Trustworthy AI: Privacy

Destiny Pounds

M.S. Student, Biomedical Data Science

Advised by Vibhuti Gupta, Ph.D.

Assistant Professor, Computer Science and Data Science

School of Applied Computational Sciences

Meharry Medical College

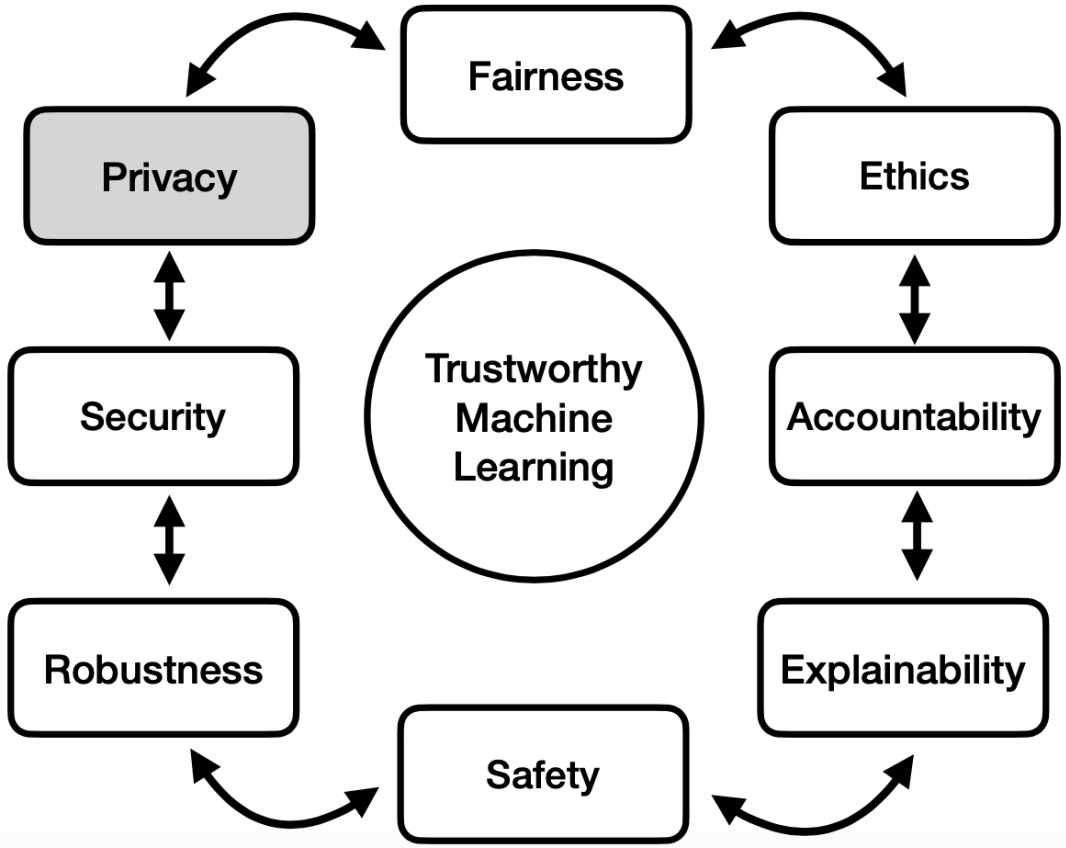
Overview

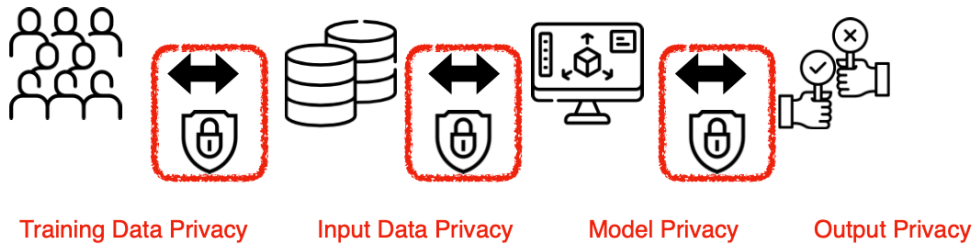
- Defining Privacy
- Privacy-Preserving Machine Learning
- PPML Techniques
  - Data Anonymization
  - Differential Privacy
  - Homomorphic Encryption
  - Secure Multi-party Computation
  - Federated Learning

