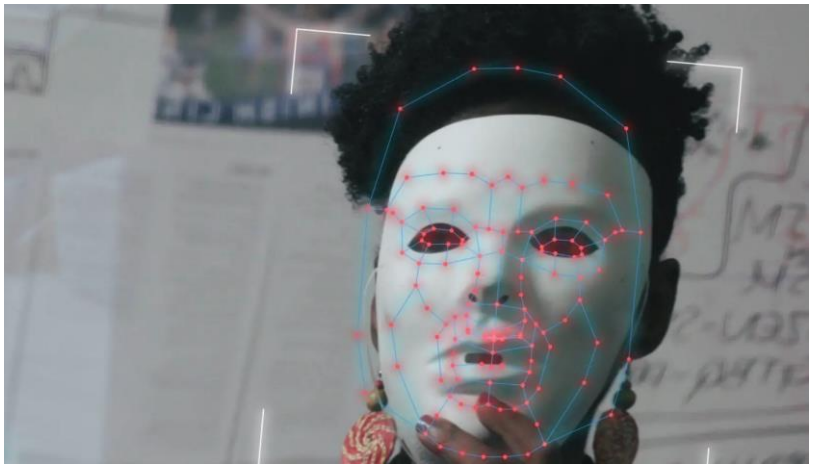
Responsible Computing: the case of data privacy

Adam Smith

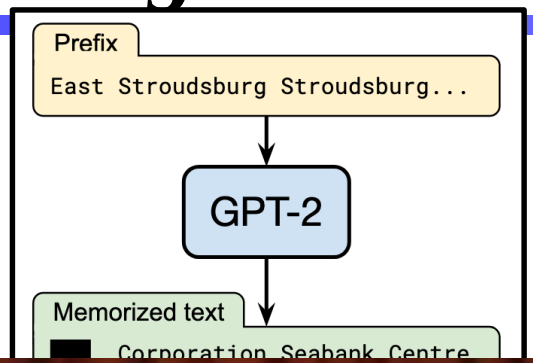
BU Computer Science/ECE/CDS

HDR Symposium Panel
October 25, 2022

What is Responsible Computing?



"Gender Shades": J. Buolamwini, T. Gebru. PMLR 81:77-91, 2018.



[Carlini
me

The Statistical Crisis in Scien

Data-dependent analysis—a "garden of forking paths"— explains why m
statistically significant comparisons don't hold up.

Andrew Gelman and Eric Loken

There is a growing realization that reported "statistically significant" claims in scientific a short mathematics test when it is expressed in two different contexts, involving either healthcare or the This *multiple comparison* well known in statistics and called "*p*-hack 2011 paper by

The "garden of forking paths". Gelman and Loken, *American Scientist*, 2014

GDPR's "Right to be Forgotten"
[Costeja González v. Google, 2014

What is Responsible Computing?

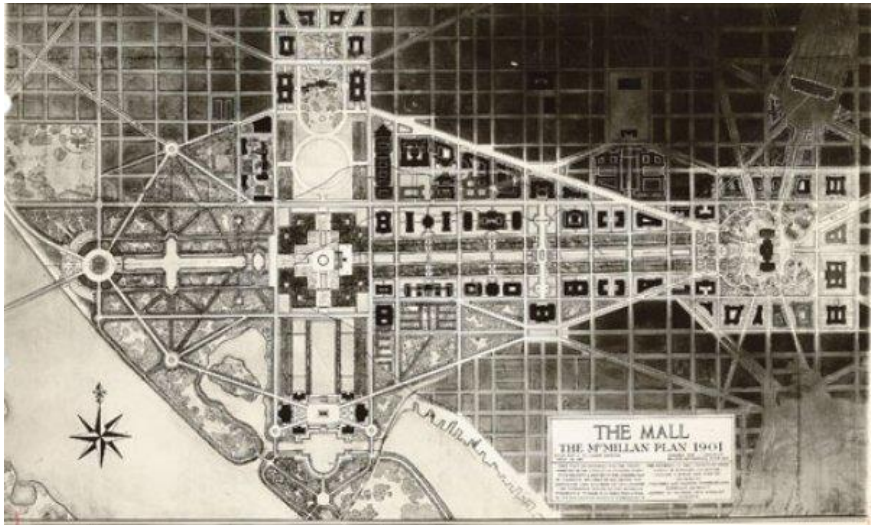
Common thread:

Failure to engineer for the whole use case

true

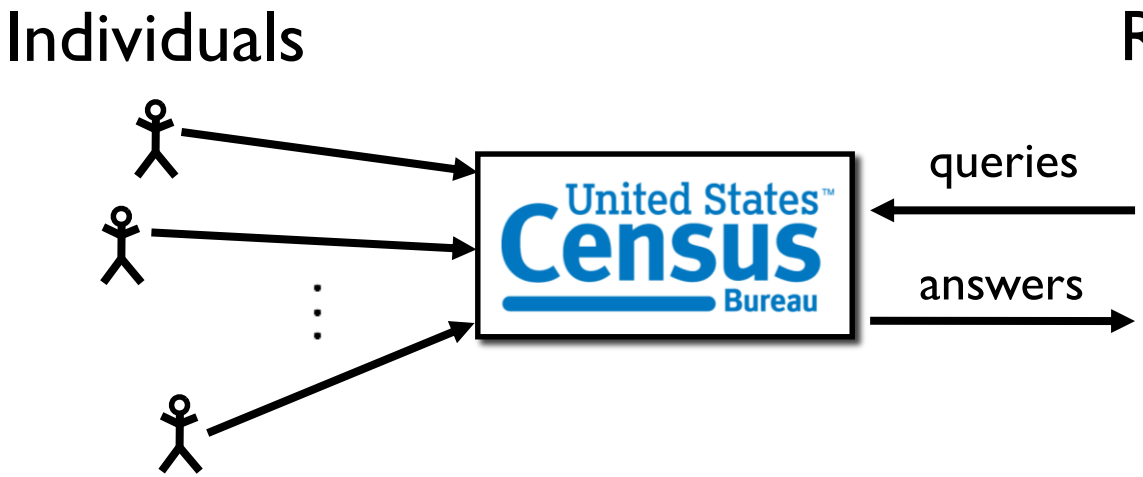
while (system is evolving):

- Understand how data-driven systems operate in context
  - How they affect people
- Formulate intermediate goals that align with final use
- Develop technical tools



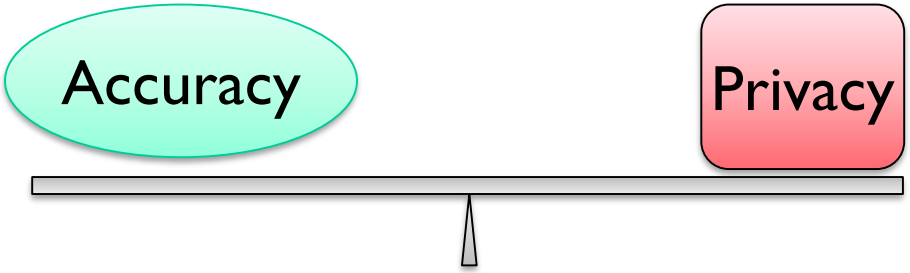
https://thumbs-media.smithsonianmag.com/filer/brief-lenfant-2.jpg__600x0_q85_upscale.jpg

Privacy in Statistical Databases



Large collections of personal information

- census data
- medical/public health
- social networks
- education
- system usage



Goal: Rigorous foundations and analysis

This talk

- Why is privacy challenging?
  - Memorization in machine learning
    [Brown, Bun, Feldman, S., Talwar 2021]

- Differential Privacy [Dwork, McSherry, Nissim, S., 2006]
  - "Privacy" as stability to small changes
  - Widely studied and deployed

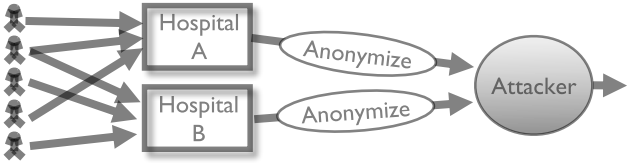
First attempt: Remove obvious identifiers

"AI recognizes blurred faces"
[McPherson Shokri Shmatikov '16]

[Gymrek McGuire Golan Halperin Erlich '13]

[Pandurangan '14]

Everything is an identifier



[Ganta Kasiviswanathan S '08]

Images: whitehouse.gov, genesandhealth.org, medium.com

Is the problem granularity?

