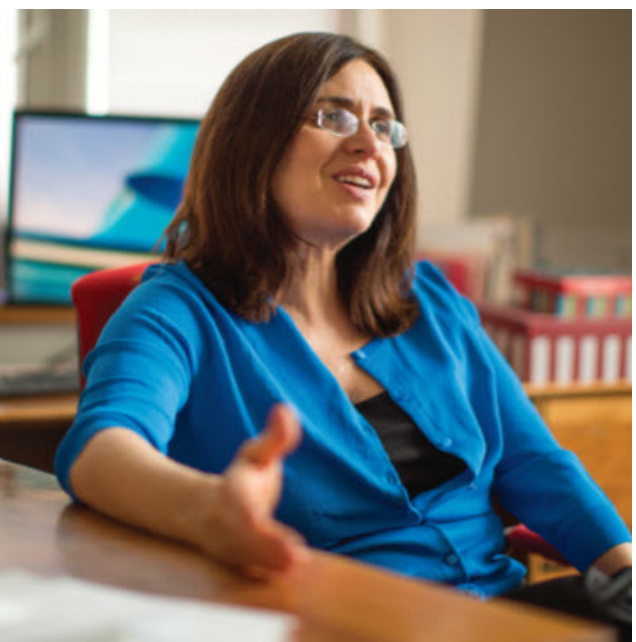
Technology Panel

- What technical tools are in our disposal for achieving privacy and security

- Privacy: Technology + Policy
  - Without Policy, technology will not be employed
  - Without Technology, policy will not be enforced



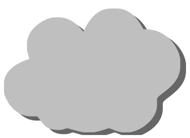
Panelists

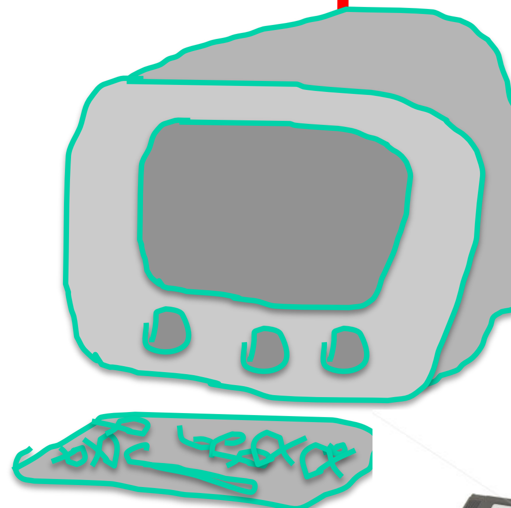
Cryptography:

