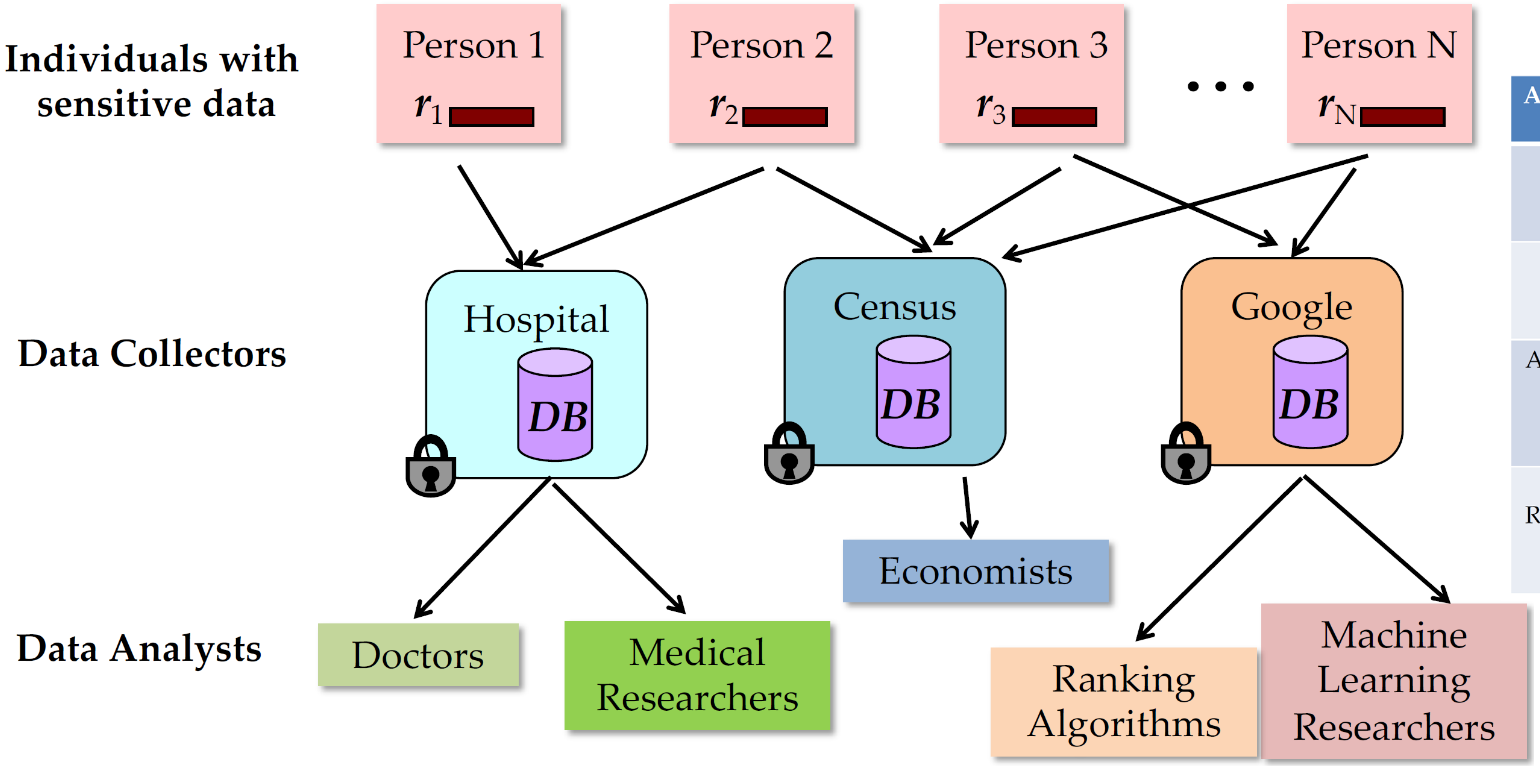
Differential Privacy

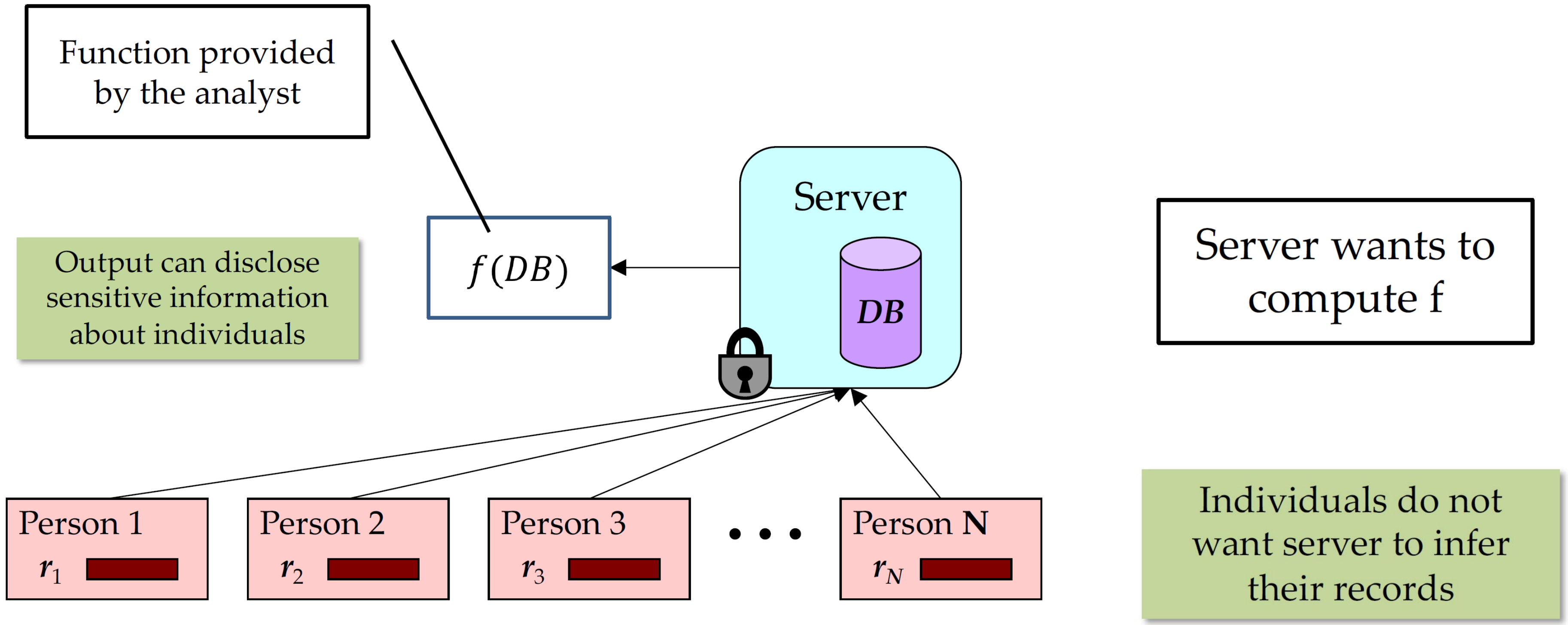
- Differential privacy in statistical databases/datasets



