Differential Privacy

Privacy attacks and definition

Marco Gaboardi
Boston University
Chalmers University
DPella

The opinions expressed in this course are mine and they do not not reflect those of the National Science Foundation or the US. Census Bureau.

Data

Data

Statistics over Data

Databases and Queries

- For simplicity we will focus on a rather abstract notion of databases and queries.
- We will describe a database as a multiset (or sometimes an histogram) and queries as functions from a database to some (often numeric) domain.
- We will usually be interested in the results of some set of queries.

Data

Name	D1	D2	D3	D4	D5	D6	D6	D7	D8	D9	D10	D11	D12	D13	D14	D15
Alice	0	0	0	1	0	0	0	0	1	0	0	1	1	0	0	1
Bob	1	0	1	1	1	0	1	0	1	0	1	0	0	1	0	0
Cynthi	0	1	0	1	1	1	0	1	0	0	0	1	0	0	1	0
Dan	1	0	1	0	0	1	1	0	1	1	0	0	0	0	1	1
Eve	0	0	0	1	1	0	1	1	0	1	0	1	0	1	0	0
Frank	0	0	1	1	0	1	1	0	1	1	0	0	1	0	1	0
Guy	1	1	0	0	1	0	1	1	1	0	1	0	1	0	0	1
Hann	0	0	0	0	1	1	0	0	0	0	0	0	1	0	0	0
Ivan	0	1	0	0	1	0	1	1	0	1	1	1	0	1	1	0
Jon	1	0	1	0	0	1	1	0	0	0	0	0	0	1	0	1
Ken	0	1	0	1	1	0	0	1	0	1	0	1	0	1	1	0
Lou	1	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0
Mike	1	1	1	0	1	1	1	1	0	0	1	0	1	0	1	0
Noa	0	1	1	0	0	0	0	1	0	0	0	1	0	0	1	0
Omer	0	1	0	1	0	1	1	0	1	0	1	0	1	0	0	1

Database

- We can think about a database as a list of records from some universe set:

$$D \in \mathcal{X}^n$$

- Sometimes we will think to them as functions

$$D(k) \in \mathcal{X}$$

- and sometimes we will write elements explicitly

$$(d_1, \ldots, d_n) \in \mathcal{X}^n$$

Counting Queries

- A counting query $q : \mathcal{X}^n \rightarrow [0,1]$ is a function counting the fraction of people in a dataset satisfying the predicate $q : \mathcal{X} \rightarrow \{0,1\}$
- In symbols:

$$q(D) = \frac{1}{n} \sum_{i=1}^{n} q(d_i)$$

- Notice that we take a normalized count, which also corresponds to the average.

Example 1

Let's consider an arbitrary universe domain $\mathcal{X}$ and let's consider the following predicate for $y \in \mathcal{X}$

$$q_y(x) = \begin{cases} 1 & \text{if } y = x \\ 0 & \text{otherwise} \end{cases}$$

we call a point function the associated counting query

$$q_y : \mathcal{X}^n \to [0, 1]$$

Example 1

Let's consider an arbitrary universe domain $\mathcal{X}$ and let's consider the following predicate for $y \in \mathcal{X}$

$$q_y(x) = \begin{cases} 1 & \text{if } y = x \\ 0 & \text{otherwise} \end{cases}$$

we call a point function the associated counting query

$$q_y : \mathcal{X}^n \to [0, 1]$$

Question: Suppose that we answer all the point function queries for $y \in \mathcal{X}$. What well know data summary do we obtain?

Example 1

$X=\{0,1\}^3$ $\quad D \in X^{10} =$

	D1	D2	D3
I1	0	0	0
I2	1	0	1
I3	0	1	0
I4	1	0	1
I5	0	0	0
I6	0	0	1
I7	1	1	0
I8	0	0	0
I9	0	1	0
I10	1	0	1

Example 1

$X = \{0,1\}^3$

$D \in X^{10} =$

	D1	D2	D3
I1	0	0	0
I2	1	0	1
I3	0	1	0
I4	1	0	1
I5	0	0	0
I6	0	0	1
I7	1	1	0
I8	0	0	0
I9	0	1	0
I10	1	0	1

$q_{000}(D) = .3$

Example 1

$X=\{0,1\}^3$

$D \in X^{10} =$

	D1	D2	D3
I1	0	0	0
I2	1	0	1
I3	0	1	0
I4	1	0	1
I5	0	0	0
I6	0	0	1
I7	1	1	0
I8	0	0	0
I9	0	1	0
I10	1	0	1

$q_{000}(D) = .3$

$q_{001}(D) = .1$

Example 1

$X=\{0,1\}^3$ $\qquad D \in X^{10} =$

	D1	D2	D3
I1	0	0	0
I2	1	0	1
I3	0	1	0
I4	1	0	1
I5	0	0	0
I6	0	0	1
I7	1	1	0
I8	0	0	0
I9	0	1	0
I10	1	0	1

$q_{000}(D) = .3$

$q_{001}(D) = .1$

$q_{010}(D) = .2$

Example 1

$X=\{0,1\}^3$ $\quad D \in X^{10} =$

	D1	D2	D3
I1	0	0	0
I2	1	0	1
I3	0	1	0
I4	1	0	1
I5	0	0	0
I6	0	0	1
I7	1	1	0
I8	0	0	0
I9	0	1	0
I10	1	0	1

$q_{000}(D) = .3$

$q_{001}(D) = .1$

$q_{010}(D) = .2$

$q_{011}(D) = 0$

Example 1

$X=\{0,1\}^3$ $\quad D \in X^{10} =$

	D1	D2	D3
I1	0	0	0
I2	1	0	1
I3	0	1	0
I4	1	0	1
I5	0	0	0
I6	0	0	1
I7	1	1	0
I8	0	0	0
I9	0	1	0
I10	1	0	1

$q_{000}(D) = .3$ $\quad q_{100}(D) = 0$

$q_{001}(D) = .1$

$q_{010}(D) = .2$

$q_{011}(D) = 0$

Example 1

$X=\{0,1\}^3$ $D \in X^{10} =$

	D1	D2	D3
I1	0	0	0
I2	1	0	1
I3	0	1	0
I4	1	0	1
I5	0	0	0
I6	0	0	1
I7	1	1	0
I8	0	0	0
I9	0	1	0
I10	1	0	1

$q_{000}(D) = .3$ $q_{100}(D) = 0$

$q_{001}(D) = .1$ $q_{101}(D) = .3$

$q_{010}(D) = .2$

$q_{011}(D) = 0$

Example I

$X = \{0,1\}^3$ $D \in X^{10} =$

	D1	D2	D3
I1	0	0	0
I2	1	0	1
I3	0	1	0
I4	1	0	1
I5	0	0	0
I6	0	0	1
I7	1	1	0
I8	0	0	0
I9	0	1	0
I10	1	0	1

$q_{000}(D) = .3$ $q_{100}(D) = 0$

$q_{001}(D) = .1$ $q_{101}(D) = .3$

$q_{010}(D) = .2$ $q_{110}(D) = .1$

$q_{011}(D) = 0$

Example 1

$X=\{0,1\}^3$ $\qquad D \in X^{10} =$

	D1	D2	D3
I1	0	0	0
I2	1	0	1
I3	0	1	0
I4	1	0	1
I5	0	0	0
I6	0	0	1
I7	1	1	0
I8	0	0	0
I9	0	1	0
I10	1	0	1

$q_{000}(D) = .3 \qquad q_{100}(D) = 0$

$q_{001}(D) = .1 \qquad q_{101}(D) = .3$

$q_{010}(D) = .2 \qquad q_{110}(D) = .1$

$q_{011}(D) = 0 \qquad q_{111}(D) = 0$

Example 1

$X=\{0,1\}^3$ $\quad D \in X^{10} =$

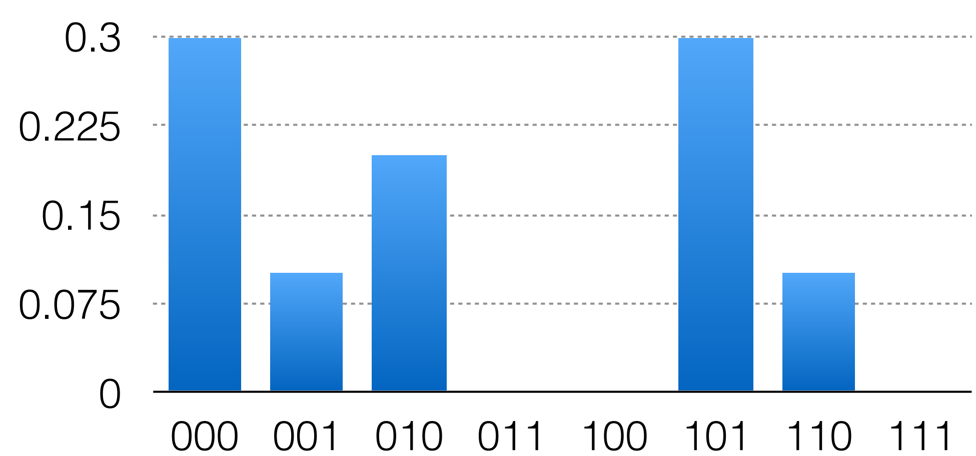
	D1	D2	D3
I1	0	0	0
I2	1	0	1
I3	0	1	0
I4	1	0	1
I5	0	0	0
I6	0	0	1
I7	1	1	0
I8	0	0	0
I9	0	1	0
I10	1	0	1

$q_{000}(D) = .3$ $\quad q_{100}(D) = 0$

$q_{001}(D) = .1$ $\quad q_{101}(D) = .3$

$q_{010}(D) = .2$ $\quad q_{110}(D) = .1$

$q_{011}(D) = 0$ $\quad q_{111}(D) = 0$



Example II

Let's consider an arbitrary ordered universe domain $\mathcal{X}$ and let's consider the following predicate for $y \in \mathcal{X}$

$$q_y(x) = \begin{cases} 1 & \text{if } x \leq y \\ 0 & \text{otherwise} \end{cases}$$

we call a threshold function the associated counting query

$$q_y : \mathcal{X}^n \to [0, 1]$$

Example II

Let's consider an arbitrary ordered universe domain $\mathcal{X}$ and let's consider the following predicate for $y \in \mathcal{X}$

$$q_y(x) = \begin{cases} 1 & \text{if } x \leq y \\ 0 & \text{otherwise} \end{cases}$$

we call a threshold function the associated counting query

$$q_y : \mathcal{X}^n \to [0,1]$$

Question: Suppose that we answer all the threshold function queries for $y \in \mathcal{X}$. What well know statistics do we obtain?

Example II

$X=\{0,1\}^3$
with order
given by the
corresponding
binary encoding.

$D \in X^{10} =$

	D1	D2	D3
I1	0	0	0
I2	1	0	1
I3	0	1	0
I4	1	0	1
I5	0	0	0
I6	0	0	1
I7	1	1	0
I8	0	0	0
I9	0	1	0
I10	1	0	1

Example II

$X = \{0,1\}^3$
with order given by the corresponding binary encoding.

$D \in X^{10} =$

	D1	D2	D3
I1	0	0	0
I2	1	0	1
I3	0	1	0
I4	1	0	1
I5	0	0	0
I6	0	0	1
I7	1	1	0
I8	0	0	0
I9	0	1	0
I10	1	0	1

$q_{000}(D) = .3$ $\quad q_{100}(D) = .6$

$q_{001}(D) = .4$ $\quad q_{101}(D) = .9$

$q_{010}(D) = .6$ $\quad q_{110}(D) = 1$

$q_{011}(D) = .6$ $\quad q_{111}(D) = 1$

Example II

$X=\{0,1\}^3$ with order given by the corresponding binary encoding.

$D \in X^{10} =$

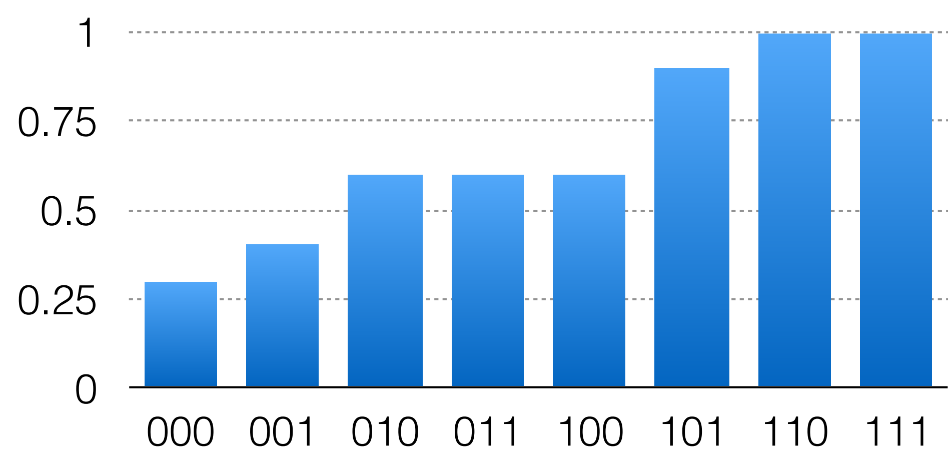
	D1	D2	D3
I1	0	0	0
I2	1	0	1
I3	0	1	0
I4	1	0	1
I5	0	0	0
I6	0	0	1
I7	1	1	0
I8	0	0	0
I9	0	1	0
I10	1	0	1

$q_{000}(D) = .3$ $\quad q_{100}(D) = .6$

$q_{001}(D) = .4$ $\quad q_{101}(D) = .9$

$q_{010}(D) = .6$ $\quad q_{110}(D) = 1$

$q_{011}(D) = .6$ $\quad q_{111}(D) = 1$



Example III

Let's consider the universe domain $\mathcal{X} = \{0,1\}^d$ and let's consider $\vec{v} \in \{1, \bar{1}, \dots, d, \vec{d}\}^k$ with $1 \leq k \leq d$ and

$$q_{\vec{v}}(x) = q_{v_1}(x) \wedge q_{v_1}(x) \wedge \cdots \wedge q_{v_k}(x)$$

where $q_j(x) = x_j$ and $q_{\bar{j}}(x) = \neg x_j$

We call a conjunction or k-way marginal the associated counting query

$$q_{\vec{v}} : \mathcal{X}^n \to [0,1]$$

Example III

Let's consider the universe domain $\mathcal{X} = \{0,1\}^d$ and let's consider $\vec{v} \in \{1, \bar{1}, \ldots, d, \vec{d}\}^k$ with $1 \leq k \leq d$ and

$$q_{\vec{v}}(x) = q_{v_1}(x) \wedge q_{v_1}(x) \wedge \cdots \wedge q_{v_k}(x)$$

where $q_j(x) = x_j$ and $q_{\bar{j}}(x) = \neg x_j$

We call a conjunction or k-way marginal the associated counting query

$$q_{\vec{v}} : \mathcal{X}^n \to [0,1]$$

Question: Which statistics does correspond to releasing conjunctions?