An Enabler of Secure Computation

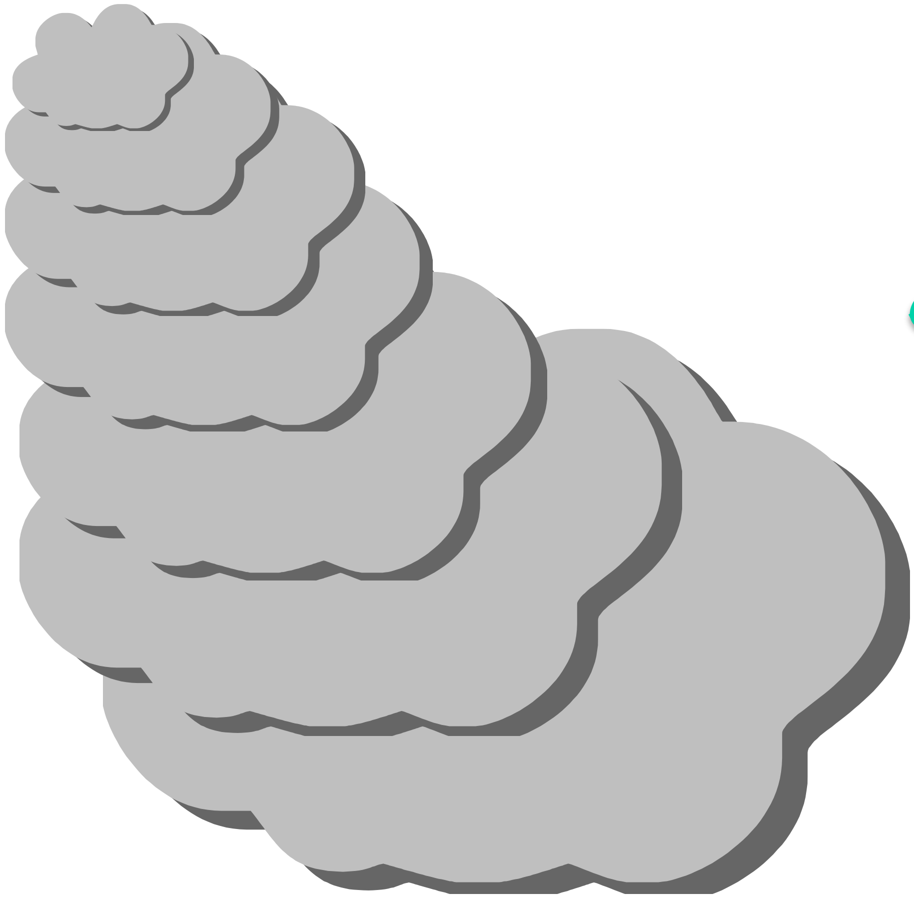
Shafi Goldwasser
CSAIL, MIT

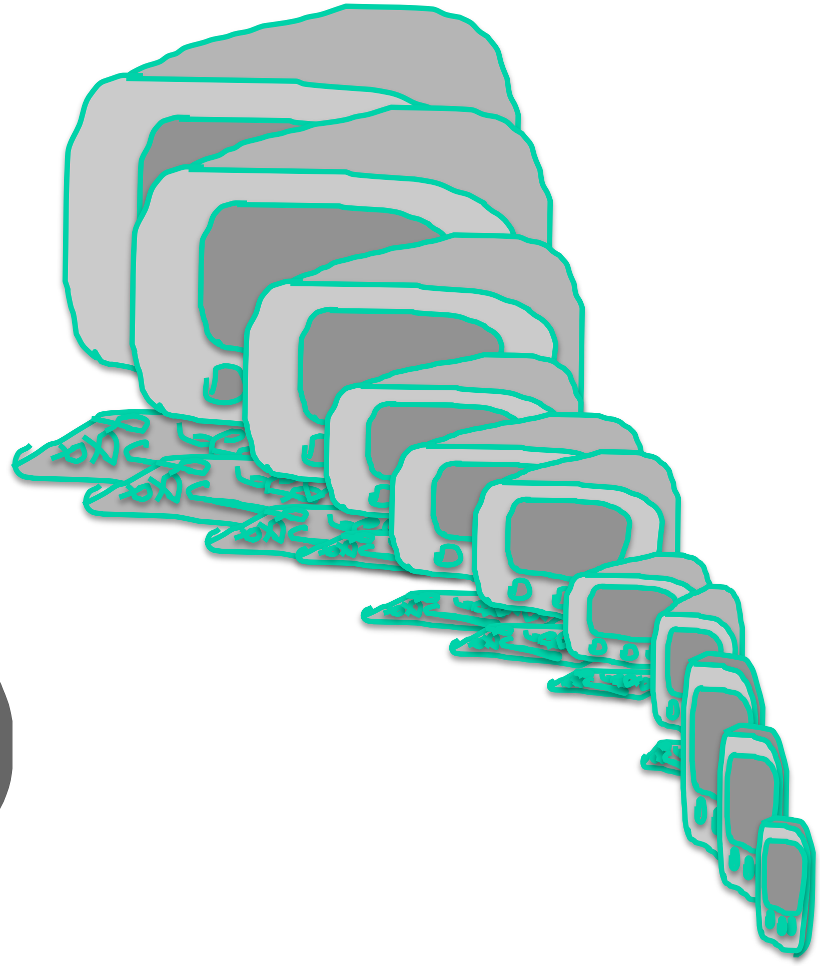
The Evolution of Computing



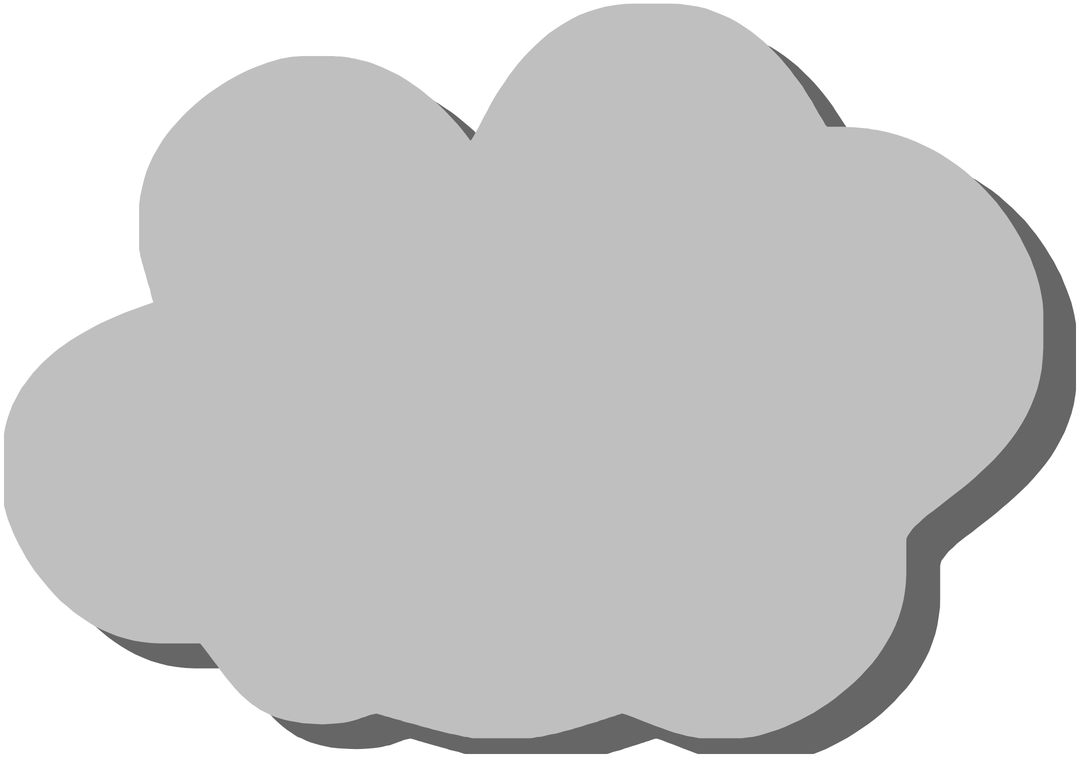


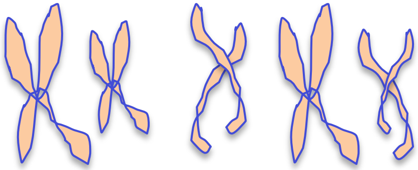
The Evolution of Computing



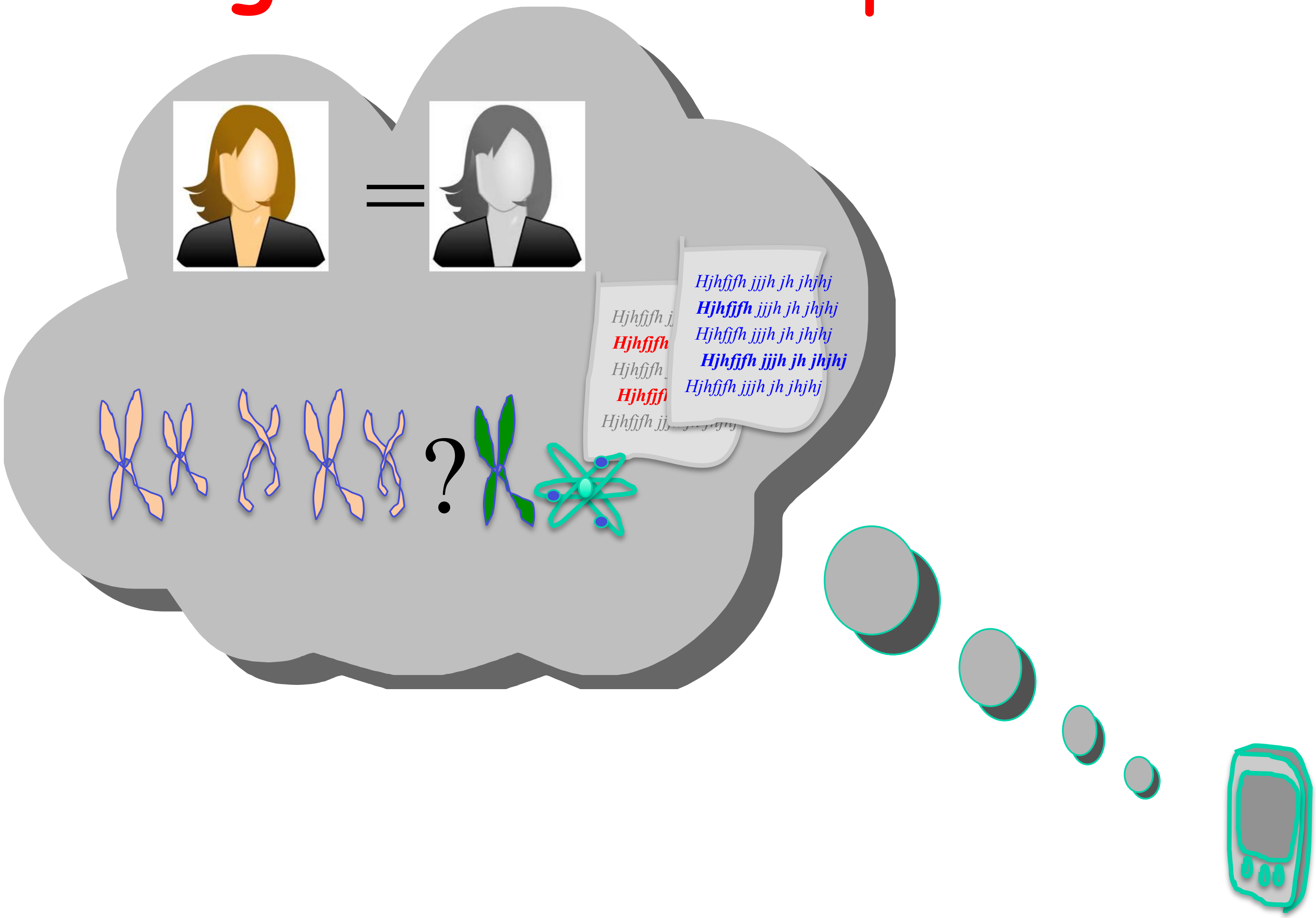


A Migration of Data



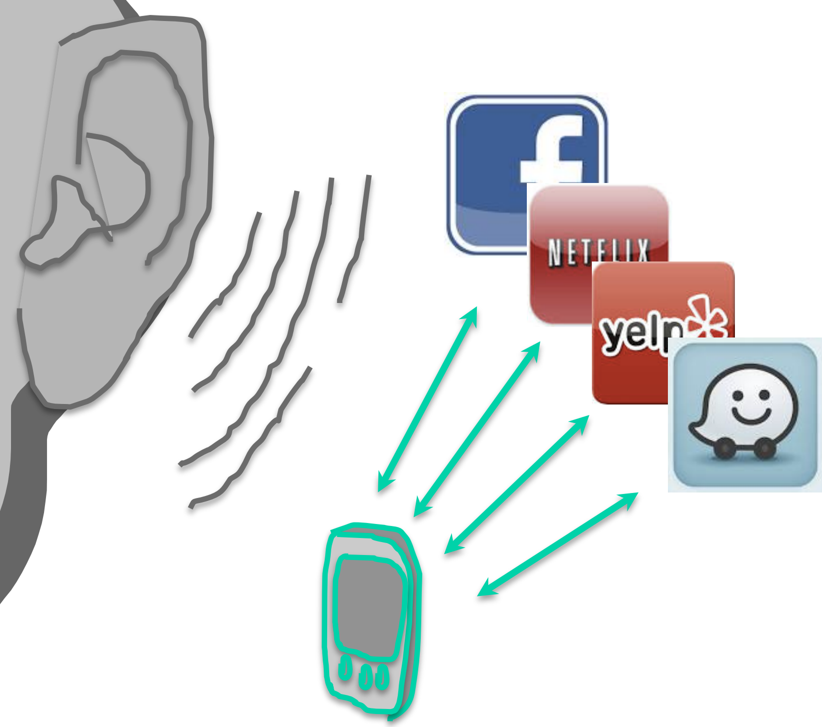


Migration of Computation

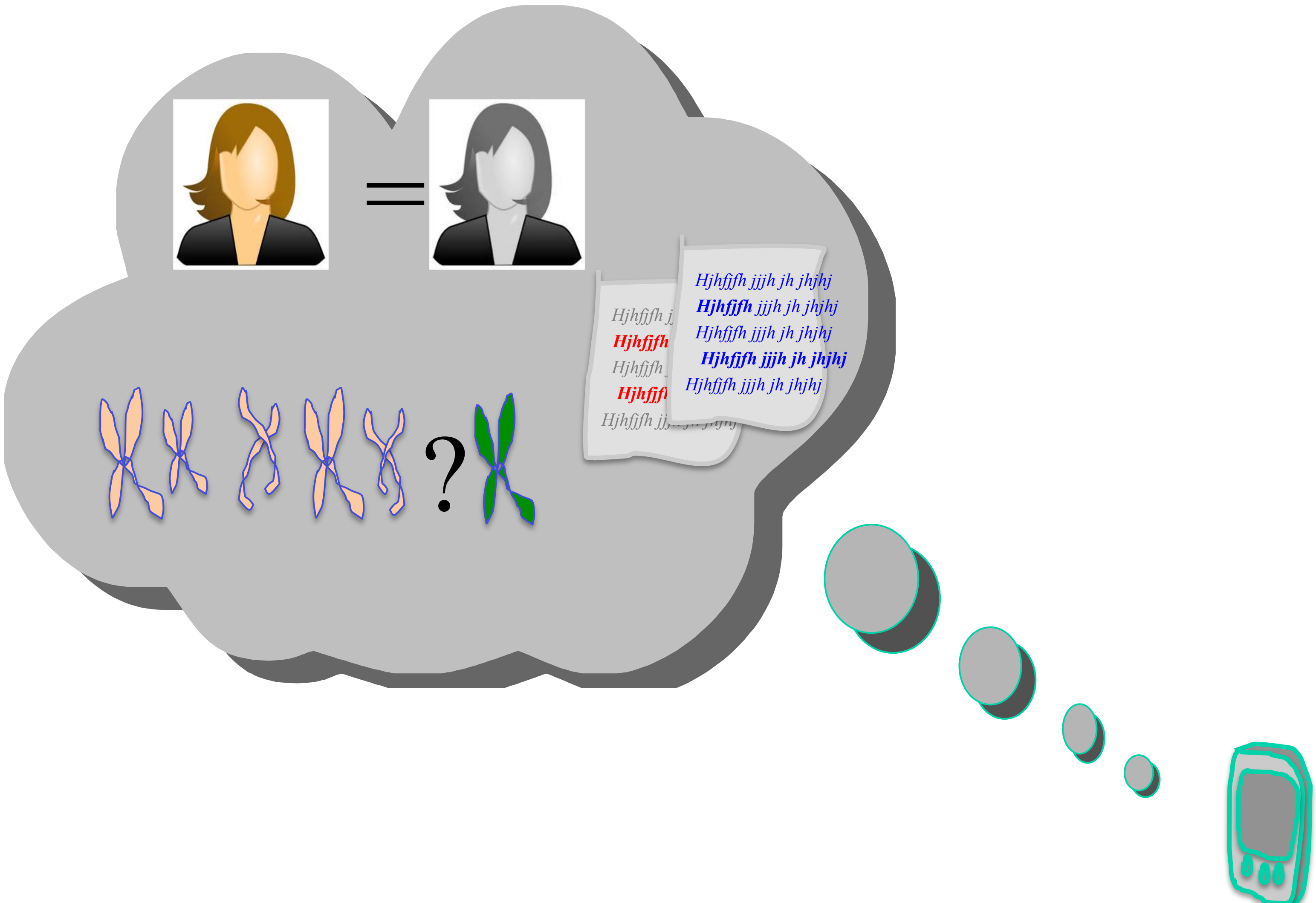


A Collection on Data about us

Medical Records
Financial Financial
Surveillance photos
Location information
Purchasing history
Browsing history
Social Interactions



