Motivating Differential Privacy

Data Science 231 - Summer 2017

Agenda

1. Anonymity as a Primer
2. Differential Privacy - A Paradigm Shift
3. “Bad” Constructs of Private Release
4. Desiderata of Private Analysis
5. Differential Privacy

Disclaimer

- These slides are meant to serve as conceptual motivation prior to your course readings
- As such, these slides are not meant to teach you everything you need to know about the topic
- Instead, they are meant to provide some background before you delve into a very mathematically dense topic

k-Anonymity as a Primer

- There have been many computational approaches to achieve privacy within a dataset

k-Anonymity as a Primer

- There have been many computational approaches to achieve privacy within a dataset
- One such method is k-Anonymity (Sweeny, Samarati 1998)

k-Anonymity as a Primer

- There have been many computational approaches to achieve privacy within a dataset
- One such method is k-Anonymity (Sweeny, Samarati 1998)
- Idea:
    - Any row in a table has at least (k-1) other identical rows

k-Anonymity as a Primer

- There have been many computational approaches to achieve privacy within a dataset
- One such method is k-Anonymity (Sweeny, Samarati 1998)
- Idea:
    - Any row in a table has at least (k-1) other identical rows
    - Equivalently, all unique tuples in a table appear at least k times

Example: 3-Anonymous Arrest Table

NAME	*AGE*	*STATE*	*IS_FELON*
REDACTED	18-24	CA	0
REDACTED	18-24	CA	0
REDACTED	18-24	CA	0
REDACTED	25-34	AL	1
REDACTED	25-34	AL	1
REDACTED	25-34	AL	1
REDACTED	45-54	AL	0
REDACTED	45-54	AL	0
REDACTED	45-54	AL	0
REDACTED	45-54	AL	0

Example: 3-Anonymous Arrest Table

NAME	***AGE***	***STATE***	***IS_FELON***
REDACTED	18-24	CA	0
REDACTED	18-24	CA	0
REDACTED	18-24	CA	0
REDACTED	25-34	AL	1
REDACTED	25-34	AL	1
REDACTED	25-34	AL	1
REDACTED	45-54	AL	0
REDACTED	45-54	AL	0
REDACTED	45-54	AL	0
REDACTED	45-54	AL	0

Example: 3-Anonymous Arrest Table

NAME	*AGE*	*STATE*	*IS_FELON*
REDACTED	18-24	CA	0
REDACTED	18-24	CA	0
REDACTED	18-24	CA	0
REDACTED	25-34	AL	1
REDACTED	25-34	AL	1
REDACTED	25-34	AL	1
REDACTED	45-54	AL	0
REDACTED	45-54	AL	0
REDACTED	45-54	AL	0
REDACTED	45-54	AL	0

Example: 3-Anonymous Arrest Table

NAME	*AGE*	*STATE*	*IS_FELON*
REDACTED	18-24	CA	0
REDACTED	18-24	CA	0
REDACTED	18-24	CA	0
REDACTED	25-34	AL	1
REDACTED	25-34	AL	1
REDACTED	25-34	AL	1
REDACTED	45-54	AL	0
REDACTED	45-54	AL	0
REDACTED	45-54	AL	0
REDACTED	45-54	AL	0

k-Anonymity as a Primer

Why have k identical tuples in a released table?

k-Anonymity as a Primer

Why have k identical tuples in a released table?

- This protects against linkage attacks

k-Anonymity as a Primer

Why have k identical tuples in a released table?

- This protects against linkage attacks
    - Any join on a k-anonymous table will be linked to k rows

k-Anonymity as a Primer

Why have k identical tuples in a released table?

- This protects against linkage attacks
  - Any join on a k-anonymous table will be linked to k rows
  - If $k > 1$, then no unique row can be singled out

k-Anonymity as a Primer

Why have k identical tuples in a released table?

- This protects against linkage attacks
  - Any join on a k-anonymous table will be linked to k rows
  - If $k > 1$, then no unique row can be singled out
  - Thus, an individual remains hidden amongst $(k-1)$ other folks

Attacks on k-Anonymity

There are problems with k-anonymity though

Attacks on k-Anonymity

There are problems with k-anonymity though

- Attacks

Attacks on k-Anonymity

There are problems with k-anonymity though

- Attacks
  - Background knowledge attacks
  - Homogeneity attacks

Attacks on k-Anonymity

There are problems with k-anonymity though

- Attacks
    - Background knowledge attacks
    - Homogeneity attacks
- Suppose Alice and Bob are roommates

Attacks on k-Anonymity

There are problems with k-anonymity though

- Attacks
    - Background knowledge attacks
    - Homogeneity attacks
- Suppose Alice and Bob are roommates
    - One day, the cops show up to their apartment and arrest Bob

Attacks on k-Anonymity

There are problems with k-anonymity though

- Attacks
    - Background knowledge attacks
    - Homogeneity attacks
- Suppose Alice and Bob are roommates
    - One day, the cops show up to their apartment and arrest Bob
    - After Bob’s trial, Alice wants to know if Bob is a felon

Attacks on k-Anonymity

There are problems with k-anonymity though

- Attacks
  - Background knowledge attacks
  - Homogeneity attacks
- Suppose Alice and Bob are roommates
  - One day, the cops show up to their apartment and arrest Bob
  - After Bob’s trial, Alice wants to know if Bob is a felon
  - Alice, being Bob’s roommate, knows that they live in Alabama and that Bob is 27 years old

Homogeneity & Background Knowledge Attack

NAME	***AGE***	***STATE***	***IS_FELON***
REDACTED	18-24	CA	0
REDACTED	18-24	CA	0
REDACTED	18-24	CA	0
REDACTED	25-34	AL	1
REDACTED	25-34	AL	1
REDACTED	25-34	AL	1
REDACTED	45-54	AL	0
REDACTED	45-54	AL	0
REDACTED	45-54	AL	0
REDACTED	45-54	AL	0

Using Alabama Knowledge

NAME	***AGE***	***STATE***	***IS_FELON***
REDACTED	18-24	CA	0
REDACTED	18-24	CA	0
REDACTED	18-24	CA	0
REDACTED	25-34	AL	1
REDACTED	25-34	AL	1
REDACTED	25-34	AL	1
REDACTED	45-54	AL	0
REDACTED	45-54	AL	0
REDACTED	45-54	AL	0
REDACTED	45-54	AL	0

Using Age = 27 Knowledge

NAME	*AGE*	*STATE*	*IS_FELON*
REDACTED	18-24	CA	0
REDACTED	18-24	CA	0
REDACTED	18-24	CA	0
REDACTED	25-34	AL	1
REDACTED	25-34	AL	1
REDACTED	25-34	AL	1
REDACTED	45-54	AL	0
REDACTED	45-54	AL	0
REDACTED	45-54	AL	0
REDACTED	45-54	AL	0

Homogeneity of the Sensitive Attribute Leads to Discovery

NAME	*AGE*	*STATE*	*IS_FELON*
REDACTED	18-24	CA	0
REDACTED	18-24	CA	0
REDACTED	18-24	CA	0
REDACTED	25-34	AL	**1**
REDACTED	25-34	AL	**1**
REDACTED	25-34	AL	**1**
REDACTED	45-54	AL	0
REDACTED	45-54	AL	0
REDACTED	45-54	AL	0
REDACTED	45-54	AL	0

Attacks on k-Anonymity

There are problems with k-anonymity though

- Attacks
    - Background knowledge attacks
    - Homogeneity attacks
- Suppose Alice and Bob are roommates
    - One day the cops show up to Alice and Bob’s place and arrest Bob
    - After Bob’s trial, Alice wants to know if Bob is a felon
    - Alice, being Bob’s roommate, knows that they live in Alabama and that Bob is 27 years old
- Hence, Alice learns that Bob is a felon

Attacks on k-Anonymity

There are problems with k-anonymity though

- Attacks
  - Background knowledge attacks
  - Homogeneity attacks
- Suppose Alice and Bob are roommates
  - One day the cops show up to Alice and Bob’s place and arrest Bob
  - After Bob’s trial, Alice wants to know if Bob is a felon
  - Alice, being Bob’s roommate, knows that they live in Alabama and that Bob is 27 years old
- Hence, Alice learns that Bob is a felon
- The lack of diversity in the sensitive attribute, along with Alice’s background knowledge, led to discovering that Bob is a felon

Beyond k-Anonymity

- These attacks on k-anonymity led to a refinement called l-diversity

Beyond k-Anonymity

- These attacks on k-anonymity led to a refinement called l-diversity
- But l-diversity had issues too

Beyond k-Anonymity

- These attacks on k-anonymity led to a refinement called l-diversity
- But l-diversity had issues too
    - Skewness attacks
    - Similarity attacks

Beyond k-Anonymity

- These attacks on k-anonymity led to a refinement called l-diversity
- But l-diversity had issues too
    - Skewness attacks
    - Similarity attacks
- This led to a further refinement called t-closeness

Beyond k-Anonymity

- These attacks on k-anonymity led to a refinement called l-diversity
- But l-diversity had issues too
    - Skewness attacks
    - Similarity attacks
- This led to a further refinement called t-closeness
- But t-closeness had issues as well

Beyond k-Anonymity

- These attacks on k-anonymity led to a refinement called l-diversity
- But l-diversity had issues too
  - Skewness attacks
  - Similarity attacks
- This led to a further refinement called t-closeness
- But t-closeness had issues as well
- And so on…

Beyond k-Anonymity

- These attacks on k-anonymity led to a refinement called l-diversity
- But l-diversity had issues too
  - Skewness attacks
  - Similarity attacks
- This led to a further refinement called t-closeness
- But t-closeness had issues as well
- And so on…
- Moral:
  - All of these conceptualizations of privacy were properties of the data
  - As such, they could be attacked via different exploits

Differential Privacy - A Paradigm Shift

Differential Privacy - A Paradigm Shift

- Recall that k-Anonymity was a property of a given dataset

Differential Privacy - A Paradigm Shift

- Recall that k-Anonymity was a property of a given dataset
- Differential Privacy (Dwork, et. al. 2006) was a paradigm shift

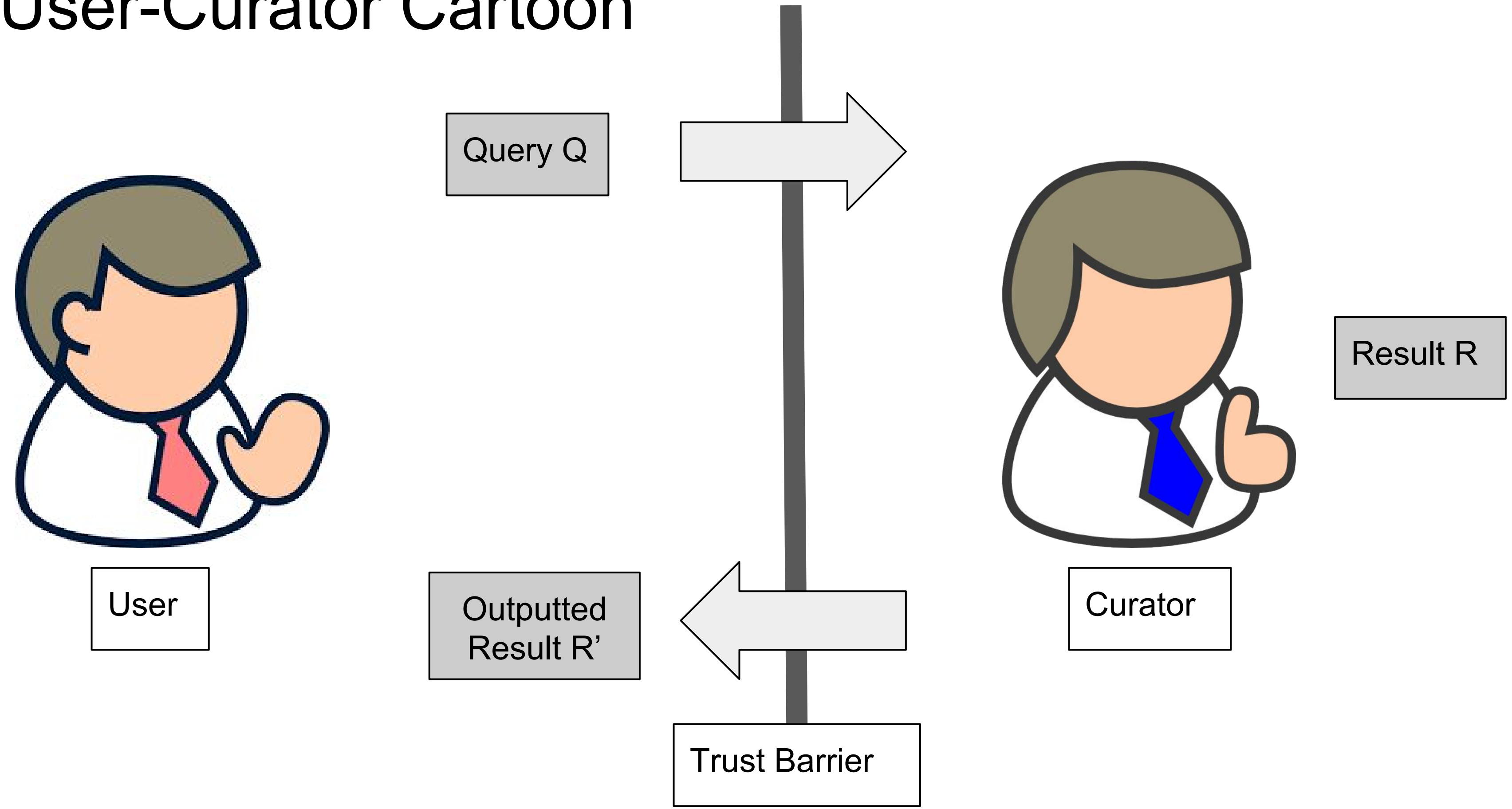
Differential Privacy - A Paradigm Shift

- Recall that k-Anonymity was a property of a given dataset
- Differential Privacy (Dwork, et. al. 2006) was a paradigm shift
    - Moved away from privacy as a property of a dataset

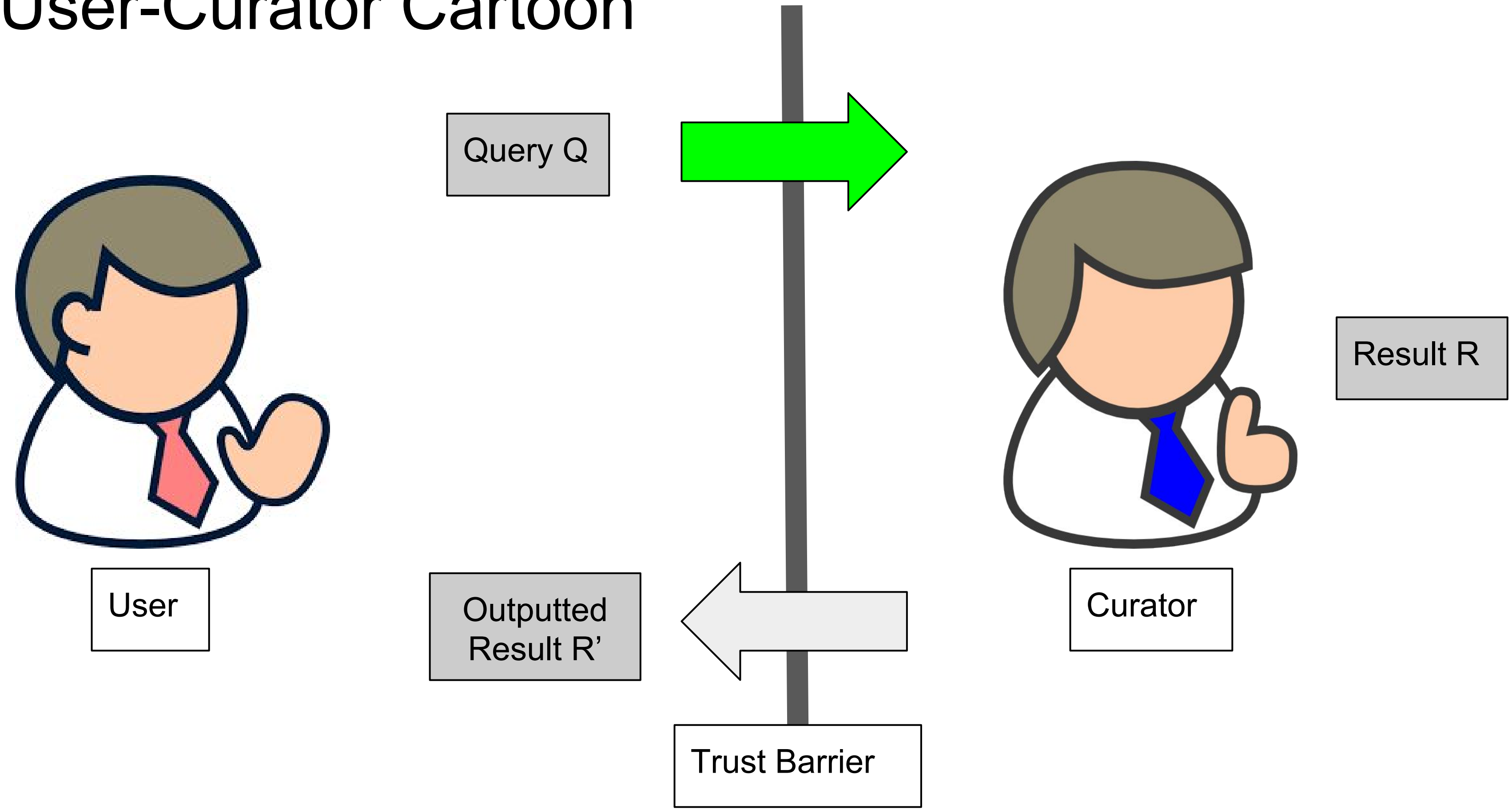
Differential Privacy - A Paradigm Shift

- Recall that k-Anonymity was a property of a given dataset
- Differential Privacy (Dwork, et. al. 2006) was a paradigm shift
    - Moved away from privacy as a property of a dataset
    - Instead, privacy as a property of a mechanism that produced a “private result”

