacy is not …

- The Millionaires Problem:

Statistical Database Privacy is not …

- Secure Multiparty Computation:
  - A set of agents each having a private input xi …
  - … Want to compute a function f(x1, x2, …, xk)
  - Each agent can learn the true answer, but must learn no other information than what can be inferred from their private input and the answer.

- Statistical Database Privacy:
  - Function output *must not disclose* individual inputs.

Statistical Database Privacy is not …

- Access Control:

Statistical Database Privacy is not …

- Access Control:
  - A set of agents want to access a set of resources (could be files or records in a database)
  - Access control rules specify who is allowed to access (*or not access*) certain resources.
  - 'Not access' usually means no information must be disclosed

- Statistical Database:
  - A single database and a single agent
  - Want to release aggregate statistics about a set of records without allowing access to individual records

Privacy Problems

- In todays systems a number of privacy problems arise:
  - Encryption when communicating data across a unsecure channel
  - Secure Multiparty Computation when different parties want to compute on a function on their private data without using a centralized third party
  - Computing on encrypted data when one wants to use an unsecure cloud for computation
  - Access control when different users own different parts of the data

- Statistical Database Privacy:
  Quantifying (and bounding) the amount of information disclosed about individual records by the output of a valid computation.

What *is* privacy?

Desiderata for a Privacy Definition

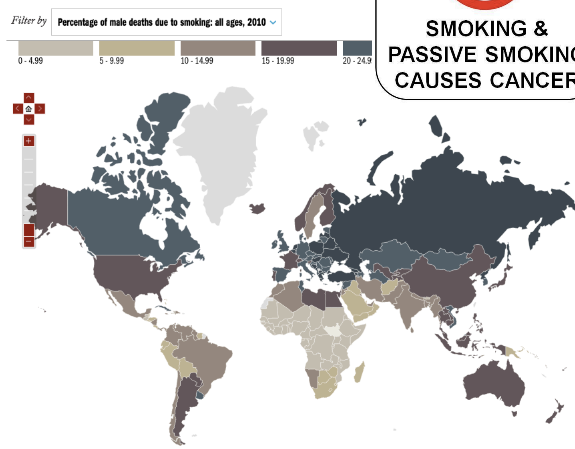
1. Resilience to background knowledge
   - A privacy mechanism must be able to protect individuals' privacy from attackers who may possess background knowledge
2. Privacy without obscurity
   - Attacker must be assumed to know the algorithm used as well as all parameters [MK15]
3. Post-processing
   - Post-processing the output of a privacy mechanism must not change the privacy guarantee [KL10, MK15]
4. Composition over multiple releases
   - Allow a graceful degradation of privacy with multiple invocations on the same data [DN03, GKS08]

Privacy Breach: Attempt 1

A privacy mechanism M(D)
that allows
an unauthorized party 
to learn sensitive information about any individual in D,

which  could not have learnt without access to M(D).

Alice



Alice has Cancer

Is this a privacy breach? *NO*

Privacy Breach: Attempt 2

A privacy mechanism M(D) that allows
an unauthorized party 
to learn sensitive information about
any individual Alice in D,

which  could not have learnt even with access to M(D)
if Alice was *not in the dataset*.

Outline

- Problem
- Differential Privacy
- Algorithms

Differential Privacy

[Dwork ICALP 2006]

For every pair of inputs that differ in one row

For every output …

D_1 D_2 O

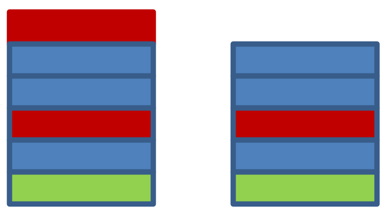
Adversary should not be able to distinguish between any D_1 and D_2 based on any O

$$\ln\left(\frac{\Pr[A(D_1) = o]}{\Pr[A(D_2) = o]}\right) \le \varepsilon, \qquad \varepsilon > 0$$

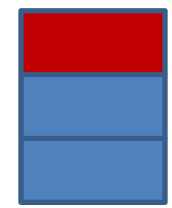
Why pairs of datasets *that differ in one row*?