US agencies buy vast quantities of personal information on the open market – a legal scholar explains why and what it means for privacy in the age of AI

Published: June 29, 2023 8:16am EDT

Privacy

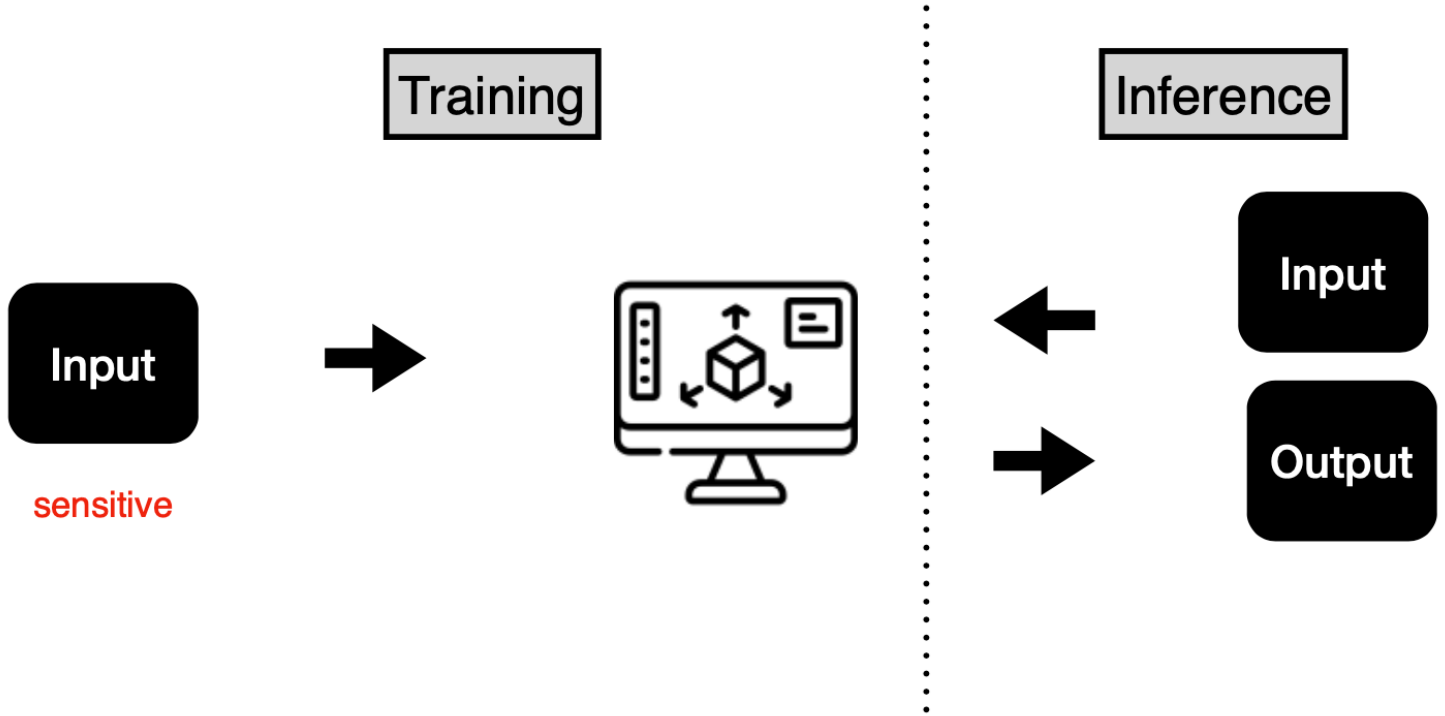




- Data privacy is a central issue to training and testing AI models, especially ones that train and infer on sensitive data.