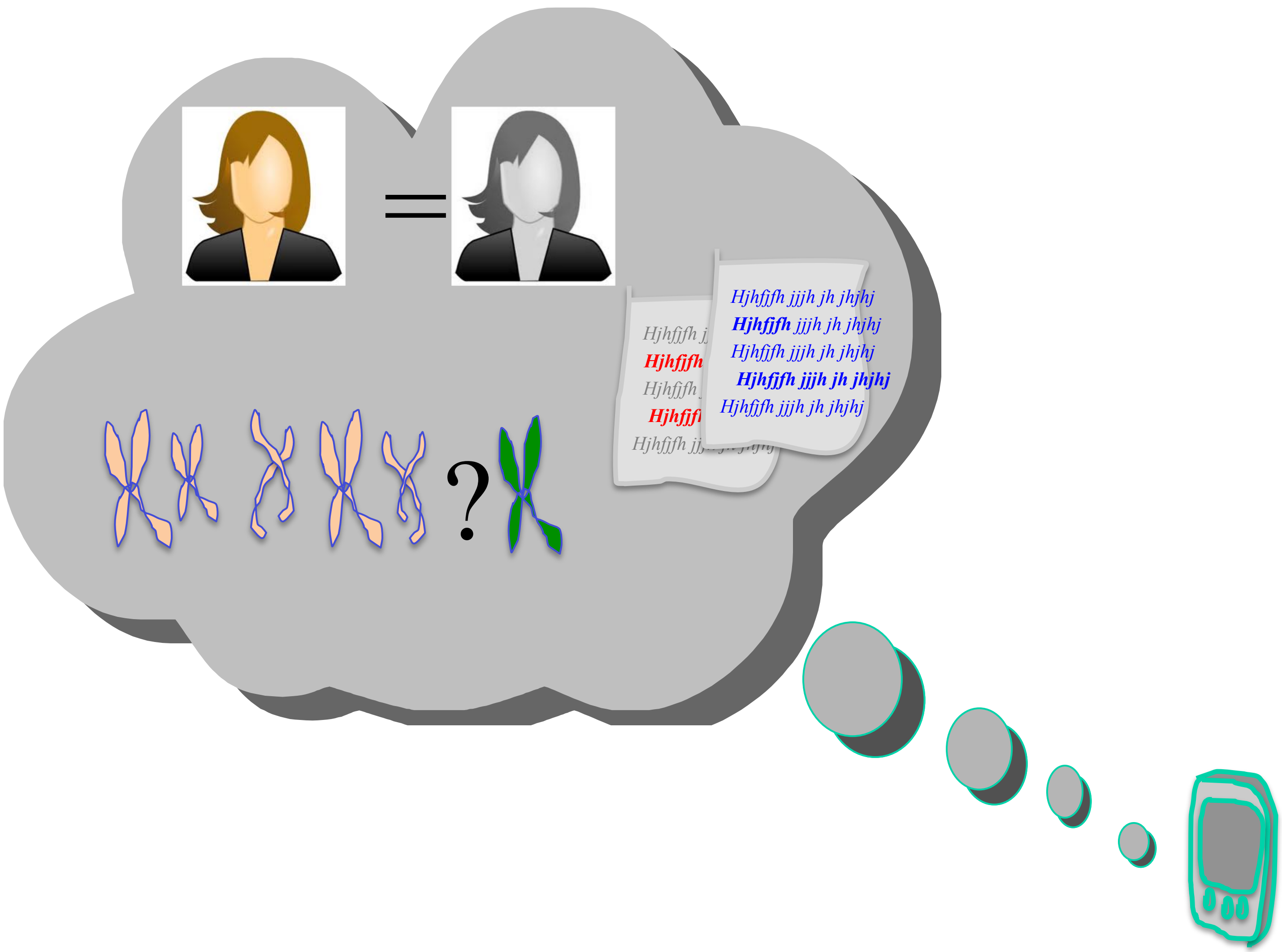
Enormous Potential Benefits in Globalization of Knowledge



Benefits

- Heath: Sharing Data Sets for research and disease control
- National Security: threat prediction, law enforcement
- Data Analytics:
  - Traffic re-routing,
  - smart energy usage,
  - economic growth by intelligent consumer targeting
  - risk predictions for financial markets

Enormous Potential Risks: Loss of Control, Misuse of Knowledge



Risks

- Loss of Control: Remote Storage & Computation threatens
  - Authenticity.
  - Correctness.
  - Availability.
- Loss of Privacy: **reveal more data than necessary to extract the benefits**
  - Loss of Anonymity
  - Loss of Fairness: profiling, price discrimination
  - Loss of competitive Edge: playing field leveled by common data

Methods Which Don't Work

- Classic Encryption: Can hide information but not process it

- Classic Anonymizing: Individual's data with identifying information removed, is still easy to recognize

Benefit vs. Risk

- Medical: Research progress vs. Patient Rights
- National Security: Surveillance vs. Liberty
- Financial: Risk Analysis vs. Market-Competition
- Economic Growth Consumer Targeting vs. Fair Pricing

Are These Contradictory Constraints?