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 Recommendations	Facebook	Friend links / profile	Another user	Recommend other users or ads to users based on social network

Statistical Database

- Statistical database query scheme



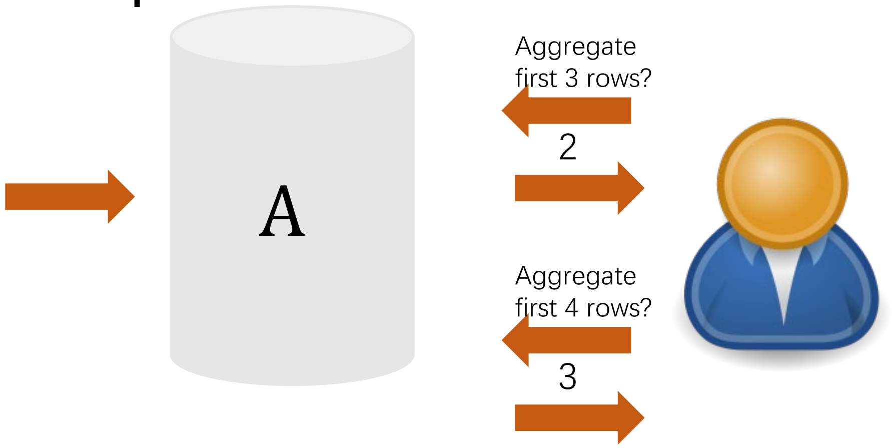
Differential Privacy

- **Differential**

Database

Name	Yes/No
Alice	1
Bob	1
Cathy	0
David	1

Answer aggregate queries

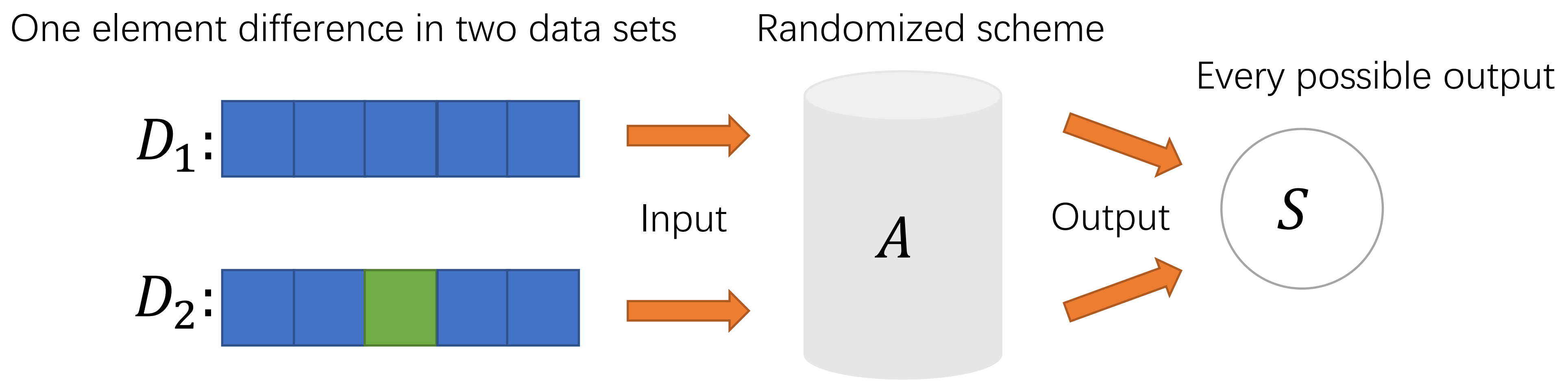


David is on row 4, Davis is 1.