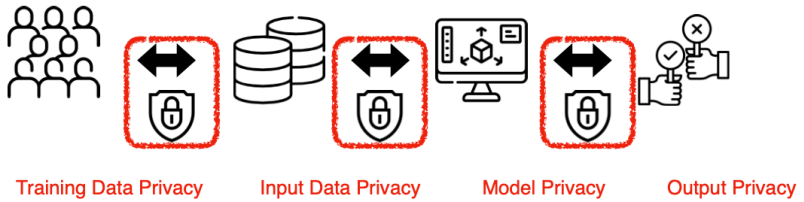
Adapted from Farnadi, G. (2022). *Trustworthy Machine Learning* [Slides]. HEC Montréal

Privacy Setting in AI/ML



We can gain useful insight about the population without knowing about individuals.

Privacy – Preserving Machine Learning

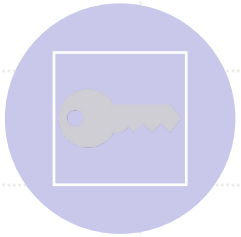


- Privacy-preserving machine learning (PPML) is a set of techniques and practices that safeguard sensitive data during training and deployment of AI models

Privacy-Preserving Techniques

DATA ANONYMIZATION

DIFFERENTIAL PRIVACY



HOMOMORPHIC ENCRYPTION

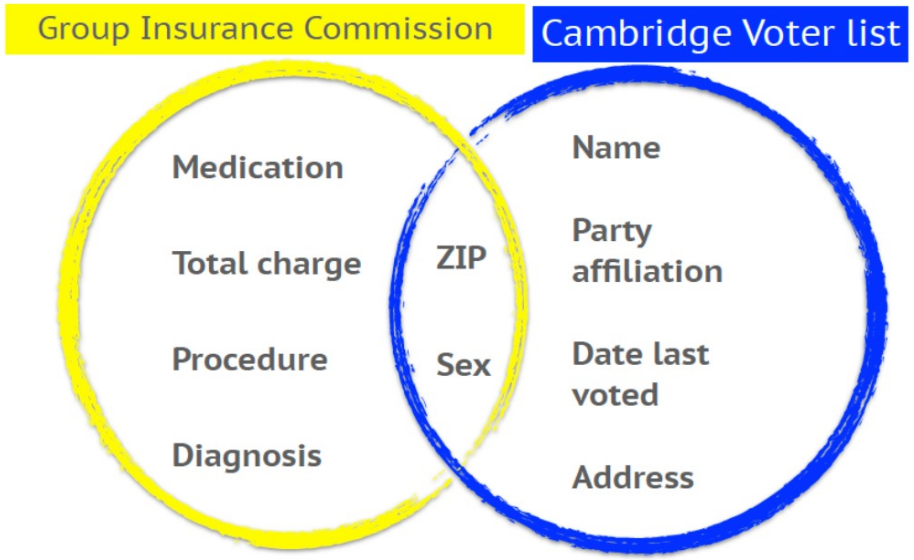
SECURE MULTI-PARTY COMPUTATION

FEDERATED LEARNING

Data Anonymization

Data anonymization is a process of modifying data in a way that eliminates connections to specific individuals.

[Sweeny 02]



Name	Sex	Party	Date Last Voted	Address	Zip
Gov. Sam Thomas	M	R	…	…	12345
Lt. Gov. Angie Stevenson	F	R	…	…	12354
Sen. Paul Childs	M	D	…	…	12346
Sen. Lisa Wells	F	D	…	…	12345
Cong. Tim Allen	M	R	…	…	12355
Cong. Rose Smith	F	D	….	…	12345