Benefit vs. Risk

- Medical: Research progress vs. Patient Rights
- National Security: Surveillance vs. Liberty
- Financial: Risk Analysis vs. Market-Competition
- Economic Growth Consumer Targeting vs. Fair Pricing

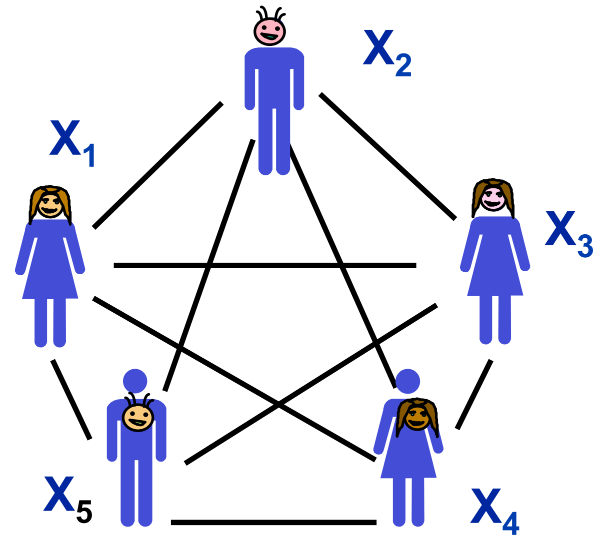
Can mathematics & technology enable us to have the best of both worlds

Cryptography 80's - today

- Host of Techniques that enable to perform computations on data without seeing the data
- Extract specific knowledge, revealing nothing extra
- Reconcile some of these seemingly contradictory "Benefits vs. Risks"

SFE: Mathematical Formulation

N distrustful parties run a protocol to extract information depending on their collective private data.



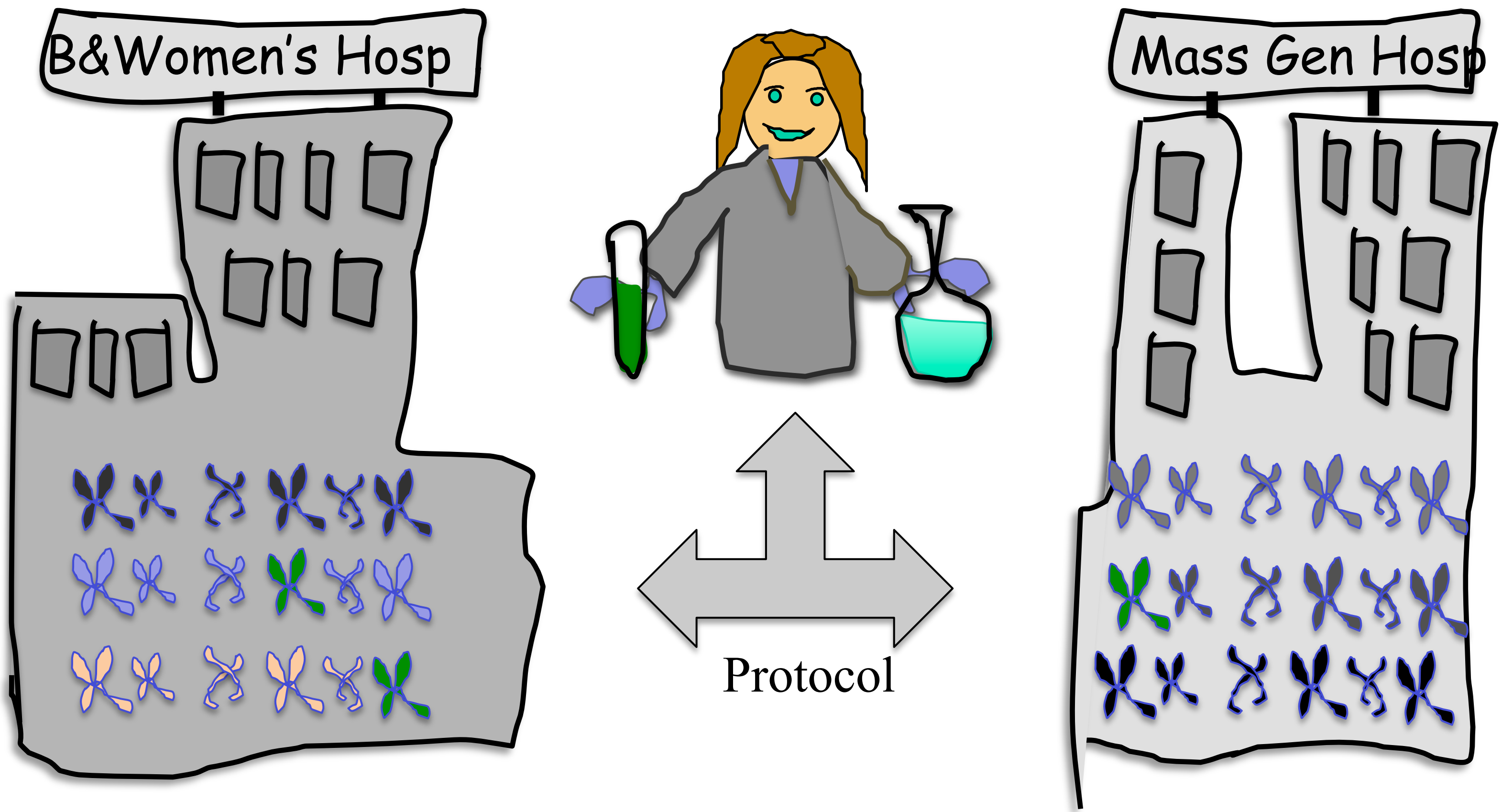
Formulate this as evaluating a function

$$f(X_1, X_2, \ldots, X_n)$$

- where X_j is private input of party j.