Example III

$X=\{0,1\}^3$

$D \in X^{10} =$

	D1	D2	D3
I1	0	0	0
I2	1	0	1
I3	0	1	0
I4	1	0	1
I5	0	0	0
I6	0	0	1
I7	1	1	0
I8	0	0	0
I9	0	1	0
I10	1	0	1

$k=2$

Example III

$X=\{0,1\}^3$

$D \in X^{10} =$

	D1	D2	D3
I1	0	0	0
I2	1	0	1
I3	0	1	0
I4	1	0	1
I5	0	0	0
I6	0	0	1
I7	1	1	0
I8	0	0	0
I9	0	1	0
I10	1	0	1

$k=2$

$q_{12}(D) = .1$

$q_{1/2}(D) = .3$

$q_{13}(D) = .3$

$q_{1/3}(D) = .1$

$q_{/12}(D) = .2$

$q_{/13}(D) = .1$

$q_{/1/2}(D) = .4$

$q_{/1/3}(D) = .5$

Example III

$X = \{0,1\}^3$

$D \in X^{10} =$

	D1	D2	D3
I1	0	0	0
I2	1	0	1
I3	0	1	0
I4	1	0	1
I5	0	0	0
I6	0	0	1
I7	1	1	0
I8	0	0	0
I9	0	1	0
I10	1	0	1

$k=2$

$q_{12}(D) = .1$

$q_{1/2}(D) = .3$

$q_{13}(D) = .3$

$q_{1/3}(D) = .1$

$q_{/12}(D) = .2$

$q_{/13}(D) = .1$

$q_{/1/2}(D) = .4$

$q_{/1/3}(D) = .5$

	D1	/D1
D2	0.1	0.2
/D2	0.3	0.4

Linear Queries

- A linear query $q : \mathcal{X}^n \to [0,1]$ is a function averaging the value of a function $q : \mathcal{X} \to [0,1]$ over the elements of the dataset.
- In symbols:

$$q(D) = \frac{1}{n} \sum_{i=1}^{n} q(d_i)$$

Sum queries

- Let's denote by $I \subseteq [n]$ a subset I of

$$\{0, \ldots, n\}$$

- A sum query $q_I : \{0,1\}^k \to \mathbb{N}^k$ is defined as

$$q_I(D) = \sum_{i \in I} d_i$$

Example

$X = \{0, 1\}^3$

$D \in X^{10} =$

	D1	D2	D3
I1	0	0	0
I2	1	0	1
I3	0	1	0
I4	1	0	1
I5	0	0	0
I6	0	0	1
I7	1	1	0
I8	0	0	0
I9	0	1	0
I10	1	0	1

Example

$X=\{0,1\}^3$

$D \in X^{10} =$

	D1	D2	D3
I1	0	0	0
I2	1	0	1
I3	0	1	0
I4	1	0	1
I5	0	0	0
I6	0	0	1
I7	1	1	0
I8	0	0	0
I9	0	1	0
I10	1	0	1

$q_{\{1,2,3\}}(D) = (1,1,1)$

Example

$X=\{0,1\}^3$ $\quad D \in X^{10} =$

	D1	D2	D3
I1	0	0	0
I2	1	0	1
I3	0	1	0
I4	1	0	1
I5	0	0	0
I6	0	0	1
I7	1	1	0
I8	0	0	0
I9	0	1	0
I10	1	0	1

$q_{\{1,2,3\}}(D) = (1,1,1)$

$q_{\{1,2,4\}}(D) = (2,0,2)$

Example

$X=\{0,1\}^3$ $D \in X^{10} =$

	D1	D2	D3
I1	0	0	0
I2	1	0	1
I3	0	1	0
I4	1	0	1
I5	0	0	0
I6	0	0	1
I7	1	1	0
I8	0	0	0
I9	0	1	0
I10	1	0	1

$$q_{\{1,2,3\}}(D) = (1,1,1)$$
$$q_{\{1,2,4\}}(D) = (2,0,2)$$
$$q_{\{5,8\}}(D) = (0,0,0)$$

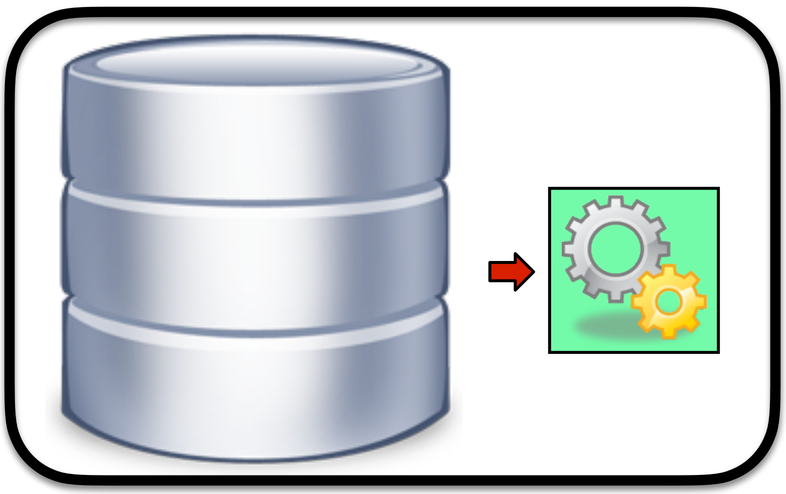
Example

$X=\{0,1\}^3$ $\quad D \in X^{10} =$

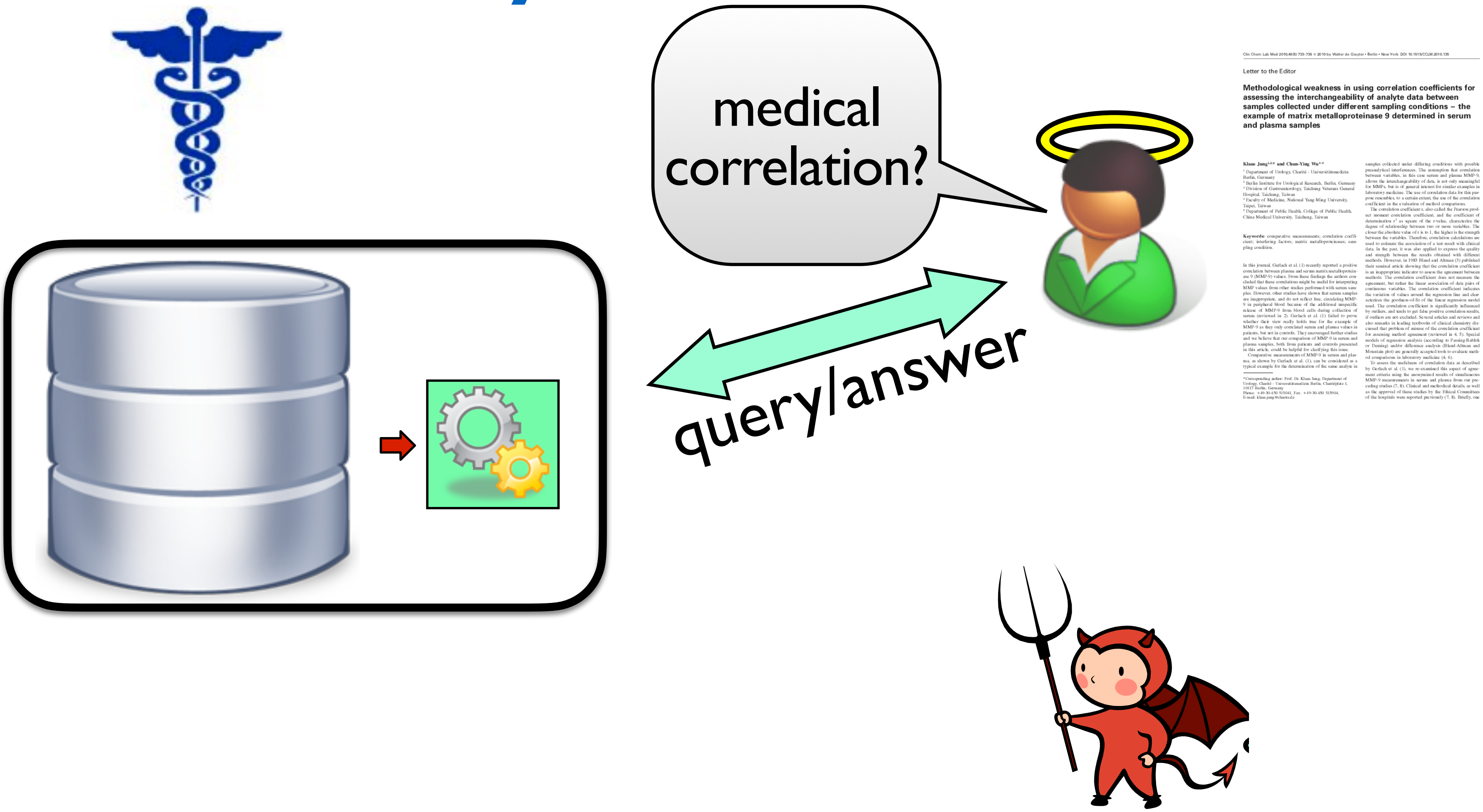
	D1	D2	D3
I1	0	0	0
I2	1	0	1
I3	0	1	0
I4	1	0	1
I5	0	0	0
I6	0	0	1
I7	1	1	0
I8	0	0	0
I9	0	1	0
I10	1	0	1

$$q_{\{1,2,3\}}(D) = (1,1,1)$$
$$q_{\{1,2,4\}}(D) = (2,0,2)$$
$$q_{\{5,8\}}(D) = (0,0,0)$$
$$q_{\{2,4,7,10\}}(D) = (4,1,3)$$