For every pair of inputs that differ in one row

For every output …



D_1 D_2



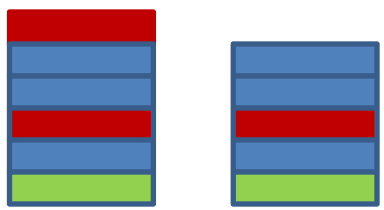
O

Simulate the presence or absence of a single record

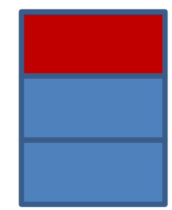
Why *all* pairs of datasets …?

For every pair of inputs that differ in one row

For every output …



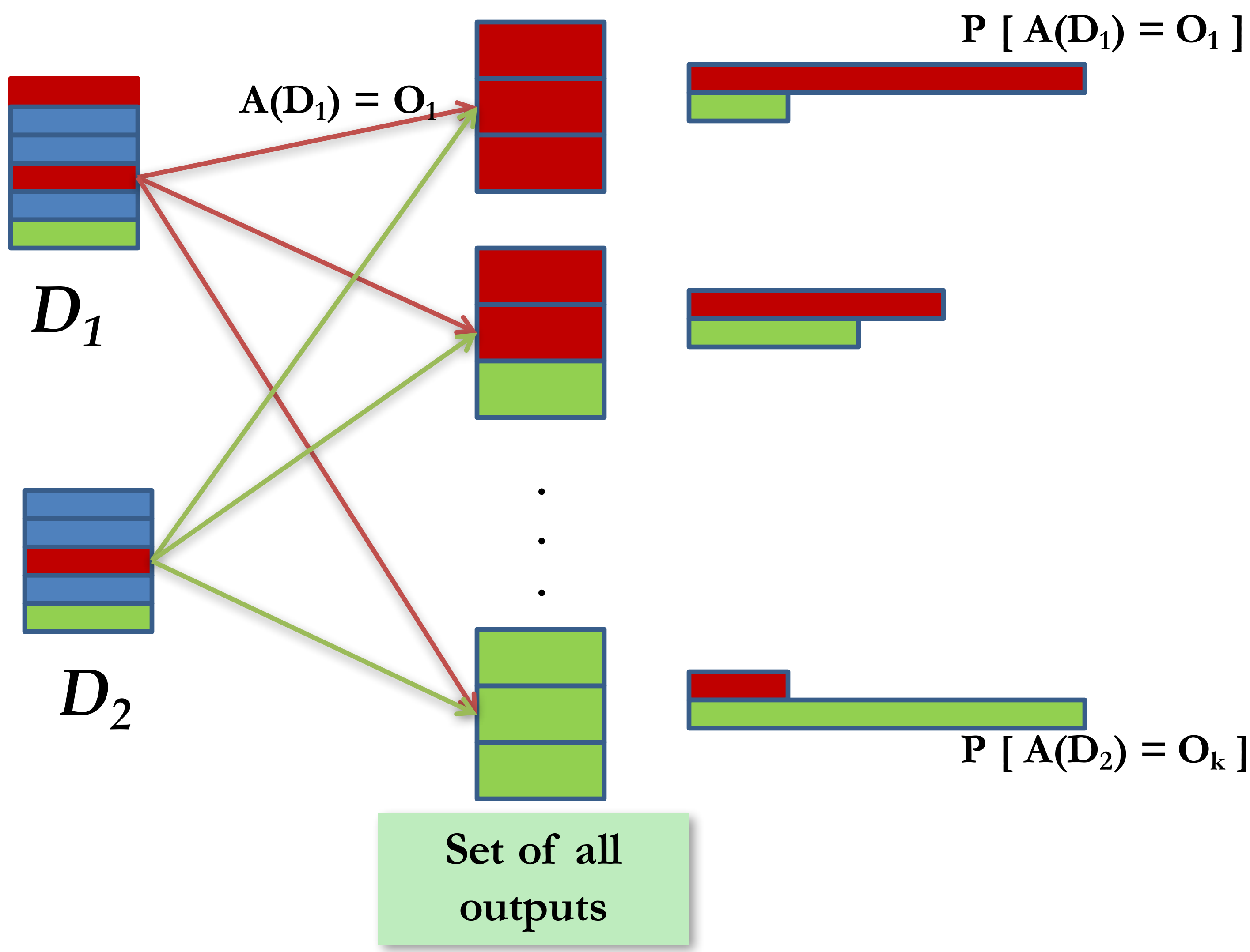
D_1 D_2



O

Guarantee holds no matter what the other records are.

Why *all* outputs?

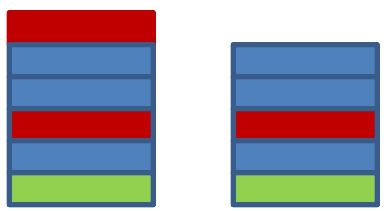


Should not be able to distinguish whether input was D_1 or D_2 no matter what the output

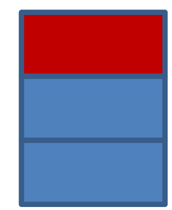
Privacy Parameter ε

For every pair of inputs that differ in one row

For every output …



D_1 D_2



O

$$\Pr[A(D_1) = o] \leq e^{\varepsilon} \Pr[A(D_2) = o]$$

Controls the degree to which D_1 and D_2 can be distinguished.
Smaller the ε more the privacy (and better the utility)