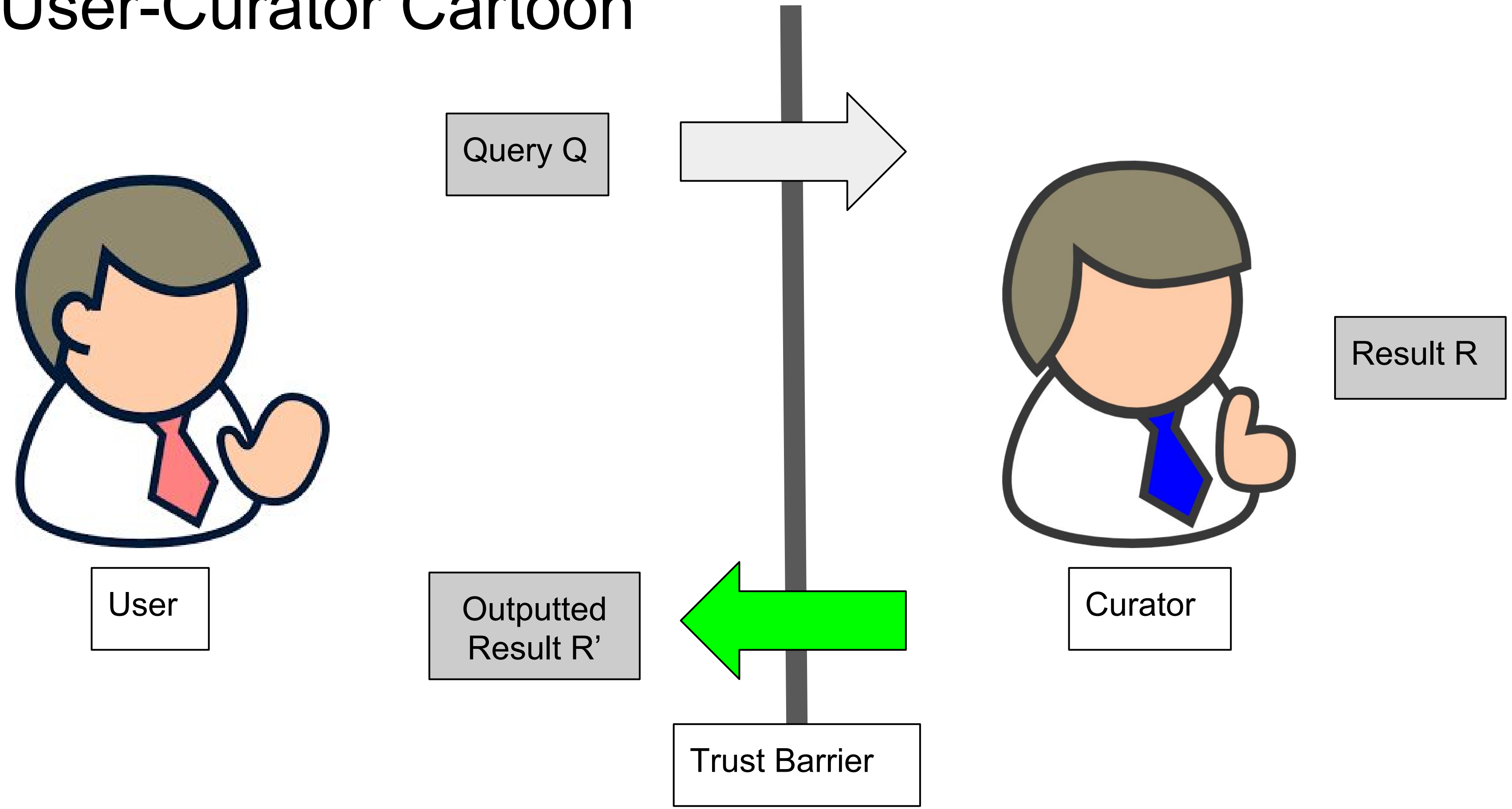
User-Curator Cartoon



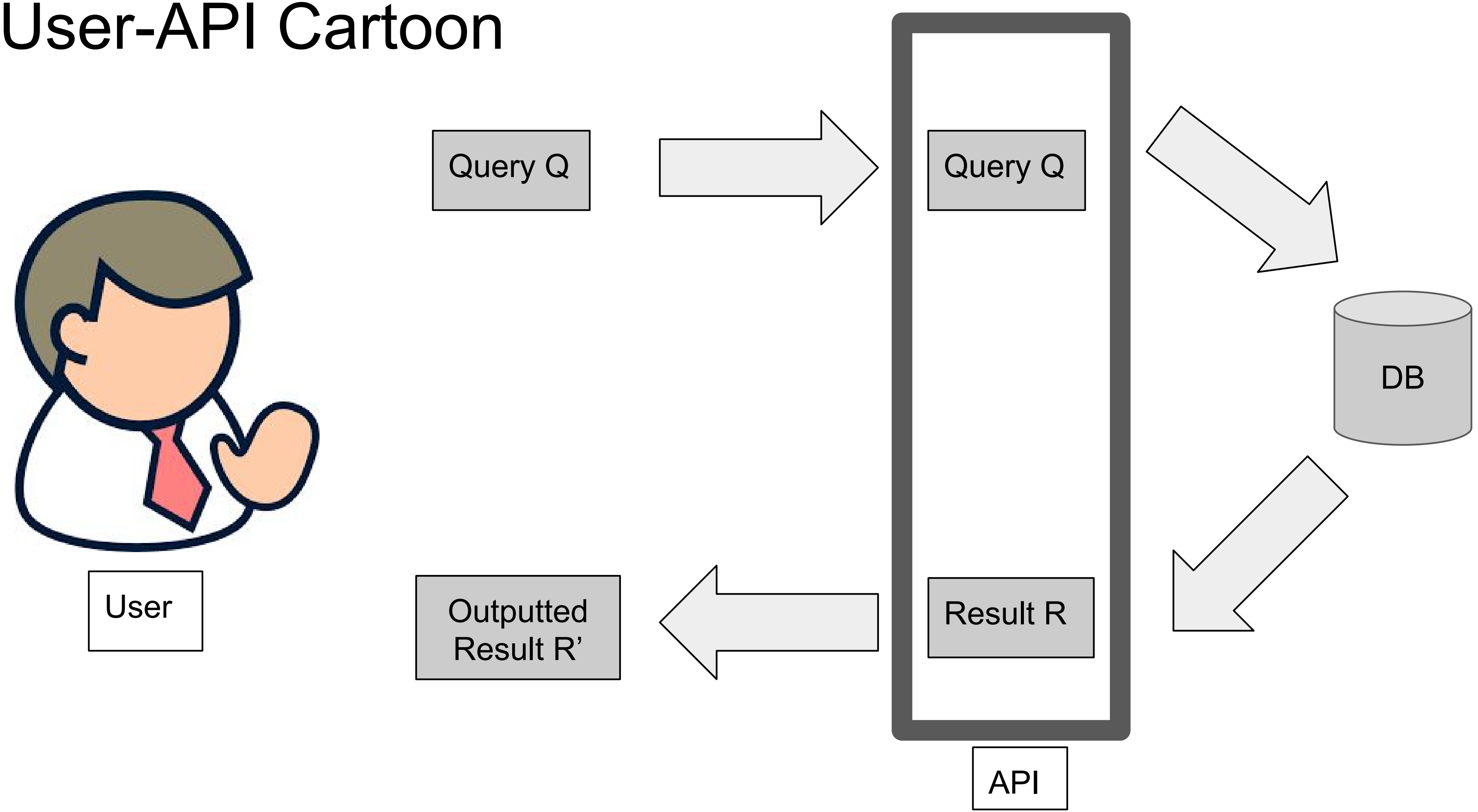
User-Curator Cartoon



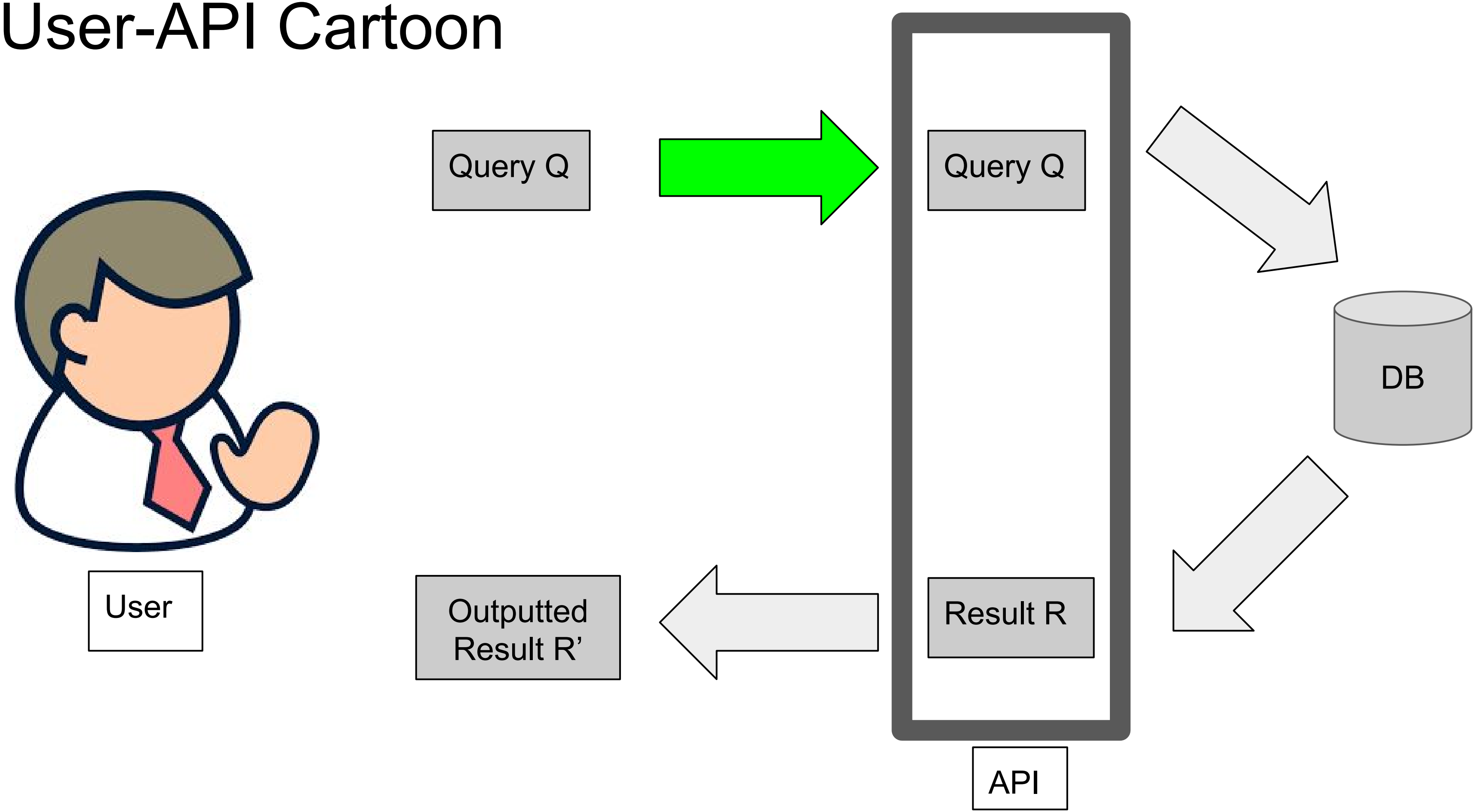
User-Curator Cartoon



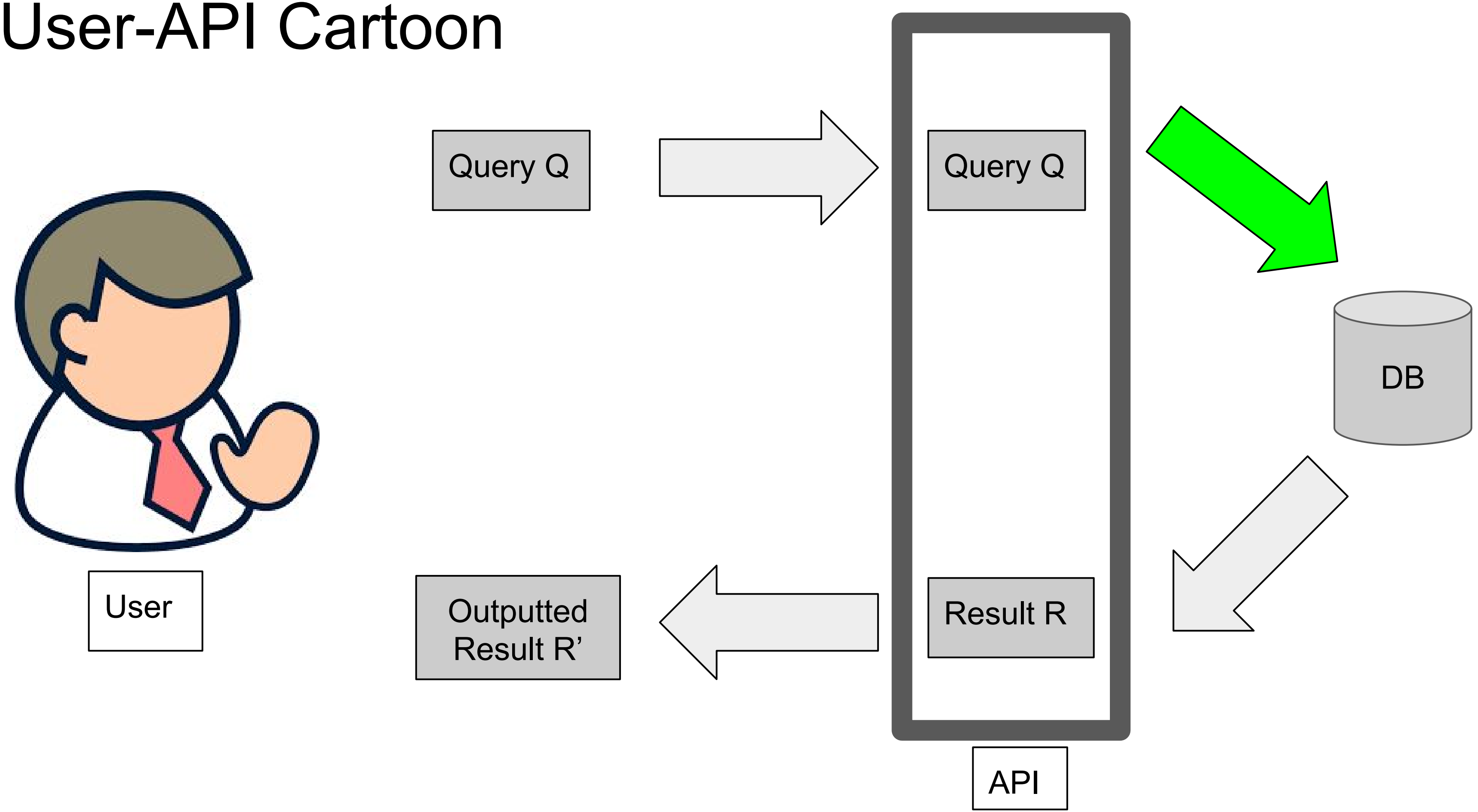
User-API Cartoon



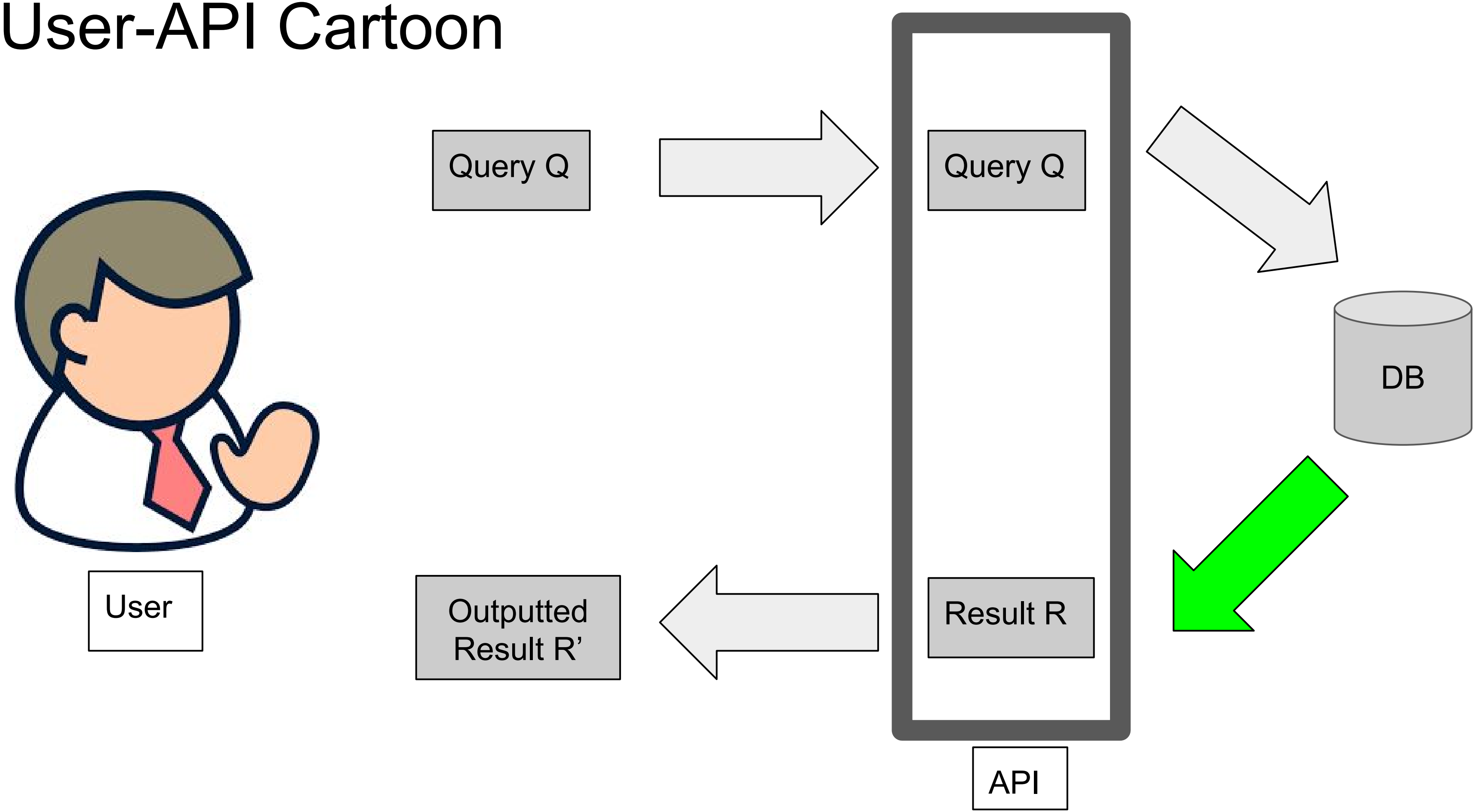
User-API Cartoon



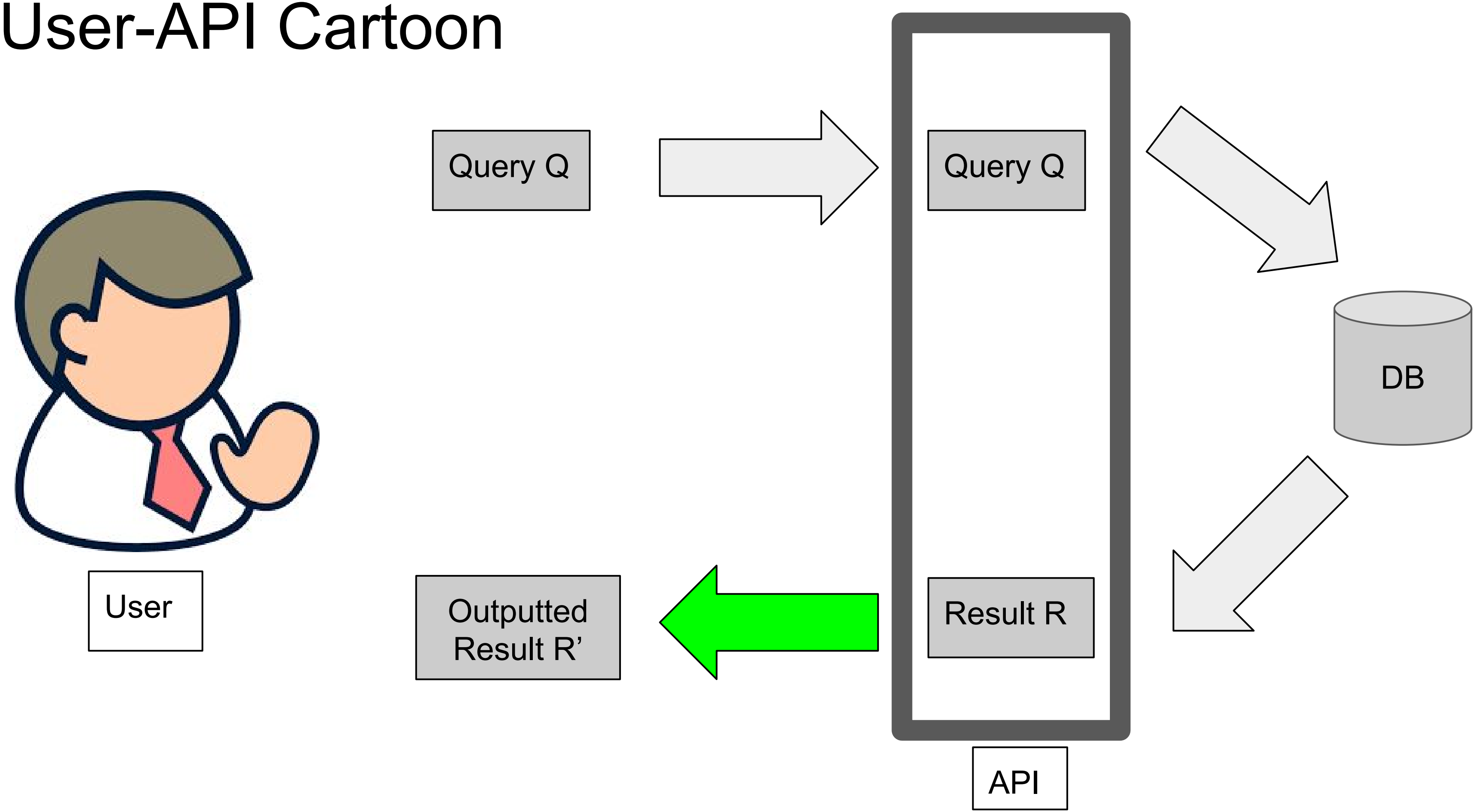
User-API Cartoon



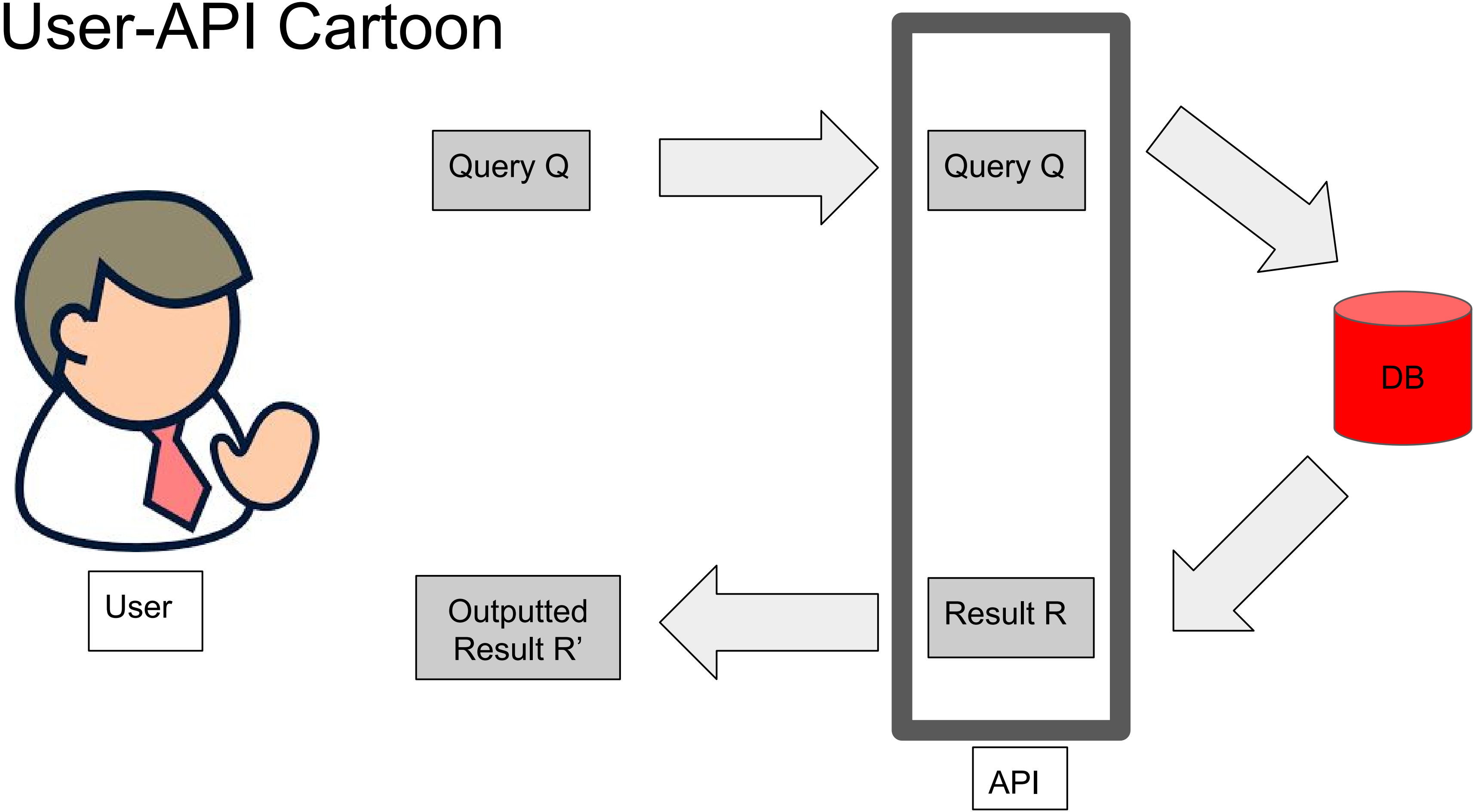
User-API Cartoon



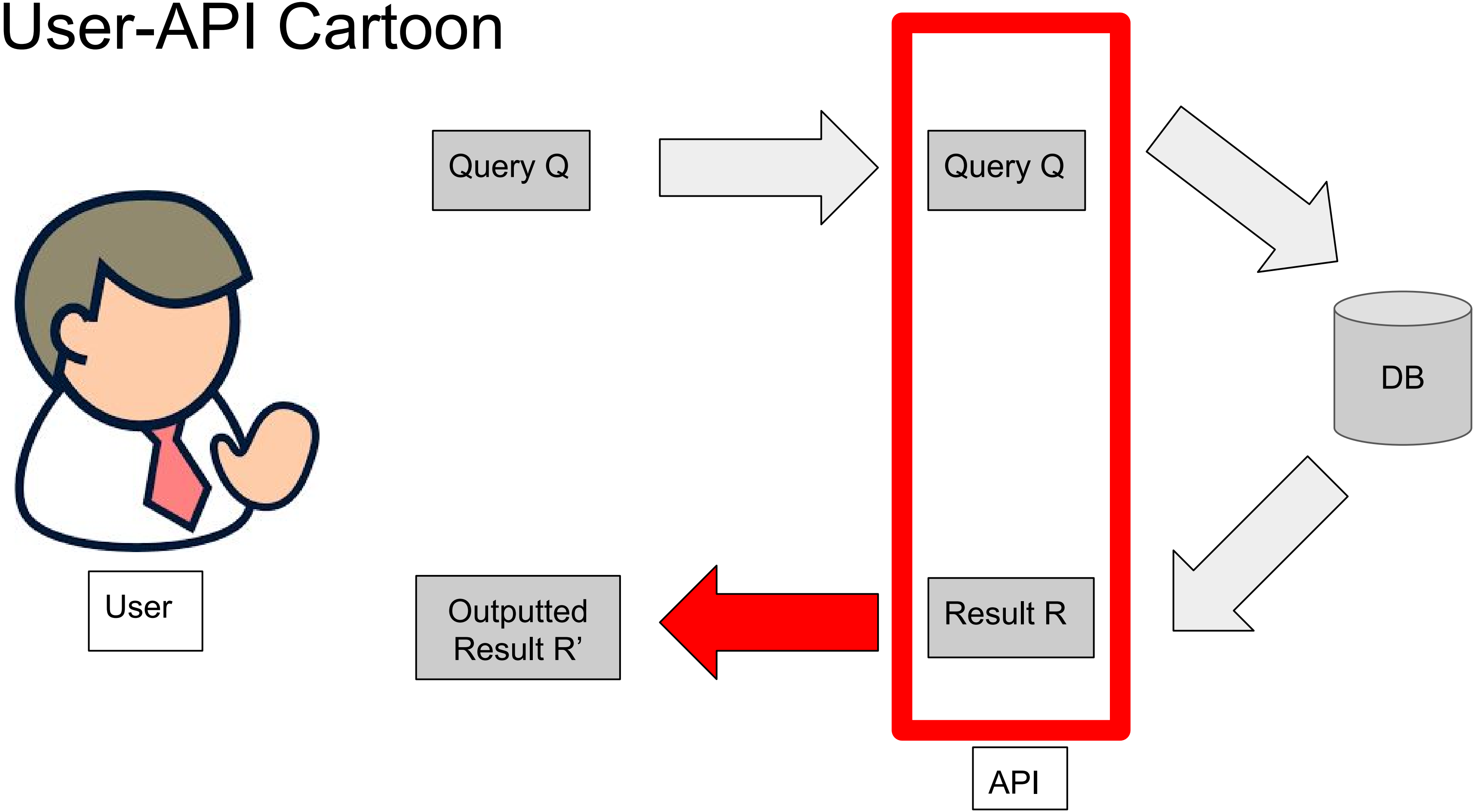
User-API Cartoon



User-API Cartoon



User-API Cartoon



Goal of Private Release

Goal of Private Release

Goal: We want to learn information about a population, while not learning about individuals specifically

Goal of Private Release

Goal: We want to learn information about a population, while not learning about individuals specifically

- Question: What should qualify as a “private release?”

Goal of Private Release

Goal: We want to learn information about a population, while not learning about individuals specifically

- Question: What should qualify as a “private release?”
- The following are 3 “bad ideas” of private release

Goal of Private Release

Goal: We want to learn information about a population, while not learning about individuals specifically

- Question: What should qualify as a "private release?"
- The following are 3 "bad ideas" of private release
  - [From a Microsoft talk by Cynthia Dwork]

Operationalizing Private Release

Bad Idea 1: Only report answers that are computed using a large number of data elements

Operationalizing Private Release

Bad Idea 1: Only report answers that are computed using a large number of data elements

- On the surface this seems reasonable

Operationalizing Private Release

Bad Idea 1: Only report answers that are computed using a large number of data elements

- On the surface this seems reasonable
- But this is vulnerable to differencing attacks

Operationalizing Private Release

Bad Idea 1: Only report answers that are computed using a large number of data elements

- On the surface this seems reasonable
- But this is vulnerable to differencing attacks
- Example: Suppose we want to learn if Nitin has been convicted of a felony

Operationalizing Private Release

Bad Idea 1: Only report answers that are computed using a large number of data elements

- On the surface this seems reasonable
- But this is vulnerable to differencing attacks
- Example: Suppose we want to learn if Nitin has been convicted of a felony
    - We cannot ask the query Q = "Is Nitin a felon?"

Operationalizing Private Release

Bad Idea 1: Only report answers that are computed using a large number of data elements

- On the surface this seems reasonable
- But this is vulnerable to differencing attacks
- Example: Suppose we want to learn if Nitin has been convicted of a felony
    - We cannot ask the query Q = “Is Nitin a felon?”
    - But we can ask both of the following queries:

Operationalizing Private Release

Bad Idea 1: Only report answers that are computed using a large number of data elements

