

# The Fundamental Limits of Data Privacy

A grayscale photograph showing a hand hovering just above a surface. On the surface, the word "Privacy" is written in a large, bold, sans-serif font. The hand is positioned as if about to touch or hover over the word, creating a sense of tension or focus on the concept of privacy.

*Privacy*

Peter Kairouz  
ECE Department  
University of Illinois at Urbana-Champaign

# Communication



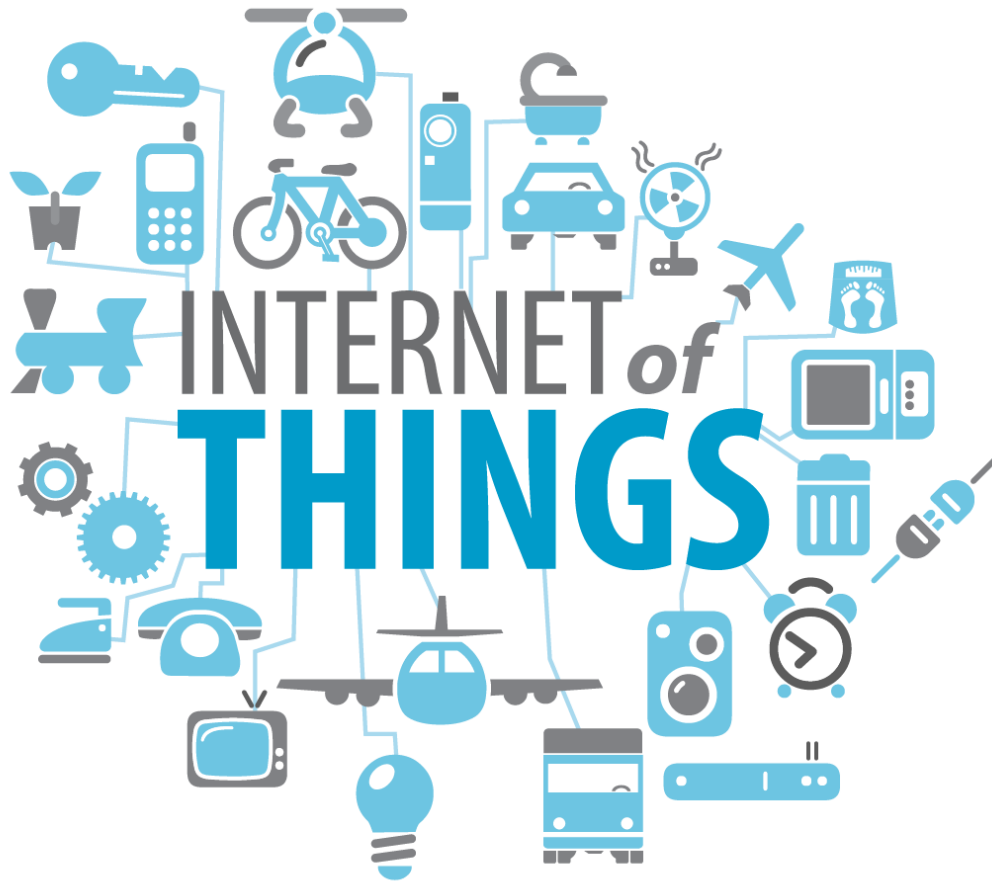
- transfer of information **from one point** in space-time **to the other**