So that parties only learn the function output, but nothing else about others inputs

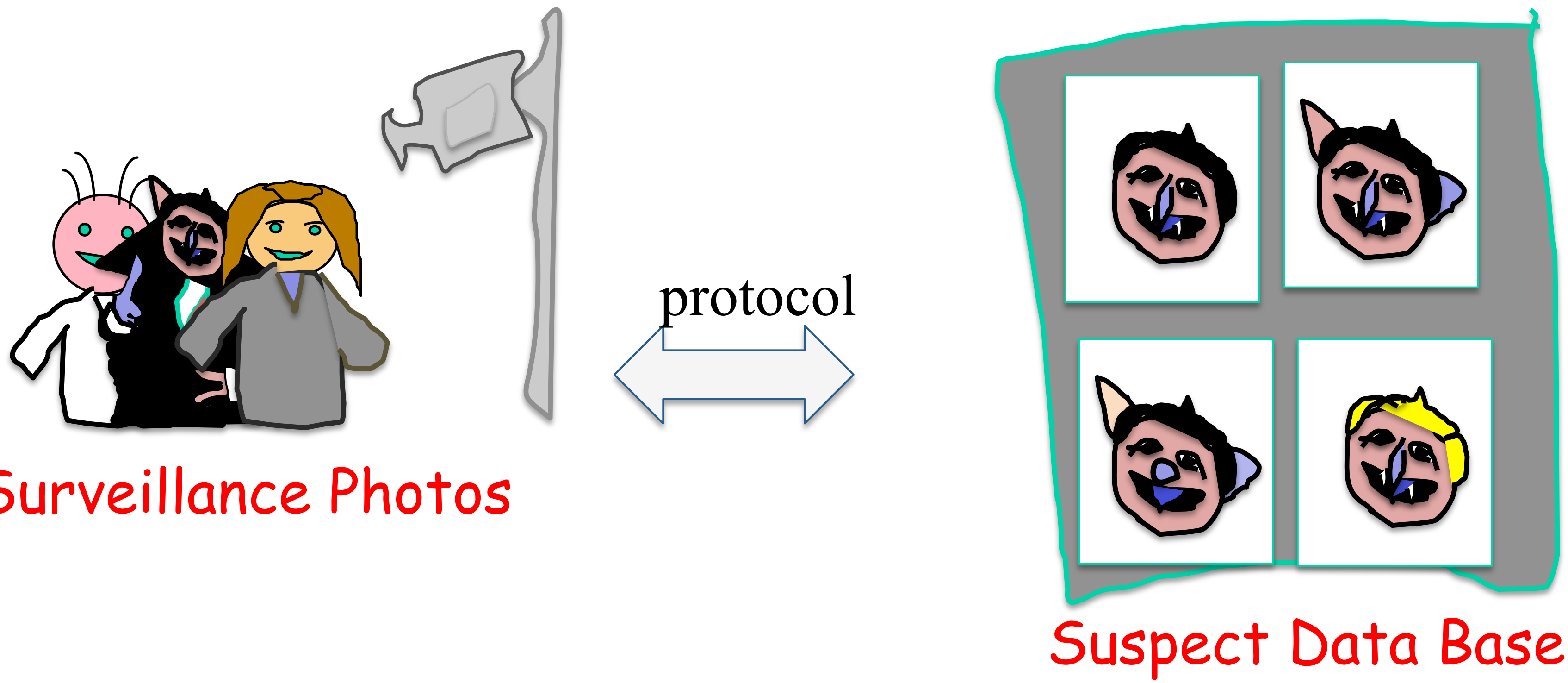
Example 1: Conduct Medical Study on Confidential Medical Data



f(B&W-DNA,MGH-DNA, Pharma) = develop drug if the green gene is prevalent in the population

N=3

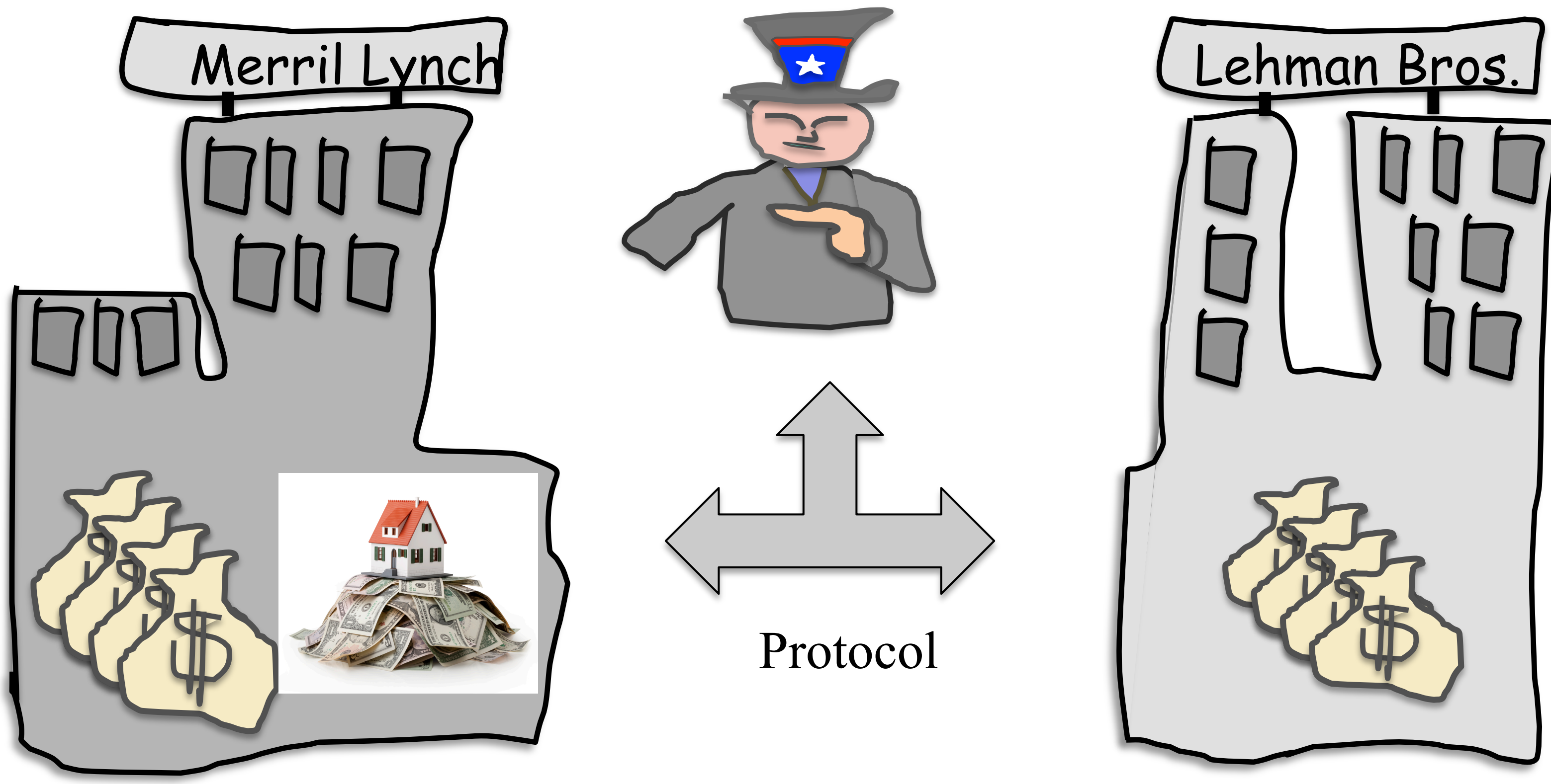
Example 2: Policing While Not Breaking Civil Liberties