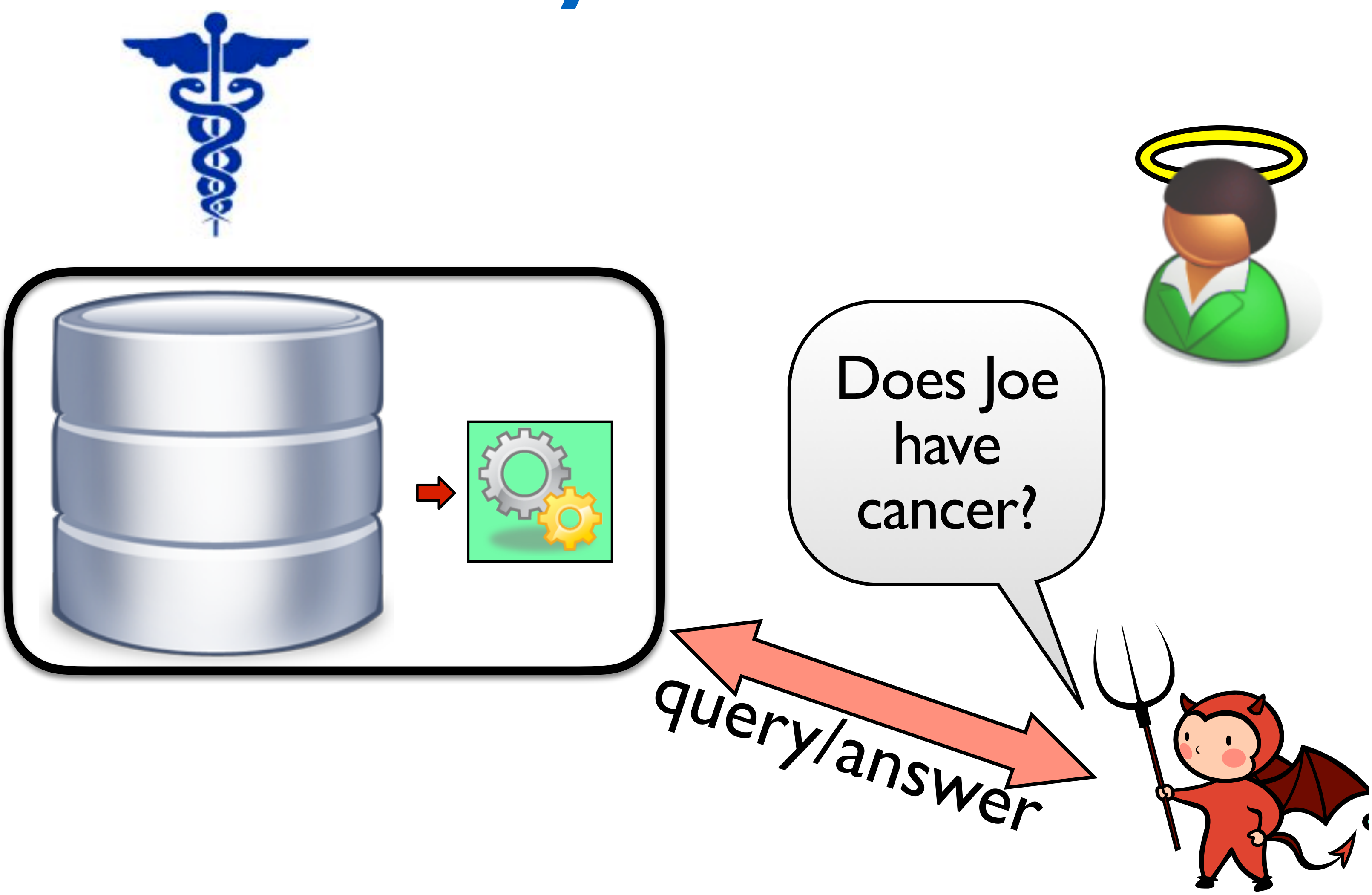
Privacy Mechanism



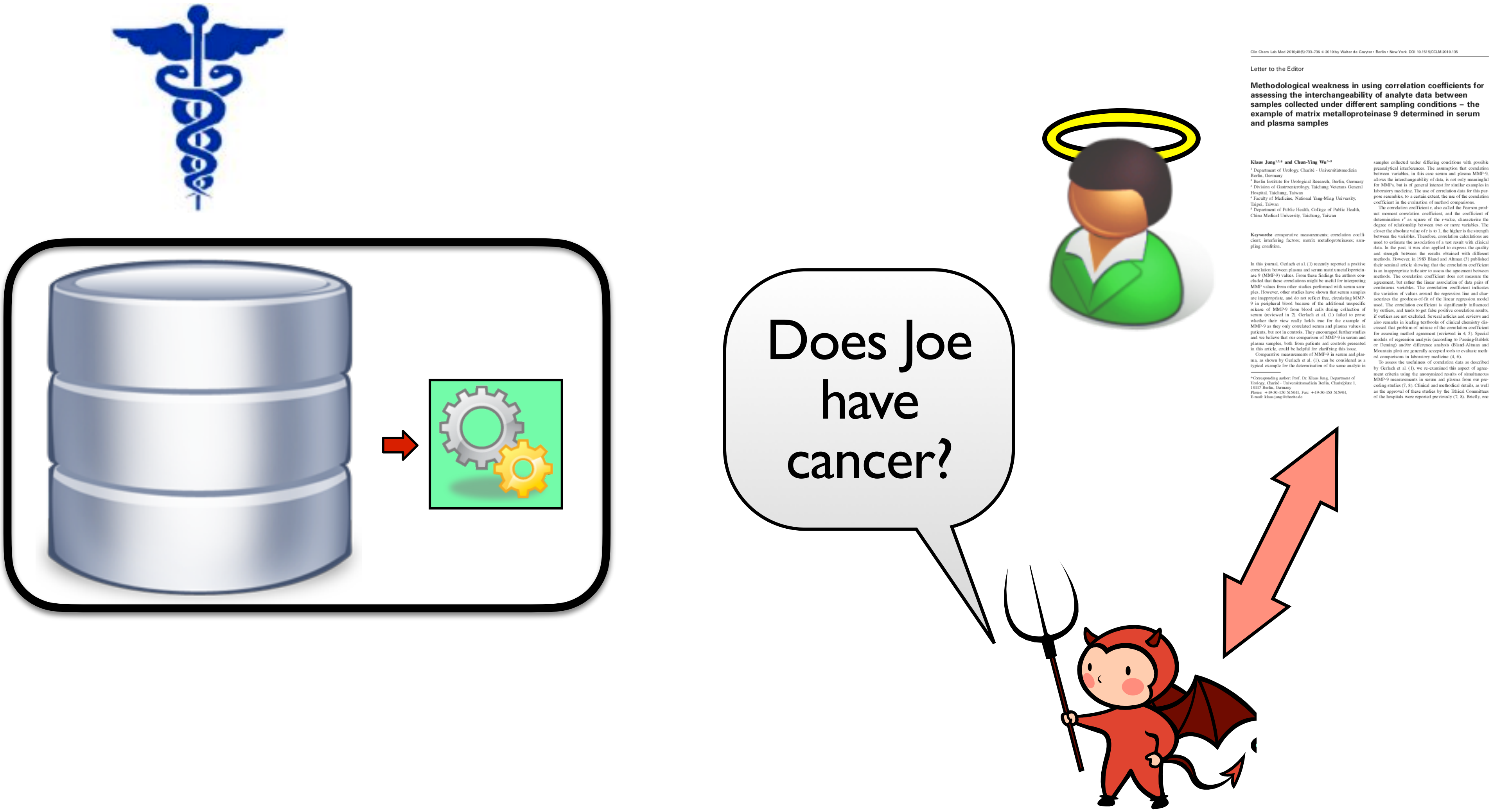
Privacy Mechanism



Privacy Mechanism

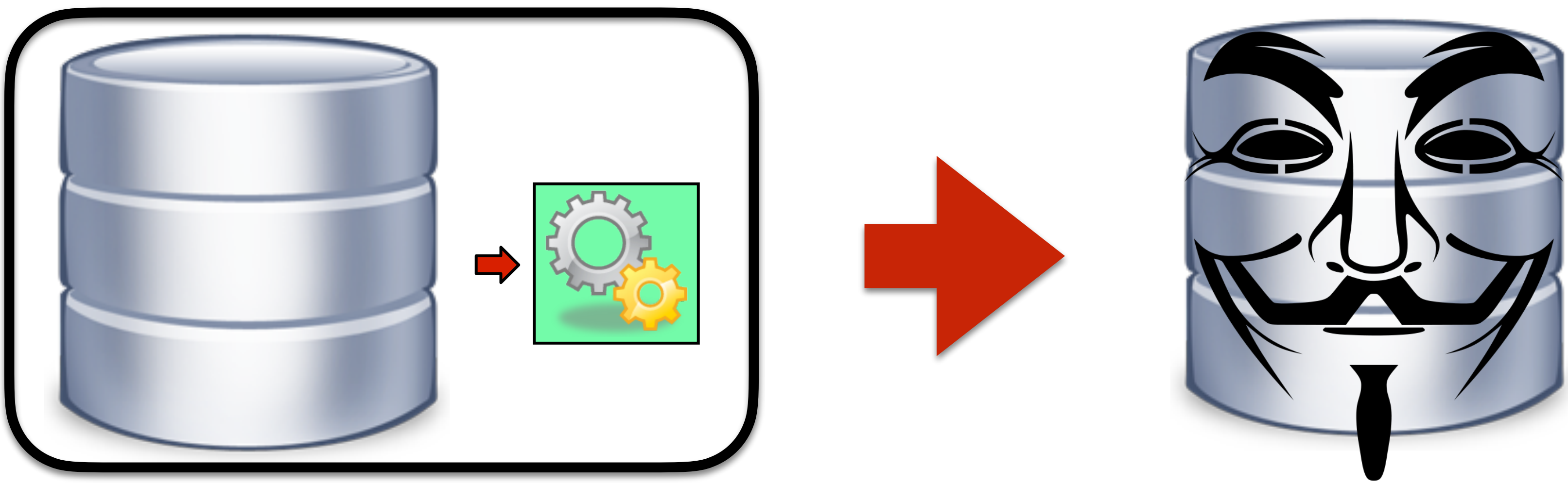


Private Mechanism



What can be a good privacy mechanism?

A natural idea: anonymizing the data



- E.g. stripping PII, guaranteeing k-anonymity, swapping, etc.

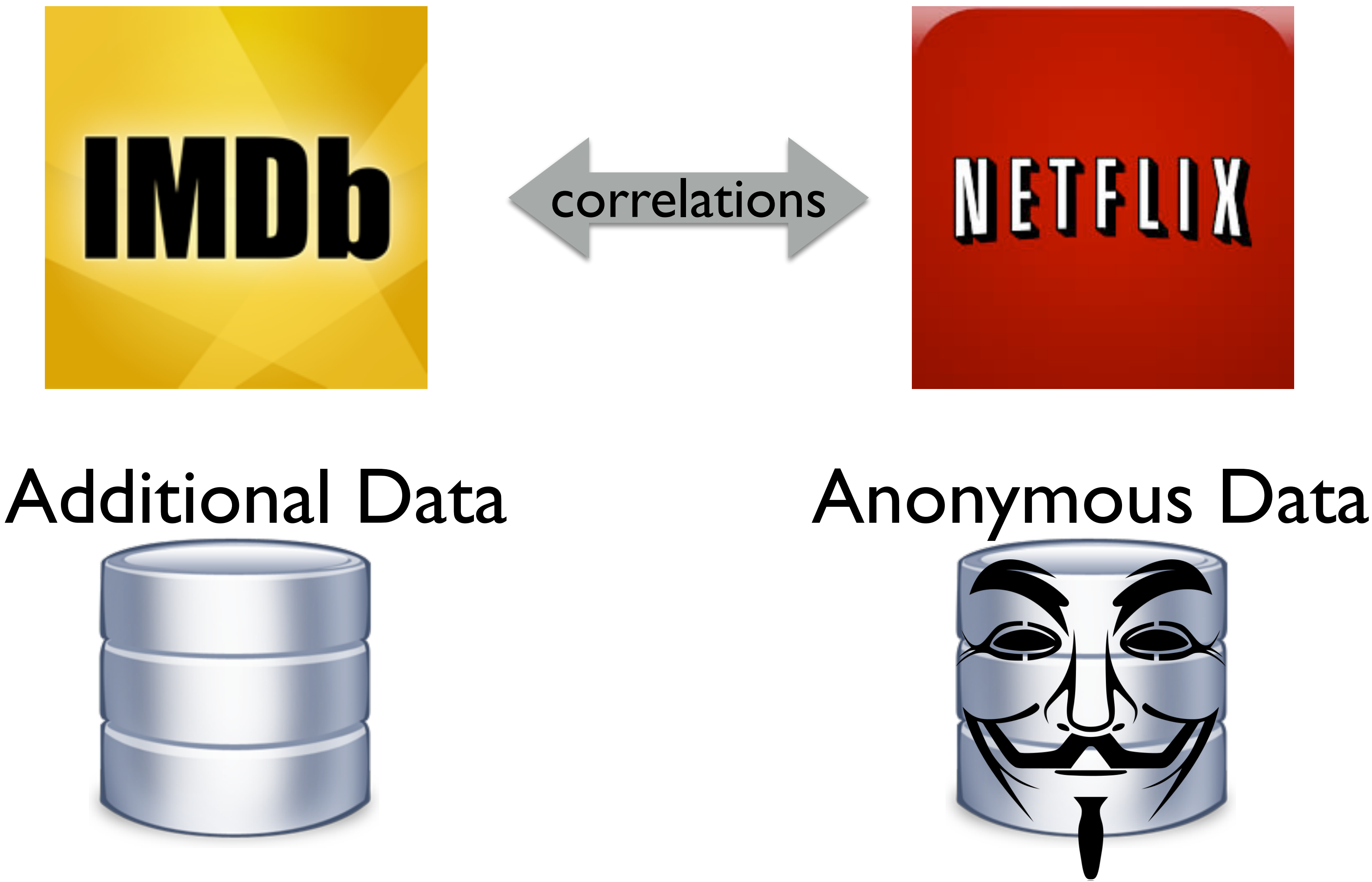
Attacks on stripping PII

(Narayanan, Shmatikov: Robust De-anonymization of Large Sparse Datasets. IEEE Symposium on Security and Privacy 2008)

Anonymous Data

Attacks on stripping PII

(Narayanan, Shmatikov: Robust De-anonymization of Large Sparse Datasets. IEEE Symposium on Security and Privacy 2008)



Attacks on Swapping

(Garfinkel, Abowd, Martindale: Understanding Database Reconstruction Attacks on Public Data. ACM Queue 16(5): 50 (2018))

United States® Census Bureau

Anonymous Data

Attacks on Swapping

(Garfinkel, Abowd, Martindale: Understanding Database Reconstruction Attacks on Public Data. ACM Queue 16(5): 50 (2018))

Commercially available data

correlations

United States® Census Bureau

Additional Data

Anonymous Data

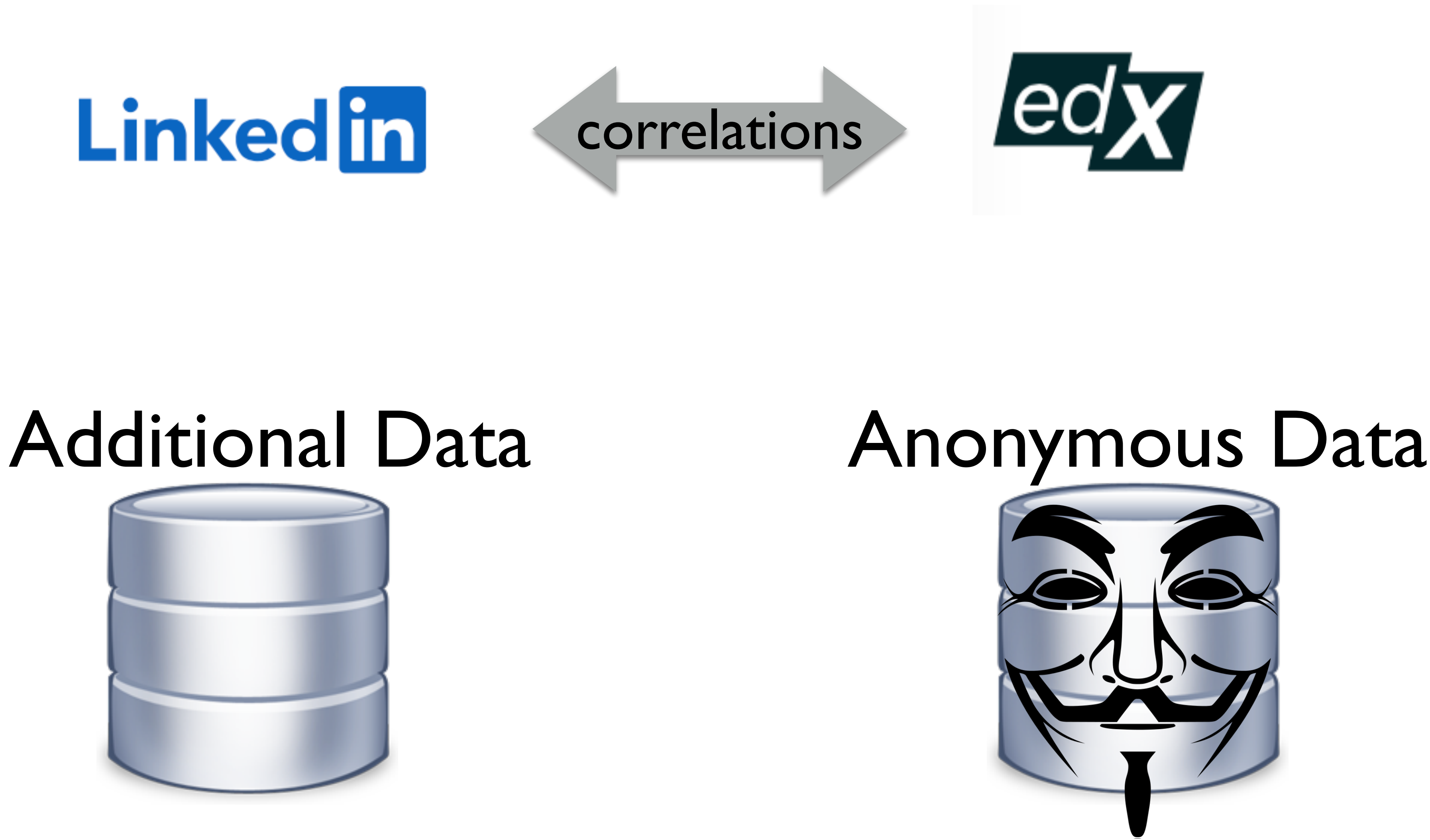
Attacks on K-anonymity

(A. Cohen: Attacks on Deidentification's Defenses. Usenix Security 2022)

Anonymous Data

Attacks on K-anonymity

(A. Cohen: Attacks on Deidentification's Defenses. Usenix Security 2022)

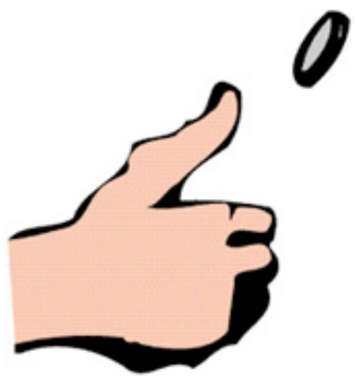


Why these anonymization techniques runs into troubles?

