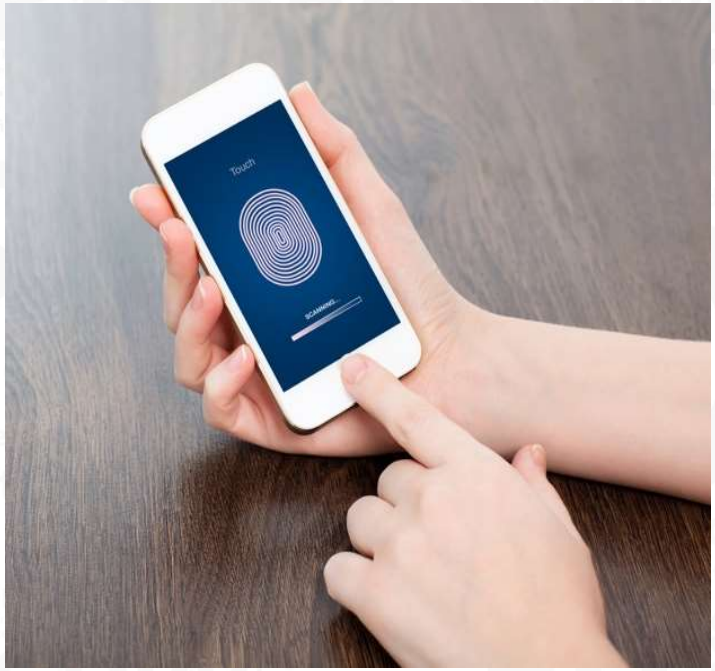
The Challenge of AI and Machine Learning: Privacy and Ethics in Data Science

Steven Greenspan, PhD

APRIL 2018

What is Privacy? How Do We Maintain it?