- On the surface this seems reasonable
- But this is vulnerable to differencing attacks
- Example: Suppose we want to learn if Nitin has been convicted of a felony
    - We cannot ask the query Q = “Is Nitin a felon?”
    - But we can ask both of the following queries:
        - Q1 = “How many felons are in the dataset?”
        - Q2 = “How many felons are in the dataset without Nitin”

Operationalizing Private Release

Bad Idea 1: Only report answers that are computed using a large number of data elements

- On the surface this seems reasonable
- But this is vulnerable to differencing attacks
- Example: Suppose we want to learn if Nitin has been convicted of a felony
    - We cannot ask the query Q = “Is Nitin a felon?”
    - But we can ask both of the following queries:
        - Q1 = “How many felons are in the dataset?”
        - Q2 = “How many felons are in the dataset without Nitin”
    - Note that Q = Q1 - Q2

Operationalizing Private Release

Bad Idea 1: Only report answers that are computed using a large number of data elements

- On the surface this seems reasonable
- But this is vulnerable to differencing attacks
- Example: Suppose we want to learn if Nitin has been convicted of a felony
    - We cannot ask the query Q = “Is Nitin a felon?”
    - But we can ask both of the following queries:
        - Q1 = “How many felons are in the dataset?”
        - Q2 = “How many felons are in the dataset without Nitin”
    - Note that Q = Q1 - Q2
    - Therefore, we can get an answer to Q despite not having access to ask Q

Operationalizing Private Release

Bad Idea 2: Add random noise to every query asked

Operationalizing Private Release

Bad Idea 2: Add random noise to every query asked

- This is not such a bad idea, but we have to be careful here

Operationalizing Private Release

Bad Idea 2: Add random noise to every query asked

- This is not such a bad idea, but we have to be careful here
- This is vulnerable to averaging attacks

Operationalizing Private Release

Bad Idea 2: Add random noise to every query asked

- This is not such a bad idea, but we have to be careful here
- This is vulnerable to averaging attacks
- Example: Suppose we want to learn if Nitin has been convicted of a felony

Operationalizing Private Release

Bad Idea 2: Add random noise to every query asked

- This is not such a bad idea, but we have to be careful here
- This is vulnerable to averaging attacks
- Example: Suppose we want to learn if Nitin has been convicted of a felony
    - Each time we ask the question “Is Nitin a felon?” we get back the true result T, plus some random noise from a distribution N

Operationalizing Private Release

Bad Idea 2: Add random noise to every query asked

- This is not such a bad idea, but we have to be careful here
- This is vulnerable to averaging attacks
- Example: Suppose we want to learn if Nitin has been convicted of a felony
    - Each time we ask the question “Is Nitin a felon?” we get back the true result T, plus some random noise from a distribution N
        - Suppose this distribution has mean 0