# Wireless communication

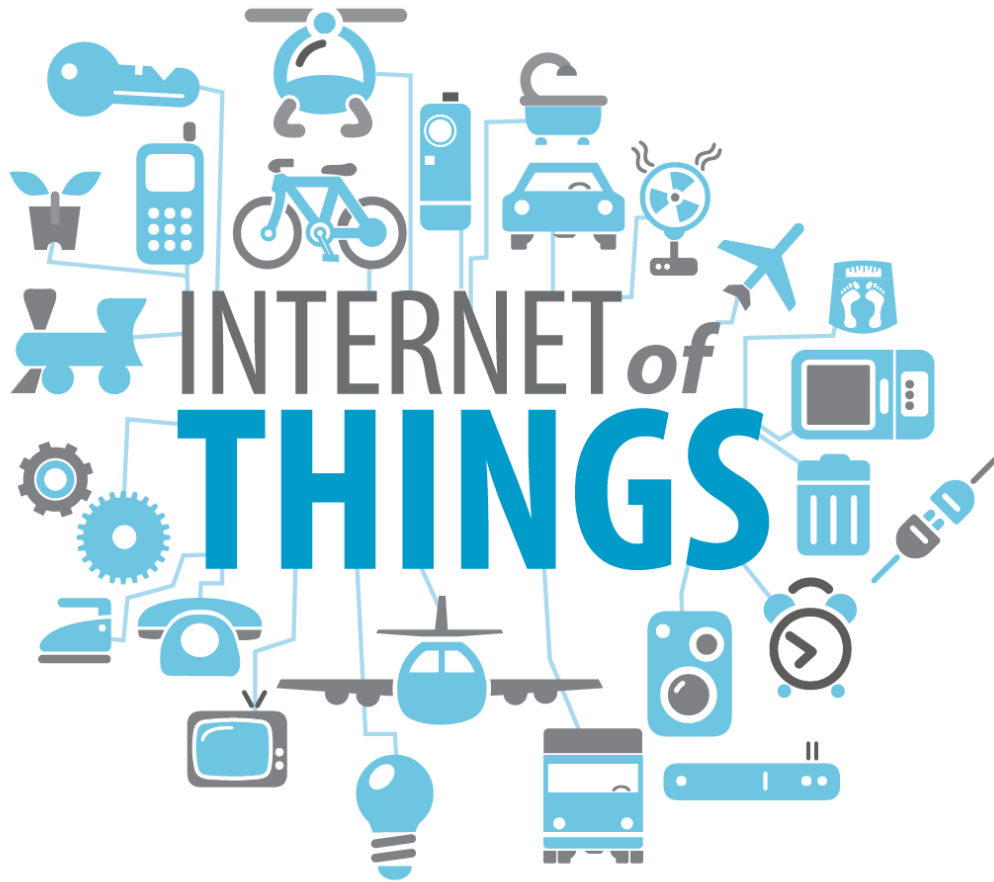


- the **fundamental limits** of wireless communication are **well understood**

# Unprecedented level of connectivity



A close-up photograph of a person's right eye, which is a striking blue color. The eye is looking directly at the camera. The eye is framed by a piece of white paper that has been torn, creating a jagged, irregular border. The background is a solid, bright white. The lighting is soft, highlighting the texture of the skin and the details of the eye.





# Recent data privacy leaks

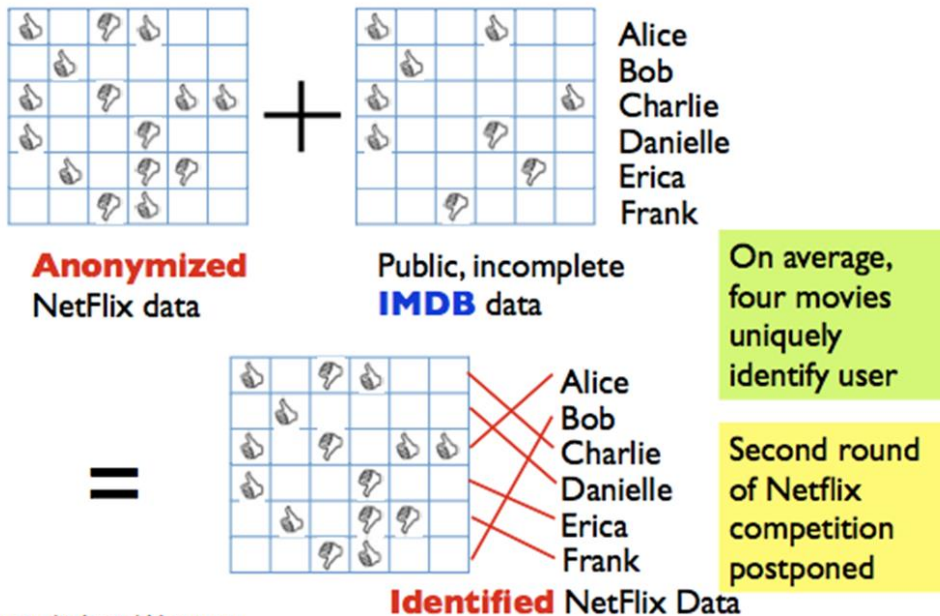
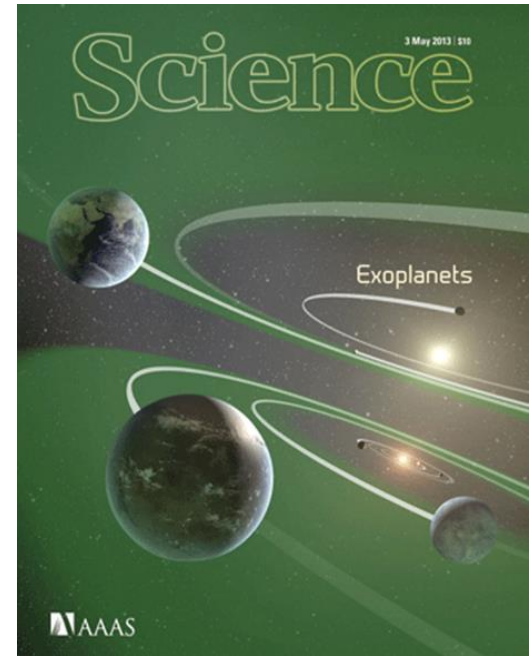