f(photos, suspects) = true only if suspect appears in them

N=2

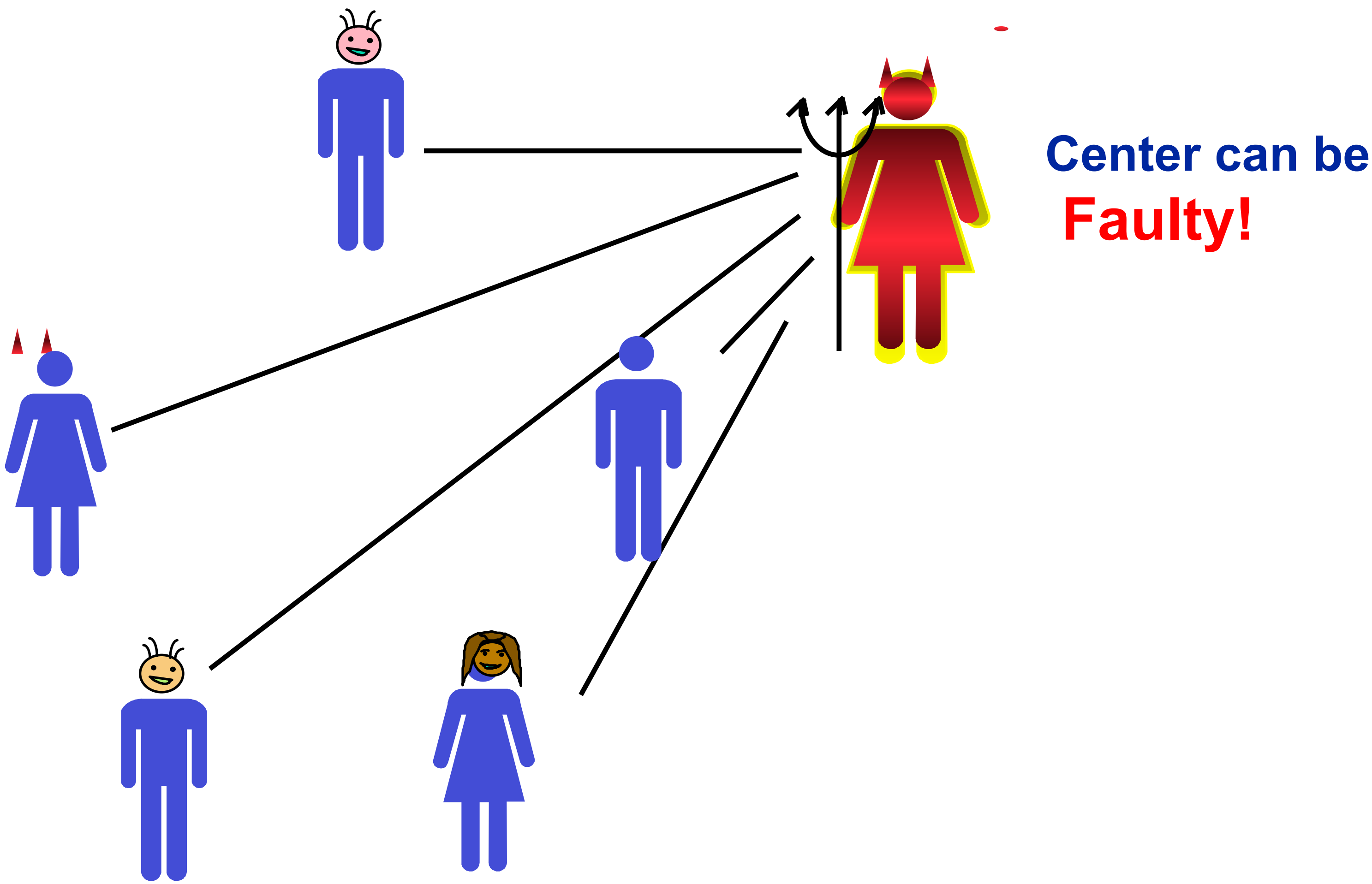
Example 3:Financial Stability of our society [AKL11]



f(MerrilLynch info, LehmanBros info, Govt) = will banks become insolvent

N=3

Trusted Center Solution?



Goal: Decentralized solution with "same" properties as solution with trusted center

Major Result
[the 80's]

Any polynomial time function f can be securely *evaluated*,
using the *cryptographic tool box*

Unconditionally, if there is an honest Majority

Assuming oblivious transfer, if no honest majority,

Major Lesson: Store All Data Distributively

- In toy examples, different entities with different goals hold different parts of the data

Not Always the case....