Desiderata for a Privacy Definition

1. Resilience to background knowledge
   - A privacy mechanism must be able to protect individuals' privacy from attackers who may possess background knowledge

2. Privacy without obscurity
   - Attacker must be assumed to know the algorithm used as well as all parameters [MK15]

3. Post-processing
   - Post-processing the output of a privacy mechanism must not change the privacy guarantee [KL10, MK15]

4. Composition over multiple releases
   - Allow a graceful degradation of privacy with multiple invocations on the same data [DN03, GKS08]

Differential Privacy

- Two equivalent definitions:

Every subset of outputs

$$\Pr[A(D_1) \in \Omega] \;\leq\; e^{\varepsilon} \Pr[A(D_2) \in \Omega]$$

Number of row additions and deletions to change X to Y

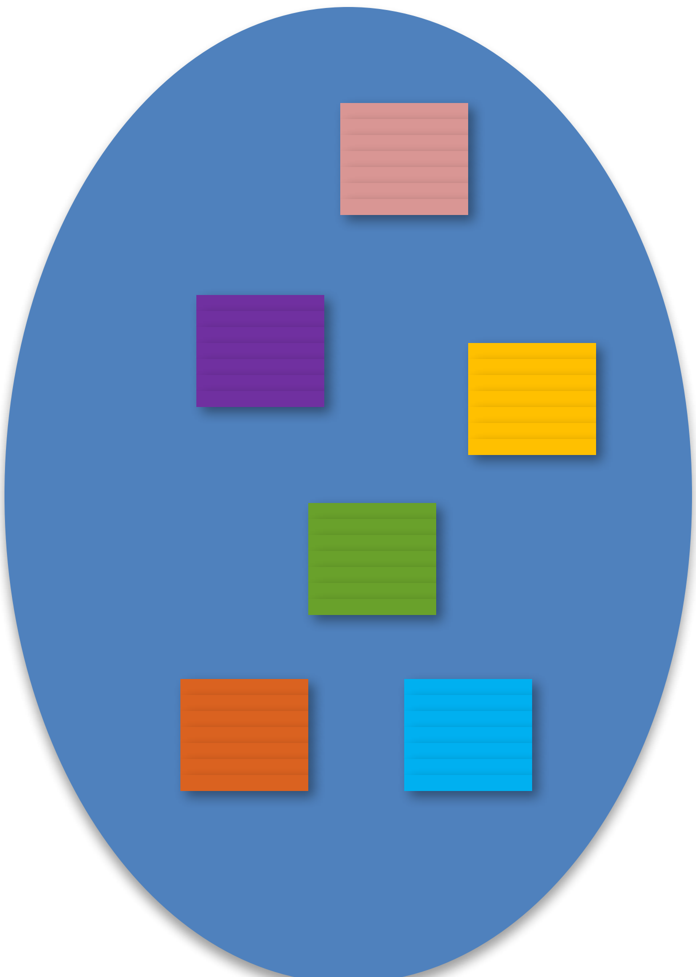
$$\Pr[A(X) \in \Omega] \;\leq\; e^{\varepsilon \cdot d(X,Y)} \Pr[A(Y) \in \Omega]$$