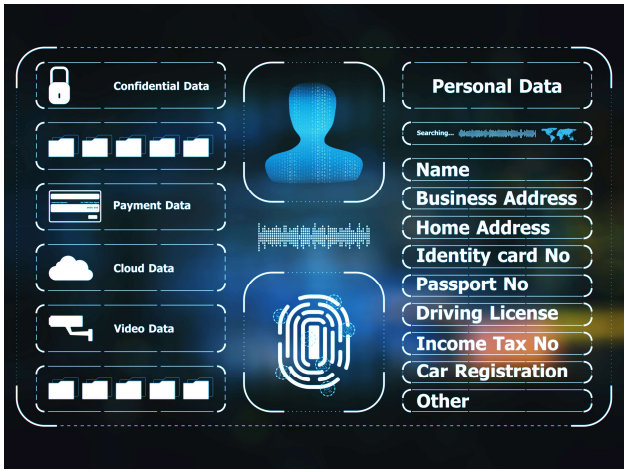


CONTROL OVER WHO SEES MY DATA

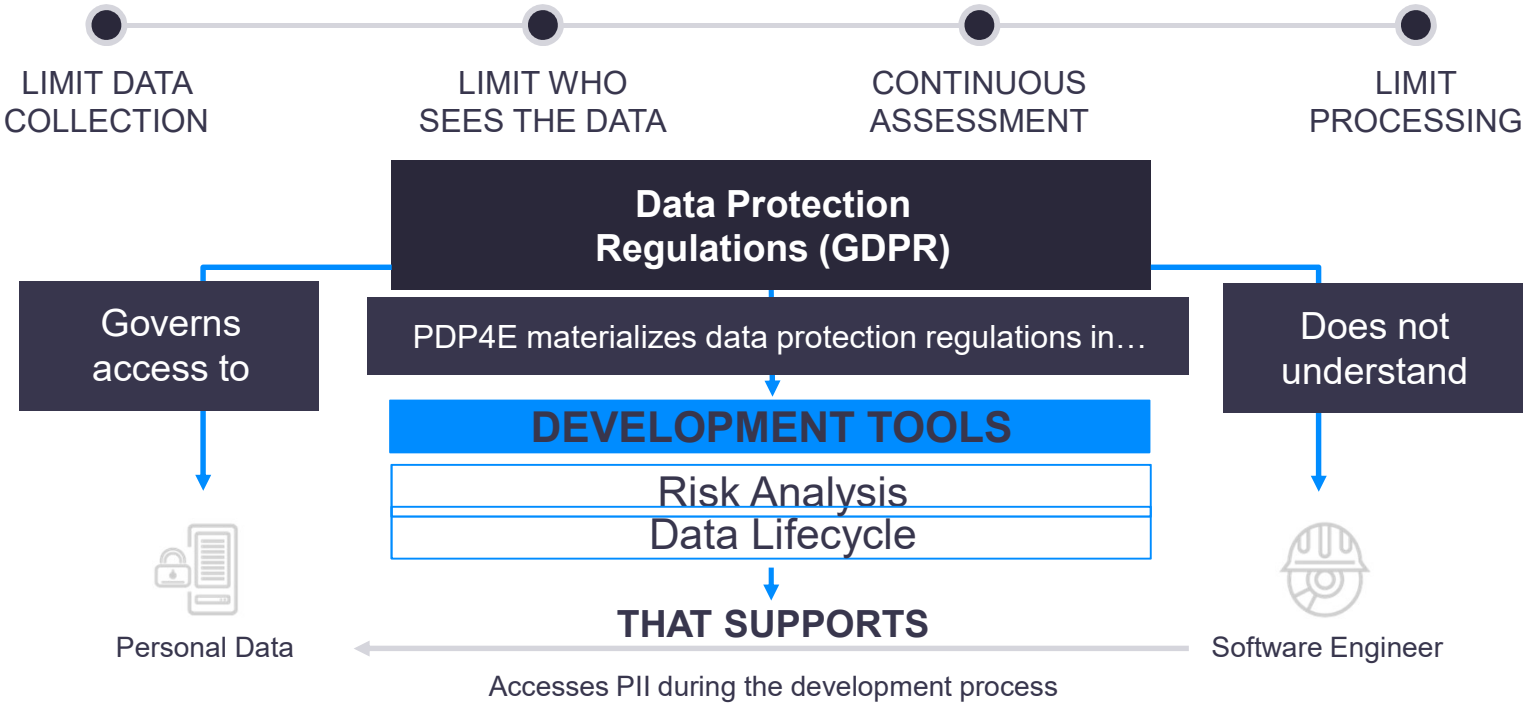
Privacy by Design:

GDPR defines "what" but not "how"

- LIMIT DATA COLLECTION
- LIMIT WHO SEES THE DATA
- CONTINUOUS ASSESSMENT
- LIMIT PROCESSING



Privacy by Design:
Privacy and Data Protection 4 Engineers



Controlling Data Access

LIMIT DATA COLLECTION | LIMIT WHO SEES THE DATA | CONTINUOUS ASSESSMENT | LIMIT PROCESSING

USER BEHAVIORS: WHAT YOU DO

CLIENT SIDE
- typing/mouse patterns
- webcam (e.g. detect person leaving desk, background movements)
- biometrics (e.g. facial recognition)
- walking patterns