Operationalizing Private Release

Bad Idea 2: Add random noise to every query asked

- This is not such a bad idea, but we have to be careful here
- This is vulnerable to averaging attacks
- Example: Suppose we want to learn if Nitin has been convicted of a felony
    - Each time we ask the question “Is Nitin a felon?” we get back the true result T, plus some random noise from a distribution N
        - Suppose this distribution has mean 0
    - Suppose we repeatedly as this question M times

Operationalizing Private Release

Bad Idea 2: Add random noise to every query asked

- This is not such a bad idea, but we have to be careful here
- This is vulnerable to averaging attacks
- Example: Suppose we want to learn if Nitin has been convicted of a felony
  - Each time we ask the question “Is Nitin a felon?” we get back the true result T, plus some random noise from a distribution N
    - Suppose this distribution has mean 0
  - Suppose we repeatedly as this question M times
  - Let R(i) be the result of the Query i when noise N(i) is drawn from N

Operationalizing Private Release

Bad Idea 2: Add random noise to every query asked

- This is not such a bad idea, but we have to be careful here
- This is vulnerable to averaging attacks
- Example: Suppose we want to learn if Nitin has been convicted of a felony
    - Each time we ask the question “Is Nitin a felon?” we get back the true result T, plus some random noise from a distribution N
        - Suppose this distribution has mean 0
    - Suppose we repeatedly as this question M times
    - Let R(i) be the result of the Query i when noise N(i) is drawn from N
    - Then R(i) = T + N(i)

Operationalizing Private Release

- If we average all of the R(i)’s, the result we get is

Operationalizing Private Release

- If we average all of the R(i)'s, the result we get is

$$\frac{1}{M}\sum_{i=1}^{M} R(i) = \frac{1}{M}\sum_{i=1}^{M}(T + N(i)) = T + \frac{1}{M}\sum_{i=1}^{M} N(i)$$

Operationalizing Private Release

- If we average all of the R(i)'s, the result we get is

$$\frac{1}{M}\sum_{i=1}^{M} R(i) = \frac{1}{M}\sum_{i=1}^{M}(T + N(i)) = T + \frac{1}{M}\sum_{i=1}^{M} N(i)$$

- When M gets large, the Law of Large Numbers implies

Operationalizing Private Release

- If we average all of the R(i)’s, the result we get is

$$\frac{1}{M}\sum_{i=1}^{M} R(i) = \frac{1}{M}\sum_{i=1}^{M}(T + N(i)) = T + \frac{1}{M}\sum_{i=1}^{M} N(i)$$

- When M gets large, the Law of Large Numbers implies

$$lim_{M\to\infty}\frac{1}{M}\sum_{i=1}^{M} N(i) = E[N] = 0$$

Operationalizing Private Release

- If we average all of the R(i)'s, the result we get is

$$\frac{1}{M}\sum_{i=1}^{M} R(i) = \frac{1}{M}\sum_{i=1}^{M}(T + N(i)) = T + \frac{1}{M}\sum_{i=1}^{M} N(i)$$

- When M gets large, the Law of Large Numbers implies

$$lim_{M\to\infty}\frac{1}{M}\sum_{i=1}^{M} N(i) = E[N] = 0$$

- Therefore

Operationalizing Private Release

- If we average all of the R(i)'s, the result we get is

$$\frac{1}{M}\sum_{i=1}^{M} R(i) = \frac{1}{M}\sum_{i=1}^{M}(T+N(i)) = T + \frac{1}{M}\sum_{i=1}^{M} N(i)$$

- When M gets large, the Law of Large Numbers implies

$$lim_{M\to\infty}\frac{1}{M}\sum_{i=1}^{M} N(i) = E[N] = 0$$

- Therefore

$$lim_{M\to\infty}\frac{1}{M}\sum_{i=1}^{M} R(i) = T$$

Operationalizing Private Release

Follow-up Question - What if we only allow queries on large set and add random noise to the results? Does this work?

Operationalizing Private Release

Follow-up Question - What if we only allow queries on large set and add random noise to the results? Does this work?

- No!

Operationalizing Private Release

Follow-up Question - What if we only allow queries on large set and add random noise to the results? Does this work?

- No!
- We can combine the differencing and averaging attacks to still learn if Nitin is a felon

Operationalizing Private Release

Bad Idea 3: Random noise with memory

Operationalizing Private Release

Bad Idea 3: Random noise with memory

- Idea
    - Keep a dictionary of every query asked with the returned result

Operationalizing Private Release

Bad Idea 3: Random noise with memory

- Idea
    - Keep a dictionary of every query asked with the returned result
    - If the query has been asked before, return the value in the dictionary

Operationalizing Private Release

Bad Idea 3: Random noise with memory

- Idea
  - Keep a dictionary of every query asked with the returned result
  - If the query has been asked before, return the value in the dictionary
  - Otherwise, compute the real answer and perturb it with noise
    - Add (query, perturbed answer) to the dictionary

Operationalizing Private Release

Bad Idea 3: Random noise with memory

- Idea
  - Keep a dictionary of every query asked with the returned result
  - If the query has been asked before, return the value in the dictionary
  - Otherwise, compute the real answer and perturb it with noise
    - Add (query, perturbed answer) to the dictionary
- Good in theory, but impractical

Operationalizing Private Release

Bad Idea 3: Random noise with memory

- Idea
  - Keep a dictionary of every query asked with the returned result
  - If the query has been asked before, return the value in the dictionary
  - Otherwise, compute the real answer and perturb it with noise
    - Add (query, perturbed answer) to the dictionary
- Good in theory, but impractical
- Fundamental CS result on undecidability

Operationalizing Private Release

Bad Idea 3: Random noise with memory

- Idea
  - Keep a dictionary of every query asked with the returned result
  - If the query has been asked before, return the value in the dictionary
  - Otherwise, compute the real answer and perturb it with noise
    - Add (query, perturbed answer) to the dictionary
- Good in theory, but impractical
- Fundamental CS result on undecidability
  - For complex query systems, it is impossible to design a single algorithm to determine if two queries are the same

Operationalizing Private Release

Moral:

These 3 “bad” ideas, show that we have to be careful about operationalizing private release

Desiderata of Private Analysis

[From Yuxiang Wang]

“I would be comfortable giving up my data for a study if…”

Desiderata of Private Analysis

[From Yuxiang Wang]

“I would be comfortable giving up my data for a study if…”

- Idea 1: My answer had no impact on the result

Desiderata of Private Analysis

[From Yuxiang Wang]

“I would be comfortable giving up my data for a study if…”

- Idea 1: My answer had no impact on the result
    - Constructing a mechanism with this property is pointless

Desiderata of Private Analysis

[From Yuxiang Wang]

“I would be comfortable giving up my data for a study if…”

- Idea 1: My answer had no impact on the result
    - Constructing a mechanism with this property is pointless
    - By induction, it follows that the output of the mechanism does not depend on anyone’s data

Desiderata of Private Analysis

[From Yuxiang Wang]

“I would be comfortable giving up my data for a study if…”

- Idea 1: My answer had no impact on the result
    - Constructing a mechanism with this property is pointless
    - By induction, it follows that the output of the mechanism does not depend on anyone’s data
- Idea 2: Someone couldn’t learn anything new about me by looking at outputted result

Desiderata of Private Analysis

[From Yuxiang Wang]

“I would be comfortable giving up my data for a study if…”

- Idea 1: My answer had no impact on the result
    - Constructing a mechanism with this property is pointless
    - By induction, it follows that the output of the mechanism does not depend on anyone’s data
- Idea 2: Someone couldn’t learn anything new about me by looking at outputted result
    - This is infeasible, given background knowledge

Desiderata of Private Analysis

[From Yuxiang Wang]

“I would be comfortable giving up my data for a study if…”

- Idea 1: My answer had no impact on the result
    - Constructing a mechanism with this property is pointless
    - By induction, it follows that the output of the mechanism does not depend on anyone’s data
- Idea 2: Someone couldn’t learn anything new about me by looking at outputted result
    - This is infeasible, given background knowledge
- So, we cannot guarantee either of these

Desiderata of Private Analysis

[From Yuxiang Wang]

“I would be comfortable giving up my data for a study if…”

- Idea 1: My answer had no impact on the result
  - Constructing a mechanism with this property is pointless
  - By induction, it follows that the output of the mechanism does not depend on anyone’s data
- Idea 2: Someone couldn’t learn anything new about me by looking at outputted result
  - This is infeasible, given background knowledge
- So, we cannot guarantee either of these
- Note that (1) & (2) are deterministic statements

Desiderata of Private Analysis

[From Yuxiang Wang]

“I would be comfortable giving up my data for a study if…”

- Idea 1: My answer had no impact on the result
    - Constructing a mechanism with this property is pointless
    - By induction, it follows that the output of the mechanism does not depend on anyone’s data
- Idea 2: Someone couldn’t learn anything new about me by looking at outputted result
    - This is infeasible, given background knowledge
- So, we cannot guarantee either of these
- Note that (1) & (2) are deterministic statements
- Let’s try a probabilistic approach instead

Desiderata of Private Analysis

I would feel comfortable giving up my information for a study if my risk of any outcome does not "change too much" by participation

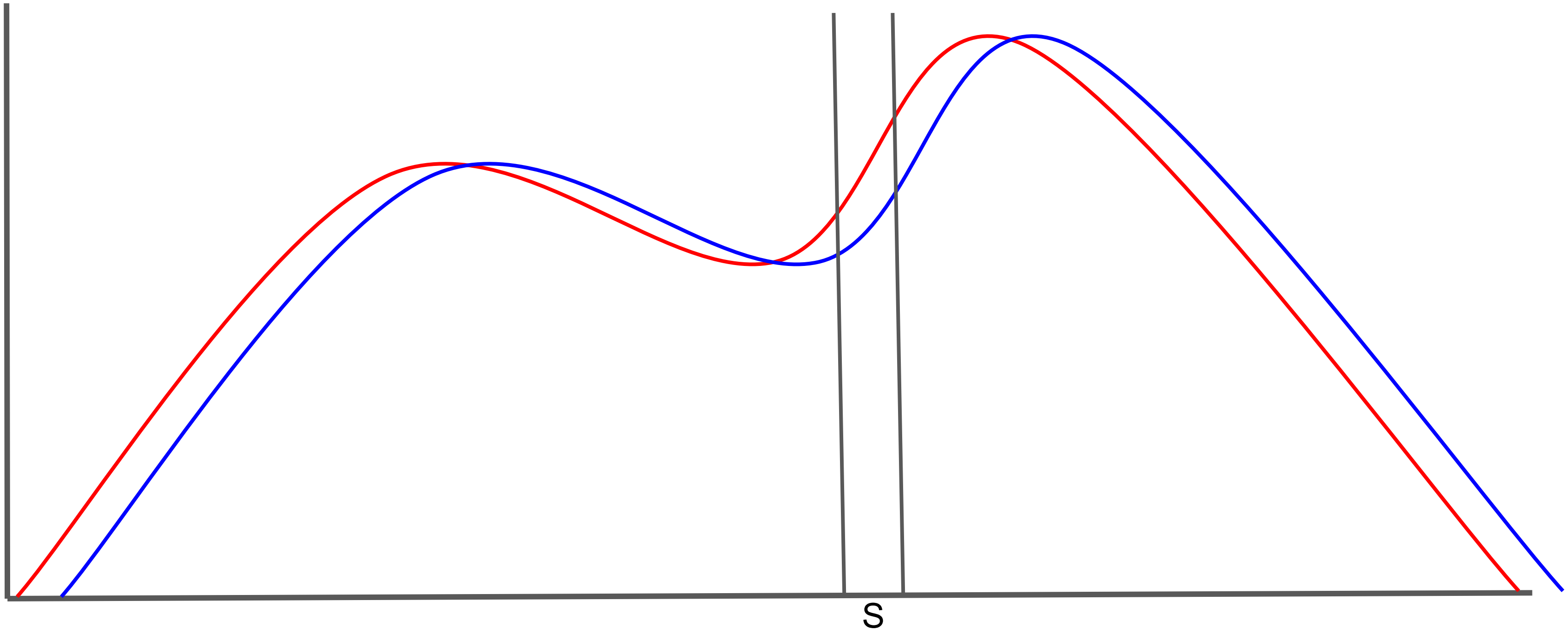
Desiderata of Private Analysis

I would feel comfortable giving up my information for a study if my risk of any outcome does not "change too much" by participation

- What do we mean by "too much?"

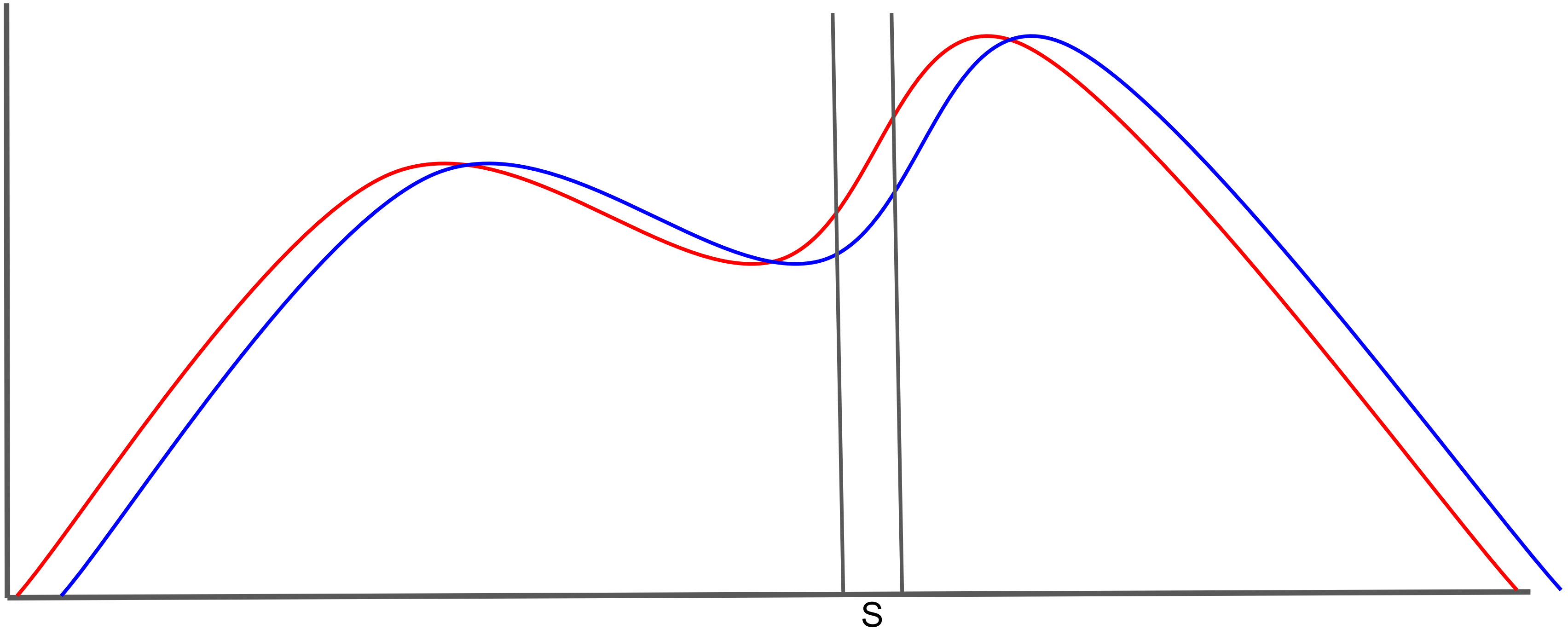
Differential Privacy - “Statistical Indistinguishability”

- Let V and V’ be two distributions



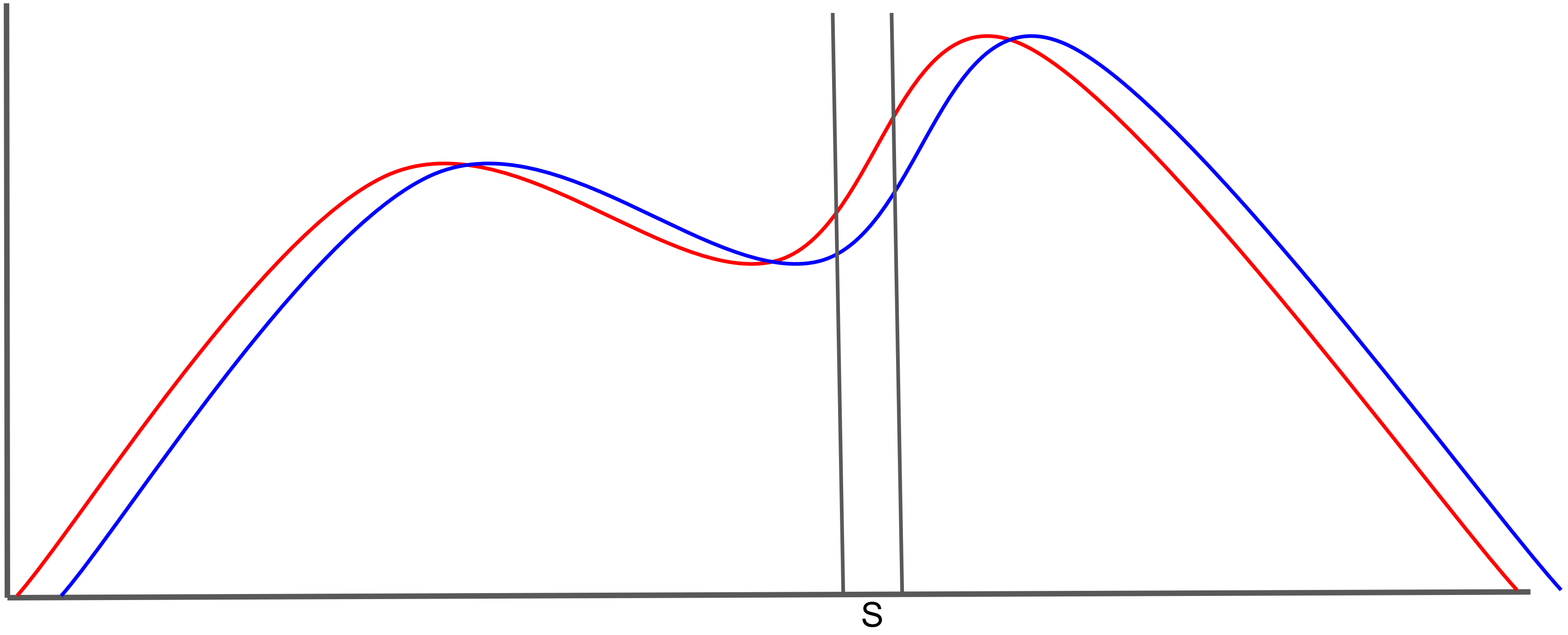
Differential Privacy - “Statistical Indistinguishability”