What if we only release aggregate information?

Statistics together may encode data

- Example: Average salary before/after resignation

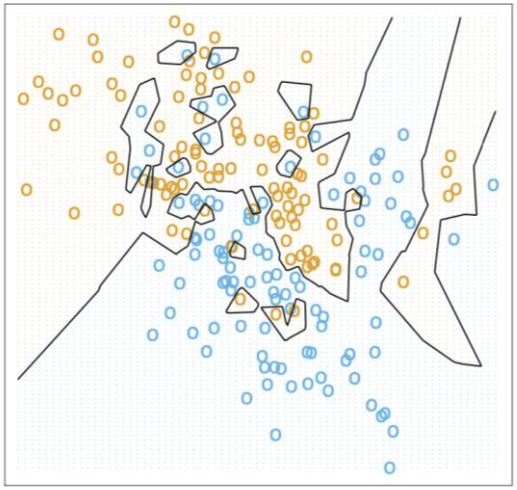
- More generally:

Too many, "too accurate" statistics reveal individual information

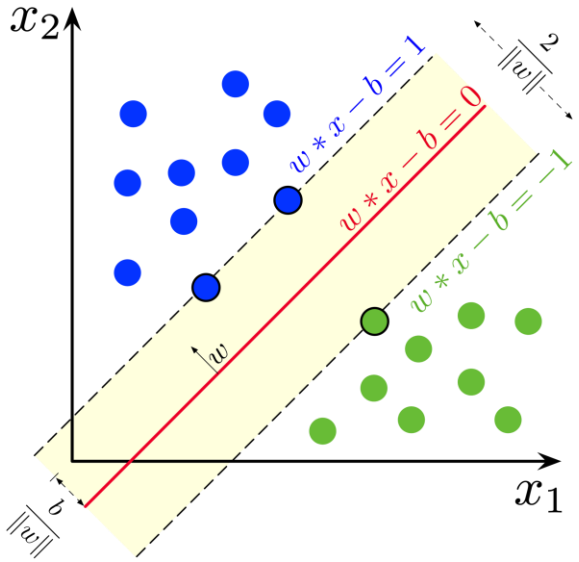
- Reconstruction attacks [Dinur Nissim 2003, …]
- Membership attacks [Homer et al, 2008, …]
- Memorization [this talk]

Cannot release everything everyone would want to know

Memorization can be explicit…

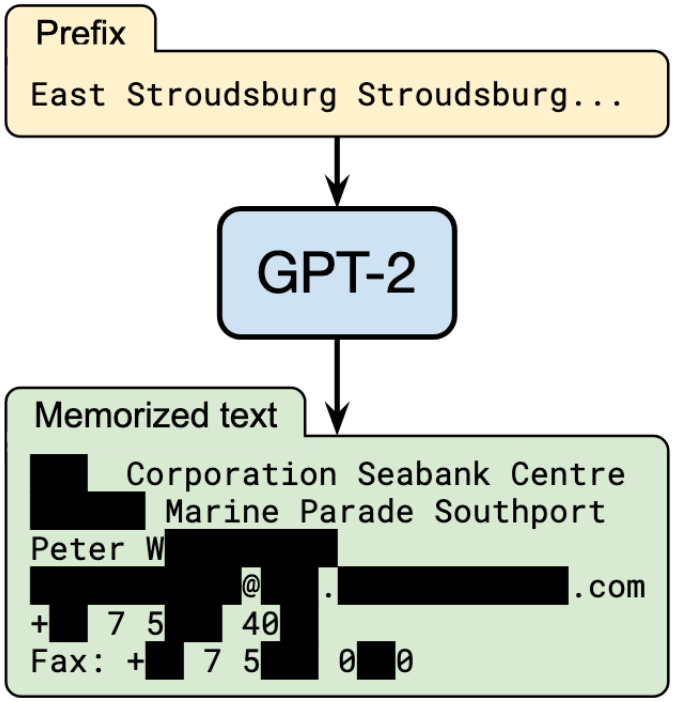


Hastie, Tibshirani, and Friedman. *The elements of statistical learning: data mining, inference, and prediction*. Springer Science & Business Media, 2009.