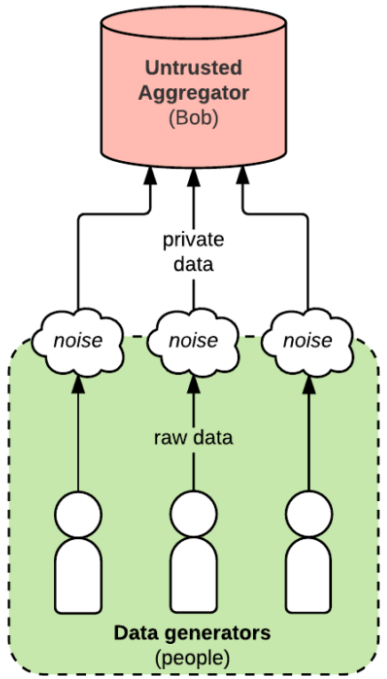
Ex. A male living in zip code 12345 was diagnosed with lung cancer. Who could it be?

Of the six people listed, three are men but only one lives within that zip code.

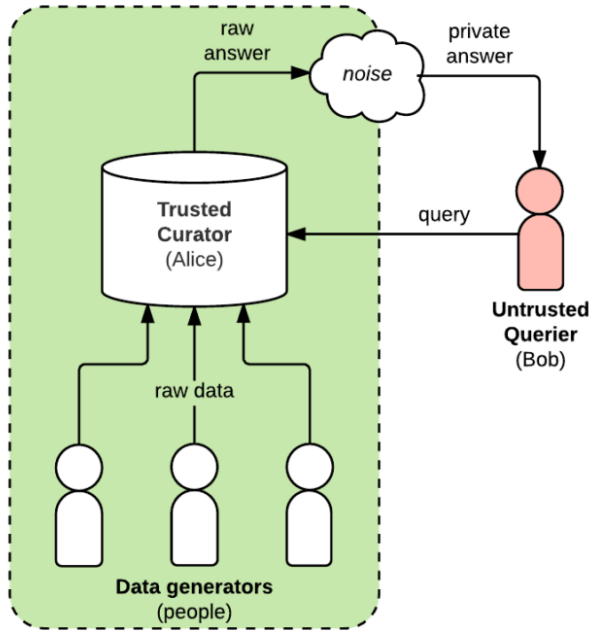
It must be Governor Sam Thomas!

Differential Privacy

- Differential privacy is a framework designed to ensure the privacy of individuals in a dataset. Noise is added to the dataset which makes it difficult for attackers to discern information that is specific to any individual.