- Let V and V’ be two distributions
- “Statistical indistinguishably” refers to the distributions V and V’ being “close enough” to one another for all outcomes S



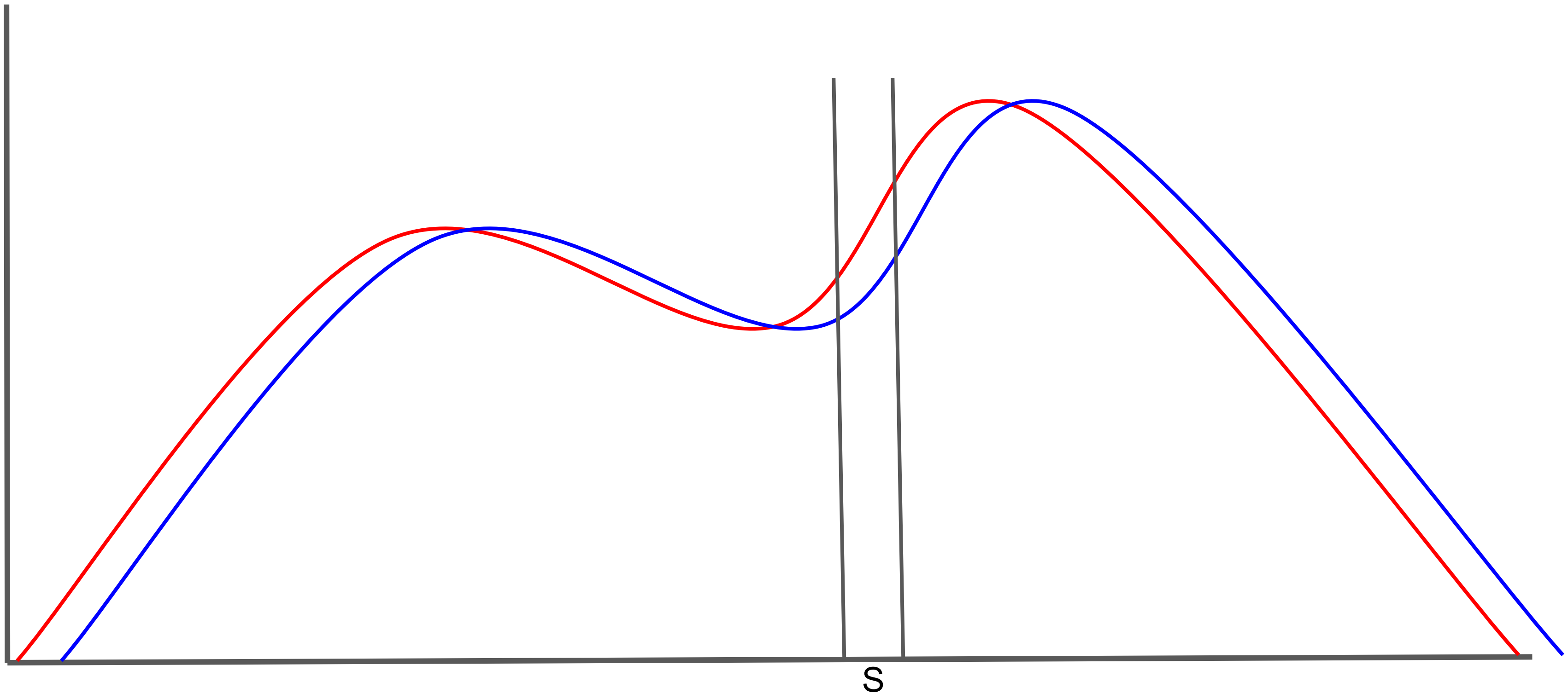
Differential Privacy - “Statistical Indistinguishability”

- Let V and V’ be two distributions
- “Statistical indistinguishably” refers to the distributions V and V’ being “close enough” to one another for all outcomes S
- For continuous distributions, sufficient to have the heights of the PDFs “close enough”



Differential Privacy - “Statistical Indistinguishability”

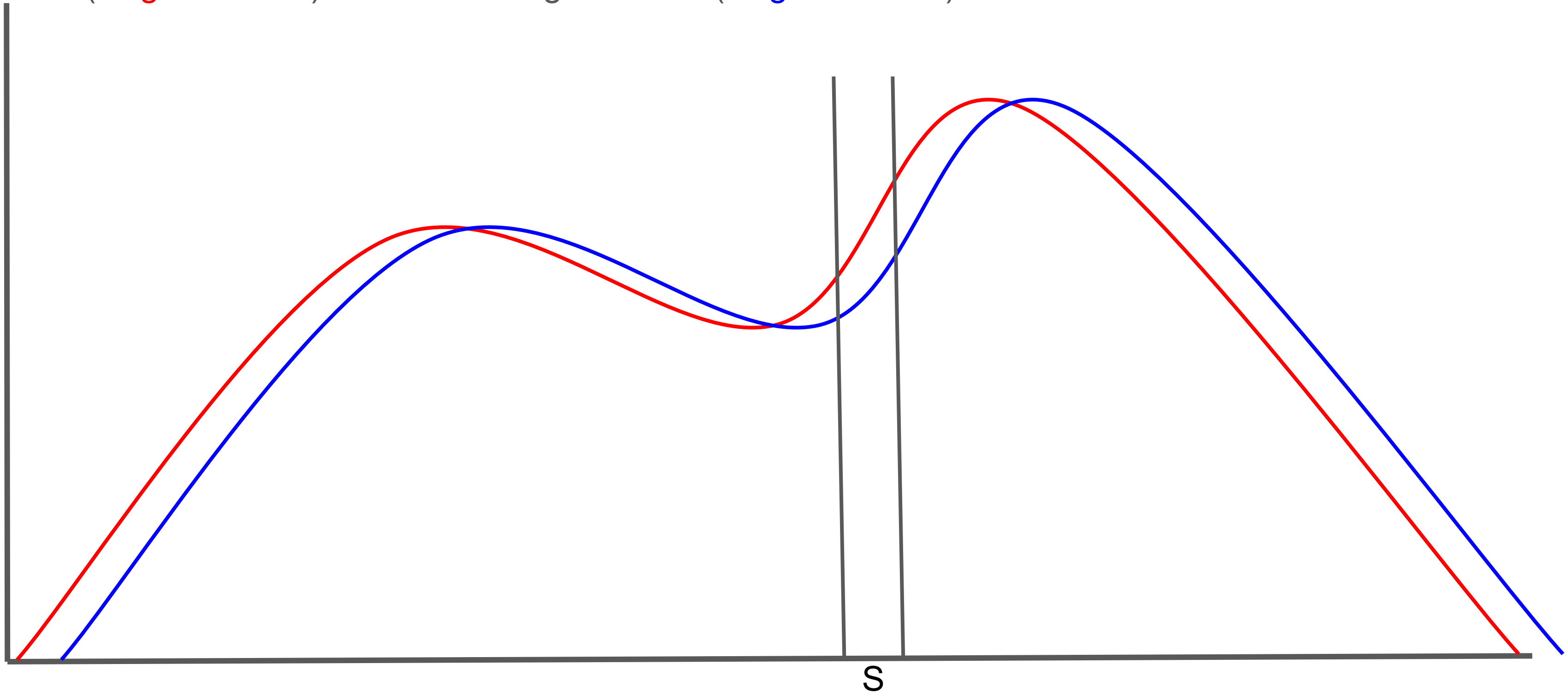
P(S Red) is “close enough” to P(S Blue)



Differential Privacy - “Statistical Indistinguishability”

P(S Red) is “close enough” to P(S Blue)

Area(Region S Red) is “close enough” to Area(Region S Blue)



Differential Privacy Formalizes Statistical Indistinguishability

A mechanism M is ε-Differentially Private if for all datasets D and D' that differ on exactly one element, and for all measurable sets S in the $Codomain(M)$,

$$P(M(D) \in S) \leq e^{\varepsilon} P(M(D') \in S)$$

Differential Privacy Formalizes Statistical Indistinguishability

A mechanism M is ε-Differentially Private if for all datasets D and D' that differ on exactly one element, and for all measurable sets S in the $Codomain(M)$,

$$P(M(D) \in S) \leq e^{\varepsilon} P(M(D') \in S)$$

By symmetry,

$$P(M(D') \in S) \leq e^{\varepsilon} P(M(D) \in S)$$

Differential Privacy Formalizes Statistical Indistinguishability

A mechanism M is ε-Differentially Private if for all datasets D and D' that differ on exactly one element, and for all measurable sets S in the $Codomain(M)$,

$$P(M(D) \in S) \leq e^{\varepsilon} P(M(D') \in S)$$

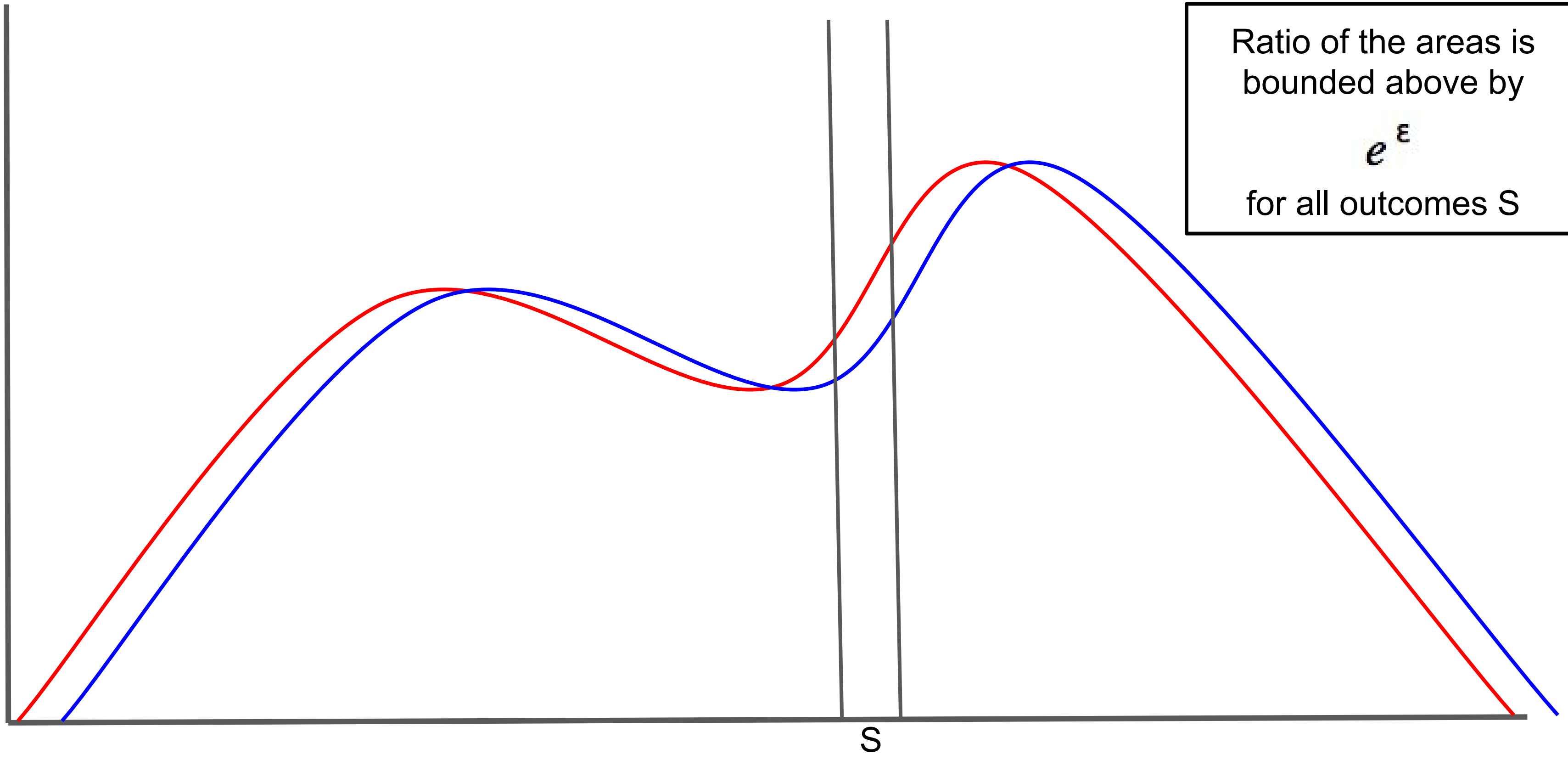
By symmetry,

$$P(M(D') \in S) \leq e^{\varepsilon} P(M(D) \in S)$$

Therefore,

$$e^{-\varepsilon} \leq \frac{P(M(D) \in S)}{P(M(D') \in S)} \leq e^{\varepsilon}$$

Differential Privacy Formalizes Statistical Indistinguishability



Ratio of the areas is bounded above by

e^{ε}

for all outcomes S

Differential Privacy in the Felon Example

Differential Privacy in the Felon Example

Suppose I am considering joining the Felon study

Differential Privacy in the Felon Example

Suppose I am considering joining the Felon study

Let D be the data when I opt-in to the study

Differential Privacy in the Felon Example

Suppose I am considering joining the Felon study

Let D be the data when I opt-in to the study

Let D' be the data when I opt-out of the study

Differential Privacy in the Felon Example

Suppose I am considering joining the Felon study

Let D be the data when I opt-in to the study
Let D' be the data when I opt-out of the study

$$P(M(D) \in S) \leq e^{\varepsilon} P(M(D') \in S)$$

Differential Privacy in the Felon Example

Suppose I am considering joining the Felon study

Let D be the data when I opt-in to the study
Let D' be the data when I opt-out of the study

$$P(M(D) \in S) \leq e^{\varepsilon} P(M(D') \in S)$$

Differential Privacy states the chance of any outcome of study cannot change by more than e^{ε}

Differential Privacy in the Felon Example

Suppose I am considering joining the Felon study

Let D be the data when I opt-in to the study
Let D' be the data when I opt-out of the study

$$P(M(D) \in S) \leq e^{\varepsilon} P(M(D') \in S)$$

Differential Privacy states the chance of any outcome of study cannot change by more than e^{ε}

My value impacts the result by no more than e^{ε}

Differential Privacy in the Felon Example

Suppose I am considering joining the Felon study

Let D be the data when I opt-in to the study
Let D' be the data when I opt-out of the study

$$P(M(D) \in S) \leq e^{\varepsilon} P(M(D') \in S)$$

Differential Privacy states the chance of any outcome of study cannot change by more than e^{ε}

My value impacts the result by no more than e^{ε}
This is the probabilistic version of "my value does not impact the result at all"

Differential Privacy in the Felon Example

We can view this in terms of harm to an individual

Differential Privacy in the Felon Example

We can view this in terms of harm to an individual

Let S be a “bad event”

Differential Privacy in the Felon Example

We can view this in terms of harm to an individual

Let S be a “bad event”

$$P(M(D) \in S) \leq e^{\varepsilon} P(M(D') \in S)$$

Differential Privacy in the Felon Example

We can view this in terms of harm to an individual

Let S be a “bad event”

$$P(M(D) \in S) \leq e^{\varepsilon} P(M(D') \in S)$$

<u>**Formally**</u>:

The harm from opting-in to a study vs opting-out is bounded above by e^{ε}

Differential Privacy in the Felon Example

We can view this in terms of harm to an individual

Let S be a "bad event"

$$P(M(D) \in S) \leq e^{\varepsilon} P(M(D') \in S)$$

<u>Formally</u>:

The harm from opting-in to a study vs opting-out is bounded above by e^{ε}

<u>Informally</u>:

The harm is almost the same regardless of participation

Differential Privacy: Claims and Non-Claims

- Differential Privacy limits the harm from participating in a study

Differential Privacy: Claims and Non-Claims

- Differential Privacy limits the harm from participating in a study
- Differential Privacy does not eliminate harms from occurring

Differential Privacy: Claims and Non-Claims

- Differential Privacy limits the harm from participating in a study
- Differential Privacy does not eliminate harms from occurring
- Differential Privacy does not prevent harm from background knowledge

Differential Privacy: Claims and Non-Claims

- Differential Privacy limits the harm from participating in a study
- Differential Privacy does not eliminate harms from occurring
- Differential Privacy does not prevent harm from background knowledge
- It limits them

Differential Privacy: Claims and Non-Claims

- Differential Privacy limits the harm from participating in a study
- Differential Privacy does not eliminate harms from occurring
- Differential Privacy does not prevent harm from background knowledge
- It limits them

Example:

Differential Privacy: Claims and Non-Claims

- Differential Privacy limits the harm from participating in a study
- Differential Privacy does not eliminate harms from occurring
- Differential Privacy does not prevent harm from background knowledge
- It limits them

Example:

- Bob is a felon

Differential Privacy: Claims and Non-Claims

- Differential Privacy limits the harm from participating in a study
- Differential Privacy does not eliminate harms from occurring
- Differential Privacy does not prevent harm from background knowledge
- It limits them

Example:

- Bob is a felon
- A study finds that felons are less likely to be hired for corporate jobs than non-felons

Differential Privacy: Claims and Non-Claims

- Differential Privacy limits the harm from participating in a study
- Differential Privacy does not eliminate harms from occurring
- Differential Privacy does not prevent harm from background knowledge
- It limits them

Example:

- Bob is a felon
- A study finds that felons are less likely to be hired for corporate jobs than non-felons
- Bob is likely to have trouble finding a corporate job, regardless if he joined the study or not

Differential Privacy: Claims and Non-Claims

- Differential Privacy limits the harm from participating in a study
- Differential Privacy does not eliminate harms from occurring
- Differential Privacy does not prevent harm from background knowledge
- It limits them

Example:

- Bob is a felon
- A study finds that felons are less likely to be hired for corporate jobs than non-felons
- Bob is likely to have trouble finding a corporate job, regardless if he joined the study or not

Limits the harm from participation

Differential Privacy: Claims and Non-Claims

- Differential Privacy limits the harm from participating in a study
- Differential Privacy does not eliminate harms from occurring
- Differential Privacy does not prevent harm from background knowledge
- It limits them

Example:

- Bob is a felon
- A study finds that felons are less likely to be hired for corporate jobs than non-felons
- Bob is likely to have trouble finding a corporate job, regardless if he joined the study or not

Limits the harm from participation

It does not eliminate the harms from occurring

First-Order Viewpoint

Using a first-order Taylor expansion,

$$e^{\varepsilon} \cong 1 + \varepsilon$$

First-Order Viewpoint

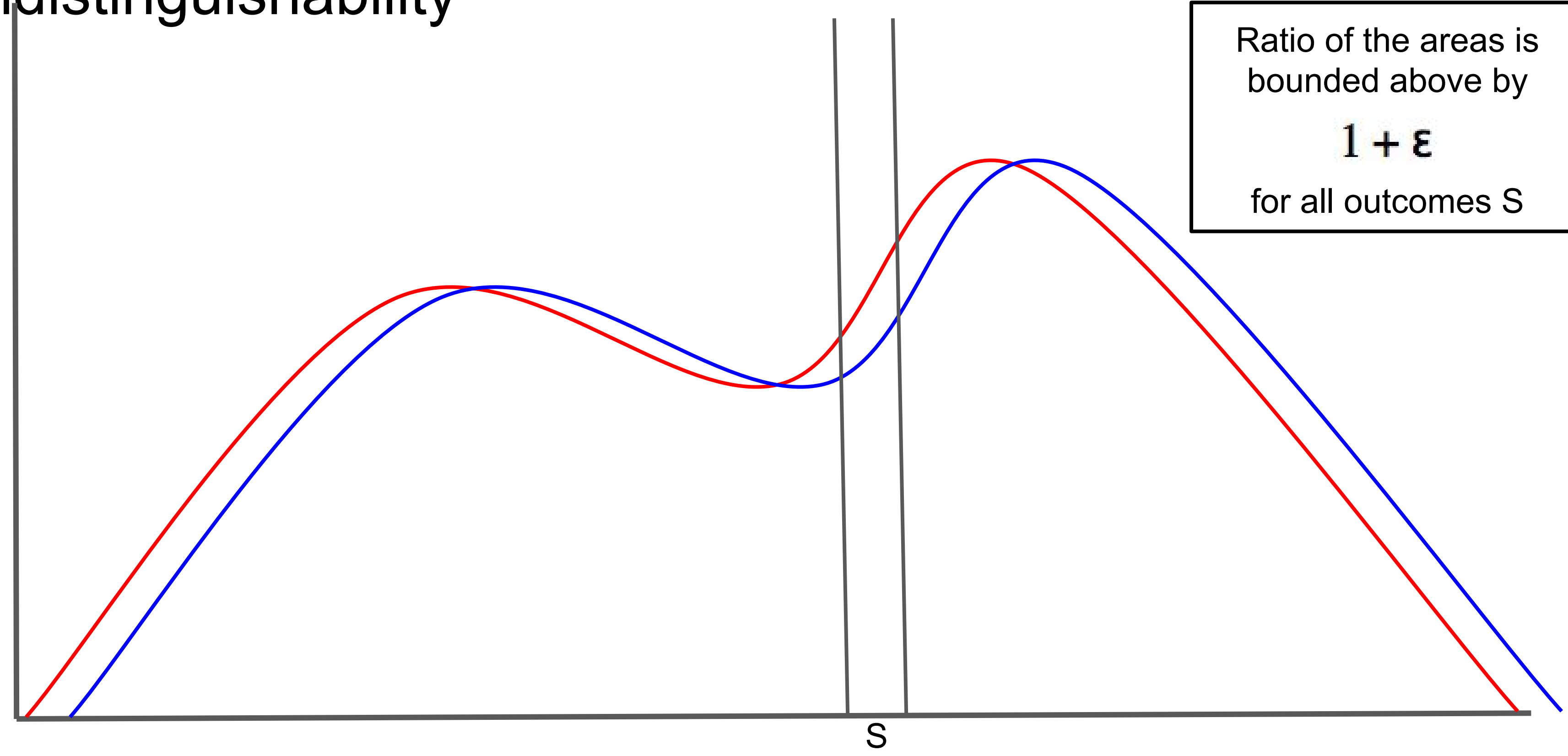
Using a first-order Taylor expansion,

$$e^{\varepsilon} \cong 1 + \varepsilon$$

Therefore,

$$1 - \varepsilon \leq \frac{P(M(D) \in S)}{P(M(D') \in S)} \leq 1 + \varepsilon$$

Differential Privacy Formalizes Statistical Indistinguishability



Ratio of the areas is bounded above by

$$1 + \varepsilon$$

for all outcomes S

First-Order Viewpoint

Using a first-order Taylor expansion,

$$e^{\varepsilon} \cong 1 + \varepsilon$$

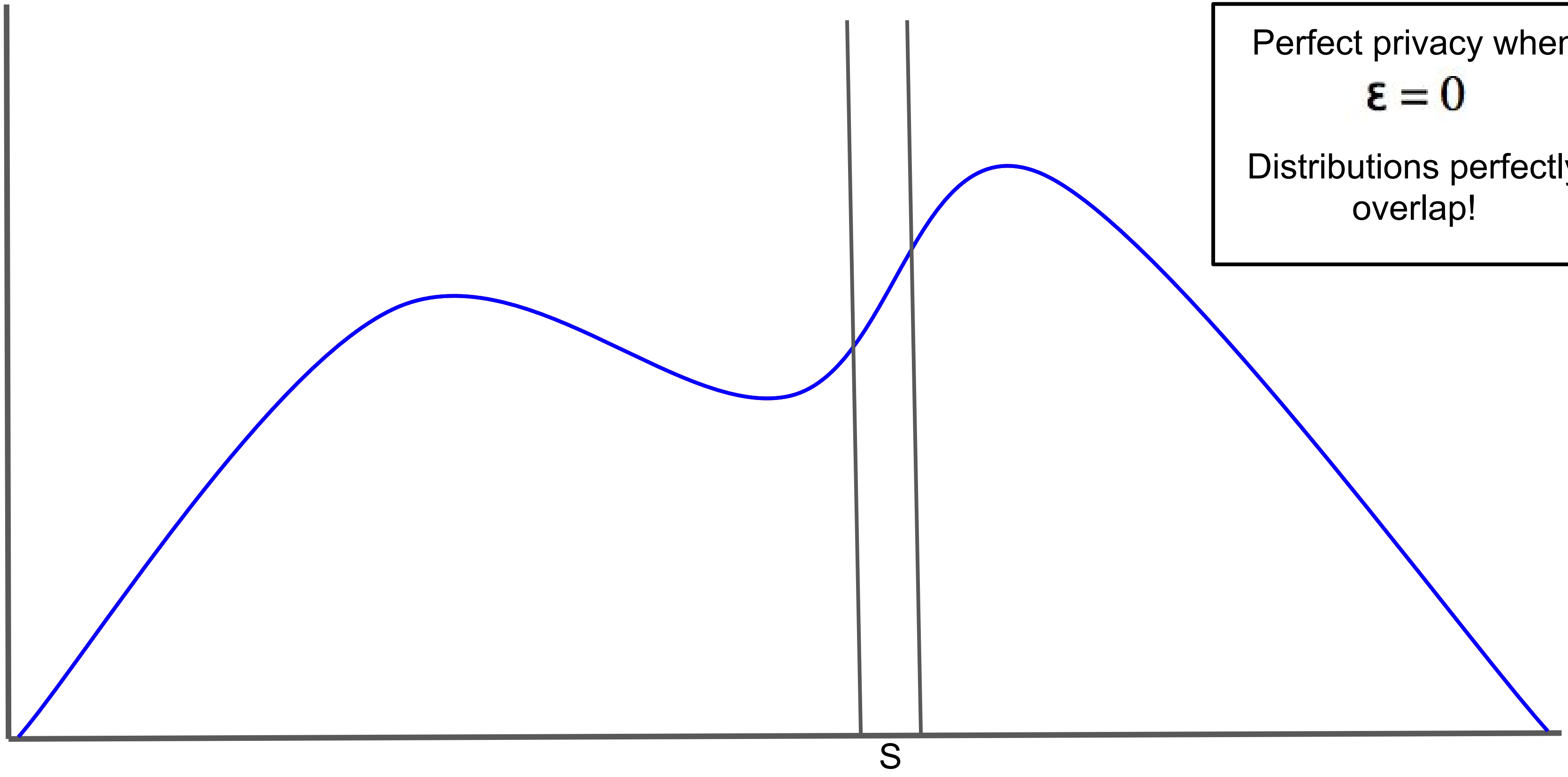
Therefore,

$$1 - \varepsilon \leq \frac{P(M(D) \in S)}{P(M(D') \in S)} \leq 1 + \varepsilon$$

When ε is very small (i.e. close to 0)

$$1 \leq \frac{P(M(D) \in S)}{P(M(D') \in S)} \leq 1$$

As epsilon goes 0, ratio of distributions approaches 1



Perfect privacy when

$$\varepsilon = 0$$

Distributions perfectly overlap!

Epsilon controls the level of privacy

- Small epsilon = More privacy

Epsilon controls the level of privacy

- Small epsilon = More privacy
  - Opt-in distribution and opt-out distribution are close to one another

Epsilon controls the level of privacy

- Small epsilon = More privacy
  - Opt-in distribution and opt-out distribution are close to one another
  - Harder for any one individual to significantly change the outcome

Epsilon controls the level of privacy

- Small epsilon = More privacy
    - Opt-in distribution and opt-out distribution are close to one another
    - Harder for any one individual to significantly change the outcome
    - Tradeoff - analysis has less utility

Epsilon controls the level of privacy

- Small epsilon = More privacy
    - Opt-in distribution and opt-out distribution are close to one another
    - Harder for any one individual to significantly change the outcome
    - Tradeoff - analysis has less utility
- Large epsilon = Less privacy

Epsilon controls the level of privacy

- Small epsilon = More privacy
    - Opt-in distribution and opt-out distribution are close to one another
    - Harder for any one individual to significantly change the outcome
    - Tradeoff - analysis has less utility
- Large epsilon = Less privacy
    - Opt-in distribution and opt-out distribution are farther apart
    - Easier for any one individual to significantly change the outcome
    - Tradeoff- analysis has more utility

Epsilon controls the level of privacy

- Small epsilon = More privacy
  - Opt-in distribution and opt-out distribution are close to one another
  - Harder for any one individual to significantly change the outcome
  - Tradeoff - analysis has less utility
- Large epsilon = Less privacy
  - Opt-in distribution and opt-out distribution are farther apart
  - Easier for any one individual to significantly change the outcome
  - Tradeoff- analysis has more utility

Epsilon allows us to quantify the amount of privacy loss from an analysis!

Epsilon controls the level of privacy

- Small epsilon = More privacy
  - Opt-in distribution and opt-out distribution are close to one another
  - Harder for any one individual to significantly change the outcome
  - Tradeoff - analysis has less utility
- Large epsilon = Less privacy
  - Opt-in distribution and opt-out distribution are farther apart
  - Easier for any one individual to significantly change the outcome
  - Tradeoff- analysis has more utility

Epsilon allows us to quantify the amount of privacy loss from an analysis!

There is a tradeoff between privacy & utility

Summarizing Differential Privacy

Mathematically:

A mechanism M is ε-Differentially Private if for all datasets D and D' that differ on exactly one element, and for all measurable sets S in the $Codomain(M)$,

$$P(M(D) \in S) \leq e^{\varepsilon} P(M(D') \in S)$$

Conceptually: "Statistical Indistinguishability" in outcomes of joining versus refraining from a study

Graphically: The probability distributions of joining versus refraining from a study are "close enough" to one another

Achieving Differential Privacy

- Many ways to achieve Differential Privacy

Achieving Differential Privacy

- Many ways to achieve Differential Privacy
- One method

Achieving Differential Privacy

- Many ways to achieve Differential Privacy
- One method
  - Add random noise to answer

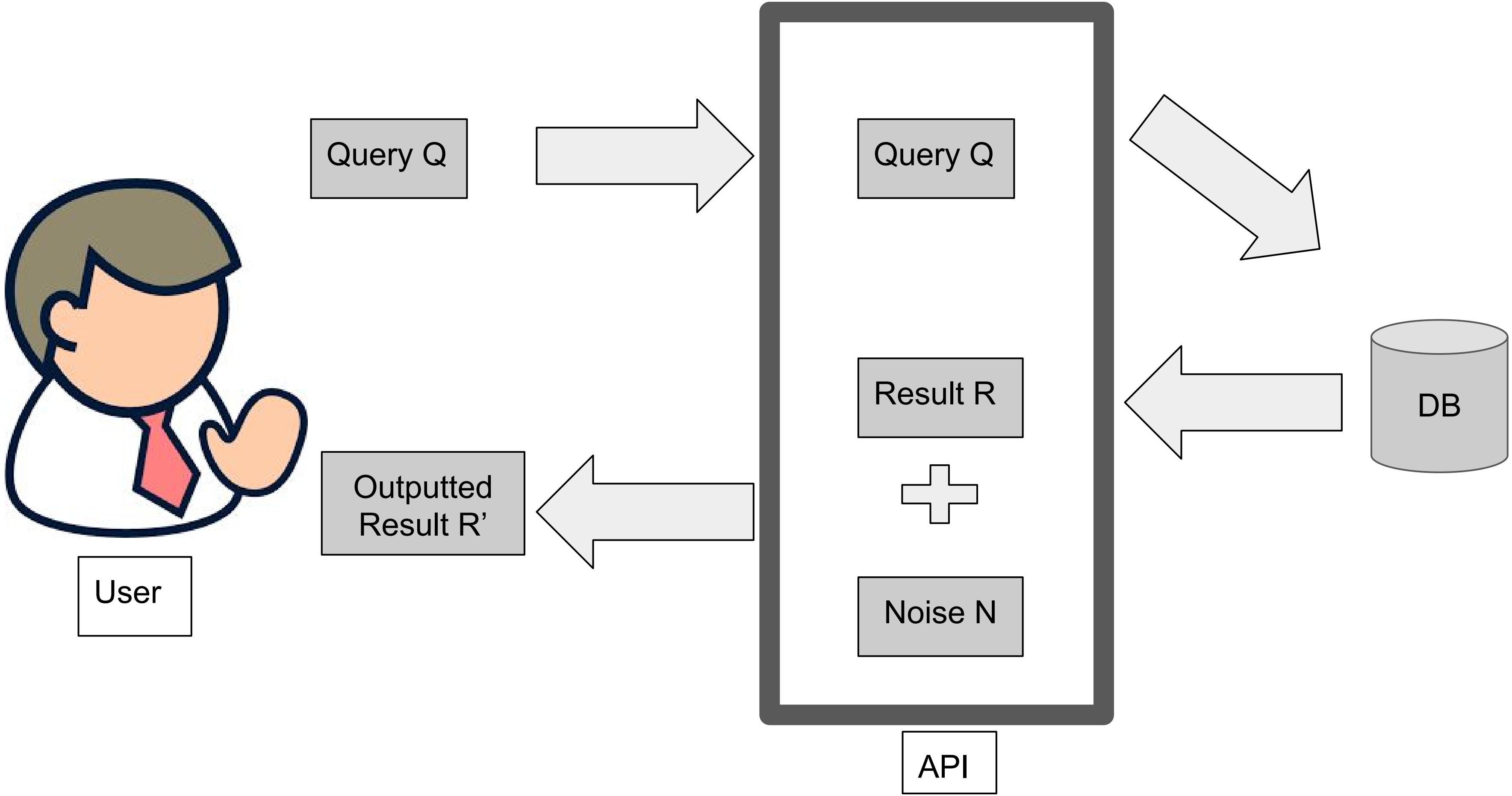
Achieving Differential Privacy

- Many ways to achieve Differential Privacy
- One method
  - Add random noise to answer
  - While limiting the number of queries that can be asked