Wikipedia, *Support vector machine* (20 August 2020)

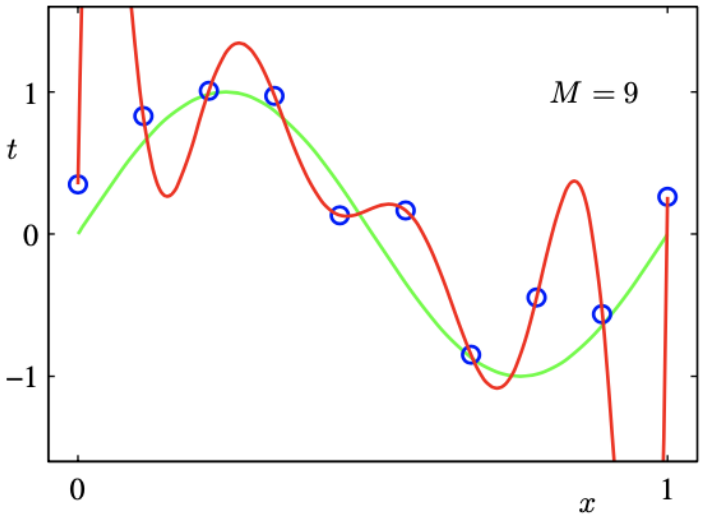
… but commonly an unintended side effect



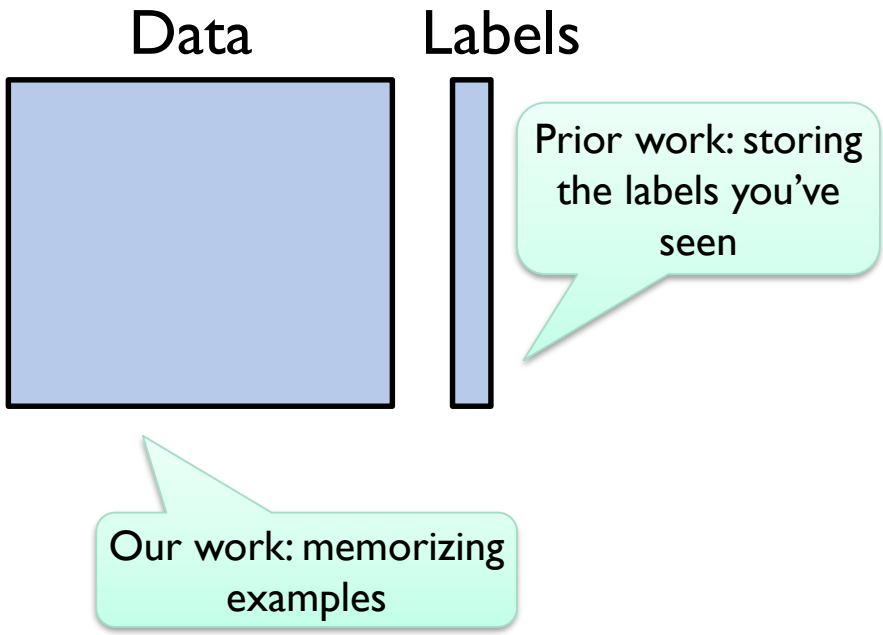
[Carlini et al. 20]

Current language models memorize irrelevant information.