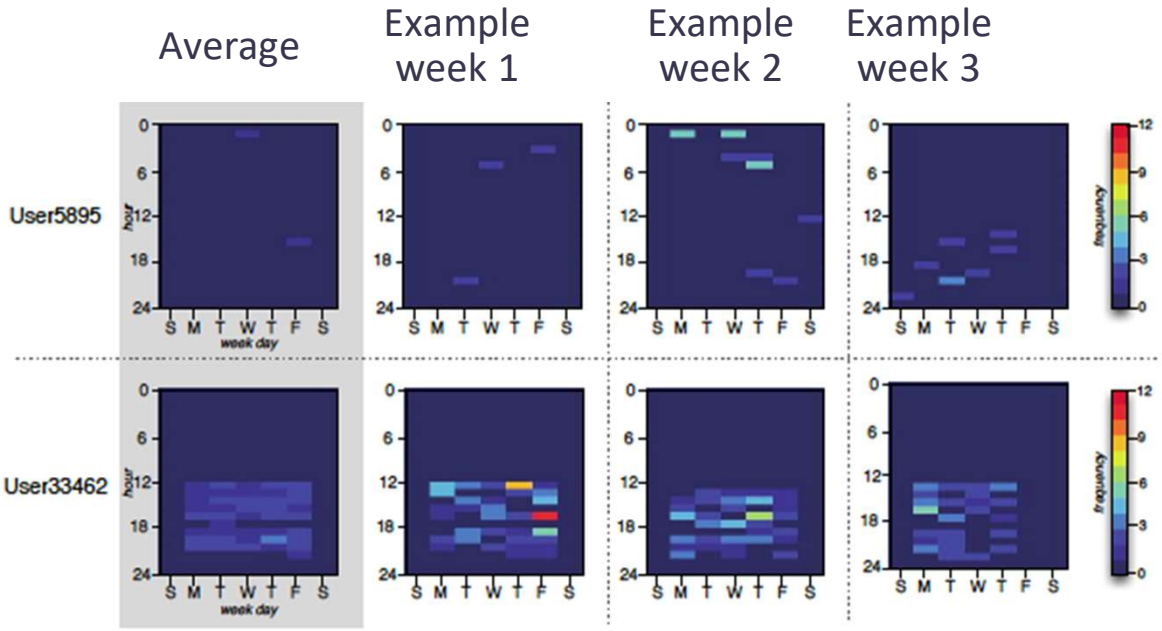
SERVER SIDE
- access pattern history - devices, data bases, …
- network/location patterns
- time of day
- building access logs

Related Project: Risk-Adaptive Continuous Authentication

We are unique in work
and play and in-between

Server side identification of anonymous and legitimate users

Classifier	Precision	Recall	TPrate	FPrate
Random Forest	98.7%	99.1%	99.1%	1.4%



Research conducted by:

Li Sun, Steve Versteeg, and Serdar Boztas

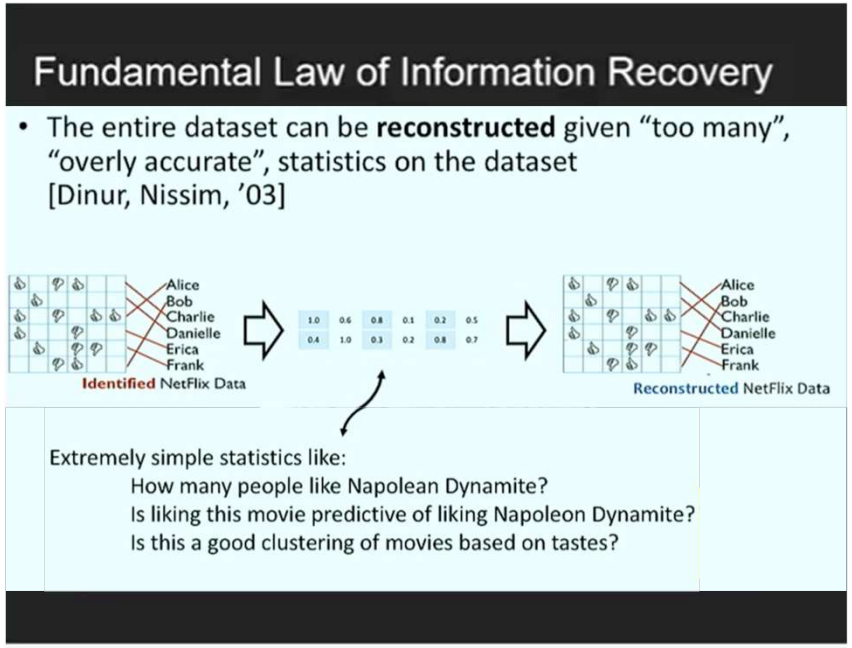
Is Control Over Who Sees My Data Sufficient?