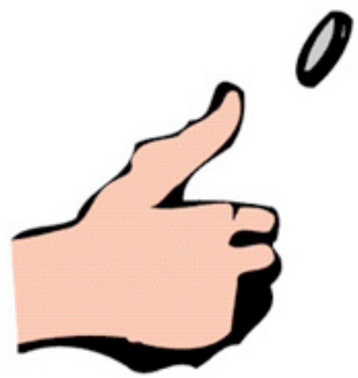
An issue that these attacks highlight is that it is difficult if not impossible to think about privacy as a property of the data.
Another issue with these anonymization notions is that they are not closed under postprocessing.

Question: How can we guarantee closure under postprocessing?

Question: How can we guarantee closure under postprocessing?

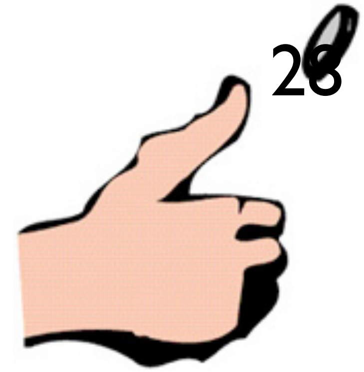


Question: How can we guarantee closure under postprocessing?



Warner, S. L. (March 1965). "Randomised response: a survey technique for eliminating evasive answer bias". Journal of the American Statistical Association. Taylor & Francis. 60 (309): 63–69.

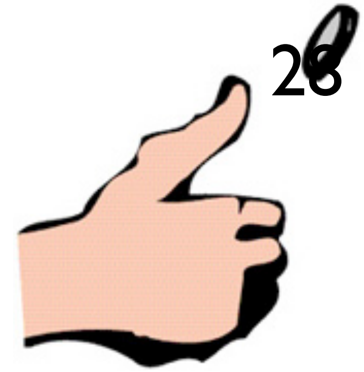
Randomized Algorithms



- Given a discrete set B the probability simplex over B, denoted $\Delta(B)$ is defined as:

$$\Delta(B) = \left\{ x \in \mathbb{R}^{|B|} : \forall i, x_i \geq 0, \text{ and } \sum_{i=1}^{|B|} x_i = 1 \right\}$$

Randomized Algorithms

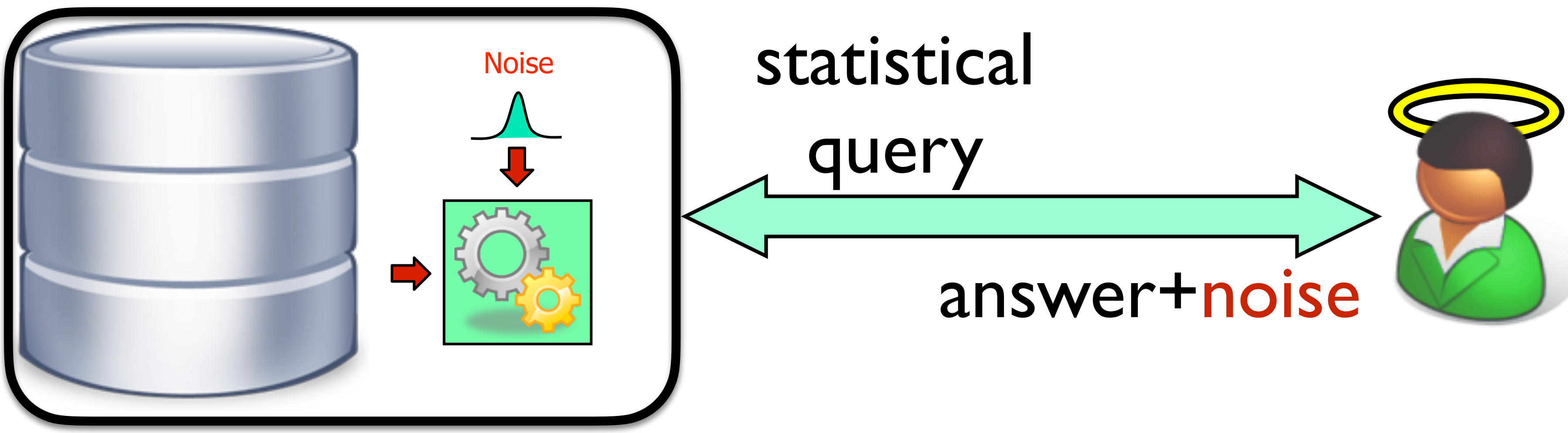


- Given a discrete set B the probability simplex over B, denoted $\Delta(B)$ is defined as:

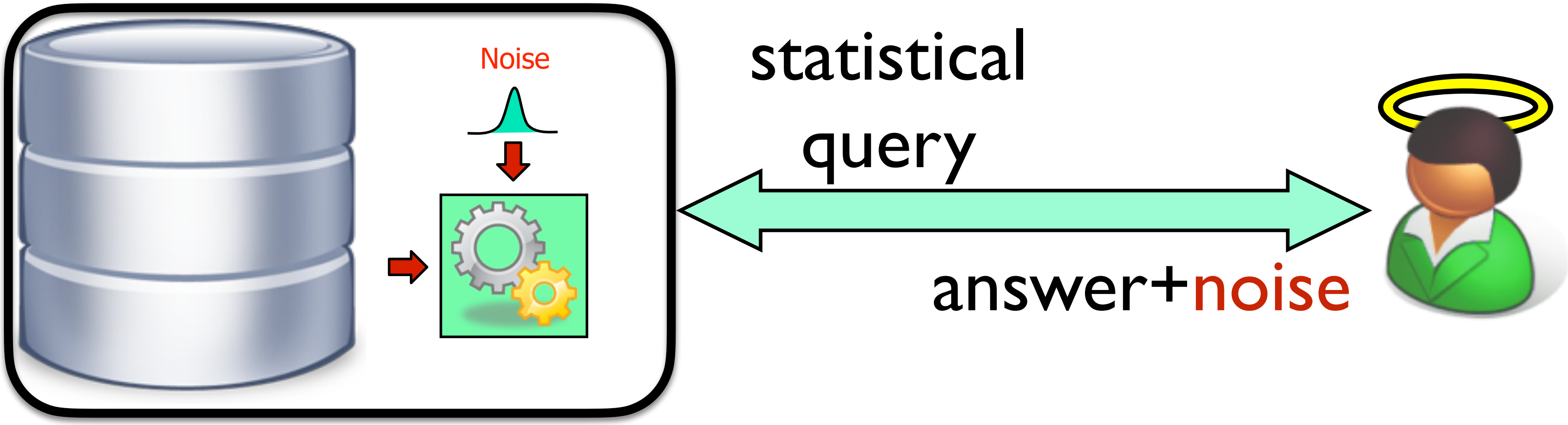
$$\Delta(B) = \left\{ x \in \mathbb{R}^{|B|} : \forall i, x_i \geq 0, \text{ and } \sum_{i=1}^{|B|} x_i = 1 \right\}$$

- A randomized algorithm $\mathcal{M}$ is an algorithm associated with a total map $M : A \to \Delta(B)$
On input $a \in A$ the algorithm outputs $\mathcal{M}(a) = b$ with probability $(M(a))_b$.
The probability space is over the coin flips of the algorithm.

Private Statistical database



Private Statistical database



Question: What kind of noise?

Sum queries

- Let's denote by $I \subseteq [n]$ a subset I of

$$\{0, \ldots, n\}$$

- A sum query $q_I : 0, 1^k \rightarrow \mathbb{N}^k$ is defined as

$$q_I(D) = \sum_{i \in I} d_i$$

Uniform Noise

- Given a query q we want to add noise to create a new randomized query:

$$q^*(D) = q(D) + Y$$

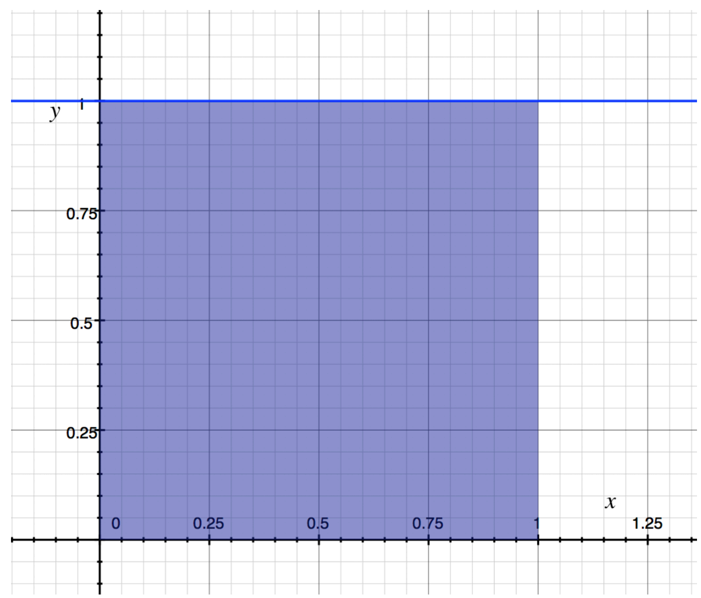
Uniform Noise

- Given a query q we want to add noise to create a new randomized query:

$$q^*(D) = q(D) + Y$$

- One way to do this is to sample Y from the uniform distribution:

$$Y \sim U[0,1]$$

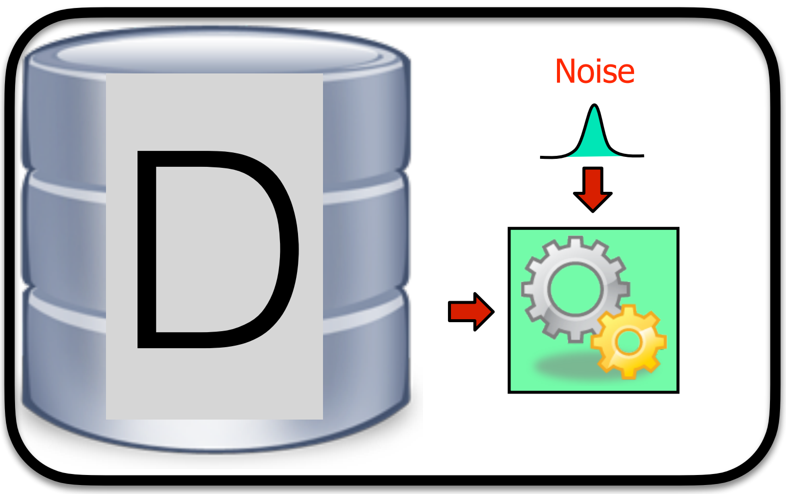


Question: Does this approach prevent privacy attacks?

Reconstruction attack

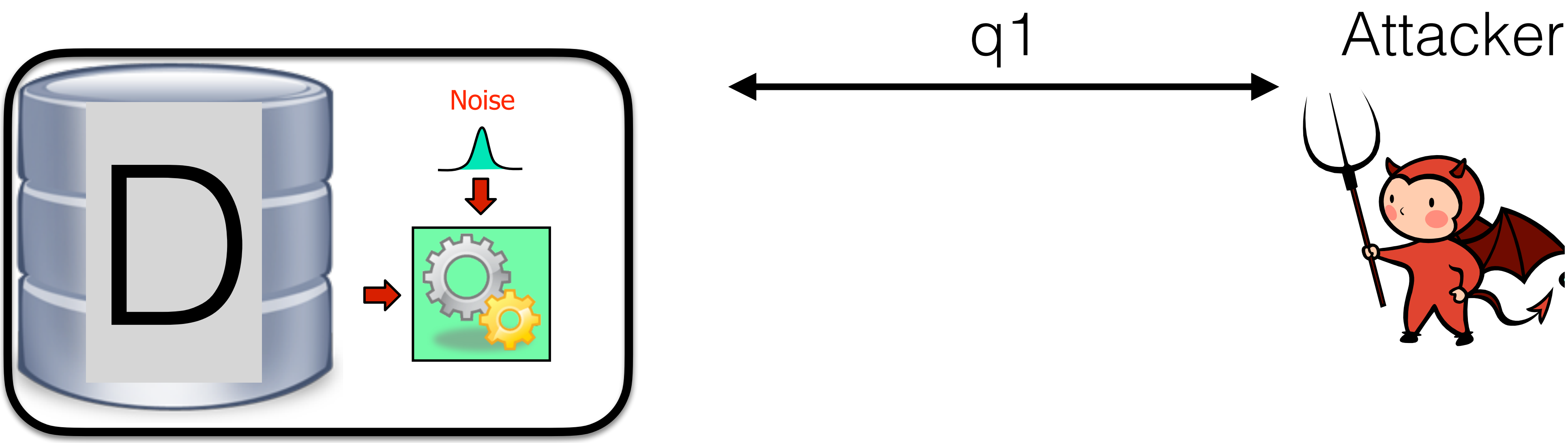
- Consider an adversary A (an algorithm) that has access to some data D through a privacy mechanism q*.
- The goal of the adversary is to output some data D’ that is as similar as possible to D.
- To output D’ the adversary can interact several times with q*.