Memorization ≠ fitting or interpolation

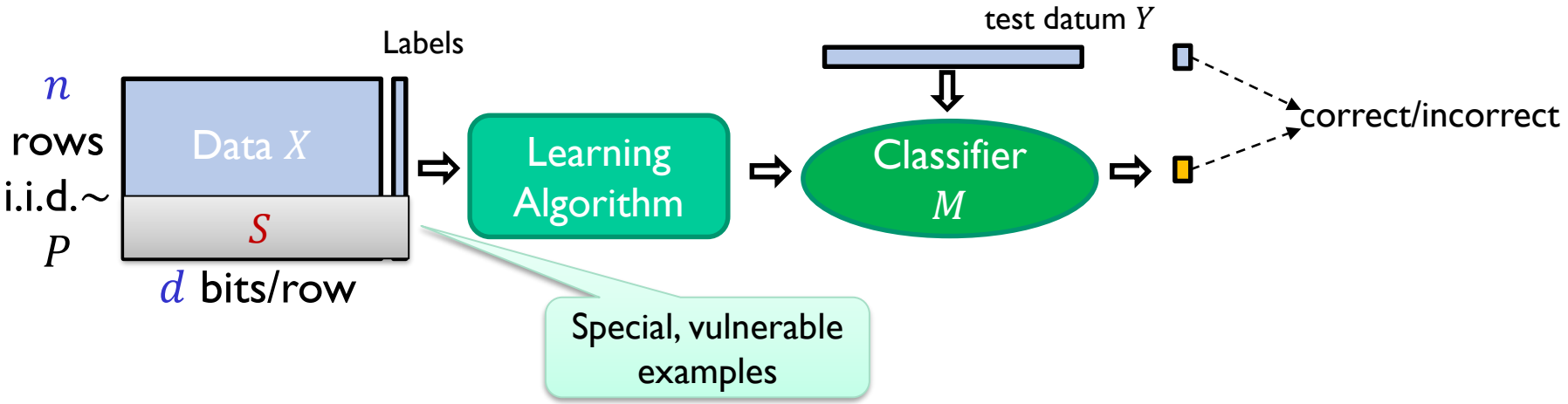


Christopher Bishop, "Pattern Recognition and Machine Learning," 2006



Memorization Can be Necessary

[Brown, Bun, Feldman, S, Talwar STOC 21]



"Theorem": There is a natural learning problem for which every data set X has a subset of rows $S \subseteq X$ such that

- S is "big": $|S| \geq n/10$ with high probability
- Every learning algorithm with low error memorizes most of S
  - If learning algorithm has error $OPT + (small)$, then

$$I(S; M|P) \geq d \cdot |S| \cdot (1 - small).$$

Understanding good generalization

Common explanations for large models

1. Expressivity
2. Optimization is easier
3. Implicit regularization leads to good generalization

Our results suggest an additional factor:

4. Large models store information whose usefulness isn't yet "understood"
   - Small subpopulations
   - Adapting to new domains

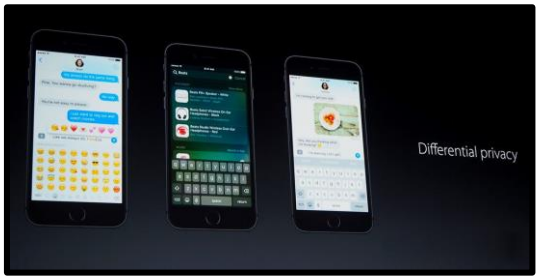
This talk

- Why is privacy challenging?
  - Memorization in machine learning
    [Brown, Bun, Feldman, S., Talwar 2021]

- Differential Privacy [Dwork, McSherry, Nissim, S., 2006]
  - "Privacy" as stability to small changes
  - Widely studied and deployed

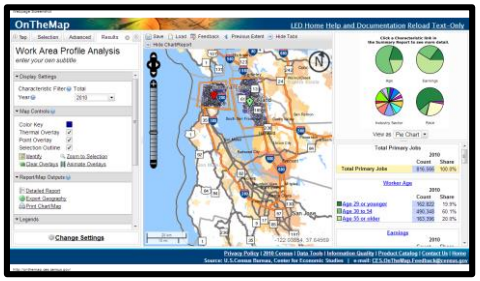
Differential Privacy [Dwork, McSherry, Nissim, S., 2006]