Differential Privacy

- **ε-Differential Privacy**: A randomized mechanism A is **ε-Differential Private**, if for every pair of input datasets that differ by one element (*neighboring datasets*), for every output S,

$$\Pr\{A(D_1) = S\} \leq e^{\varepsilon} \times \Pr\{A(D_2) = S\}.$$

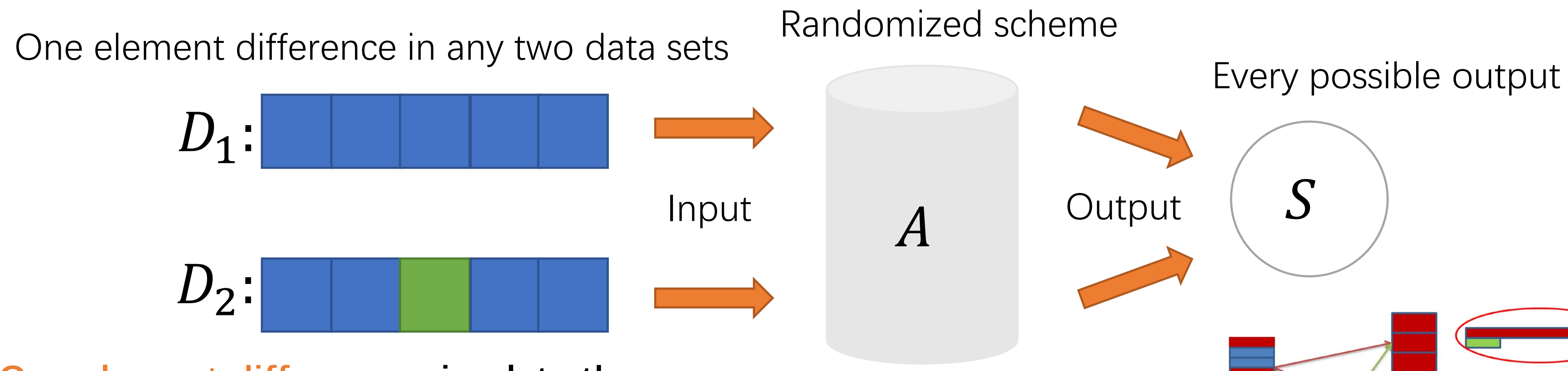


Cynthia Dwork, Frank McSherry, Kobbi Nissim, and Adam Smith. 2006. Calibrating noise to sensitivity in private data analysis. In Proceedings of the Third conference on Theory of Cryptography (TCC'06).

Differential Privacy

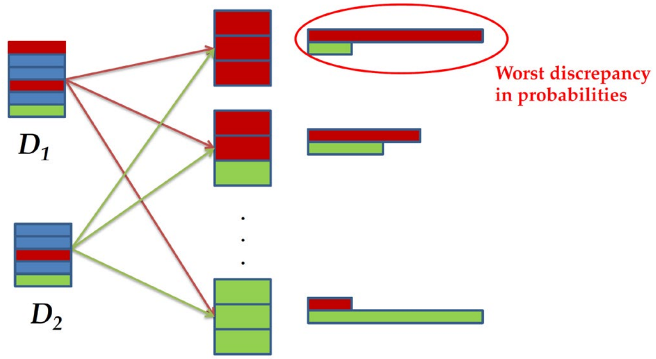
- ***ε*-Differential Privacy**

$$\Pr\{A(D_1) = S\} \le e^{\varepsilon} \times \Pr\{A(D_2) = S\}.$$



One element difference: simulate the presence/absence/change of a record

Every pair: guarantee holds no matter what the other records are.



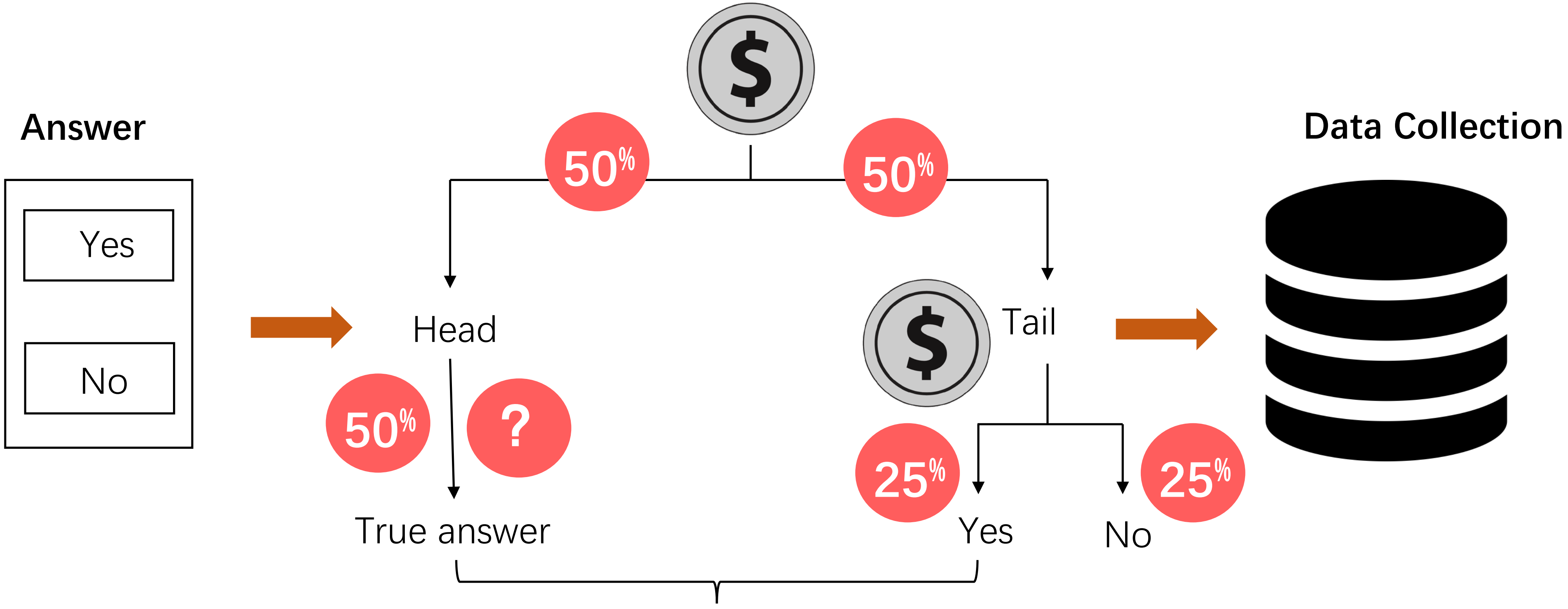
Differential Privacy

- Resilience to **background knowledge**
  - A privacy mechanism must be able to protect individuals' privacy from attackers who may possess background knowledge
- Privacy **without obscurity**
  - Attacker must be assumed to know the algorithm used as well as all parameters
- **Post-processing**
  - Post-processing the output of a privacy mechanism must not change the privacy guarantee