Reconstruction attack

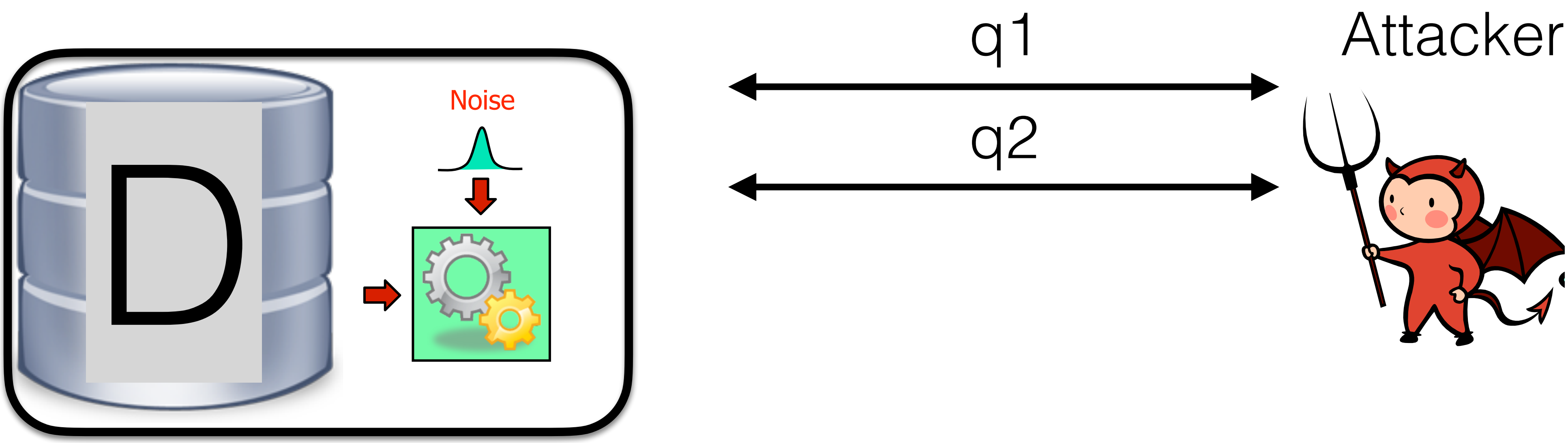


Attacker

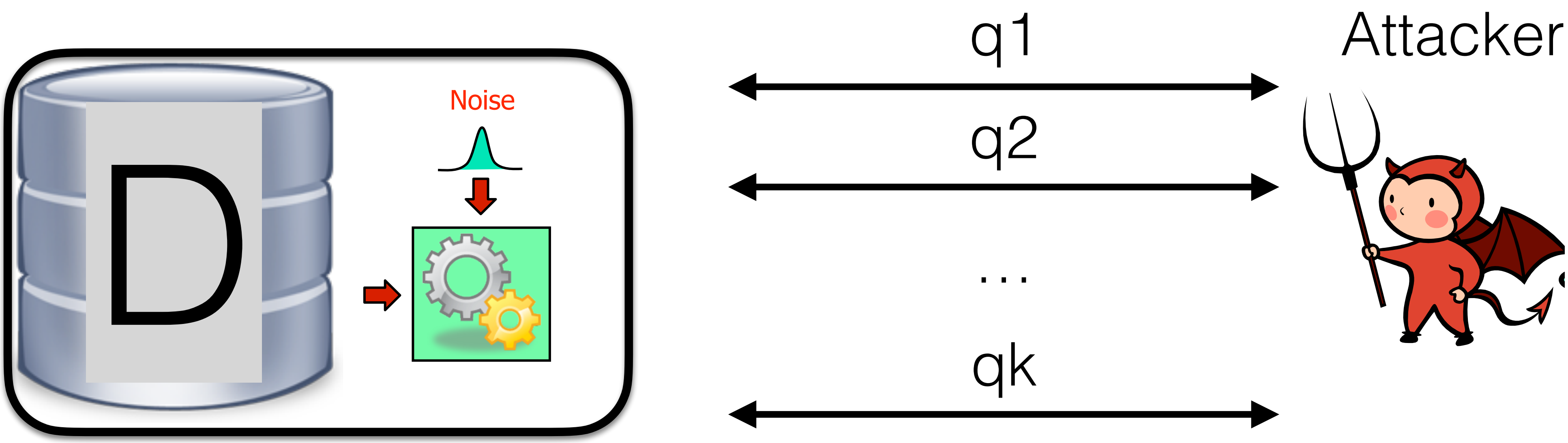
Reconstruction attack



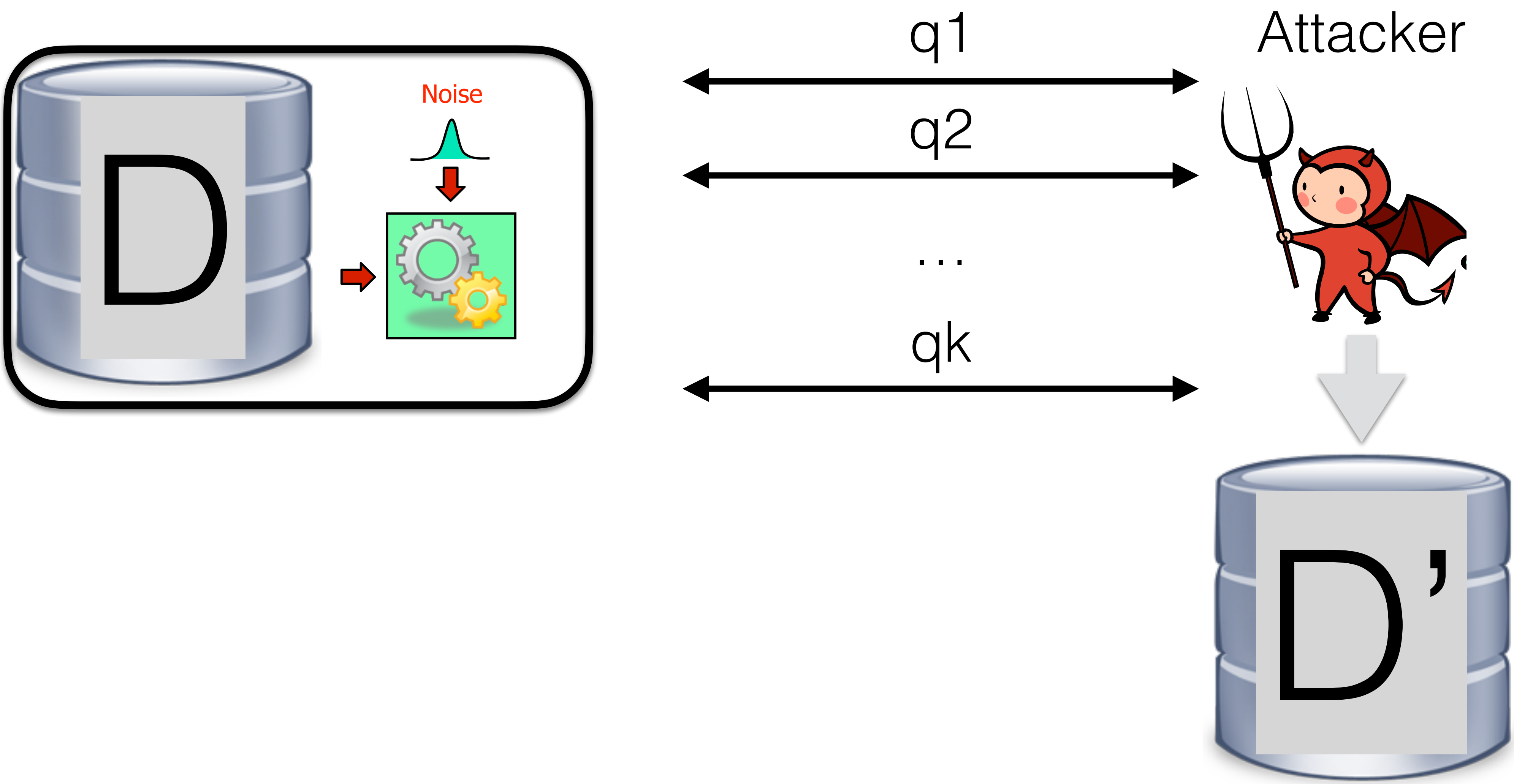
Reconstruction attack



Reconstruction attack



Reconstruction attack



Reconstruction attack

We say that the attacker wins if

$$d(D, D') \sim 0$$

In our case we can use Hamming distance

Additive Noise Perturbation

- We say that M is a privacy mechanism obtained by adding noise if for every query q, M creates a new randomized query:

$$q^*(D) = q(D) + Y$$

Additive Noise Perturbation

- We say that M is a privacy mechanism obtained by adding noise if for every query q, M creates a new randomized query:

$$q^*(D) = q(D) + Y$$

- We say that a mechanism M add noise within perturbation E iff for every q and every D:

$$|q^*(D)-q(D)| \leq E$$

Reconstruction attack with exponential adversary

Let $M:\{0,1\}^n \rightarrow R$ be a privacy mechanism adding noise within E perturbation. Then there is an adversary that can reconstruct the database within 4E positions.

[DinurNissim'02]

Reconstruction attack with exponential adversary

Let $M:\{0,1\}^n \rightarrow R$ be a privacy mechanism adding noise within **E=o(n)** perturbation. Then Then there is an adversary that can reconstruct the database with constant error and running in exponential time.

[DinurNissim'02]

Reconstruction attack with polynomial adversary

Let $M:\{0,1\}^n \rightarrow R$ be a privacy mechanism adding noise within **$E=o(\sqrt{n})$** perturbation. Then there is an adversary that can reconstruct the database with constant error running in polynomial time and **answering n queries.**

[DinurNissim'02, DworkYekhanin'08]

Number of queries

A privacy mechanism can answer with perturbation $\sqrt{n}$ at most a number of queries sublinear in n.

Number of queries

A privacy mechanism can answer with perturbation $\sqrt{n}$ at most a number of queries sublinear in n.

Question: Why error $\sqrt{n}$ is a good reference?

Sample error

- Suppose that a database contains n individuals drawn uniformly at random from a population of size N>>n.
- Suppose we are interested in a medical condition that affects a fraction p of the population.
- Then we expect the number of individuals in the dataset with condition p is

$$np \pm \Theta(\sqrt{n})$$

- The sampling error is of the order of $\sqrt{n}$.

Sample error

- Suppose that a database contains n individuals drawn uniformly at random from a population of size N>>n.
- Suppose we are interested in a medical condition that affects a fraction p of the population.
- Then we expect the number of individuals in the dataset with condition p is

$$np \pm \Theta(\sqrt{n})$$

- The sampling error is of the order of $\sqrt{n}$.

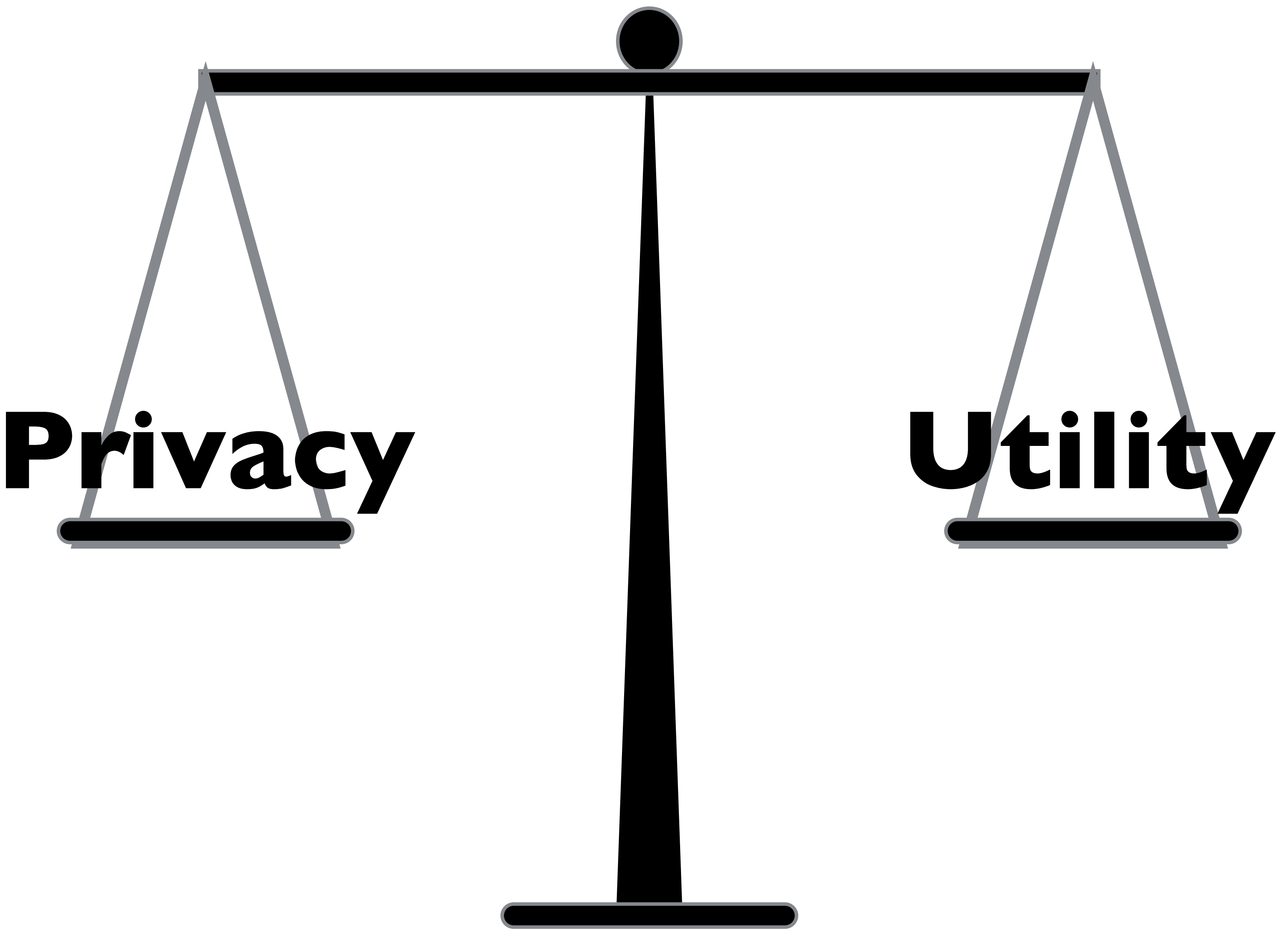
We would like the noise we introduce for privacy to be comparable to the sampling error.

Foundamental Law of Information Reconstruction

The release of too many overly accurate statistics gives privacy violations.