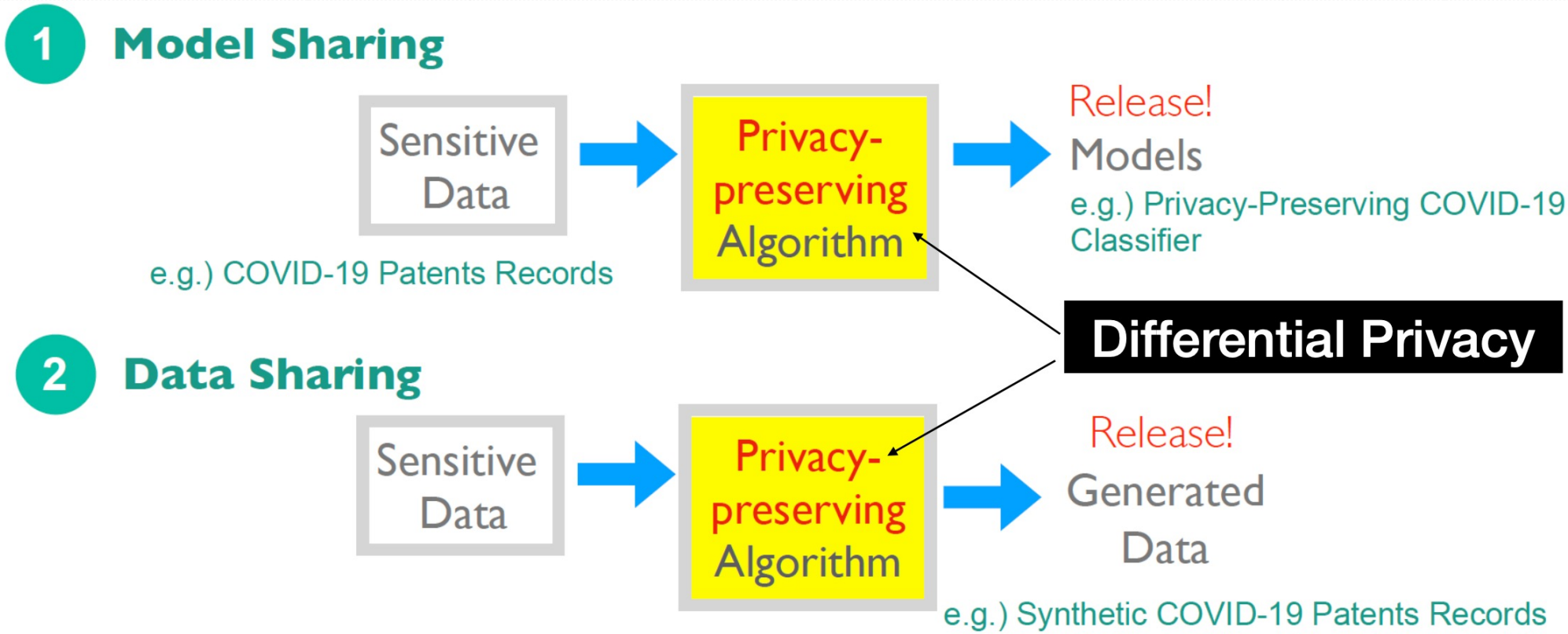


Local privacy



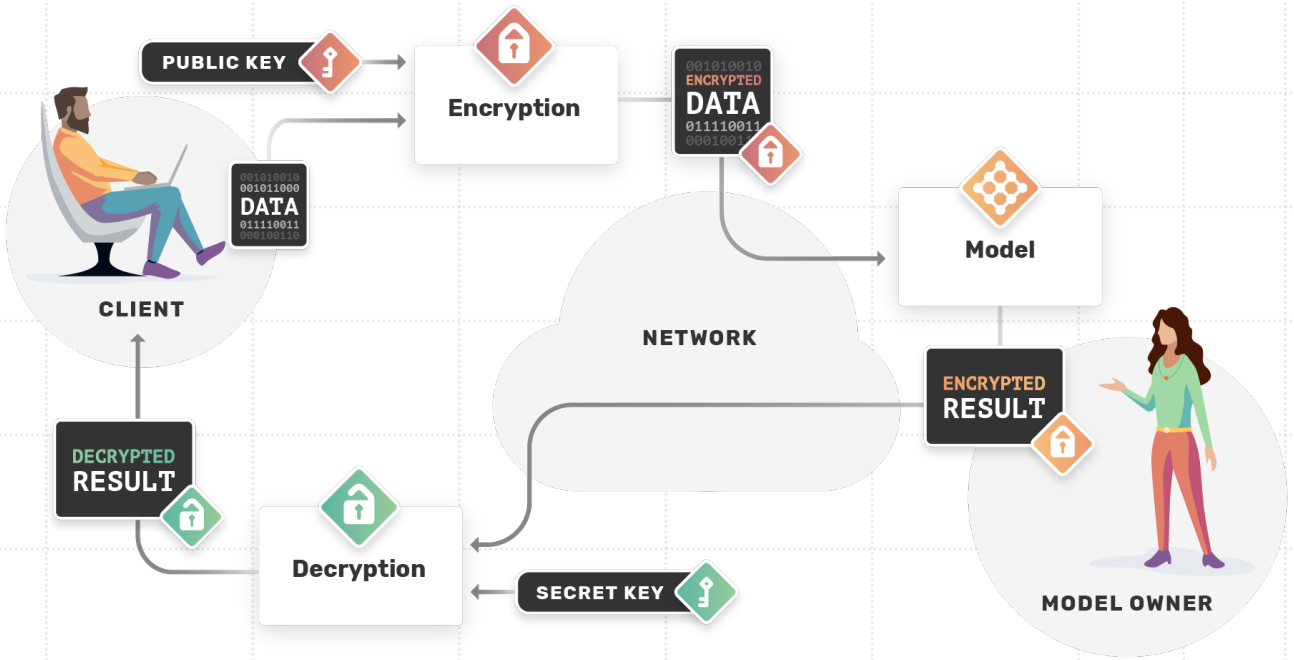
Global privacy

Privacy Settings: Single Data Source

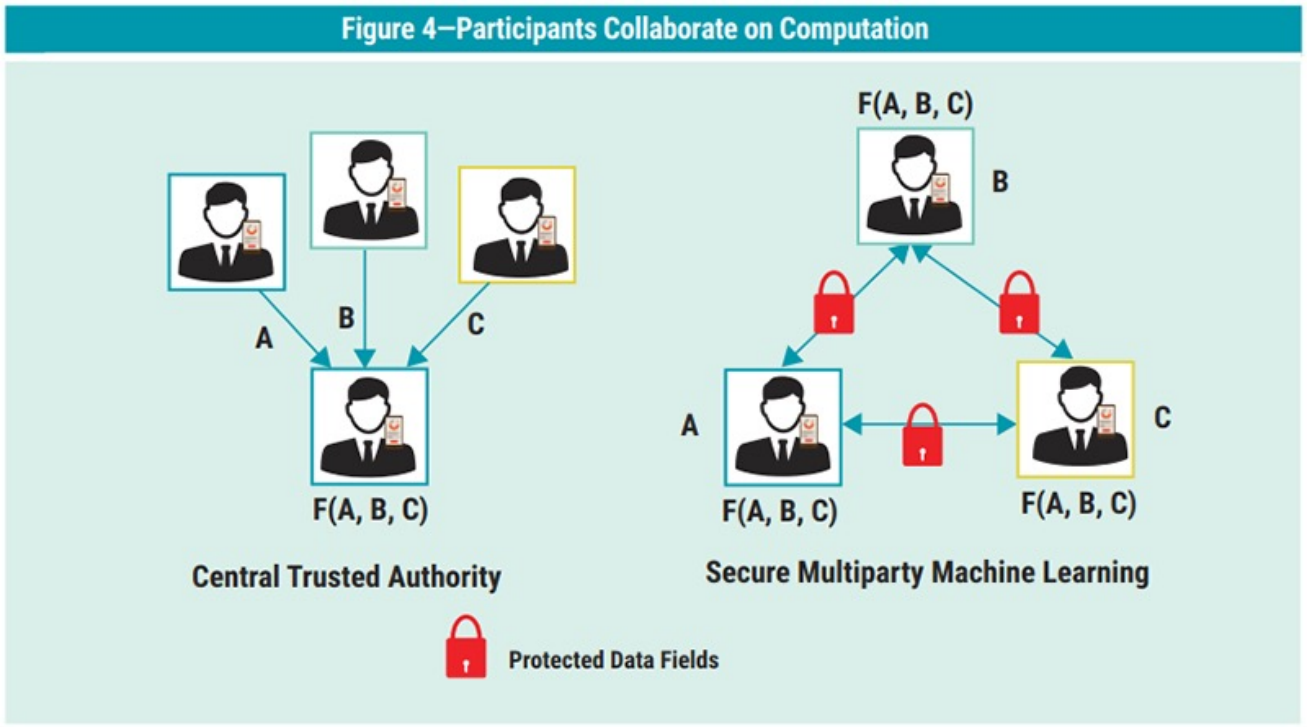


Homomorphic Encryption

- Homomorphic encryption is conversion of data into a coded format that still allows it to be manipulated like original data without compromising encryption.