Image credit: Arvind Narayanan

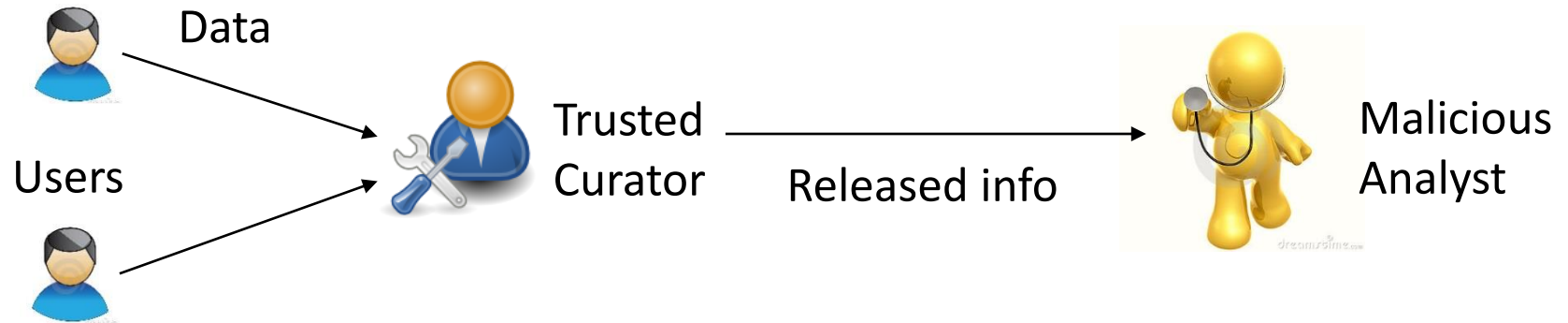
11



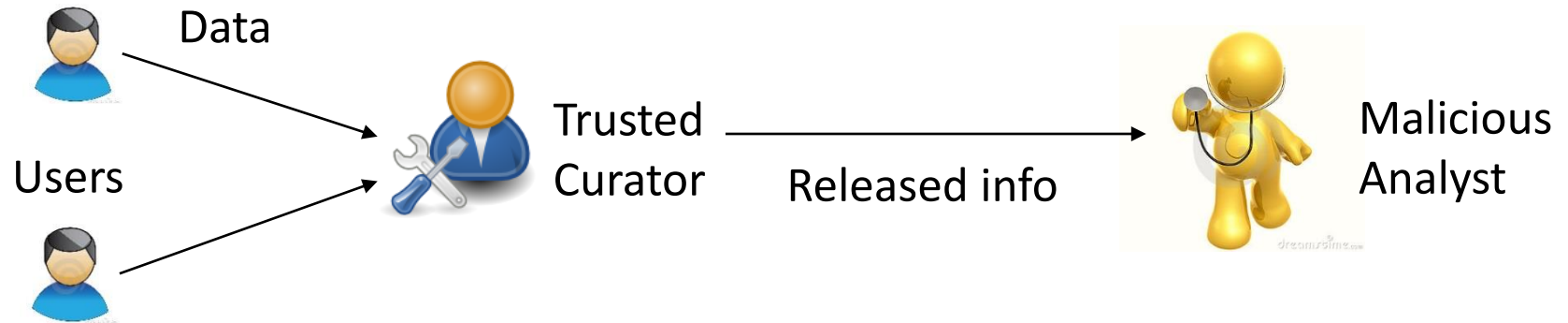
**de-anonymizing** Netflix data, **identifying** personal genomes



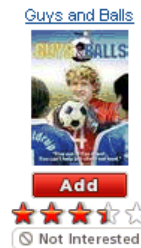
# Global privacy model



# Global privacy model



## Other Movies You Might Enjoy



Eiken has been added to your Queue at position 2.

This movie is available now.

[Move To Top Of My Queue](#)

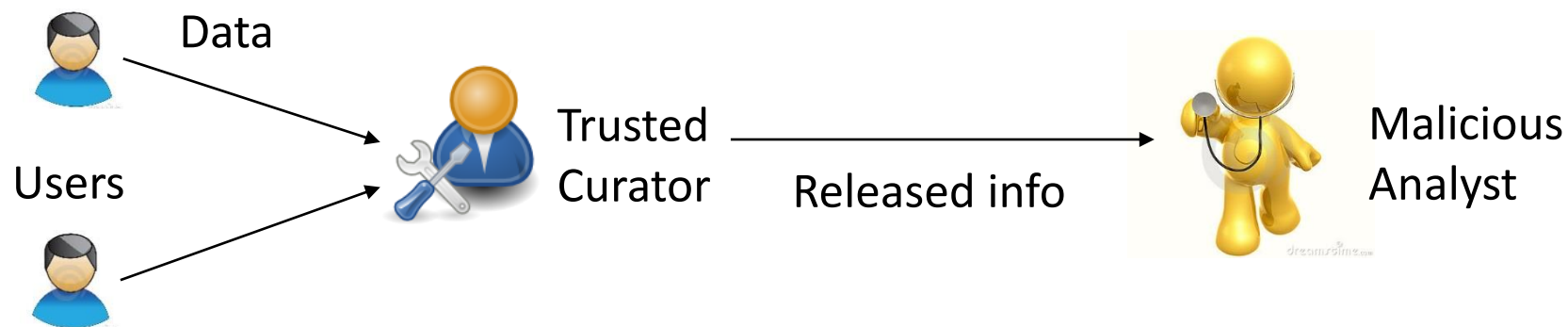
---

[Continue Browsing](#) [Visit your Queue >](#)