WHAT ELSE CAN THEY INFER ABOUT ME?

- Anonymization does not guarantee privacy!
- Some use cases for sharing data
  - Medical Researchers
  - Customized User Experience
  - Bank Fraud

Anonymization is not sufficient
Fundamental Law of Information Recovery

- Anonymized data can be re-identified by **linking PII with external sources of information**
  - Analyzing 15 months of anonymized mobile phone data for about 1.5 million users, found that it took very few pieces of data to uniquely identify 95 percent of the users – that is, trace the activity to a specific anonymous individual. (*How hard is it to 'de-anonymize' cellphone data? (2013),* MIT & Universite Catholique de Louvain, Belgium)
  - In August 2016, Australia's federal Department of Health published medical billing records of about 2.9 million Australians online. University of Melbourne team found that patients can be re-identified, without decryption, through a process of linking the unencrypted parts of the record with known information about the individual. (*The Simple Process of Re-Identifying Patients on Public Health Records (2017),* U Melbourne)
- An entire dataset can be reconstructed given "too many", **"overly accurate "** statistics on the data set. (Dinur, Nissim, 2003)



Sharing Data While Preventing Inferences that Violate GDPR

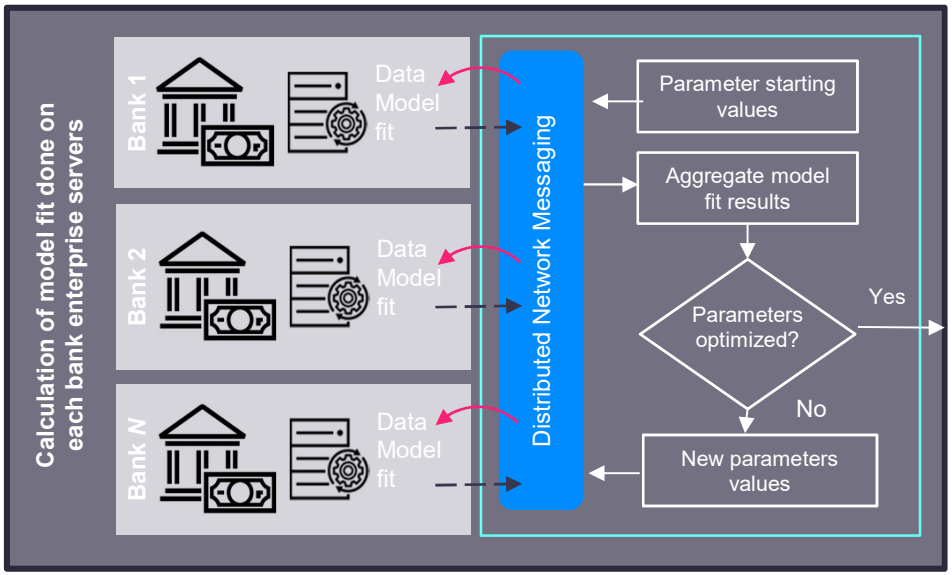
LIMIT DATA COLLECTION

LIMIT WHO SEES THE DATA

CONTINUOUS ASSESSMENT

LIMIT PROCESSING

ANOTHER PRIVACY DEFINITION:
Analyst has no way of knowing whether or not a particular person was included in the database