Outline

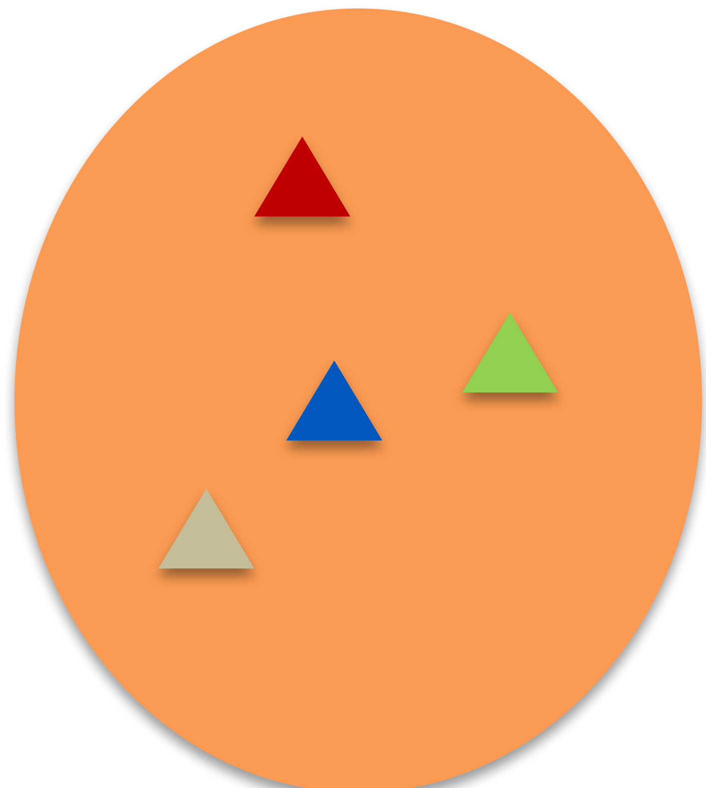
- Problem
- Differential Privacy
- Algorithms