# Global privacy model



National Institutes  
of Health



# Local privacy model



# Local privacy model



Google

facebook®



Microsoft

# Local privacy model



*Car rental*

Economy 2 door sedan  
Hertz rental car reservation

Name	Booking Number
Mr. John Smith	E12345678
Thu, 18 Apr, 2013 11:40	Fri, 26 Apr, 2013 21:50

Hertz San Diego  
987 Harbor Dr, San Diego, CA 92101

[Get directions](#)

[Manage reservation](#)

[View email](#)

## Next Appointment

Agency Meeting  
11:30 AM  
Ninth Ave, New York, NY 10011

[Email guests](#)

## Flights

Delta Air Lines  
flight 8772  
from DeltaAirLines@e.delta.com

**Status: Scheduled** / Fri, Nov 29, 2013  
Depart San Francisco International  
SFO 11:45 PM  
Terminal 1

63

SCATTERED CLOUDS  
5mph  
10%

TUE	WED	THU	FRI
68° 48°	67° 44°	65° 48°	57° 46°

## Hotels

The Connaught Hotel  
Carlos Place, Mayfair, London W1K 2AL, United Kingdom  
Check in from 12:00pm today



[Call](#)

[Hotel information](#)

[View email](#)

## Packages

Chromecast  
Shipped - 1 min ago

1 item from  
Amazon.com  
Estimated arrival  
Tuesday, May 13



57 mins to work  
Normal traffic on US - 101

[Navigate / 57 mins via US - 101](#)

## Restaurant Reservations

Broder  
2508 SE Clinton St, Portland, OR 97202  
**Reservation in 1 hour**  
Travel time walking 45 minutes



[Get directions](#)

[Manage reservation](#)

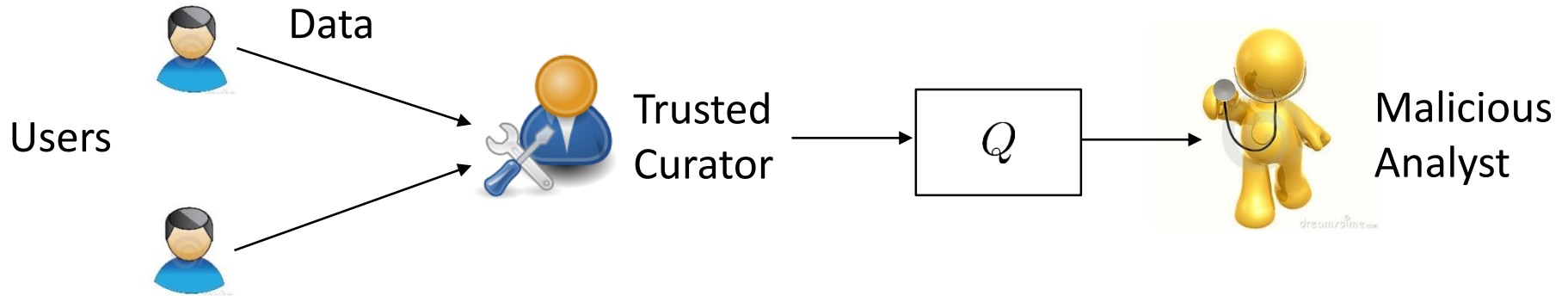
[View email](#)

## Friends' Birthdays



Conrad  
Kent

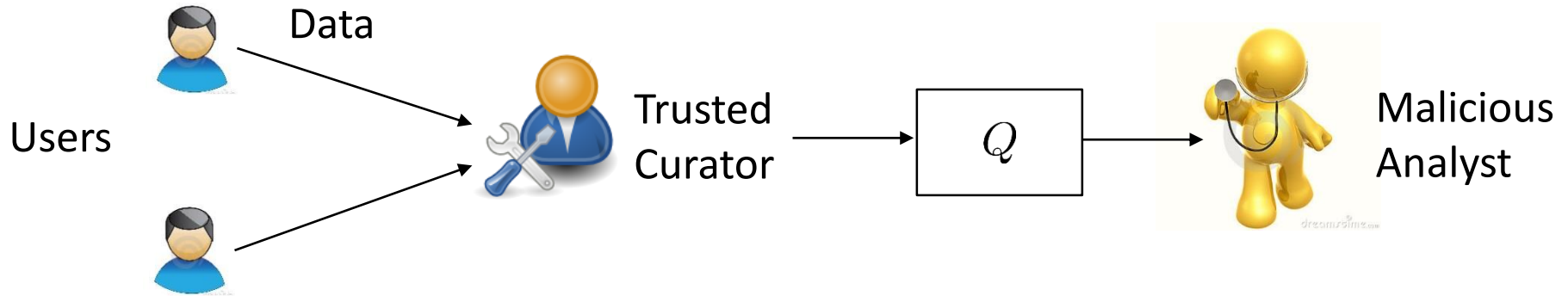
# Differential privacy



- $Q$  is a privacy mechanism
- privacy enforced by imposing **differential privacy** parametrized by  $\epsilon$