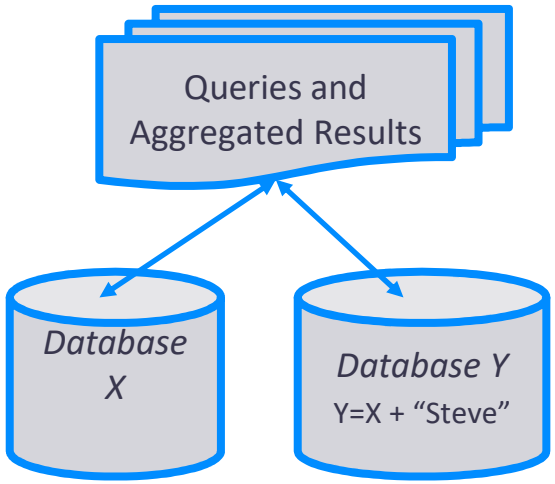
Applying differential privacy to distributed deep learning

Related Project: Privacy-Preserving Multiparty Analytics

Differential Privacy

Differential privacy essentially masks the contribution of any single individual, making it impossible to infer any information specific to an individual, including whether the individual's information was used at all.

DP resists linkage attacks and auxiliary information,
DP supplies a quantifiable measure of harm incurred by individuals and composes nicely.

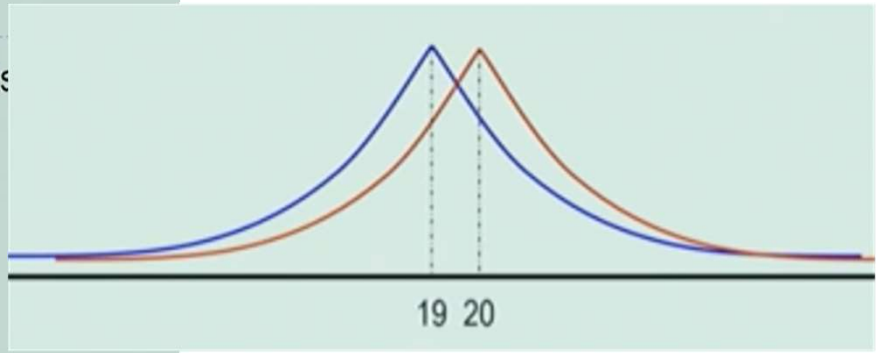


Differential Privacy

M gives ϵ-differential privacy if for all pairs of data sets differing in the data of one person, and all events S

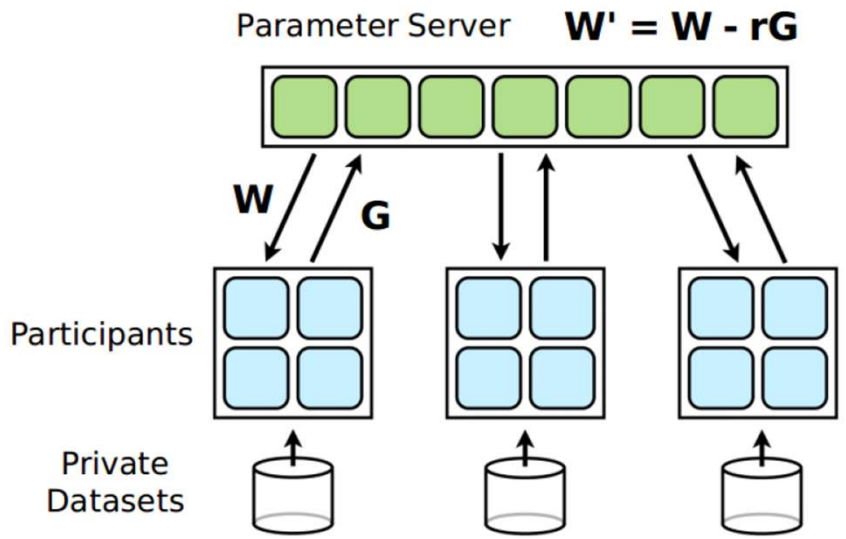
$$\Pr[M(x) \in S] \leq e^{\epsilon} \Pr[M(y) \in S]$$