- Many current deployments



Apple

Google

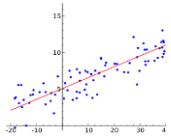


US Census

- Burgeoning field of research

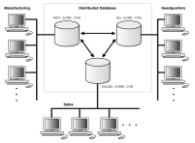
Algorithms

Crypto, security



Statistics, learning

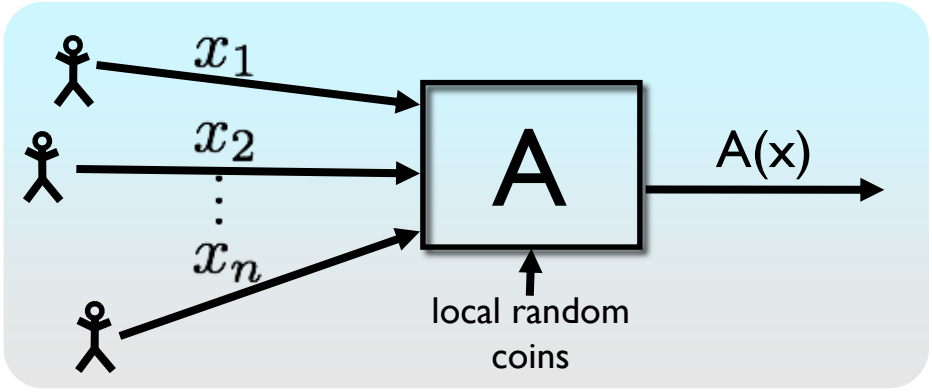
Game theory, economics

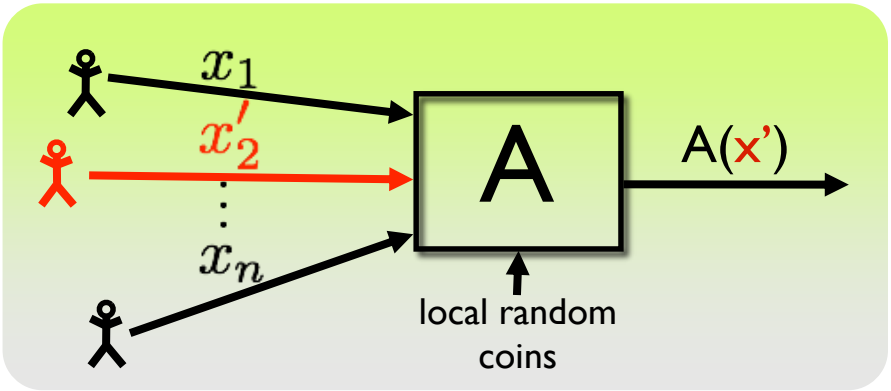


Databases, programming languages

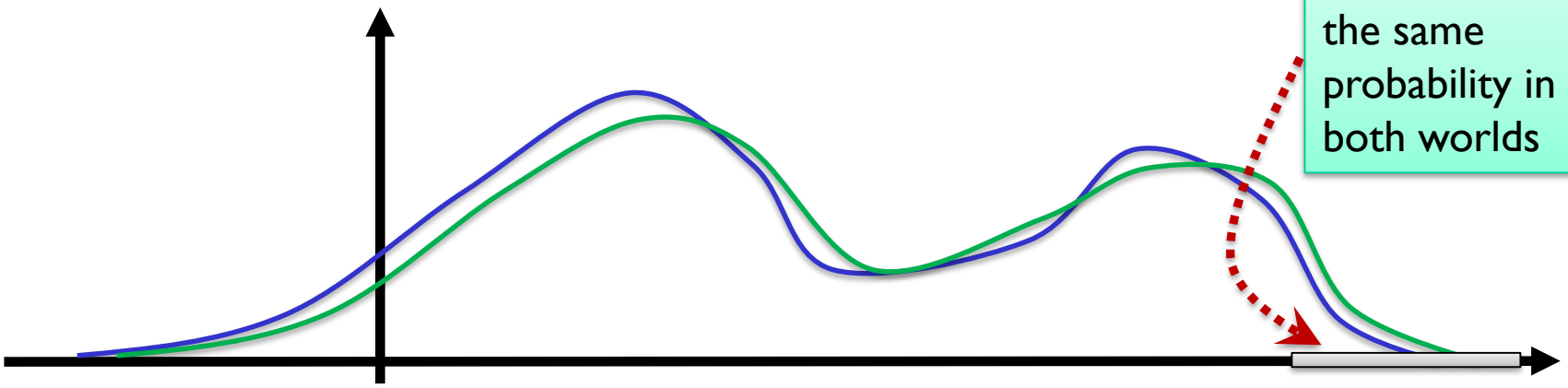
Law, policy

Differential Privacy [Dwork, McSherry, Nissim, S., 2006]