[DinurNissim02]

Privacy vs Utility



Quantitative notions of Privacy

- The impossibility results discussed above suggest a quantitative notion of privacy,
- A notion where the privacy loss depends on the number of queries that are allowed.

Quantitative notions of Privacy

- The impossibility results discussed above suggest a quantitative notion of privacy,
- A notion where the privacy loss depends on the number of queries that are allowed.

What can this notion be?

Let’s take inspiration from semantics security.

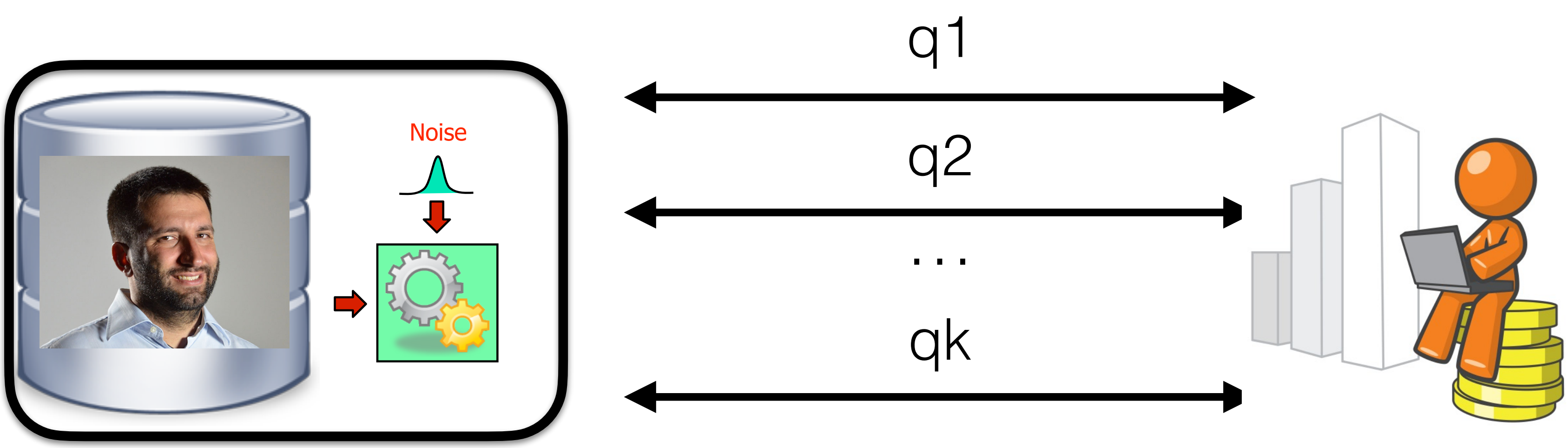
- The analyst learn the same after the analysis as what she would have learnt if I didn’t contribute my data.

Let’s take inspiration from semantics security.

- The analyst learn the same after the analysis as what she would have learnt if I didn’t contribute my data.

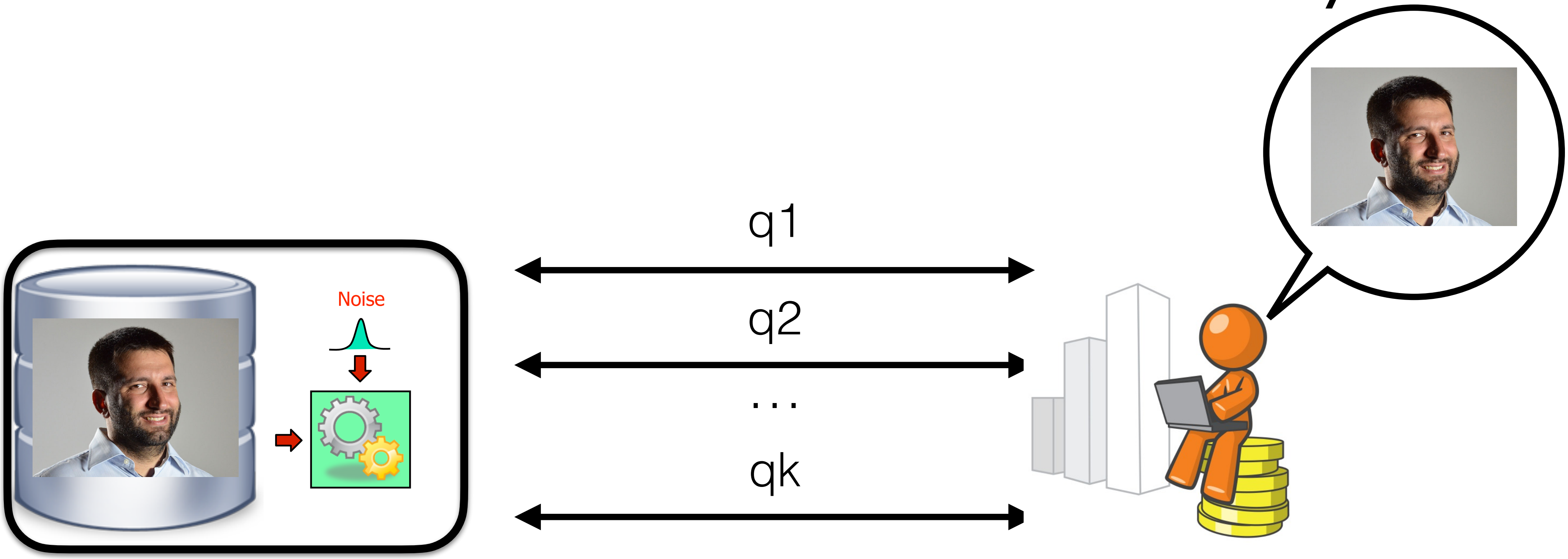
Let's take inspiration from semantics security.

- The analyst learn the same after the analysis as what she would have learnt if I didn't contribute my data.



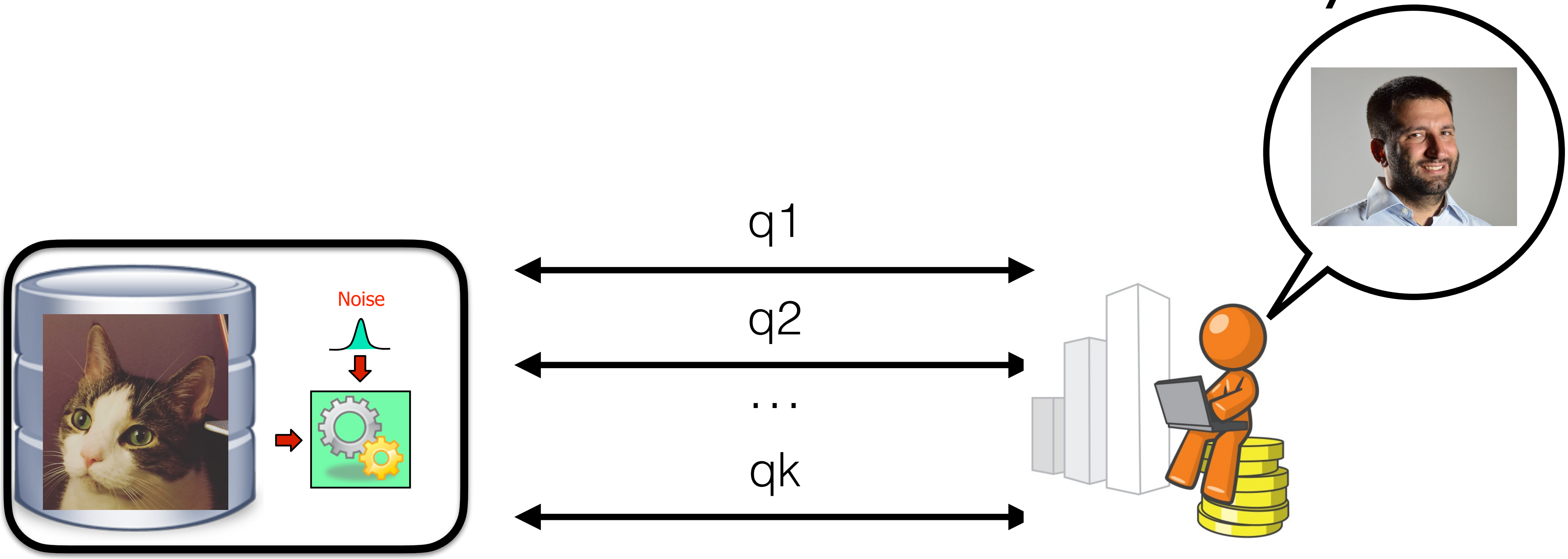
Let's take inspiration from semantics security.

- The analyst learn the same after the analysis as what she would have learnt if I didn't contribute my data.



Let's take inspiration from semantics security.

- The analyst learn the same after the analysis as what she would have learnt if I didn't contribute my data.



Privacy-preserving data analysis?

Prior Knowledge

~

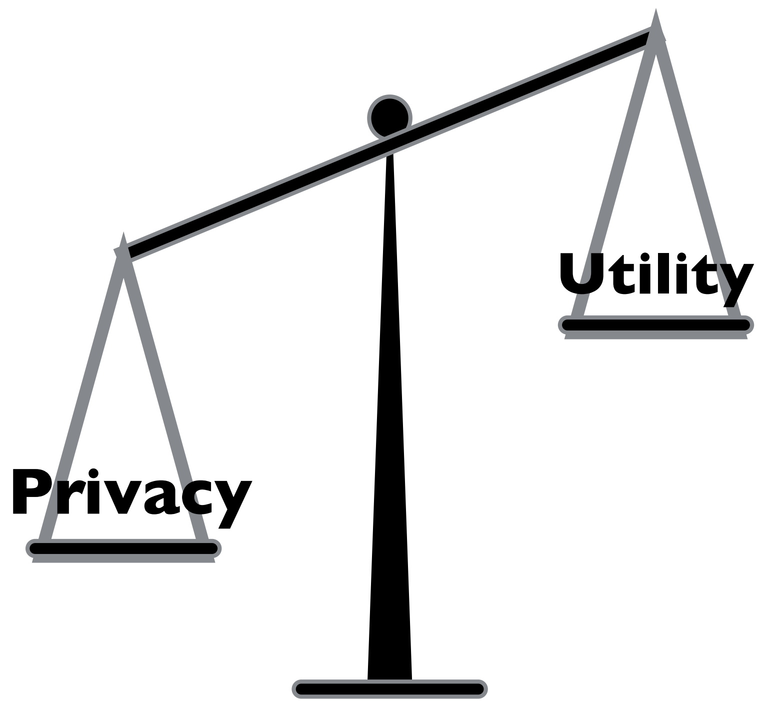
Posterior Knowledge

Privacy-preserving data analysis?

Privacy-preserving data analysis?

Question: What is the problem with this requirement?

Privacy-preserving data analysis?



If nothing can be learned about an individual, then nothing at all can be learned at all!

[DworkNaor10]

Let's take inspiration from semantics security v2.

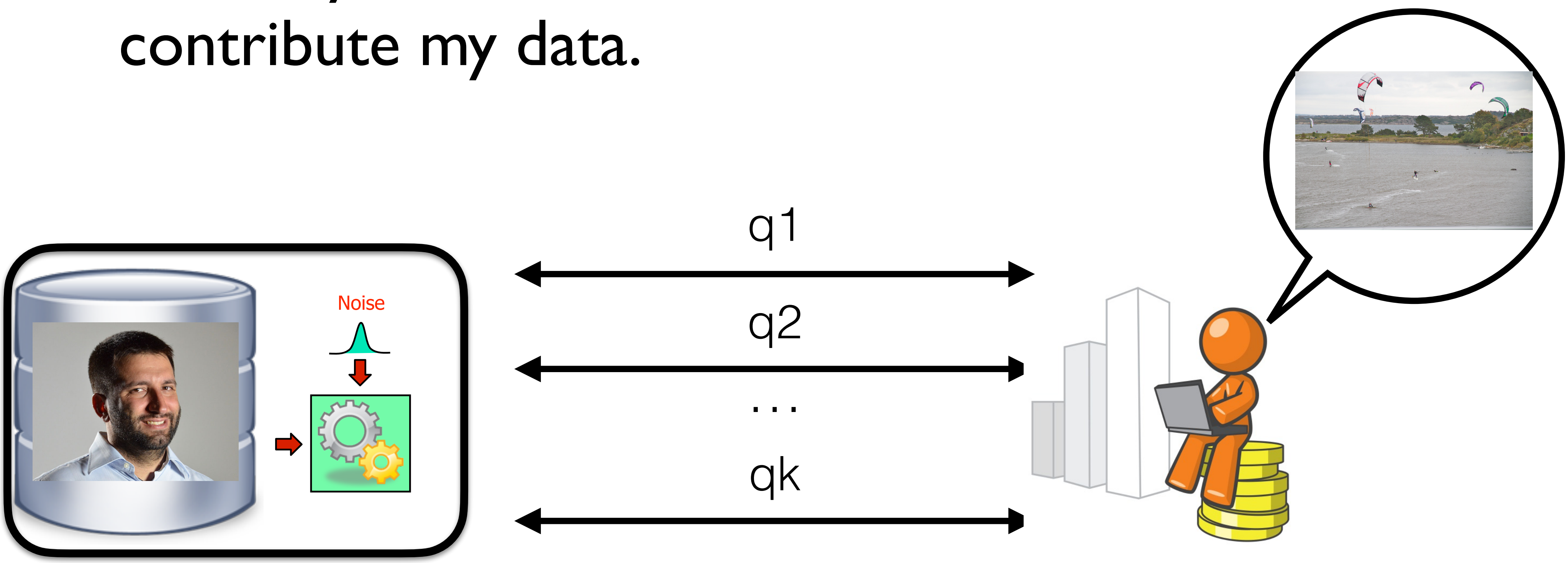
- The analyst learn almost the same about me after the analysis as what she would have learnt if I didn't contribute my data.

Let's take inspiration from semantics security v2.

- The analyst learn almost the same about me after the analysis as what she would have learnt if I didn't contribute my data.