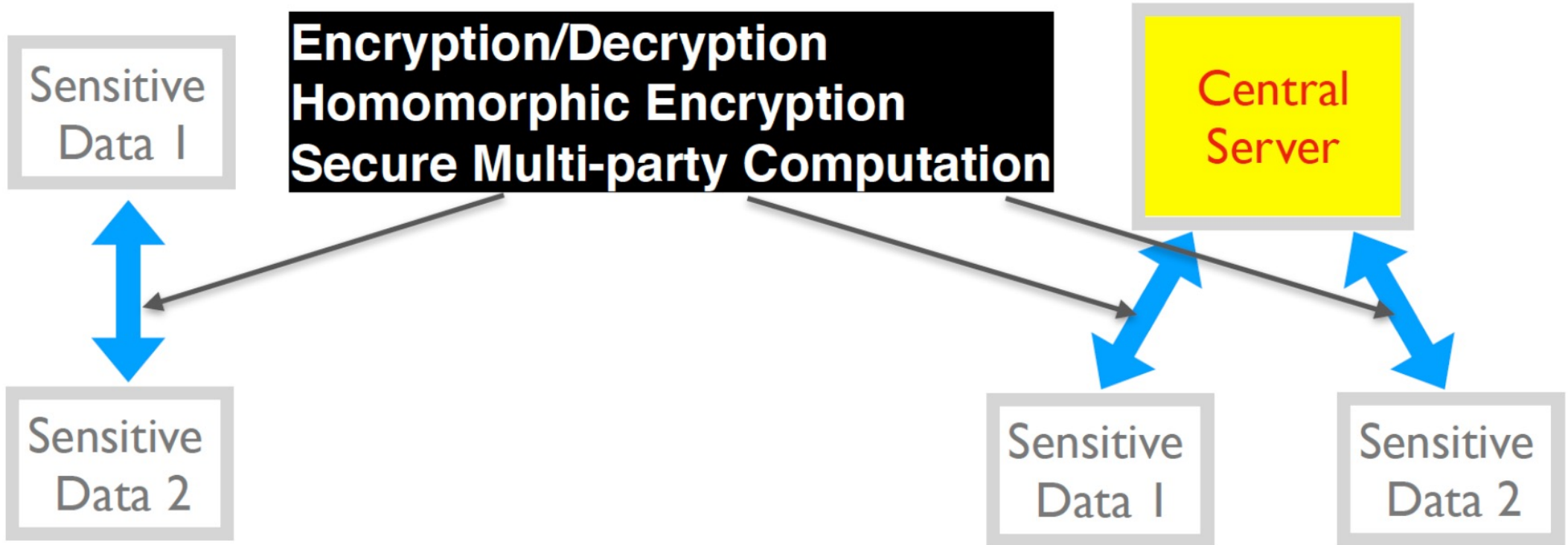


Secure Multi-party Computation



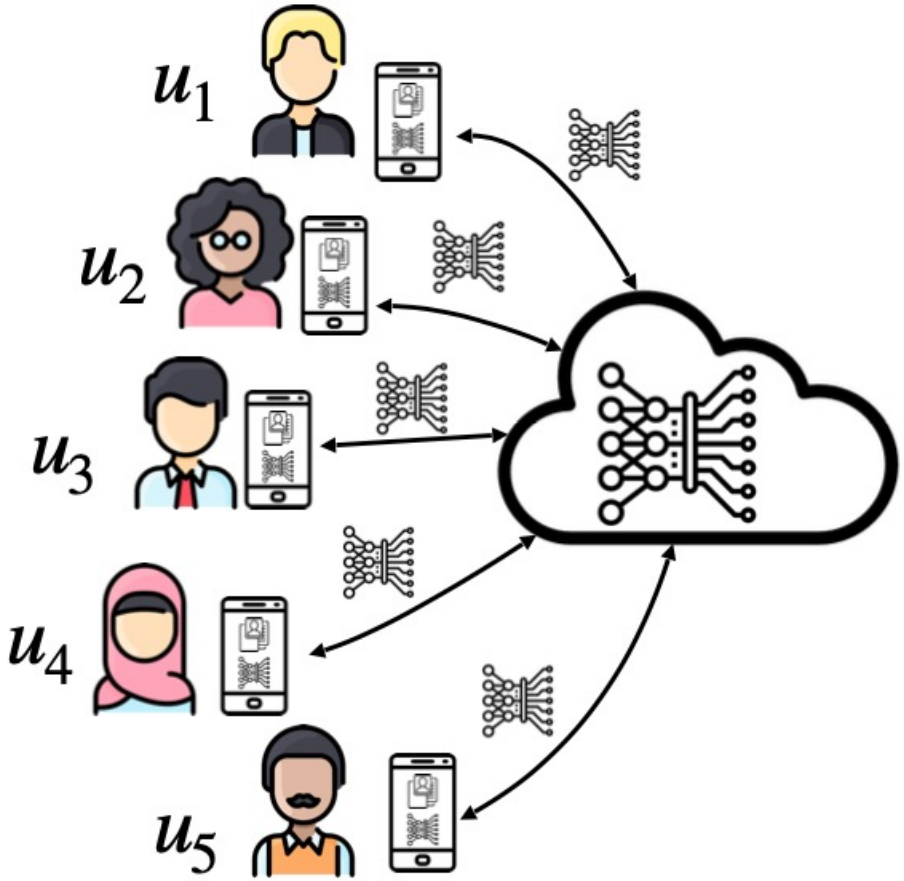
- Secure multi-party computations is a technique that allows multiple parties to collectively perform computations on their combined data while ensuring the privacy of individual information.

Privacy Settings: Multiple Data Sources



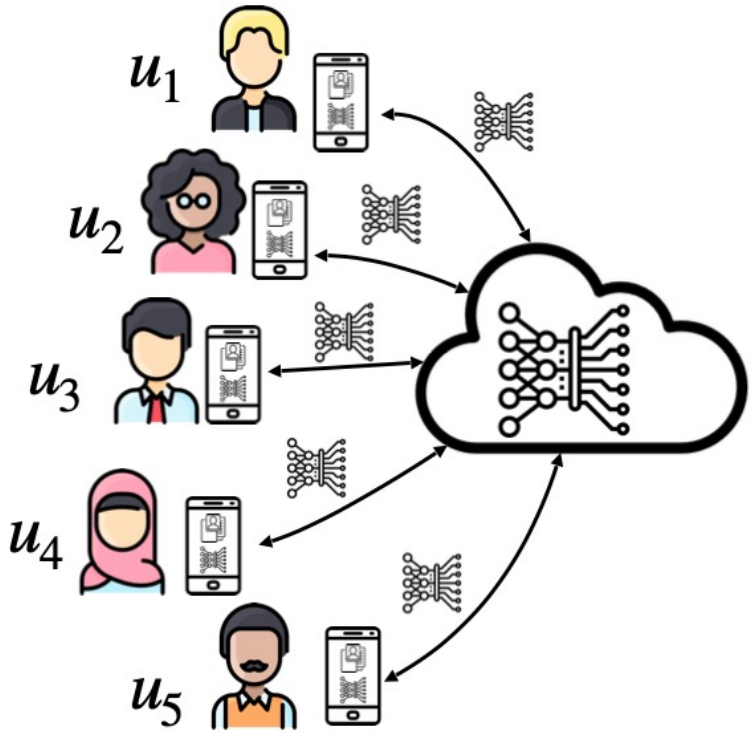
Federated Learning

- Federated learning is a machine learning setting where multiple entities (clients) collaborate in solving a problem, under the coordination of a central server or service provider.



Federated Learning

Data is generated locally and remains decentralized. Each client stores its own data and cannot read the data of other clients.



A central server/service coordinates training, but never sees raw data.

Thank You

Please send us your questions at:

vgupta@mmc.edu and

dpounds24@email.mmc.edu