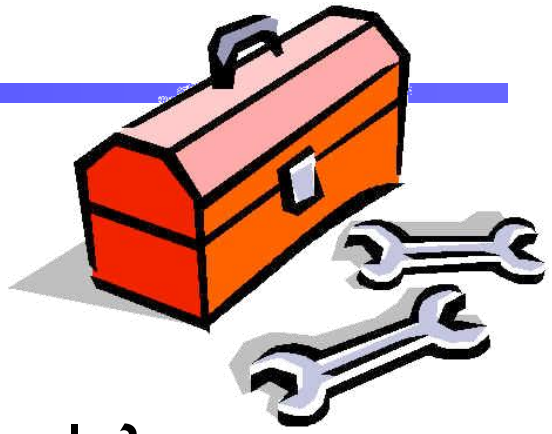




- A thought experiment
  - Change one person's data (or add or remove them)
  - Will the **distribution of outputs** change much?



Research on differential privacy



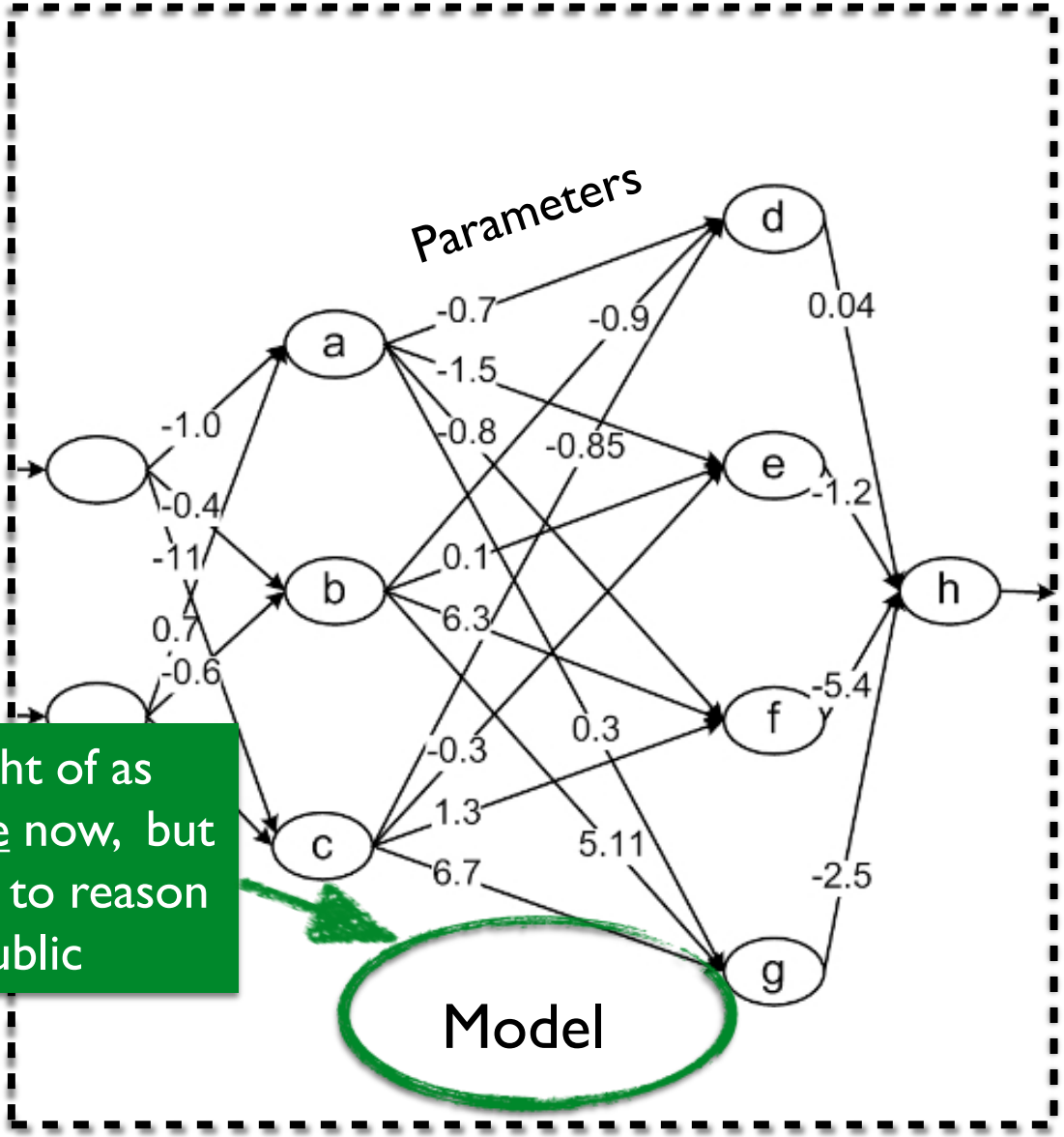
- Definitions
  - Pinning down "privacy"
- Algorithms: what can we compute privately?
  - Fundamental techniques
  - Specific applications
- Attacks: "Cryptanalysis" for data privacy
  - Impossibility results
- Implications for other areas
  - Interactive machine learning and statistical analysis