# Differential privacy



$\epsilon$  controls the level of privacy

large  $\epsilon$ , low privacy

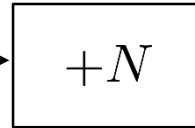
small  $\epsilon$ , high privacy

# **Global Privacy Model**

# The Laplace mechanism

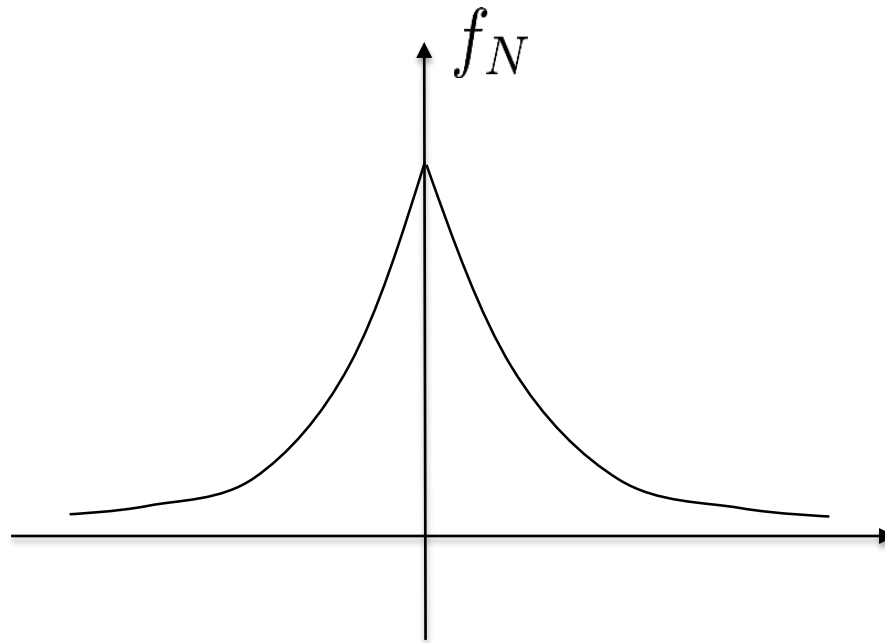


Trusted  
Curator



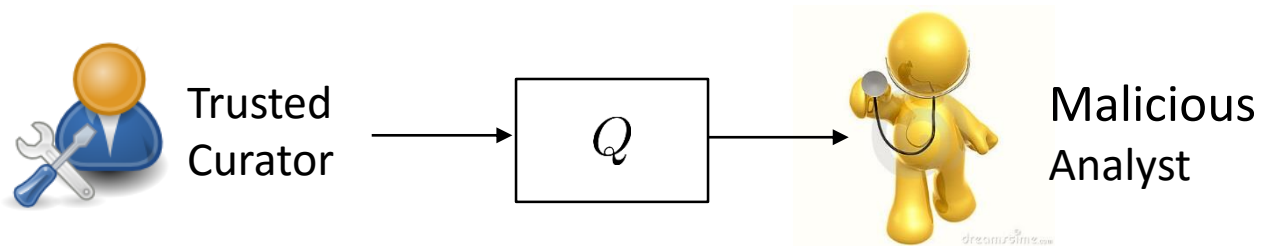
Malicious  
Analyst

Laplace Mechanism



# What would Shannon do?

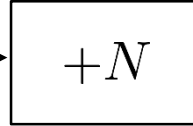
- there is a **fundamental tradeoff** between **privacy** and **utility**



# Data independent noise is optimal



Trusted  
Curator



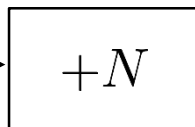
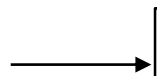
Malicious  
Analyst



# Staircase mechanisms are optimal

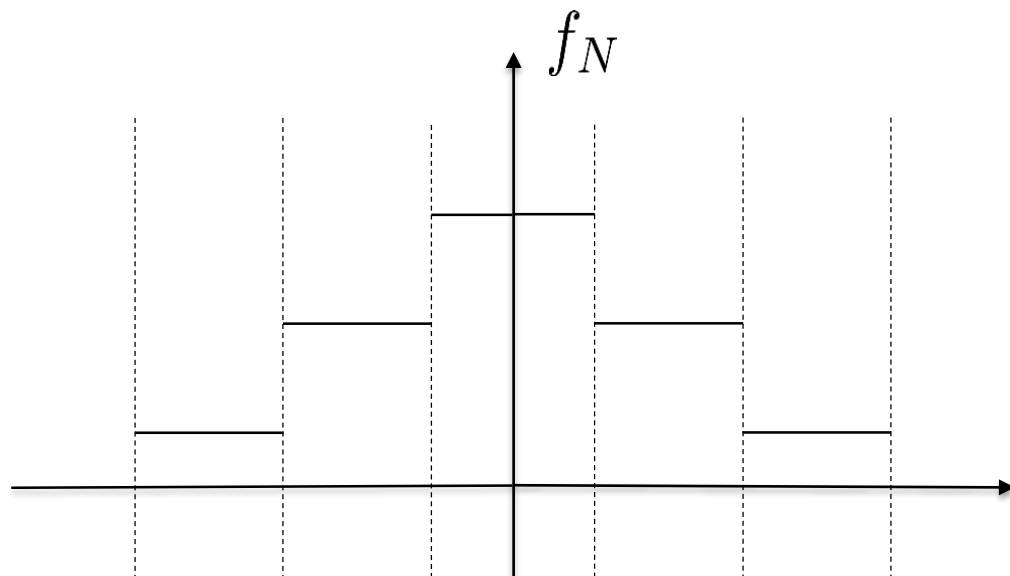


Trusted  
Curator



Malicious  
Analyst

## Staircase Mechanism



# Staircase mechanisms are optimal



differential privacy

Scholar

About 2,560,000 results (0.03 sec)