Non-trivial deterministic Algorithms do not satisfy differential privacy

Space of all inputs

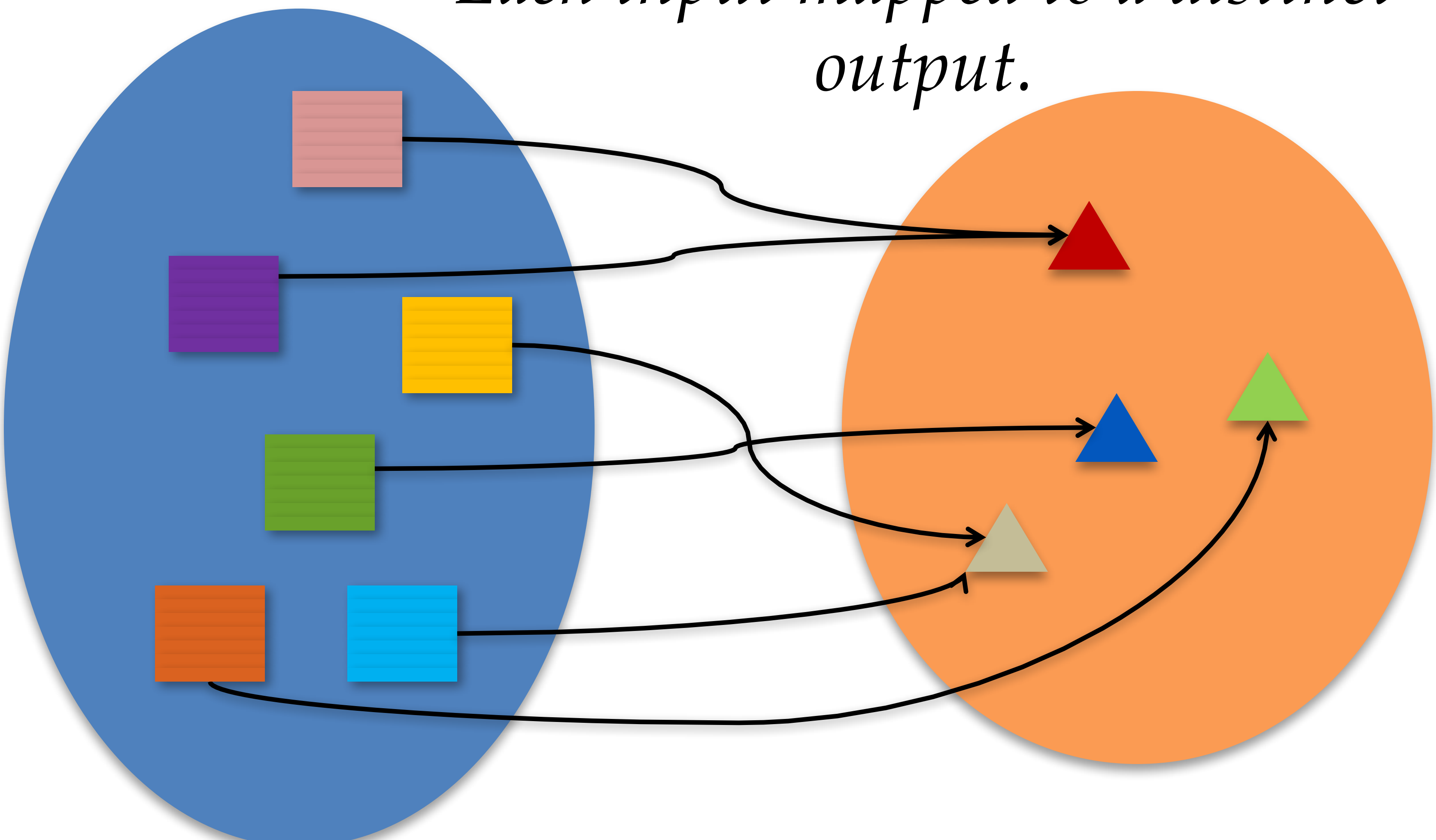


Space of all outputs (at least 2 distinct ouputs)