Frontier: Deep Learning with DP

McSherry Williams 09,
Chaudhuri, Song, Sarwate 13,
Bassily, S., Thakurta 2014,
Abadi et al 2016, …

Sensitive Data

Deep Learning



Revealed now, but should be hidden

Thought of as private now, but better to reason as if public

Example: https://github.com/tensorflow/privacy

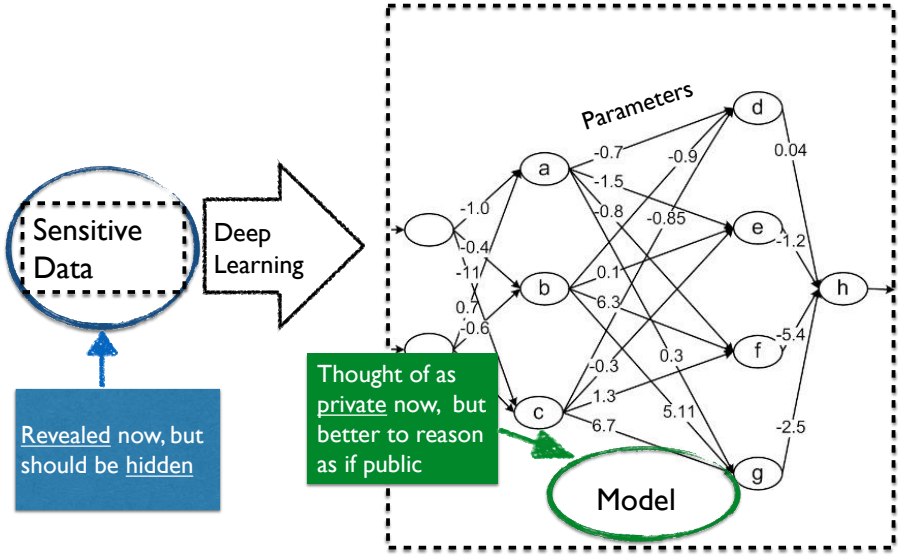
Frontier: Deep Learning with DP

Main technique

- Noisy SGD

Challenges

- Tools for inference
  - Constraints complicate interpretation
- Computational advances
  - How best to leverage huge advances in optimization?
  - (Often privacy requires convergence)
- Tighter analysis of privacy (and other) properties
- General-purpose algorithms



What are we missing?

- Technical principles for personal ownership of personal data
  - What does it mean to control use?
- Whom do privacy technologies empower?
  - Big tech?
- Painless processes for tech-policy dialogue
  - (I don't want to read your court opinions, and you don't want to read my papers.)