Articles

Case law

My library

## Differential privacy

[C Dwork](#) - Automata, languages and programming, 2006 - Springer

Abstract In 1977 Dalenius articulated a desideratum for statistical databases: nothing about an individual should be learnable from the database that cannot be learned without access to the database. We give a general impossibility result showing that a formalization of ...

Cited by 1744 Related articles All 22 versions Web of Science: 293 Cite Save

Any time

Since 2016

Since 2015

Since 2012

Custom range...

## Differential privacy: A survey of results

[C Dwork](#) - Theory and applications of models of computation, 2008 - Springer

Abstract Over the past five years a new approach to **privacy**-preserving data analysis has born fruit [13, 18, 7, 19, 5, 37, 35, 8, 32]. This approach differs from much (but not all!) of the related literature in the statistics, databases, theory, and cryptography communities, in that ...

Cited by 749 Related articles All 24 versions Cite Save

Sort by relevance

Sort by date

## Mechanism design via differential privacy

[F McSherry](#), [K Talwar](#) - ... of Computer Science, 2007. FOCS'07. ..., 2007 - [ieeexplore.ieee.org](#)

Abstract We study the role that **privacy**-preserving algorithms, which prevent the leakage of specific information about participants, can play in the design of mechanisms for strategic agents, which must encourage players to honestly report information. Specifically, we ...

Cited by 573 Related articles All 24 versions Cite Save

☒ include patents

☒ include citations

## Differential privacy via wavelet transforms

[X Xiao](#), [G Wang](#), [J Gehrke](#) - Knowledge and Data Engineering, ..., 2011 - [ieeexplore.ieee.org](#)

Abstract—**Privacy** preserving data publishing has attracted considerable research interest in recent years. Among the existing solutions, **e-differential privacy** provides the strongest **privacy** guarantee. Existing data publishing methods that achieve **e-differential privacy**, ...

☒ Create alert

# **Local Privacy Model**

# Local privacy



have you ever used illegal drugs?

# Local privacy



have you ever used illegal drugs?



say **yes**



answer **truthfully**



# What would Shannon do?

- there is a **fundamental tradeoff** between **privacy** and **utility**



# Main result: **binary data**

- for binary data:



**lie** w.p.  $\frac{1}{e^\epsilon + 1}$

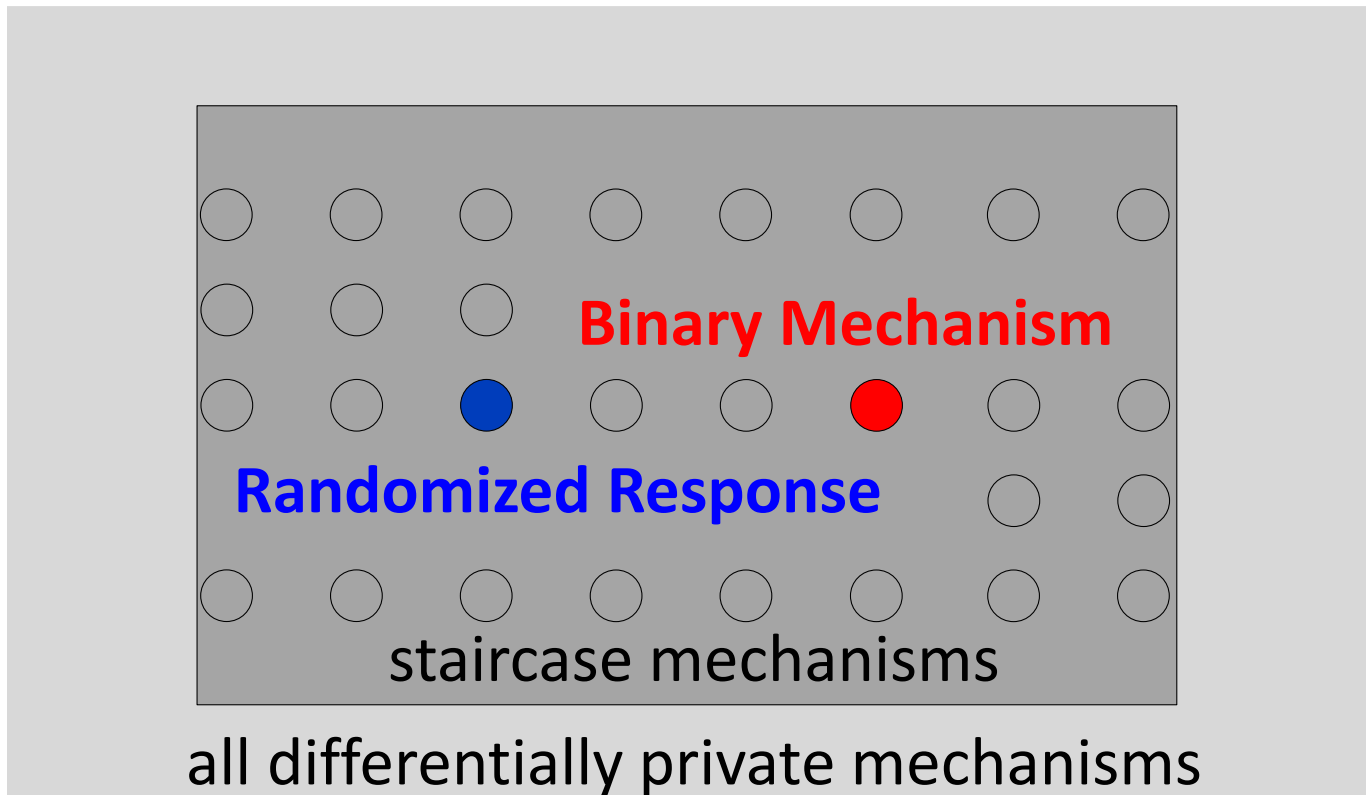


**say the truth** w.p.  $\frac{e^\epsilon}{e^\epsilon + 1}$

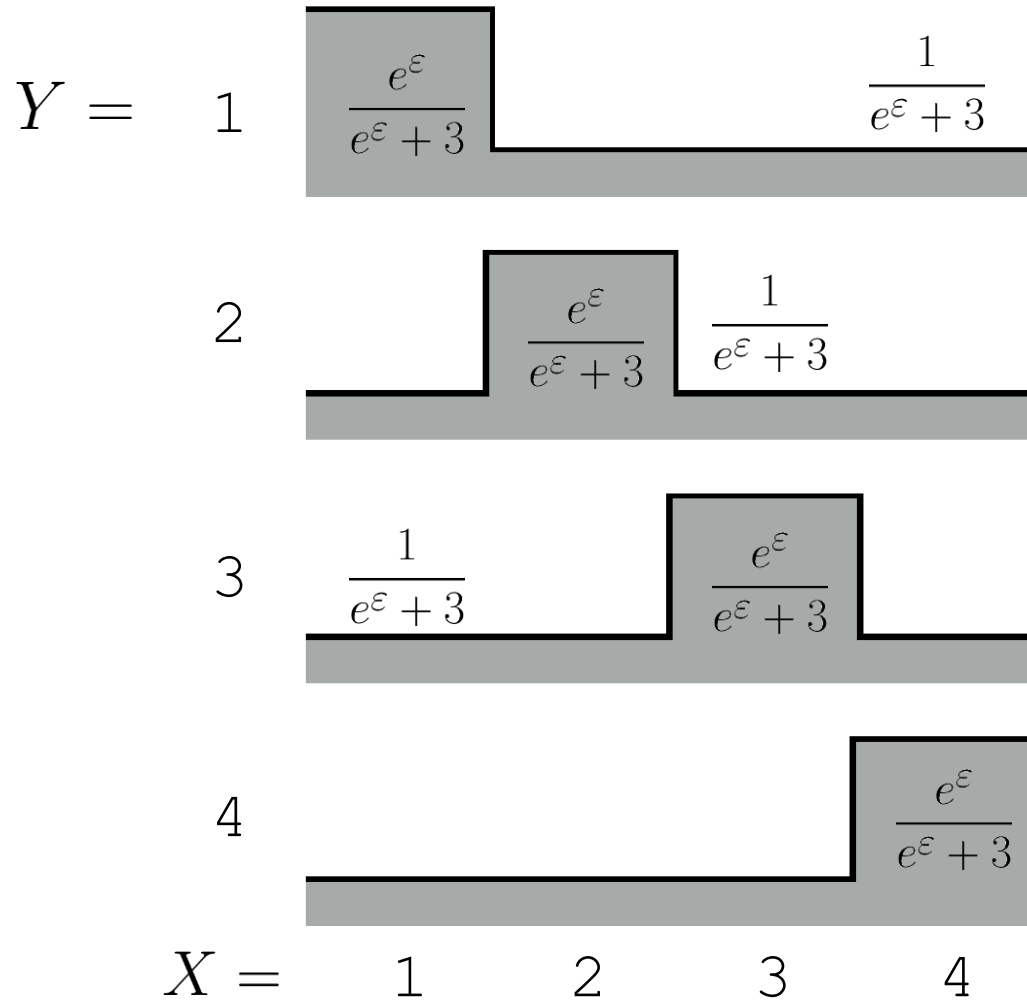
- optimal for **all utilities obeying the data processing inequality**

# Main result: **general data**

- for general k-ary data:

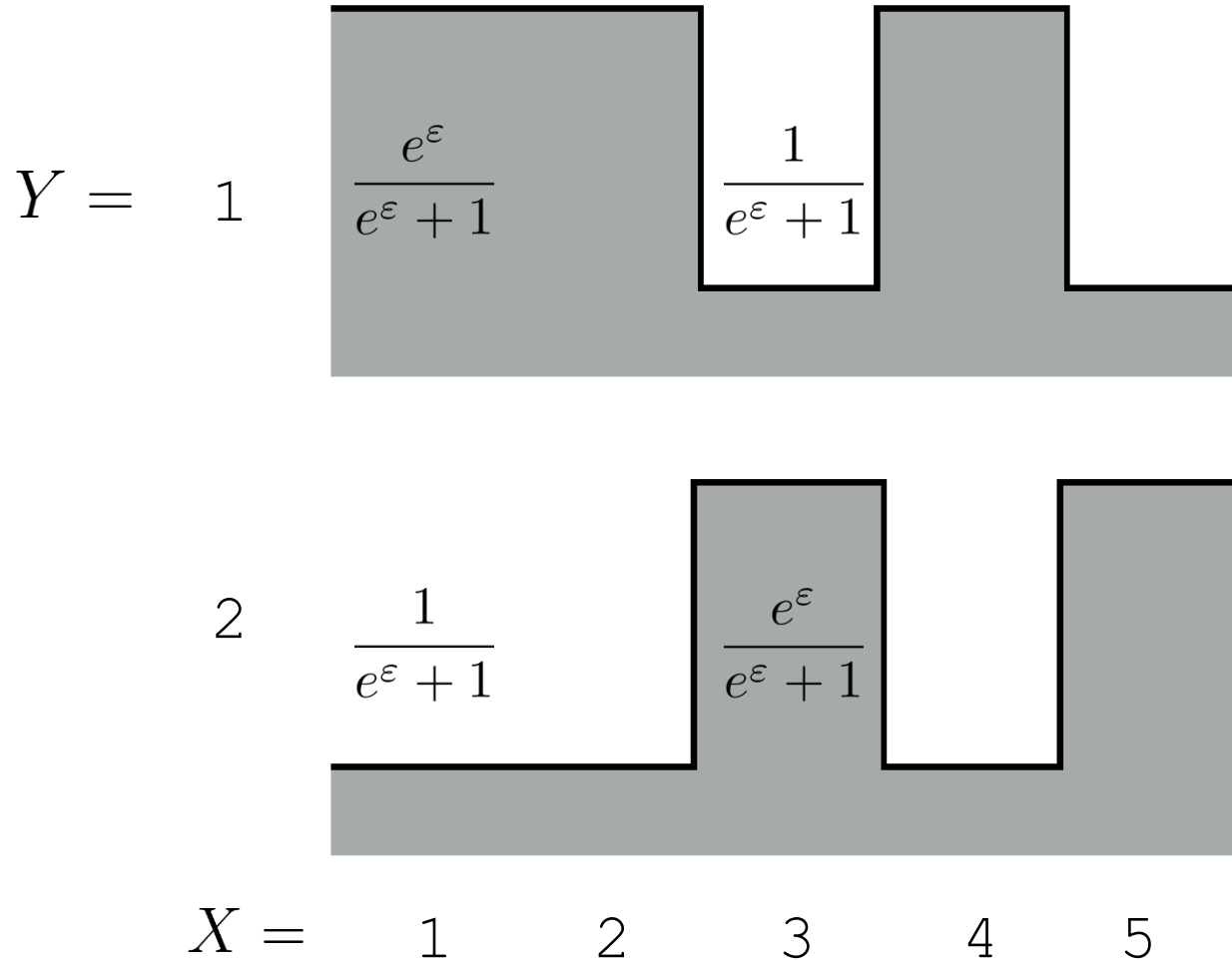


# Randomized Response



- optimal in the **low privacy regime**

# Binary mechanism



- optimal in the **high privacy regime**



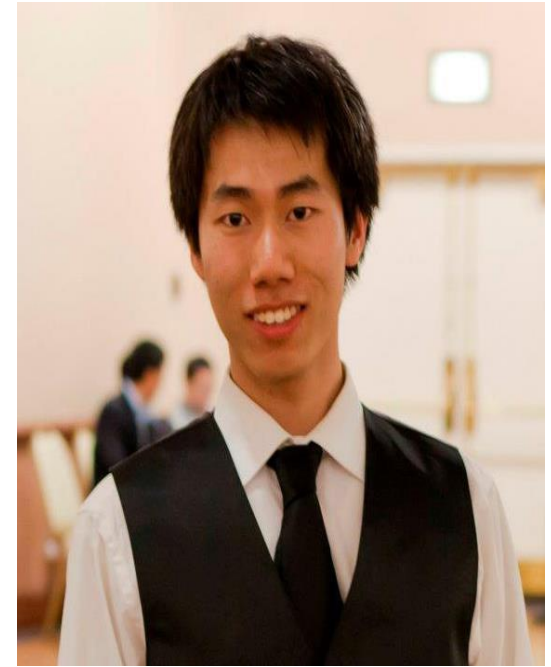
# Acknowledgments



**Sewoong Oh**



**Pramod Viswanath**



**Quan Geng**