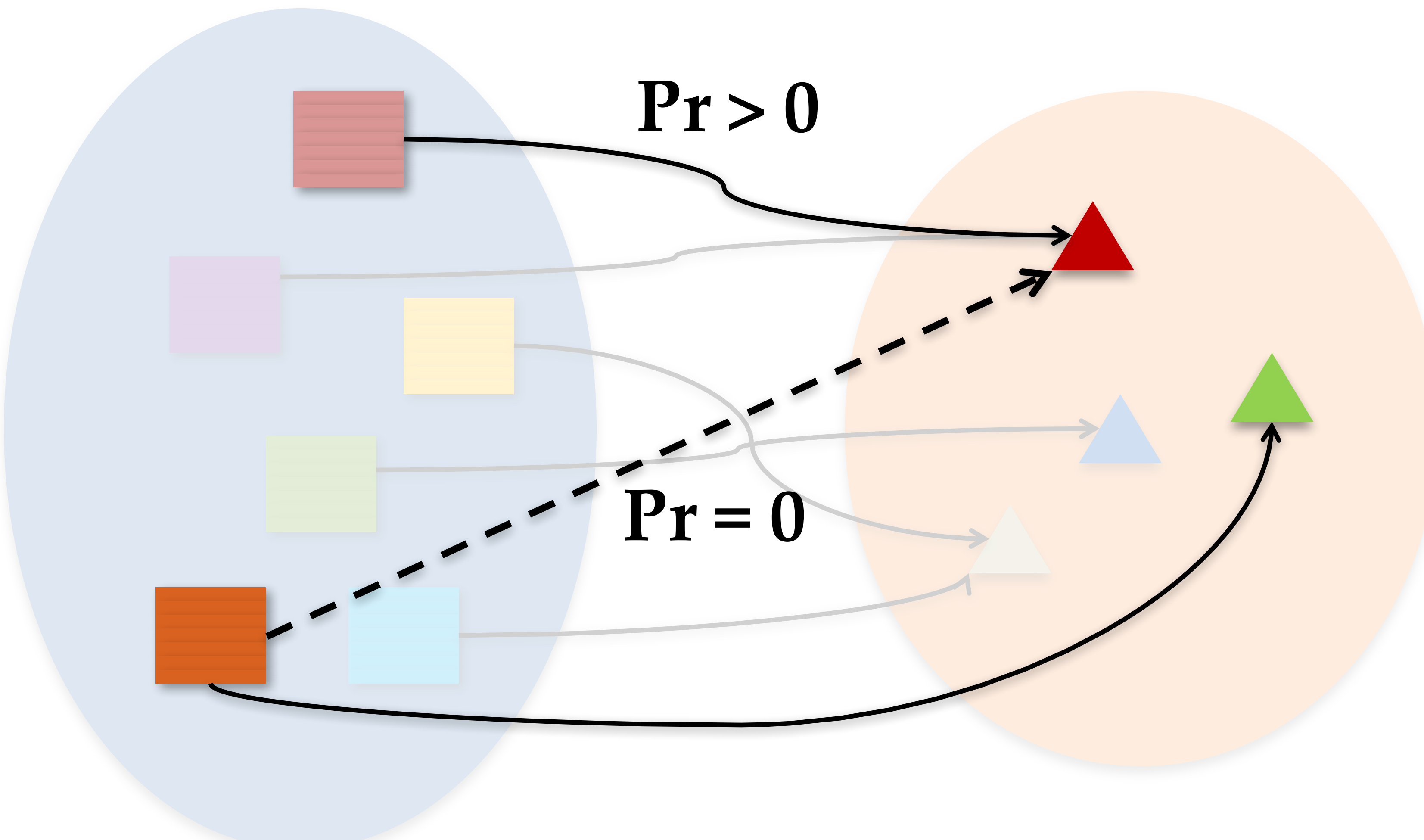


Non-trivial deterministic Algorithms do not satisfy differential privacy

Each input mapped to a distinct output.



There exist two inputs that differ in one entry mapped to different outputs.



Random Sampling …

… also does not satisfy differential privacy

Input Output

D_1 D_2 O

$\Pr[D_2 \rightarrow O] = 0$ implies $\log\left(\frac{\Pr[D_1 \rightarrow O]}{\Pr[D_2 \rightarrow O]}\right) = \infty$

[W 65]

Randomized Response (a.k.a. local randomization)

D

Disease (Y/N)
Y
Y
N
Y
N
N

With probability p,
Report true value

With probability 1-p,
Report flipped value

O

Disease (Y/N)
Y
N
N
N
Y
N

Differential Privacy Analysis

- Consider 2 databases D, D′ (of size M) that differ in the j^{th} value
  - $D[j] \neq D'[j]$. But, $D[i] = D'[i]$, for all $i \neq j$

- Consider some output O

$$\frac{P(D \to O)}{P(D' \to O)} \leq e^{\varepsilon} \Leftrightarrow \frac{1}{1+e^{\varepsilon}} < p < \frac{e^{\varepsilon}}{1+e^{\varepsilon}}$$

Next class

- Basic Algorithmic Primitives
- Composition