Randomness introduced by M



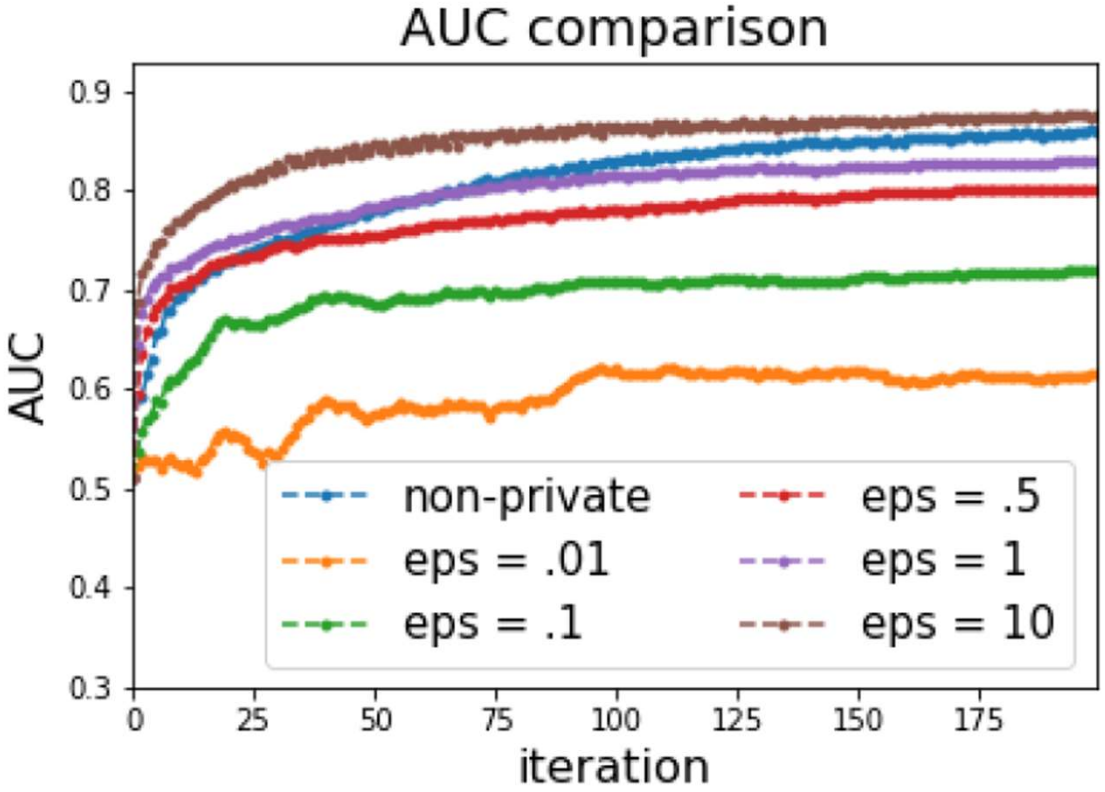
Current Status – technical progress

- Theoretical contributions
  - Extended the application of differential privacy to deep learning architectures
  - Developed evaluation methodology
  - New techniques for improving utility and efficiency



Current Status – technical progress

- Theoretical contributions
  - End-to-end privacy guarantee on distributed deep learning
  - Better utility by applying
    - mini-batch sub-sampling,
    - focusing on parameters with large gradients
  - Efficient learning process with less communication bandwidth requirement



Creating a Framework and Tools for Privacy by Design

LIMIT DATA COLLECTION

LIMIT WHO SEES THE DATA

CONTINUOUS ASSESSMENT

LIMIT PROCESSING

We are creating technology that help software engineers

- Continuously assess GDPR compliance during development and after deployment
- Provide continuous control over who sees the data
- Prevent data inferences that violate privacy policies

Steven Greenspan
VP & Research Scientist
Steven.Greenspan@ca.com

www.linkedin.com/in/steveg72

Research.ca.com