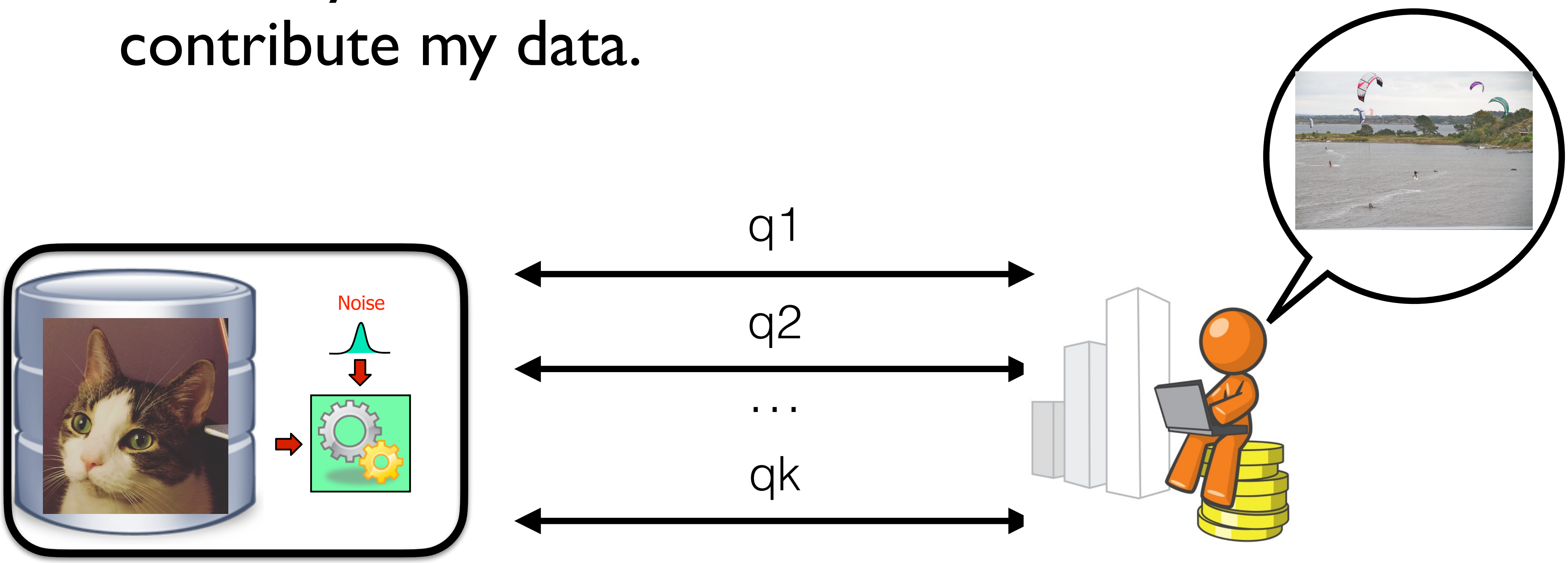
Let's take inspiration from semantics security v2.

- The analyst learn almost the same about me after the analysis as what she would have learnt if I didn't contribute my data.



Let's take inspiration from semantics security v2.

- The analyst learn almost the same about me after the analysis as what she would have learnt if I didn't contribute my data.



Adjacent databases

- We can formalize the concept of contributing my data or not in terms of a notion of distance between datasets.
- Given two datasets D, D'$\in\{0,1\}^n$, their distance is defined as:

$$D\Delta D' = |\{k \leq n \mid D(k) \neq D'(k)\}|$$

- We will call two datasets adjacent when $D\Delta D'=1$ and we will write D~D'.

(ε,δ)-Differential Privacy

Definition

Given $\varepsilon,\delta \geq 0$, a probabilistic query $Q: X^n \rightarrow R$ is (ε,δ)-differentially private iff

for all adjacent database b_1, b_2 and for every $S \subseteq R$:

$$\Pr[Q(b_1) \in S] \leq \exp(\varepsilon)\Pr[Q(b_2) \in S] + \delta$$

(ε,δ)-Differential Privacy

A query returning a probability distribution

Definition

Given $\varepsilon,\delta \geq 0$, a probabilistic query $Q: X^n \rightarrow R$ is (ε,δ)-differentially private iff for all adjacent database b_1, b_2 and for every $S \subseteq R$:

$$\Pr[Q(b_1) \in S] \leq \exp(\varepsilon)\Pr[Q(b_2) \in S] + \delta$$

(ε,δ)-Differential Privacy

Privacy parameters

Definition

Given $\boxed{\varepsilon,\delta} \geq 0$, a probabilistic query $Q: X^n \rightarrow R$ is (ε,δ)-differentially private iff for all adjacent database b_1, b_2 and for every $S \subseteq R$:

$$\Pr[Q(b_1) \in S] \leq \exp(\varepsilon) \Pr[Q(b_2) \in S] + \delta$$

(ε,δ)-Differential Privacy

Definition

Given $\varepsilon, \delta \geq 0$, a probabilistic query $Q: X^n \rightarrow R$ is (ε, δ)-differentially private iff for all adjacent database b_1, b_2 and for every $S \subseteq R$:

$$\Pr[Q(b_1) \in S] \leq \exp(\varepsilon) \Pr[Q(b_2) \in S] + \delta$$

a quantification over all the databases

(ε,δ)-Differential Privacy

Definition

Given $\varepsilon,\delta \geq 0$, a probabilistic query $Q: X^n \rightarrow R$ is (ε,δ)-differentially private iff for all adjacent database b_1, b_2 and for every $S \subseteq R$:

$$\Pr[Q(b_1) \in S] \leq \exp(\varepsilon)\Pr[Q(b_2) \in S] + \delta$$

a notion of adjacency or distance

(ε,δ)-Differential Privacy

Definition

Given $\varepsilon,\delta \geq 0$, a probabilistic query $Q: X^n \rightarrow R$ is (ε,δ)-differentially private iff for all adjacent database b_1, b_2 and for every $S \subseteq R$:

$$\Pr[Q(b_1) \in S] \leq \exp(\varepsilon)\Pr[Q(b_2) \in S] + \delta$$

and over all the possible outcomes

ε-Differential Privacy

Definition

Given $\varepsilon \geq 0$, a probabilistic query $Q: X^n \rightarrow R$ is ε-differentially private iff for all adjacent database b_1, b_2 and for every $S \subseteq R$:

$$\Pr[Q(b_1) \in S] \leq \exp(\varepsilon) \Pr[Q(b_2) \in S]$$

ε-Differential Privacy

Definition

Given $\varepsilon \geq 0$, a probabilistic query $Q: X^n \rightarrow R$ is ε-differentially private iff for all adjacent database b_1, b_2 and for every $S \subseteq R$:

$$\Pr[Q(b_1) \in S] \leq \exp(\varepsilon) \Pr[Q(b_2) \in S]$$

Let's substitute a concrete instance:

$$\Pr[Q(b \cup \{x\}) \in S] \leq \exp(\varepsilon) \Pr[Q(b \cup \{y\}) \in S]$$

ε-Differential Privacy

Definition

Given $\varepsilon \geq 0$, a probabilistic query $Q: X^n \to R$ is ε-differentially private iff for all adjacent database b_1, b_2 and for every $S \subseteq R$:

$$\Pr[Q(b_1) \in S] \leq \exp(\varepsilon)\Pr[Q(b_2) \in S]$$

Let's substitute a concrete instance:

$$\Pr[Q(b \cup \{x\}) \in S] \leq \exp(\varepsilon)\Pr[Q(b \cup \{y\}) \in S]$$

Let's use the two quantifiers:

$$\exp(-\varepsilon)\Pr[Q(b \cup \{y\}) \in S] \leq \Pr[Q(b \cup \{x\}) \in S] \leq \exp(\varepsilon)\Pr[Q(b \cup \{y\}) \in S]$$

ε-Differential Privacy

Definition

Given $\varepsilon \geq 0$, a probabilistic query $Q: X^n \to R$ is ε-differentially private iff for all adjacent database b_1, b_2 and for every $S \subseteq R$:

$$\Pr[Q(b_1) \in S] \leq \exp(\varepsilon) \Pr[Q(b_2) \in S]$$

Let's substitute a concrete instance:

$$\Pr[Q(b \cup \{x\}) \in S] \leq \exp(\varepsilon) \Pr[Q(b \cup \{y\}) \in S]$$

Let's use the two quantifiers:

$$\exp(-\varepsilon) \Pr[Q(b \cup \{y\}) \in S] \leq \Pr[Q(b \cup \{x\}) \in S] \leq \exp(\varepsilon) \Pr[Q(b \cup \{y\}) \in S]$$

And for $\varepsilon \to 0$

$$(1-\varepsilon) \Pr[Q(b \cup \{y\}) \in S] \leq \Pr[Q(b \cup \{x\}) \in S] \leq (1+\varepsilon) \Pr[Q(b \cup \{y\}) \in S]$$

ε-Differential Privacy

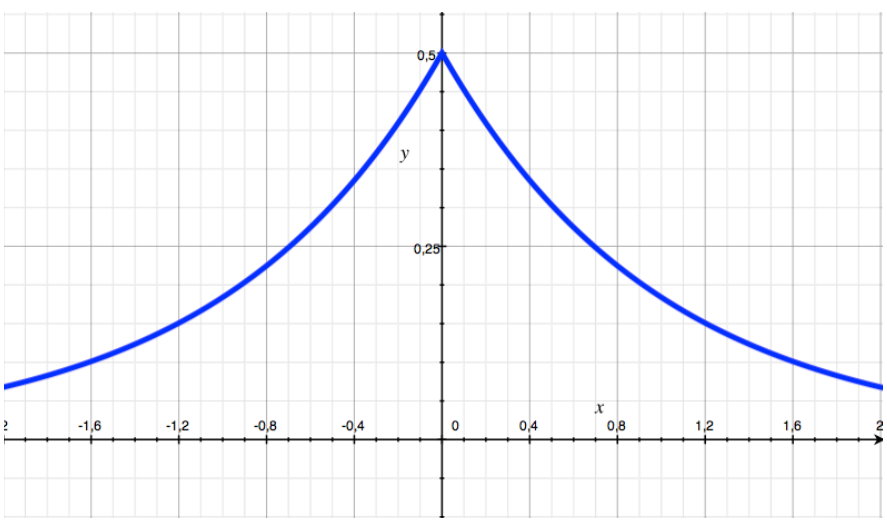
In general we can think about the following quantity as the privacy loss incurred by observing r on the databases b and b'.

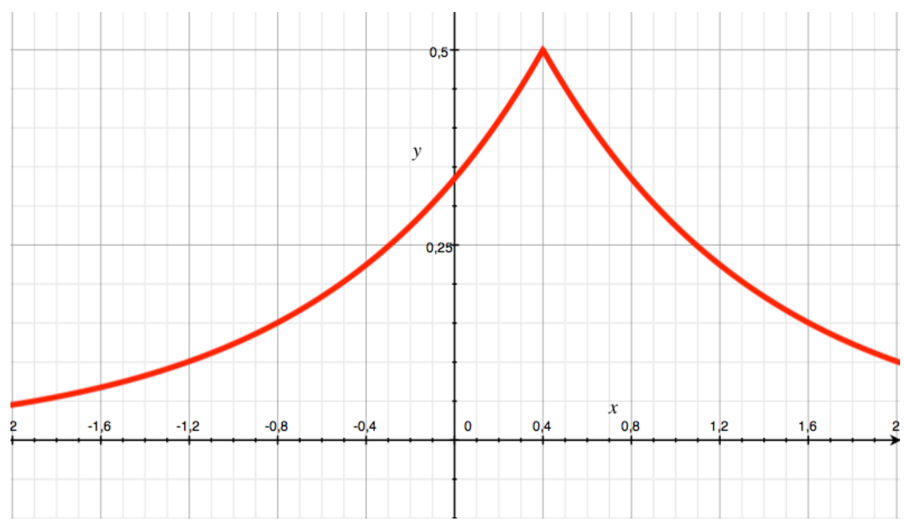
$$L_{b,b'}(r) = \log \frac{\Pr[Q(b)=r]}{\Pr[Q(b')=r]}$$