Differential Privacy Mechanisms

Randomized Response

- **Randomized response mechanism**
  - Survey the distribution of a sensitive attribute in the customers without revealing sensitive information



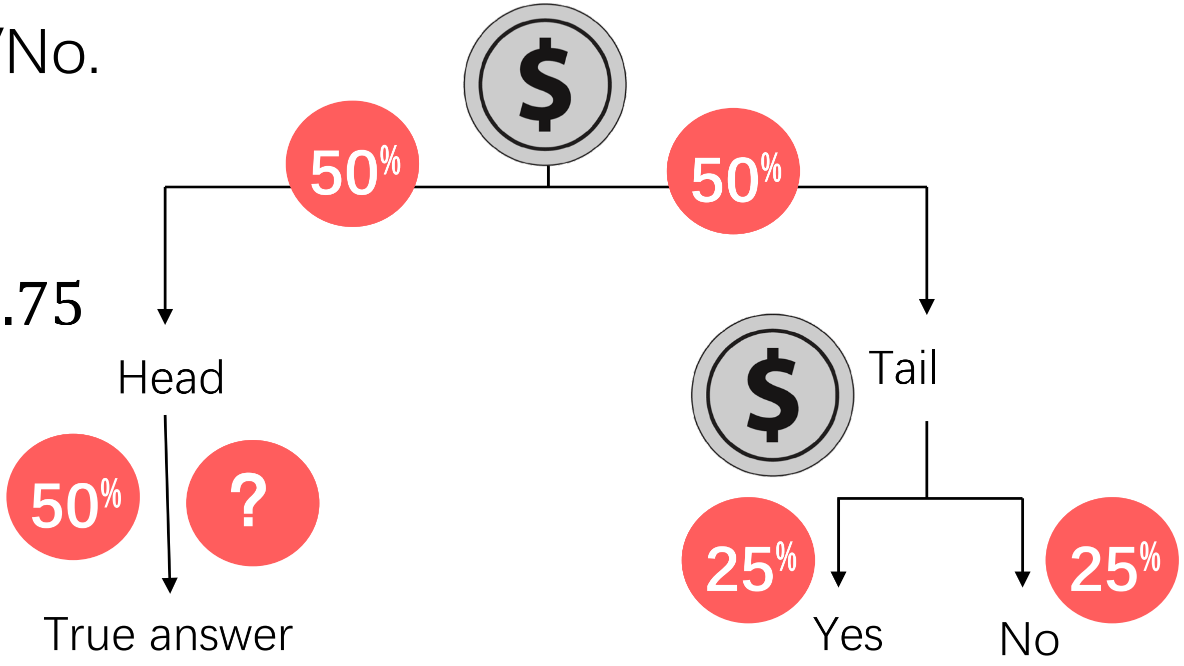
Randomized Response

- What is the privacy it guarantees in the framework of differential privacy?
  - Consider two neighboring dataset different in one row, Yes and No.
  - Two possible output for this row: Yes/No.
  - Analyze the probability.

$$\Pr\{A(\text{No}) = \text{Yes}\} = 0.25, \Pr\{A(\text{Yes}) = \text{Yes}\} = 0.75$$

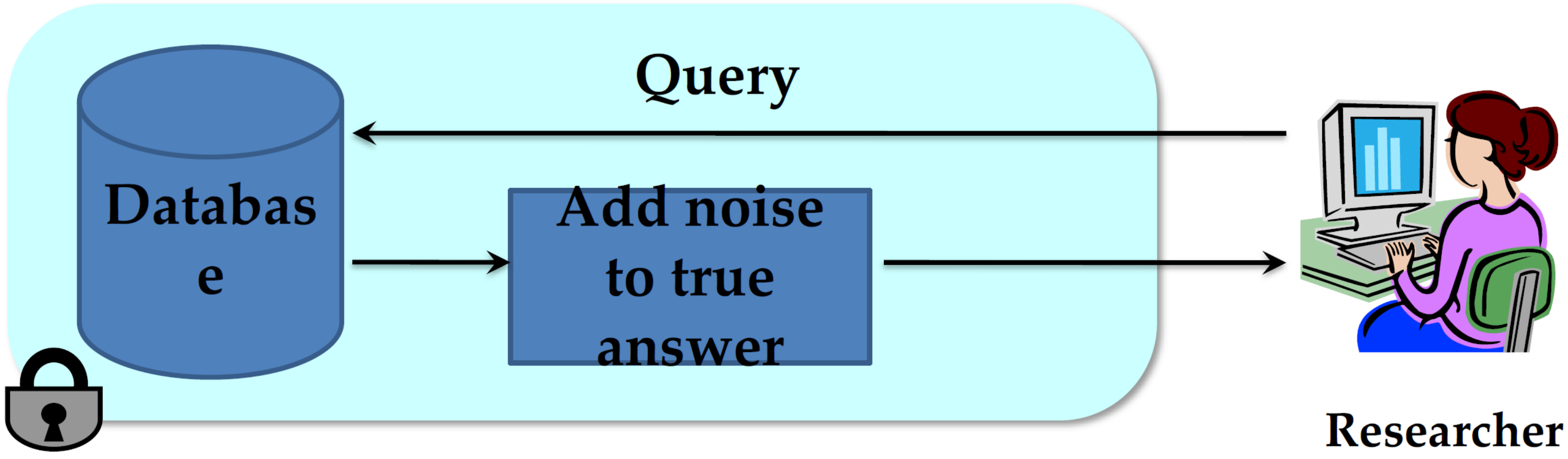
$$\Pr\{A(\text{No}) = \text{Yes}\} \leq e^{\varepsilon} \times \Pr\{A(\text{Yes}) = \text{Yes}\},$$

$$\varepsilon = \log 3.$$



Output Randomization

- Add noise to answers such that
  - Each answer does not leak too much information about the database.
  - Noisy answers are close to the original answers.