- **By design can** store data so no single entity has entire data or power

SFE: Theory and Practice

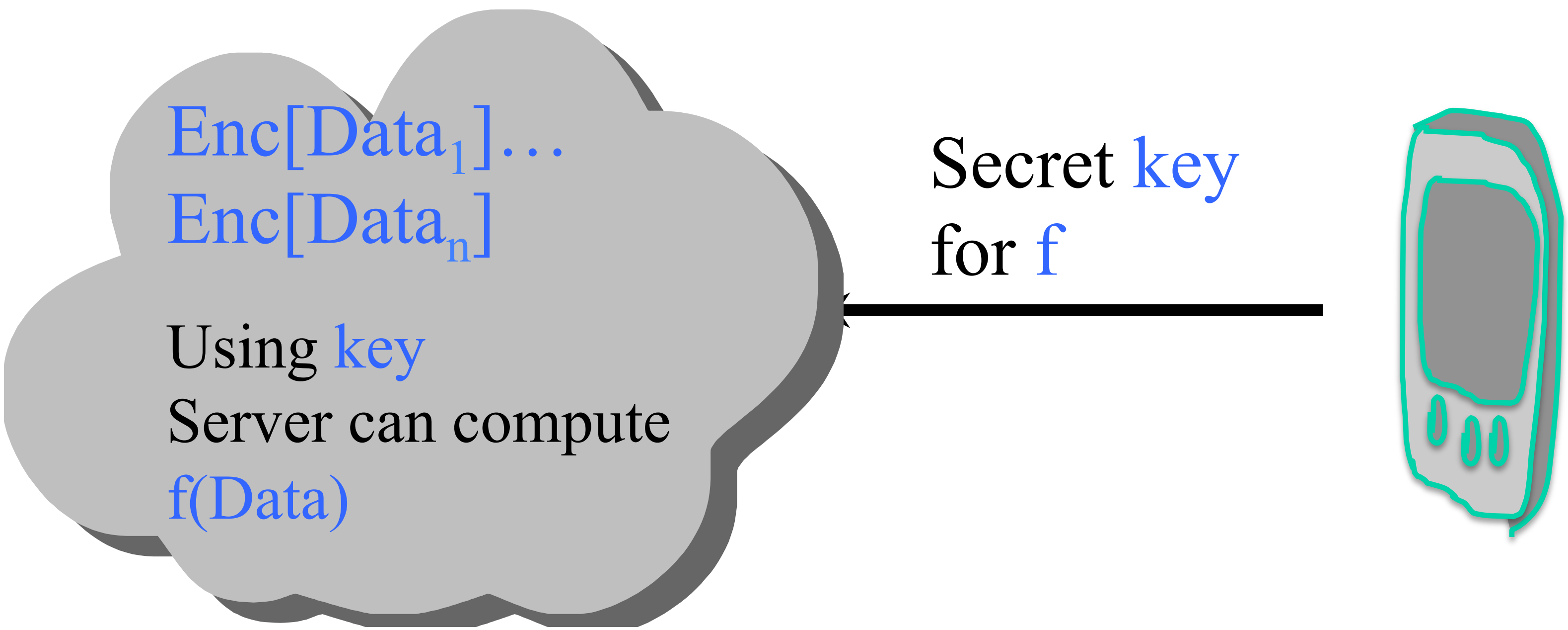
- [80's] Proof of Concept: Great Theory, for general functions, but impractical
- In recent times [Lindell2013 Survey]: optimized implementations for simple classes but useful functions and relaxed security, achieve impressive practical performance, much more work needed

But:

- Requires Interaction
- Not robust to an "insider" leaking all

Truly Amazing Progress: Computing on Encrypted Data

- Fully Homomorphic Encryption [Gentry2009]
- Functional Encryption [SW05, GVW13, GKPVZ13, GGHRSW13, GGSJ14]



- Theory to Practice: Research to be done

From Data to Programs

- Browsing
- Searching
- Social Interactions
- General Programs

Goal: keep *which* programs you run private?

Promise: Program Obfuscation Methods

Secure Function Evaluation ≠ Privacy

- Given f, SFE shows how to compute f(data) revealing nothing extra on data

But

- f(data) itself may reveal too much
- f_1(data),…,…,f_n(data) may reveal too much
- f($data_1$,…,$data_n$) can reveal $data_j$, if {$data_i$} is chosen maliciously

Differential Privacy Research: Which classes of functions are safe to compute?

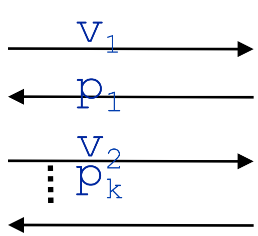
A Combination of Privacy and Secure Computation

- A two-stage process:
  - Decide that the function/algorithm should be computed – an issue of **privacy**
  - Apply secure computation techniques to compute it securely – **security**

Security Definition: Simulation Paradigm

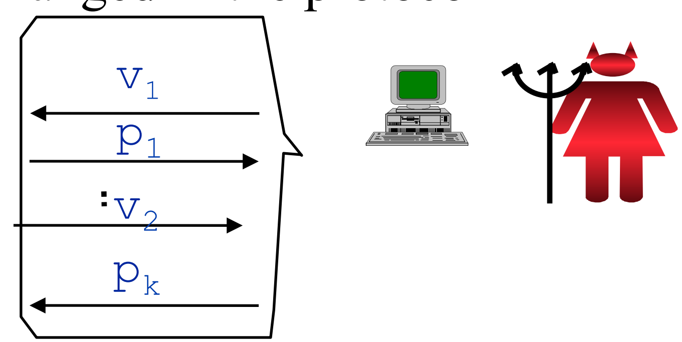
Given your own inputs and the output of the computation, can generate "simulated view" of the protocol which is computationally indistinguishable from "real view" .

Distribution of
REAL Messages
Exchanged in the protocol



$\approx$

SIMULATED
Messages
Exchanged in the protocol



Loss of Privacy is Complex

- Different Entities Collect and Protect Data in an Un-coordinated fashion
- Unforeseen Cross Referencing of information held by different entities on the same individual, causes greater privacy loss
- Aggregate information on many can reveals information on a single individual