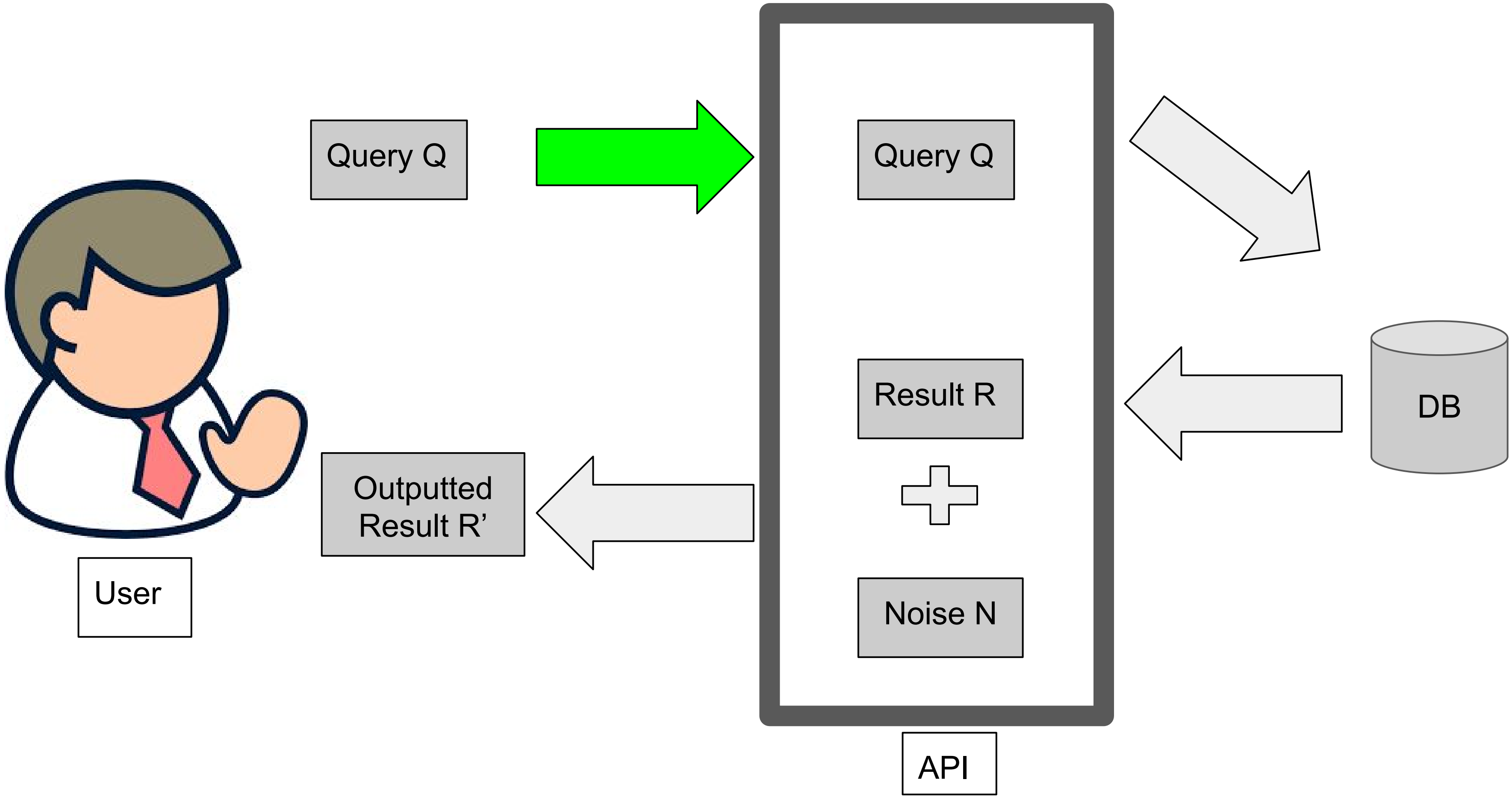
Achieving Differential Privacy

- Many ways to achieve Differential Privacy
- One method
    - Add random noise to answer
    - While limiting the number of queries that can be asked
        - Privacy budget