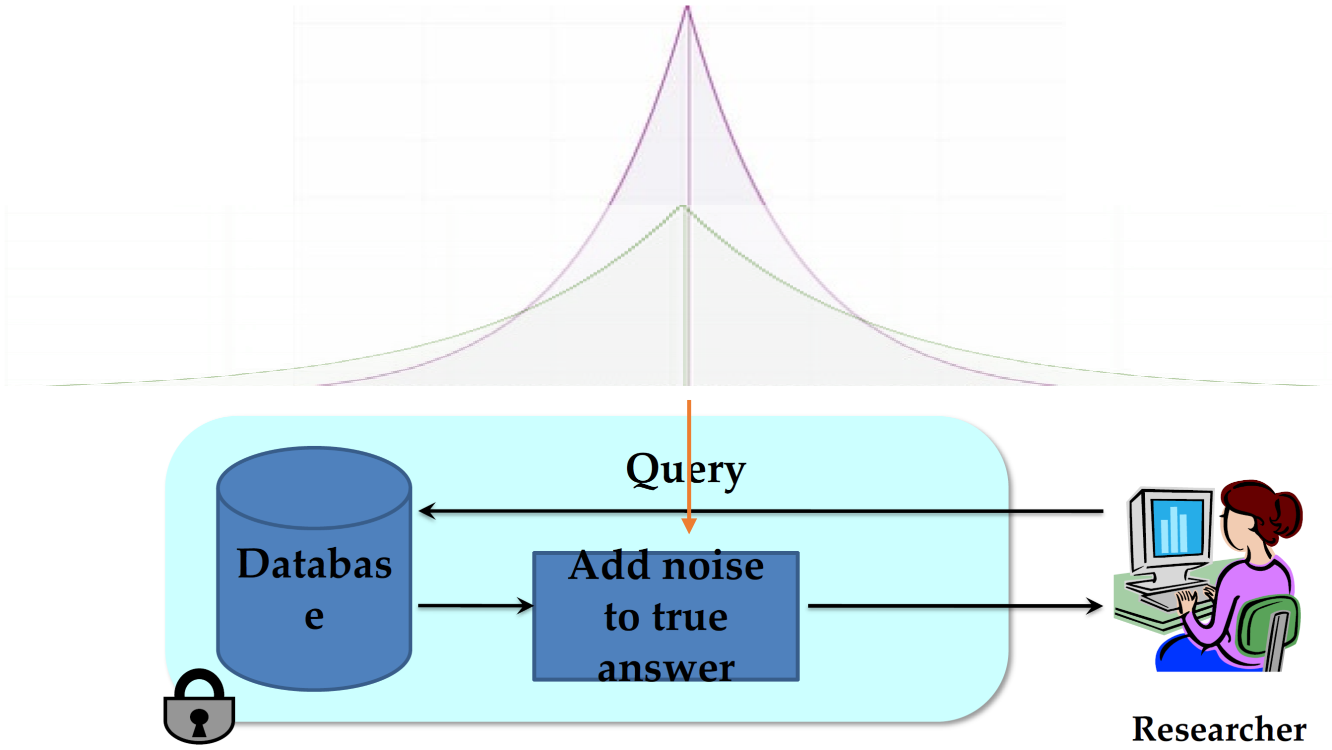
Laplace Mechanism

- To achieve differential privacy, we need add to the true answer, noise following Laplace distribution:

  - $Lap(b) = \frac{1}{2b} \exp\left(-\frac{|x|}{b}\right).$

    Scale

  - Mean = 0
  - Variance = $2b^2$



Laplace Mechanism

- How much noise for privacy?
- **Sensitivity:** let $\mathfrak{D}$ be a collection of datasets, function $f\colon \mathfrak{D} \to \mathbb{R}$, the L_1-sensitivity of f is:

$$\Delta f = \max_{x,y \text{ are neighboring datasets}} \|f(x) - f(y)\|_1.$$

- **E.g.**
  - Sensitivity for COUNT: 1
  - Sensitivity for SUM: max of the elements added.

Laplace Mechanism

- How much noise for privacy?
- **Theorem**： we add noise following $\mathrm{Lap}\left(\frac{\Delta \mathrm{f}}{\varepsilon}\right)$ to the true answer, we can achieve ε-differential privacy.

Theorem： we add noise following $\mathrm{Lap}\left(\frac{\Delta \mathrm{f}}{\varepsilon}\right)$ to the true answer, we can achieve ε-differential privacy.