ε-Differential Privacy

Q : db => R   probabilistic

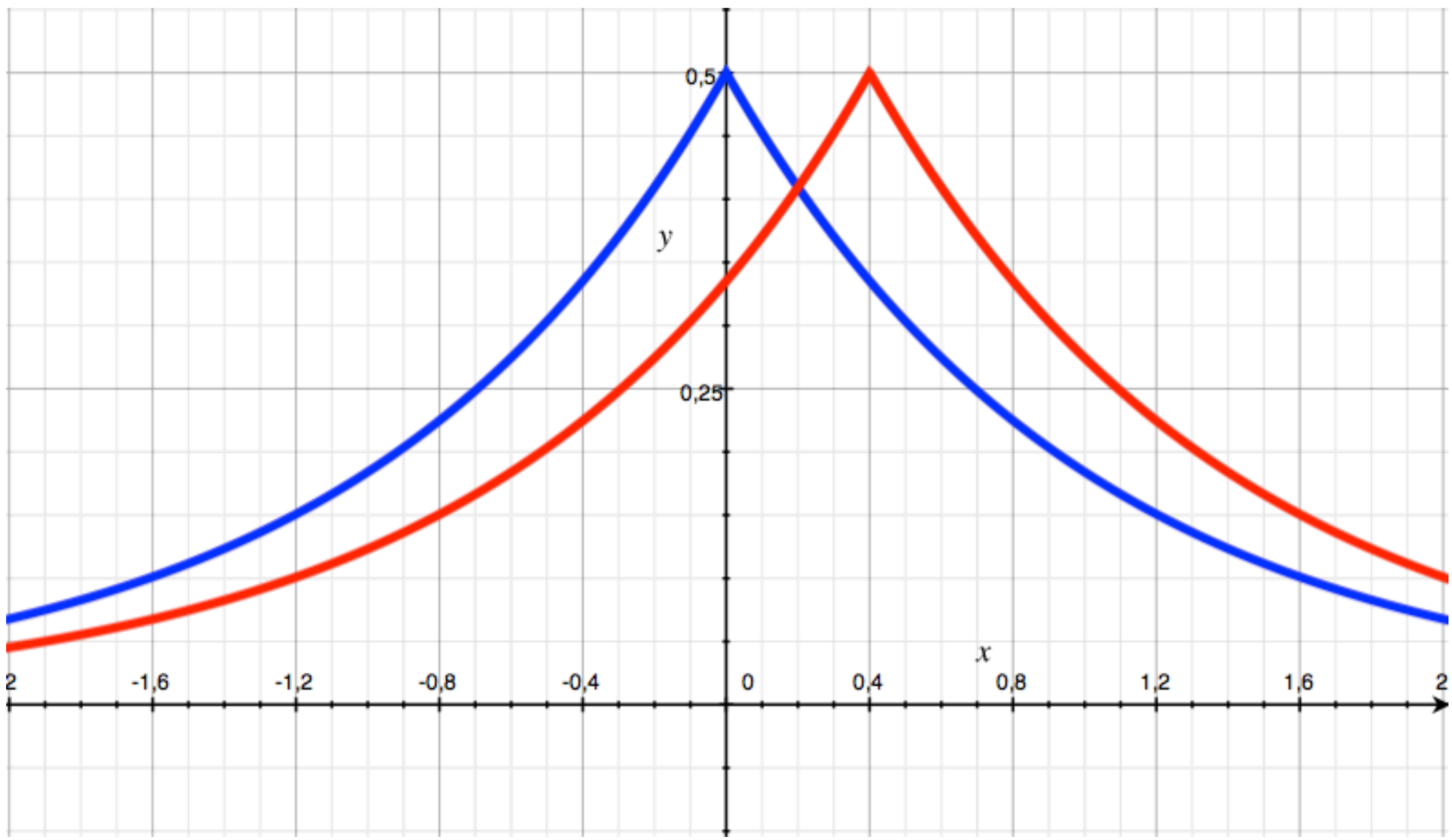
$Q(b \cup \{x\})$ $Q(b \cup \{y\})$





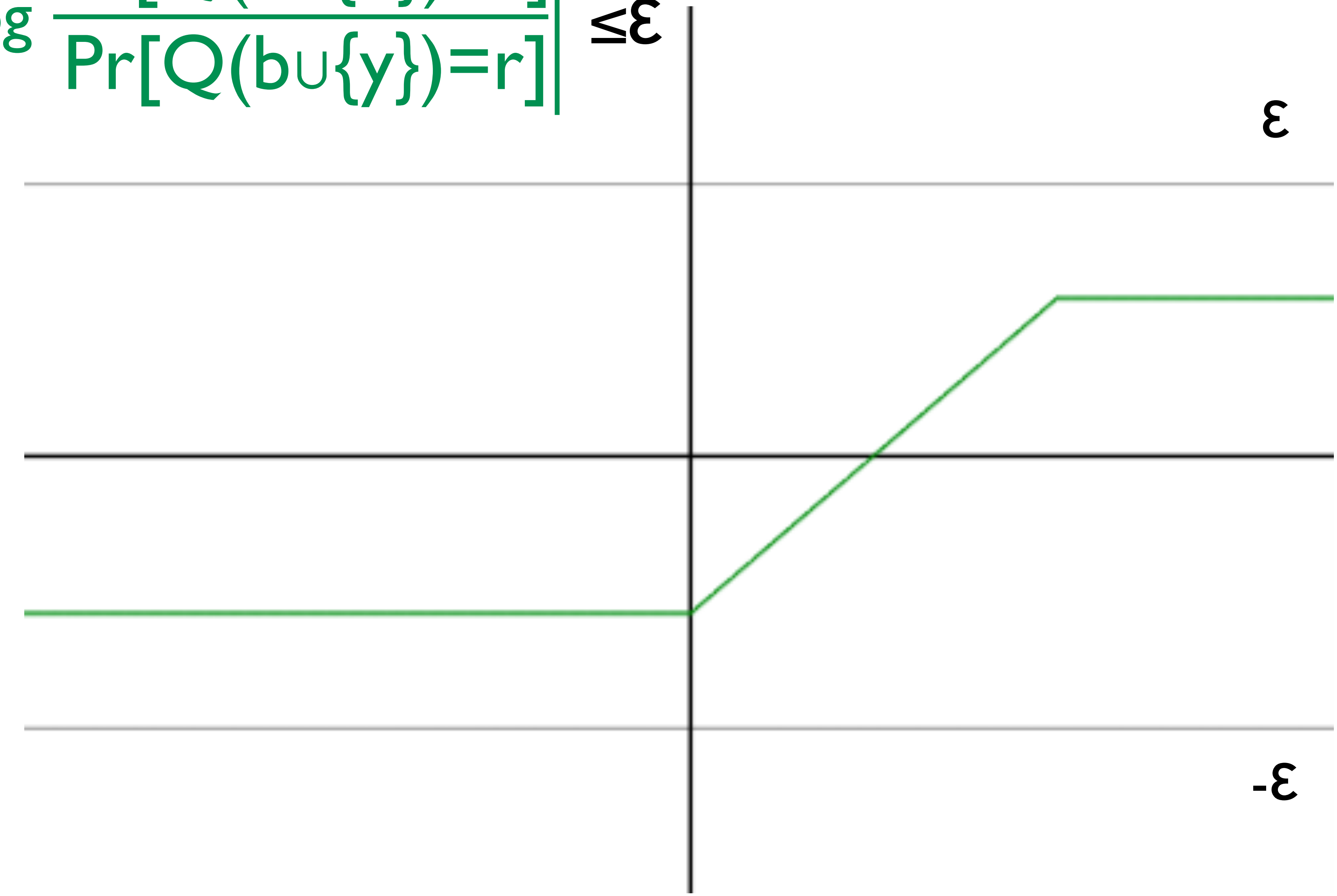
ε-Differential Privacy

$$d(Q(b \cup \{x\}), Q(b \cup \{y\})) \leq \varepsilon$$



ε-Differential Privacy

$$\left|\log \frac{\Pr[Q(b \cup \{x\})=r]}{\Pr[Q(b \cup \{y\})=r]}\right| \leq \varepsilon$$



(ε,δ)-Differential Privacy

Definition

Given $\varepsilon,\delta \geq 0$, a probabilistic query $Q: X^n \rightarrow R$ is (ε,δ)-differentially private iff for all adjacent database b_1, b_2 and for every $S \subseteq R$:

$$\Pr[Q(b_1) \in S] \leq \exp(\varepsilon)\Pr[Q(b_2) \in S] + \delta$$

(ε,δ)-Differential Privacy

Definition
Given $\varepsilon,\delta \geq 0$, a probabilistic query $Q: X^n \rightarrow R$ is (ε,δ)-differentially private iff for all adjacent database b_1, b_2 and for every $S \subseteq R$:

$$\Pr[Q(b_1)\in S] \leq \exp(\varepsilon)\Pr[Q(b_2)\in S] + \delta$$

Similarly, we have

$$\log \frac{\Pr[Q(b_1)\in S] - \delta}{\Pr[Q(b_2)\in S]} \leq \varepsilon$$

$$-\varepsilon \leq \log \frac{\Pr[Q(b_1)\in S] + \delta}{\Pr[Q(b_2)\in S]}$$

(ε,δ)-Differential Privacy

Definition

Given $\varepsilon,\delta \geq 0$, a probabilistic query $Q: X^n \rightarrow R$ is (ε,δ)-differentially private iff for all adjacent database b_1, b_2 and for every $S \subseteq R$:

$$\Pr[Q(b_1) \in S] \leq \exp(\varepsilon)\Pr[Q(b_2) \in S] + \delta$$

Similarly, we have

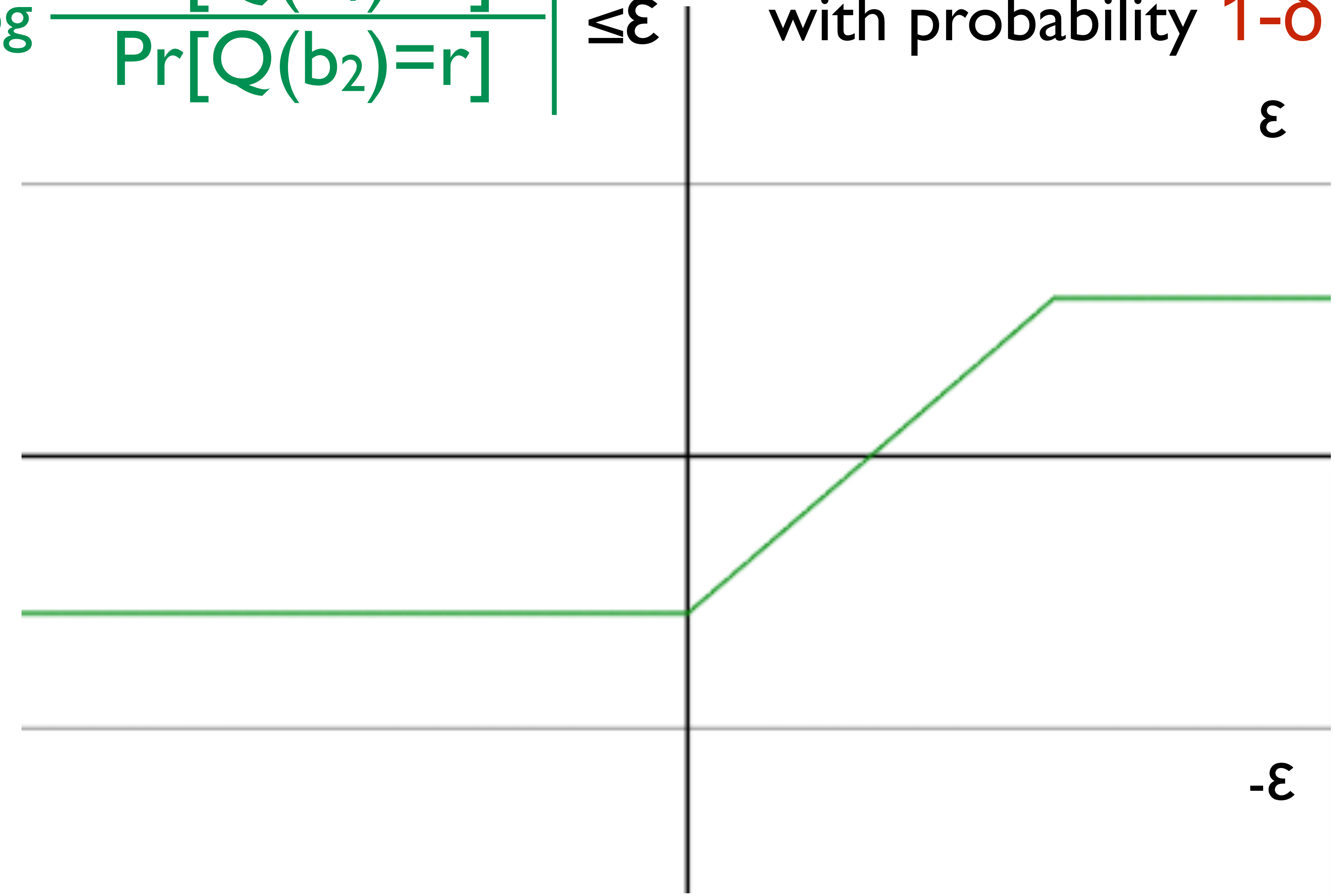
$$\log \frac{\Pr[Q(b_1) \in S] - \delta}{\Pr[Q(b_2) \in S]} \leq \varepsilon$$

$$-\varepsilon \leq \log \frac{\Pr[Q(b_1) \in S] + \delta}{\Pr[Q(b_2) \in S]}$$

Probability of failure

(ε,δ)-Differential Privacy

$$\left|\log \frac{\Pr[Q(b_1)=r]}{\Pr[Q(b_2)=r]}\right| \leq \varepsilon$$ with probability $1-\delta$



The rest of the class

- Understanding some basic methods to guarantee differential privacy and how they provide an answer for the privacy vs utility trade-off.
- Looking at how we can formally support differential privacy using EasyCrypt.

Summary

- Statistical queries and databases,
- Additive noise perturbation,
- Reconstruction attack,
- Fundamental Law of Information Reconstruction,
- Differential privacy

Reconstruction attack with exponential adversary

Let $M:\{0,1\}^n \rightarrow R$ be a privacy mechanism adding noise within E perturbation. Then there is an adversary that can reconstruct the database within 4E positions.

[DinurNissim'02]

Proof

Query phase: For each $S \subseteq [n]$ let $a_S^* = q_S^*(D)$.

Rule out phase: For each $D' \in \{0,1\}^n$:
if there exists S such that $|q_S(D') - a_S^*| > E$ then rule out D'.

Output phase: Output a database D' that was not ruled out.

Proof

Query phase: For each $S \subseteq [n]$ let $a_S^* = q_S^*(D)$.

Rule out phase: For each $D' \in \{0,1\}^n$:
if there exists S such that $|q_S(D') - a_S^*| > E$ then rule out D'.

Output phase: Output a database D' that was not ruled out.

Notice that since for the real database we clearly have

$$|q_S(D) - q_S^*(D)| \leq E$$

the procedure clearly return a candidate output in an exponential number of steps.

Proof

Query phase: For each $S \subseteq [n]$ let $a_S^* = q_S^*(D)$.

Rule out phase: For each $D' \in \{0,1\}^n$:
if there exists S such that $|q_S(D') - a_S^*| > E$ then rule out D'.

Output phase: Output a database D' that was not ruled out.

Notice that since for the real database we clearly have

$$|q_S(D) - q_S^*(D)| \leq E$$

the procedure clearly return a candidate output in an exponential number of steps.

We now want to show that $d_H(D,D') \leq 4E$

Proof

Let 's consider D to be the real dataset and D' to be the outputted one. Consider the sets of indices

$$R = \{ i \mid D(i)=0 \} \qquad \text{and} \qquad T = \{ i \mid D(i)=1 \}$$

Proof

Let 's consider D to be the real dataset and D' to be the outputted one. Consider the sets of indices

$$R = \{ i \mid D(i)=0 \} \quad \text{and} \quad T = \{ i \mid D(i)=1 \}$$

Since D' was not ruled out we have

$$|q_S^*(D)-q_S(D')| \le E$$

but by definition we also have

$$|q_S^*(D)-q_S(D)| \le E$$

so by triangle inequality $|q_S(D)-q_S(D')| \le 2E$.

Since $q_R(D)=0$, we have that on the indices in R the Hamming distance between D and D' is at most 2E.

Proof

Let 's consider D to be the real dataset and D' to be the outputted one. Consider the sets of indices

$$R = \{ i \mid D(i)=0 \} \quad \text{and} \quad T = \{ i \mid D(i)=1 \}$$

Since D' was not ruled out we have

$$|q_S^*(D)-q_S(D')| \leq E$$

but by definition we also have

$$|q_S^*(D)-q_S(D)| \leq E$$

so by triangle inequality $|q_S(D)-q_S(D')| \leq 2E$.
Since $q_R(D)=0$, we have that on the indices in R the Hamming distance between D and D' is at most 2E.

We can apply a similar reasoning to T. So overall D and D' differ in at most 4E positions.