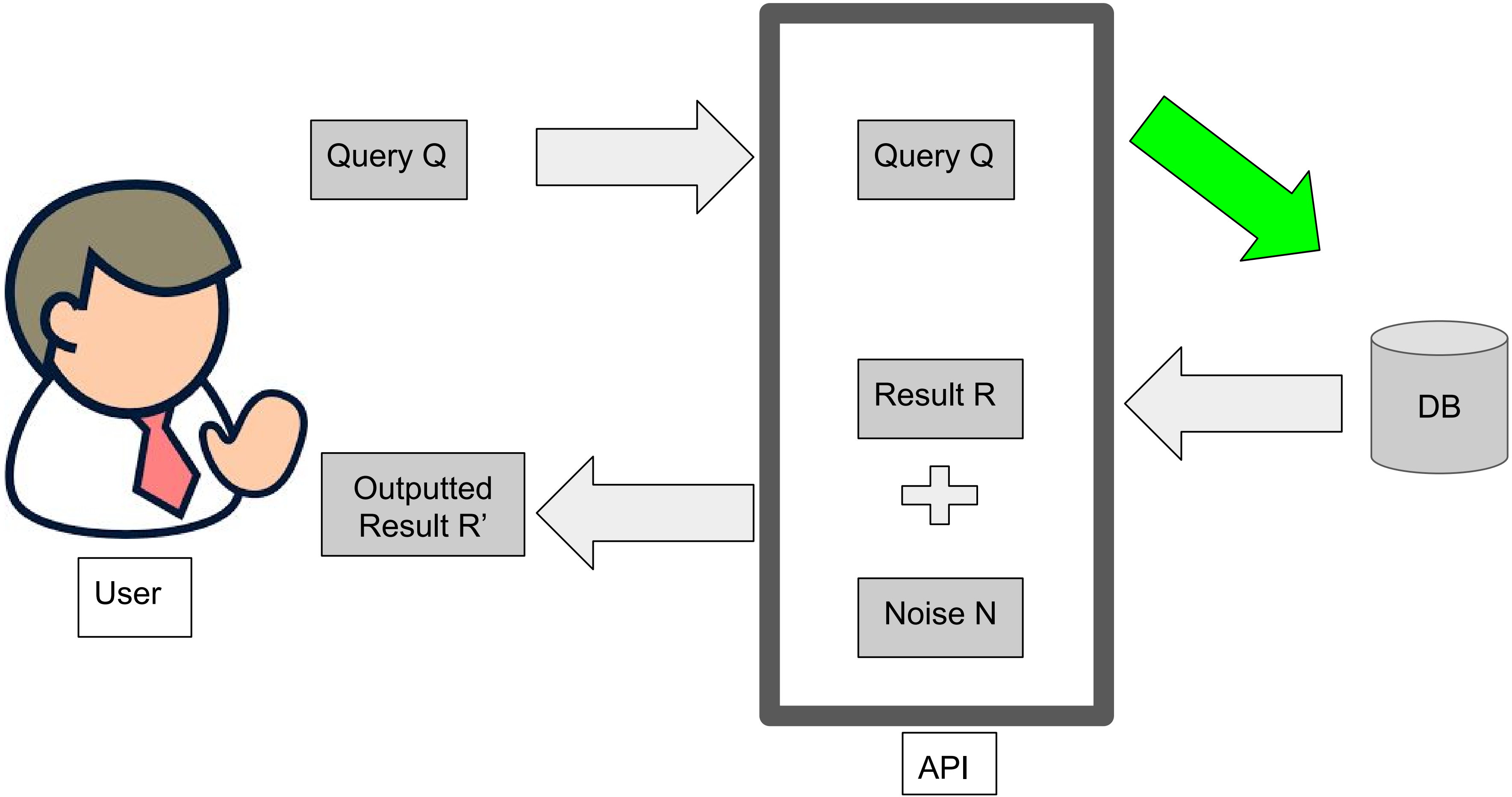
Random Noise User-API Cartoon



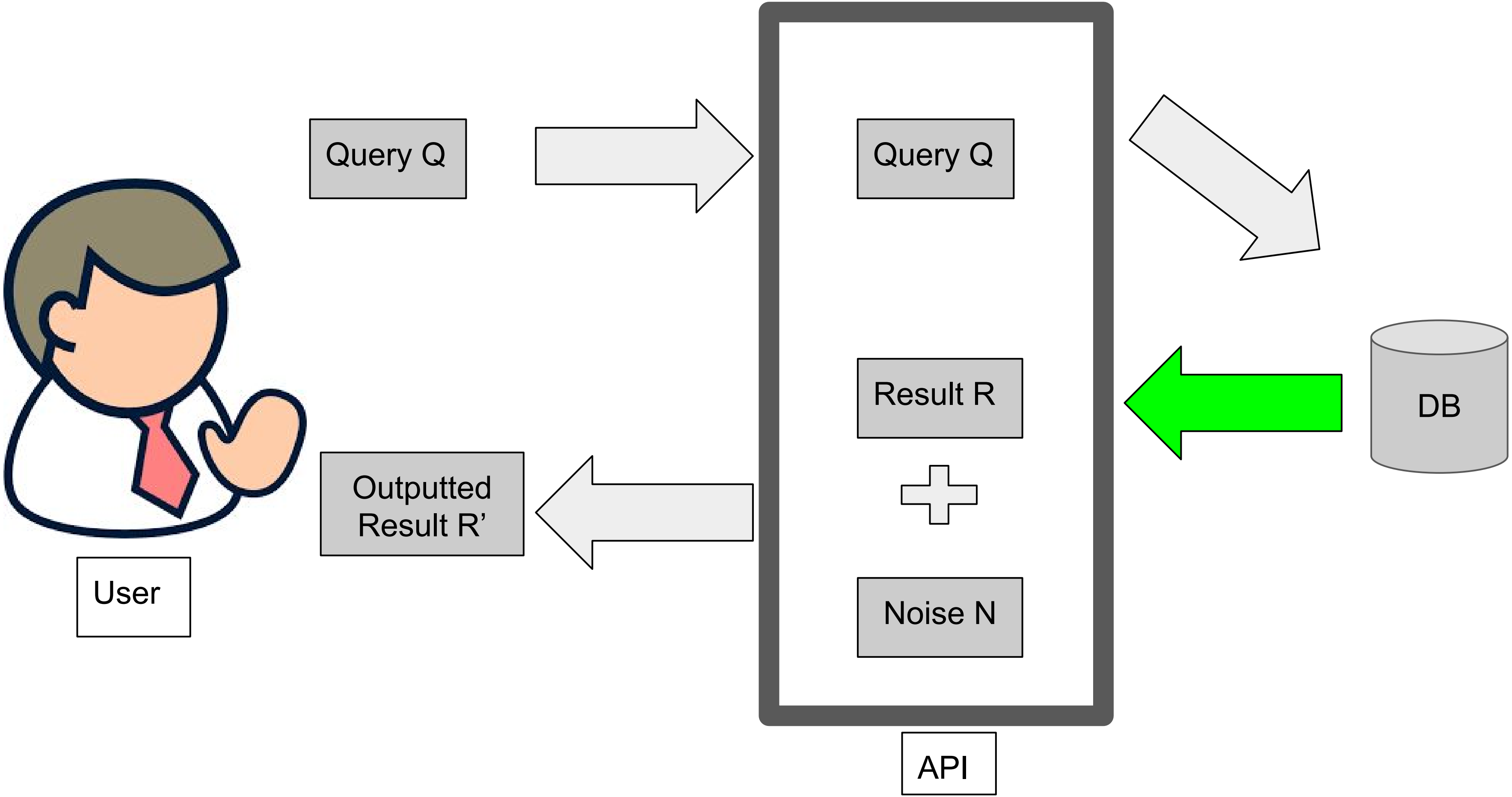
Random Noise User-API Cartoon



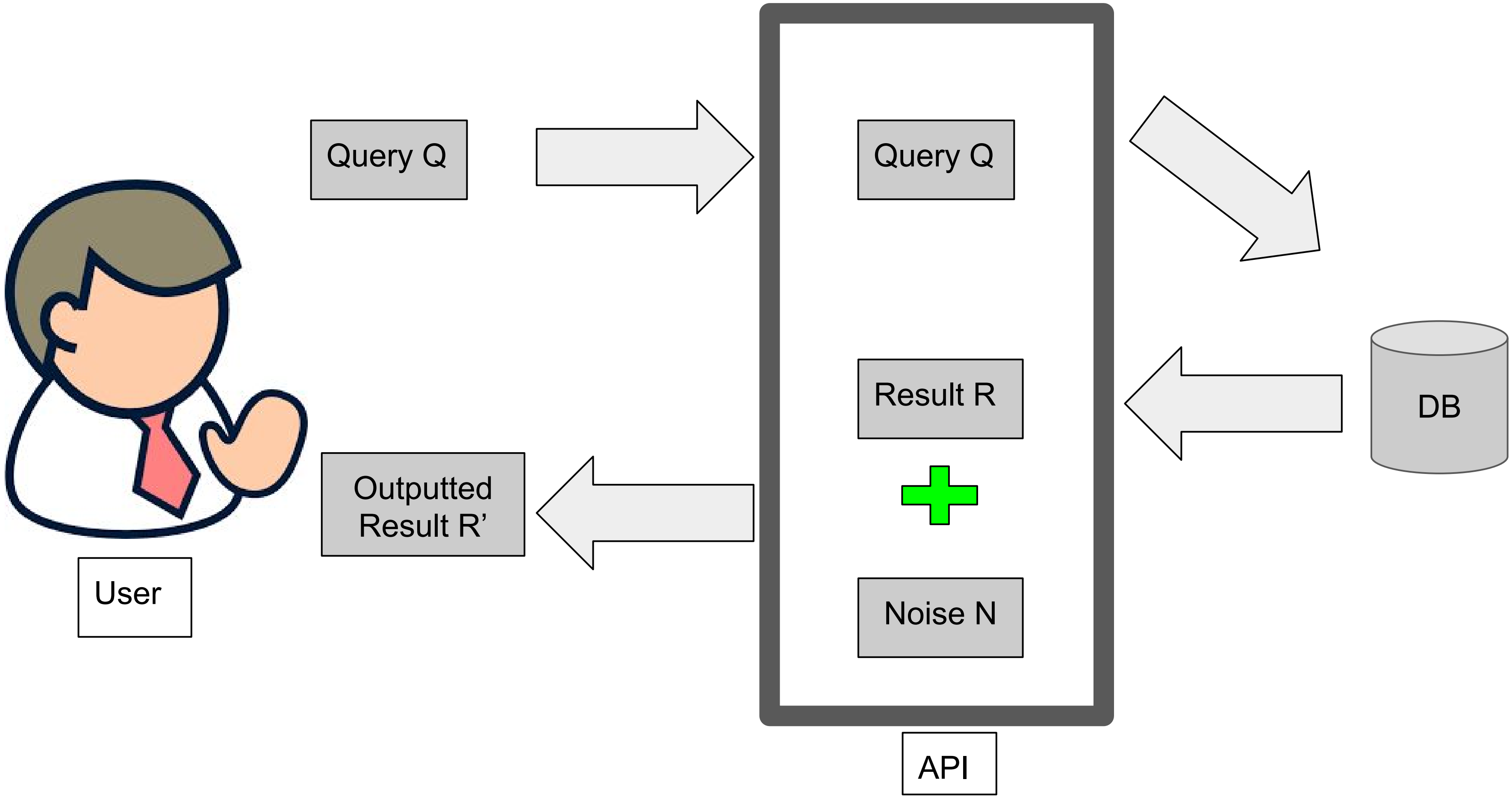
Random Noise User-API Cartoon



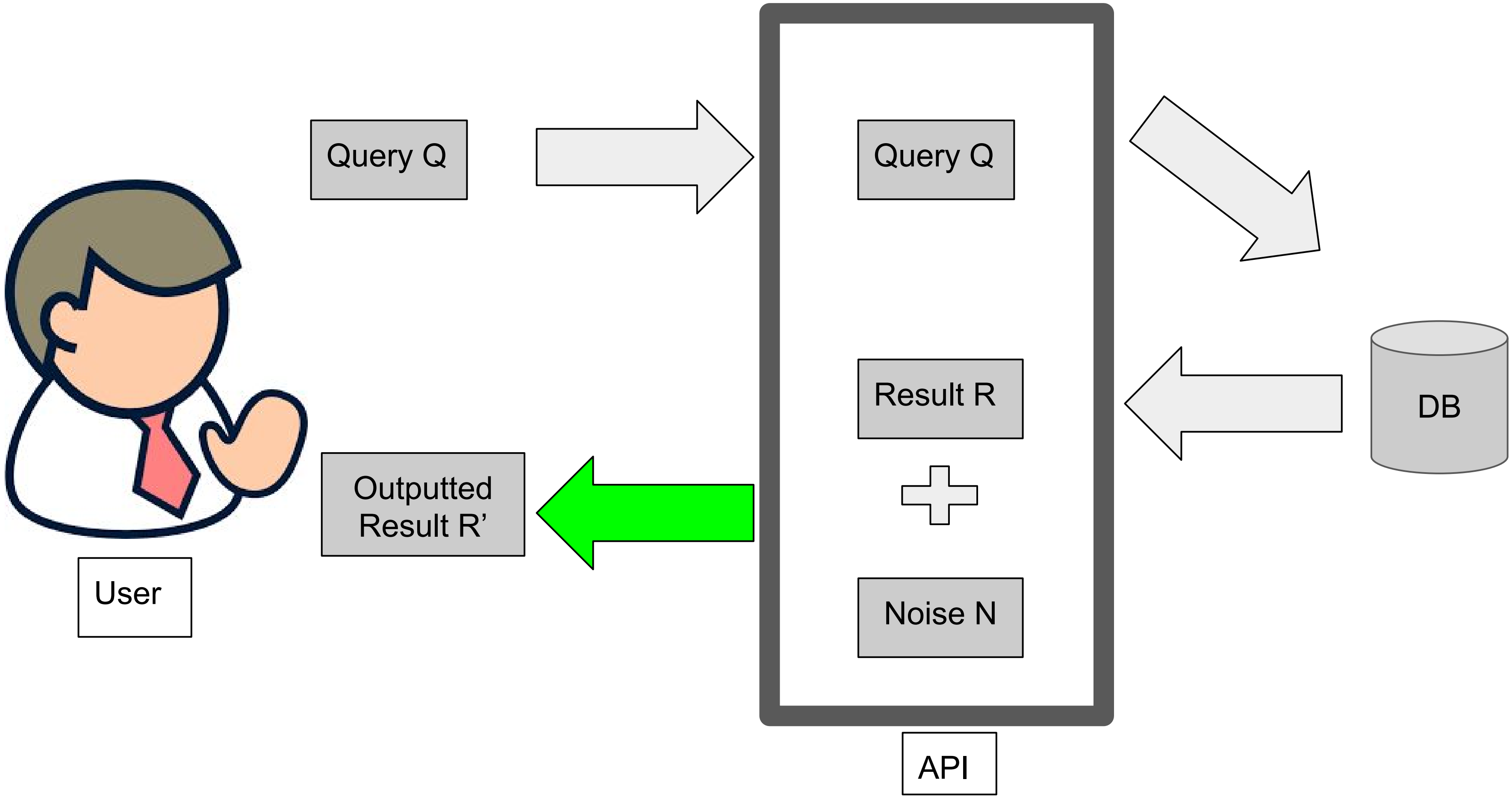
Random Noise User-API Cartoon



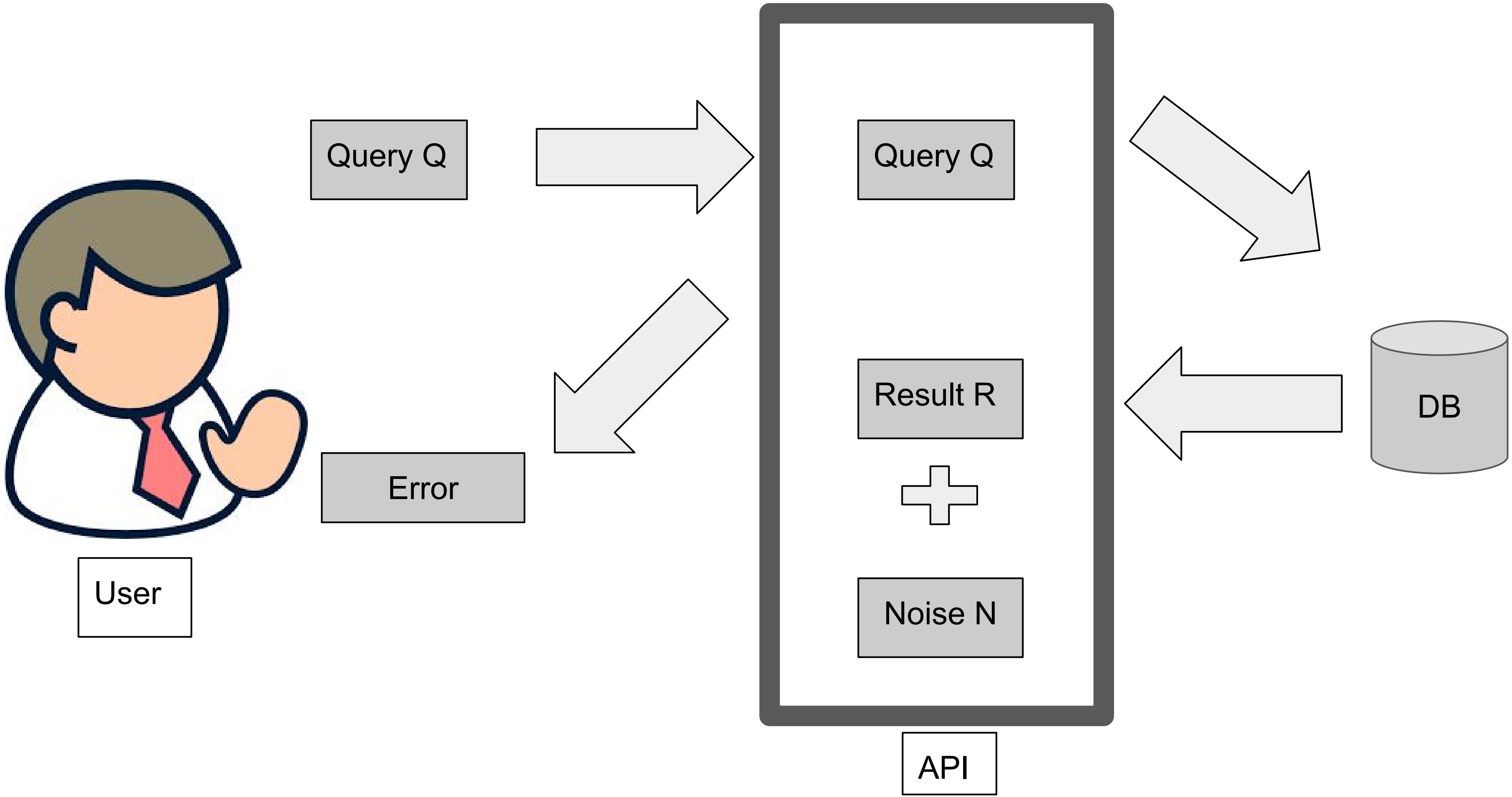
Random Noise User-API Cartoon



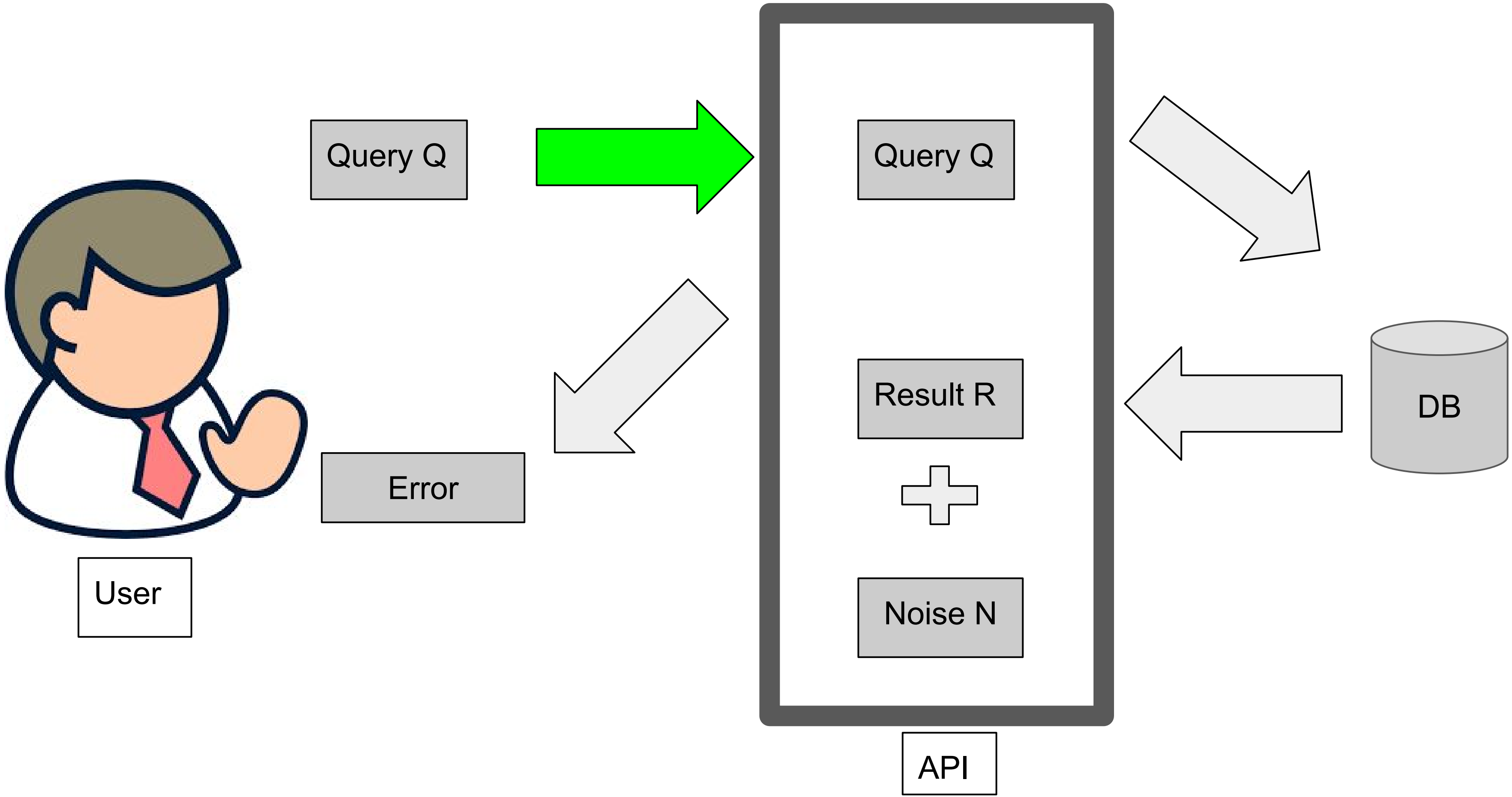
Random Noise User-API Cartoon



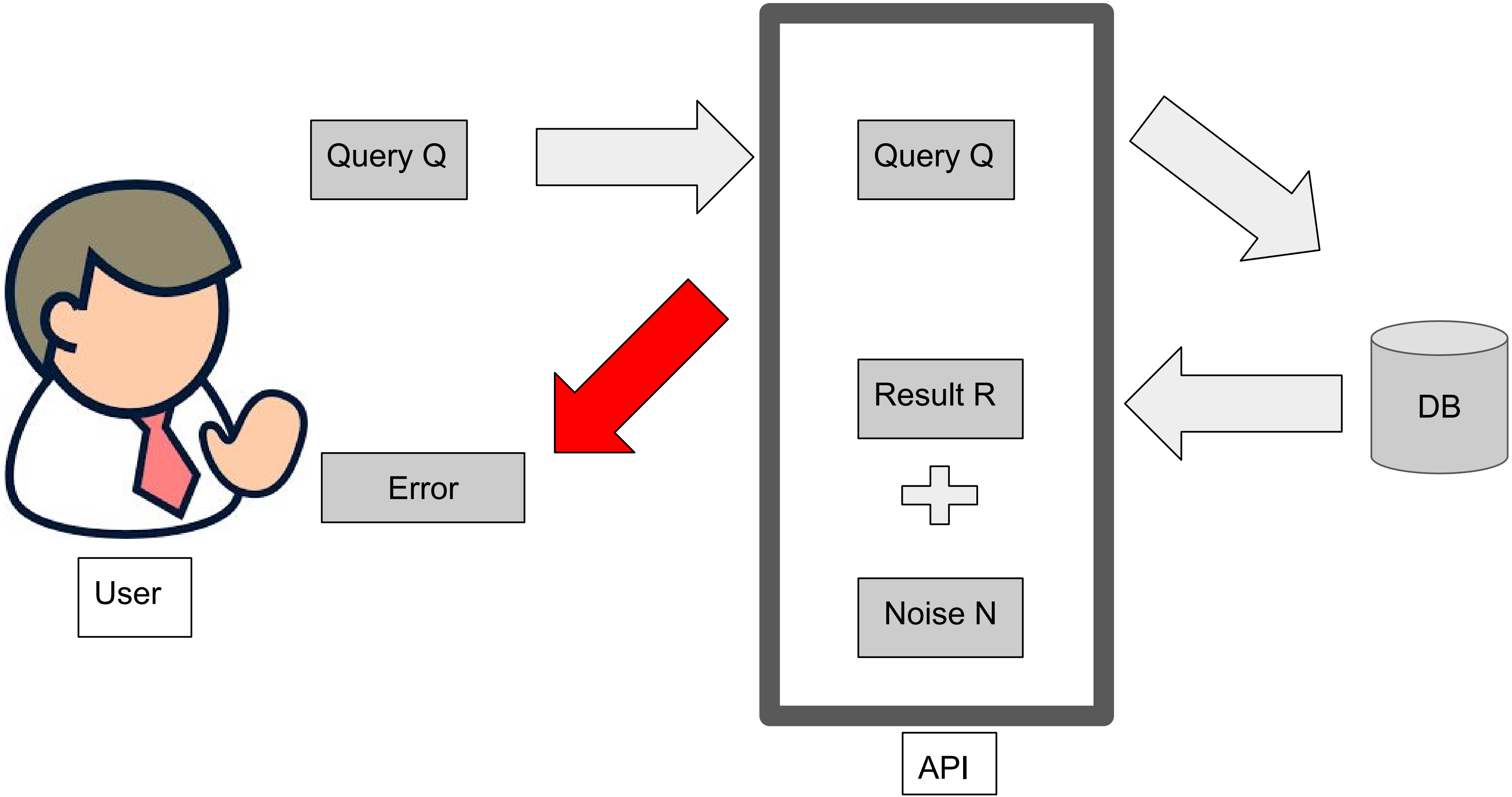
Too many queries are not allowed



Too many queries are not allowed



Too many queries are not allowed



Achieving Differential Privacy

- Common concern

Achieving Differential Privacy

- Common concern
    - “You are giving me the wrong answer”

Achieving Differential Privacy

- Common concern
  - “You are giving me the wrong answer”
- Response 1

Achieving Differential Privacy

- Common concern
  - “You are giving me the wrong answer”
- Response 1
  - This is a choice we make

Achieving Differential Privacy

- Common concern
  - “You are giving me the wrong answer”
- Response 1
  - This is a choice we make
  - We give up some utility (validity) to gain privacy

Achieving Differential Privacy

- Common concern
  - “You are giving me the wrong answer”
- Response 1
  - This is a choice we make
  - We give up some utility (validity) to gain privacy
  - More privacy -> Less utility

Achieving Differential Privacy

- Common concern
  - “You are giving me the wrong answer”
- Response 1
  - This is a choice we make
  - We give up some utility (validity) to gain privacy
  - More privacy -> Less utility
- Response 2

Achieving Differential Privacy

- Common concern
  - "You are giving me the wrong answer"
- Response 1
  - This is a choice we make
  - We give up some utility (validity) to gain privacy
  - More privacy -> Less utility
- Response 2
  - We already have sampling error in our estimates

Achieving Differential Privacy

- Common concern
  - "You are giving me the wrong answer"
- Response 1
  - This is a choice we make
  - We give up some utility (validity) to gain privacy
  - More privacy -> Less utility
- Response 2
  - We already have sampling error in our estimates
  - Sampling error is a randomness error

Achieving Differential Privacy

- Common concern
    - "You are giving me the wrong answer"
- Response 1
    - This is a choice we make
    - We give up some utility (validity) to gain privacy
    - More privacy -> Less utility
- Response 2
    - We already have sampling error in our estimates
    - Sampling error is a randomness error
    - Just like the noise from Differential Privacy

Recap

- Goal: To learn information about a population, while not learning about individuals specifically

Recap

- Goal: To learn information about a population, while not learning about individuals specifically
- Differential Privacy

Recap

- Goal: To learn information about a population, while not learning about individuals specifically
- Differential Privacy
    - Is a property of your mechanism

Recap

- Goal: To learn information about a population, while not learning about individuals specifically
- Differential Privacy
  - Is a property of your mechanism
  - Is not the property of a given dataset

Recap

- Goal: To learn information about a population, while not learning about individuals specifically
- Differential Privacy
    - Is a property of your mechanism
    - Is not the property of a given dataset
- One way to achieve Differential Privacy is with Random Noise

Recap

- Goal: To learn information about a population, while not learning about individuals specifically
- Differential Privacy
  - Is a property of your mechanism
  - Is not the property of a given dataset
- One way to achieve Differential Privacy is with Random Noise
  - Need to prevent against too many queries (Privacy Budget)

Recap

- Goal: To learn information about a population, while not learning about individuals specifically
- Differential Privacy
  - Is a property of your mechanism
  - Is not the property of a given dataset
- One way to achieve Differential Privacy is with Random Noise
  - Need to prevent against too many queries (Privacy Budget)
- There is a tradeoff between Privacy and Utility

Recap

- Goal: To learn information about a population, while not learning about individuals specifically
- Differential Privacy
    - Is a property of your mechanism
    - Is not the property of a given dataset
- One way to achieve Differential Privacy is with Random Noise
    - Need to prevent against too many queries (Privacy Budget)
- There is a tradeoff between Privacy and Utility
    - This is controlled by epsilon

Recap

- Goal: To learn information about a population, while not learning about individuals specifically
- Differential Privacy
    - Is a property of your mechanism
    - Is not the property of a given dataset
- One way to achieve Differential Privacy is with Random Noise
    - Need to prevent against too many queries (Privacy Budget)
- There is a tradeoff between Privacy and Utility
    - This is controlled by epsilon
        - Increase epsilon -> More Utility

Recap

- Goal: To learn information about a population, while not learning about individuals specifically
- Differential Privacy
    - Is a property of your mechanism
    - Is not the property of a given dataset
- One way to achieve Differential Privacy is with Random Noise
    - Need to prevent against too many queries (Privacy Budget)
- There is a tradeoff between Privacy and Utility
    - This is controlled by epsilon
        - Increase epsilon -> More Utility
        - Decrease epsilon -> More Privacy

Recap

- Goal: To learn information about a population, while not learning about individuals specifically
- Differential Privacy
    - Is a property of your mechanism
    - Is not the property of a given dataset
- One way to achieve Differential Privacy is with Random Noise
    - Need to prevent against too many queries (Privacy Budget)
- There is a tradeoff between Privacy and Utility
    - This is controlled by epsilon
        - Increase epsilon -> More Utility
        - Decrease epsilon -> More Privacy
    - More on this in your course readings