- Proof:
  - Assume that the output for both **datasets $\boldsymbol{x}, \boldsymbol{y}$** is the same, denoted as $\boldsymbol{z}$.

$$\frac{p_x(z)}{p_y(z)} = \frac{\exp\left(-\frac{\epsilon|f(x)-z|}{\Delta f}\right)}{\exp\left(-\frac{\epsilon|f(y)-z|}{\Delta f}\right)}$$

$$= \exp\left(\frac{\epsilon(|f(y)-z| - |f(x)-z|)}{\Delta f}\right)$$

$$\leq \exp\left(\frac{\epsilon|f(y)-f(x)|}{\Delta f}\right)$$

$$\leq \exp(\epsilon)$$

Laplace Mechanism

- **Utility**
  - Error: $E(\text{true answer} - \text{noise answer})^2$

$$=Var(\text{Lap}\left(\frac{\Delta\text{f}}{\varepsilon}\right))= 2\left(\frac{\Delta\text{f}}{\varepsilon}\right)^2$$

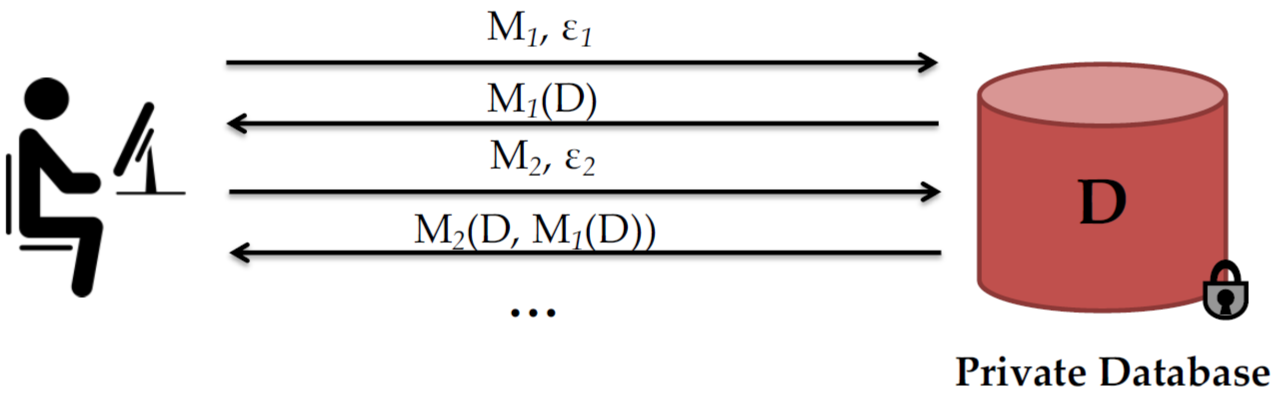
Laplace Mechanism vs Randomized Response

- Same ε-differential privacy.

- Laplace mechanism assumes data collected is trusted
- Randomized Response does not require data collected to be trusted
  - Also called a Local Algorithm, since each record is perturbed

Composition Theorem

- **Sequential Composition**



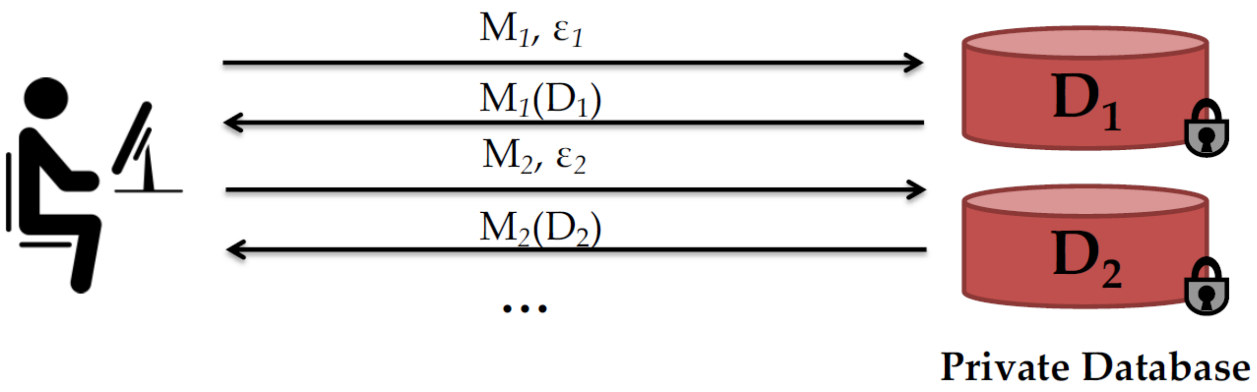
- If $M_1, M_2, \ldots, M_k$ are algorithms that access a private database D such that each M_i satisfies ε_i -differential privacy,

  then the combination of their outputs satisfies ε-differential privacy with

$$\varepsilon = \varepsilon_1 + \ldots + \varepsilon_k$$

Composition Theorem

- **Parallel Composition**



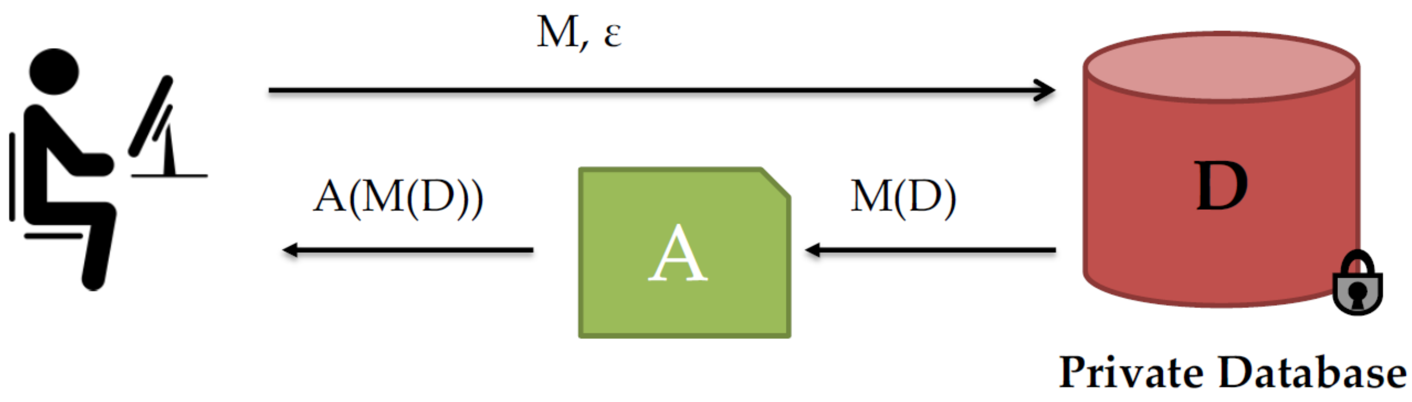
- If M_1, M_2, …, M_k are algorithms that access are algorithms that access disjoint databases D_1, D_2, …, D_k such that each M_i satisfies ε_i -differential privacy,

  then the combination of their outputs satisfies ε-differential privacy with

$$\varepsilon = \max(\varepsilon_1 , \ldots , \varepsilon_k)$$

Composition Theorem

- **Postprocessing**

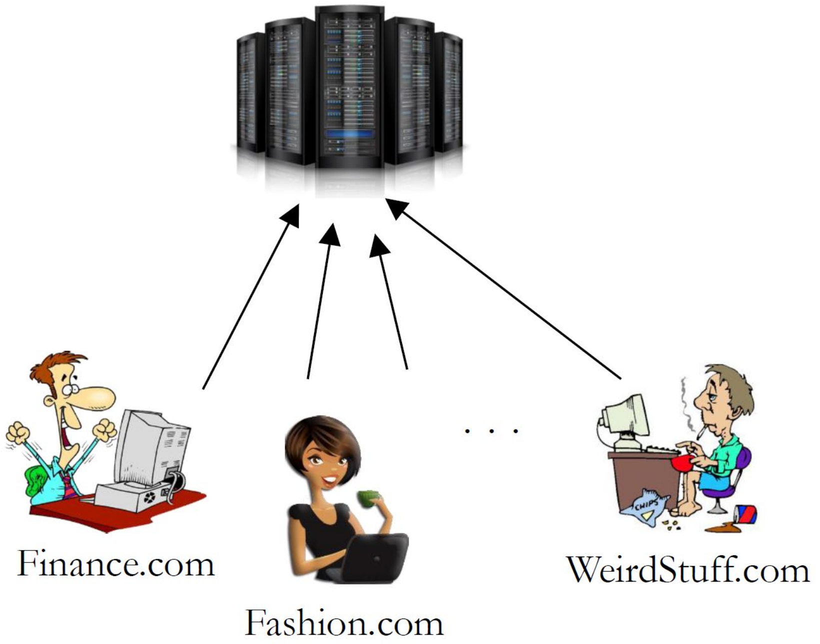


- If M is an ε-differentially private algorithm, any additional post-processing $A \circ M$ also satisfies ε-differential privacy.

Differential Privacy Applications

Differential Privacy in Chrome

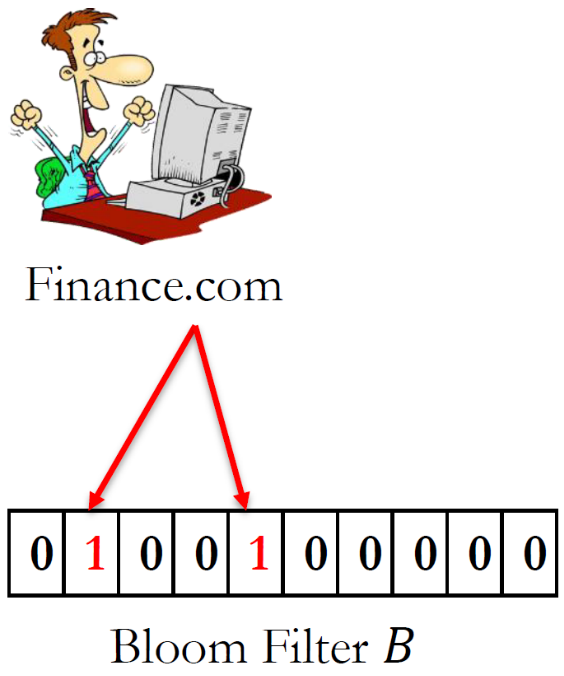
- **Problem:** What are the frequent unexpected Chrome homepage domains?
  - To learn malicious software that change Chrome setting without users' consent.
  - Protect user privacy.



Erlingsson, Úlfar, Vasyl Pihur, and Aleksandra Korolova. "Rappor: Randomized aggregatable privacy-preserving ordinal response." CCS 2014.

Client Input Perturbation

- Step 1: Use **Bloom filter**. h hash functions to hash input website string to k-bit vector



Randomized Response

- Step 2: Perturb B to fake Bloom Filter B' with randomized response, with a probability parameter f.

Finance.com

$$B'_i = \begin{cases} 1, & \text{with probability } \frac{1}{2}f \\ 0, & \text{with probability } \frac{1}{2}f \\ B_i, & \text{with probability } 1 - f \end{cases}$$

0	1	0	0	1	0	0	0	0	0

Bloom Filter B

$\Rightarrow$

0	1	1	0	0	0	0	1	0	0

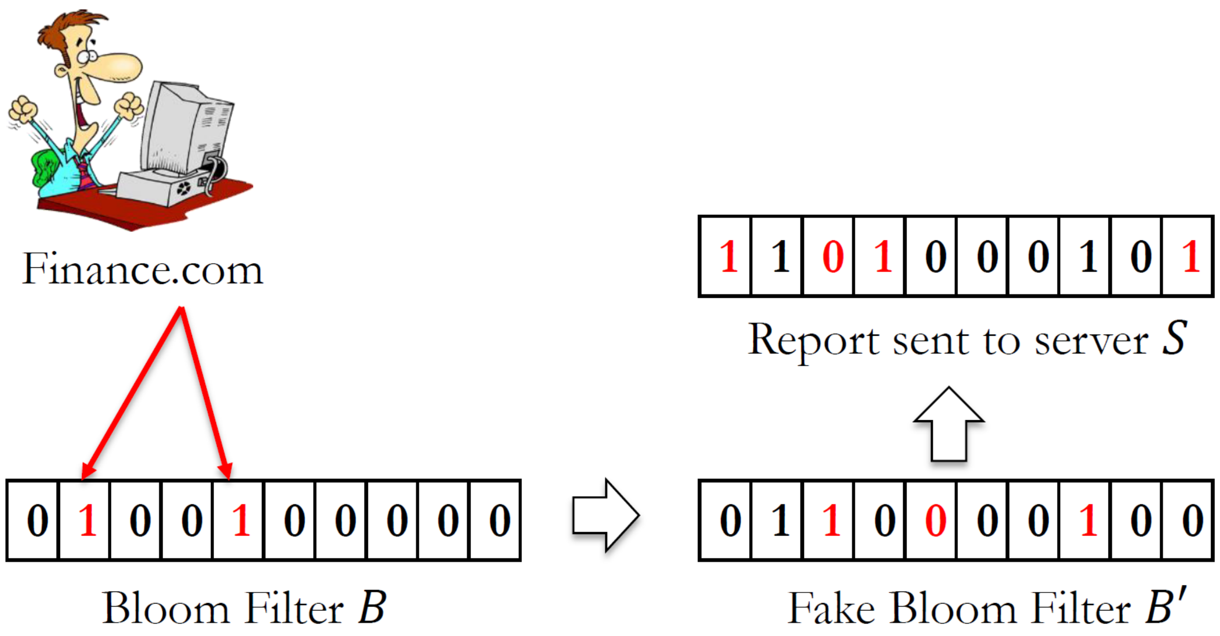
Fake Bloom Filter B'

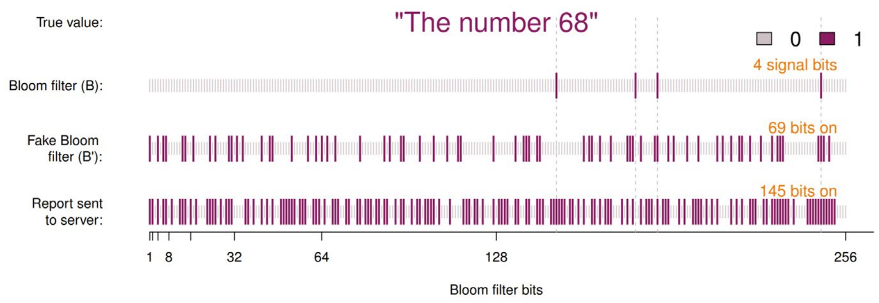
Instantaneous Randomized Response

- Step 3: another randomized response $B' \rightarrow S$
  - Flip the bit 1 with probability p
  - Flip the bit 0 with probability q

Why randomize two times?

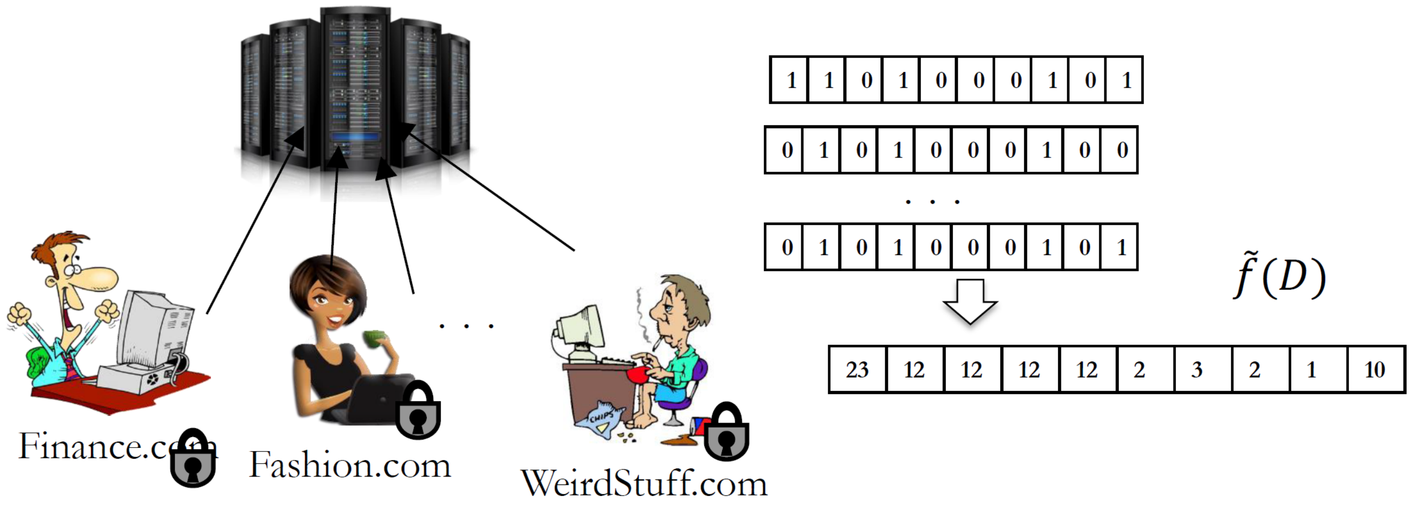
- Chrome collects information each day
- Want perturbed values to look different on different days to avoid linking





Server Report Decoding

- Estimate bit frequency from report



Differential privacy

- Definitions
  - Guarantee anyone's privacy
- Mechanisms:
  - Randomized Response
  - Laplace
- Applications