EPI RQ4 Research Update: Privacy Preserving Distributed Machine Learning

Saba Amiri
s.amiri@uva.nl

Supervisor: Adam Belloum

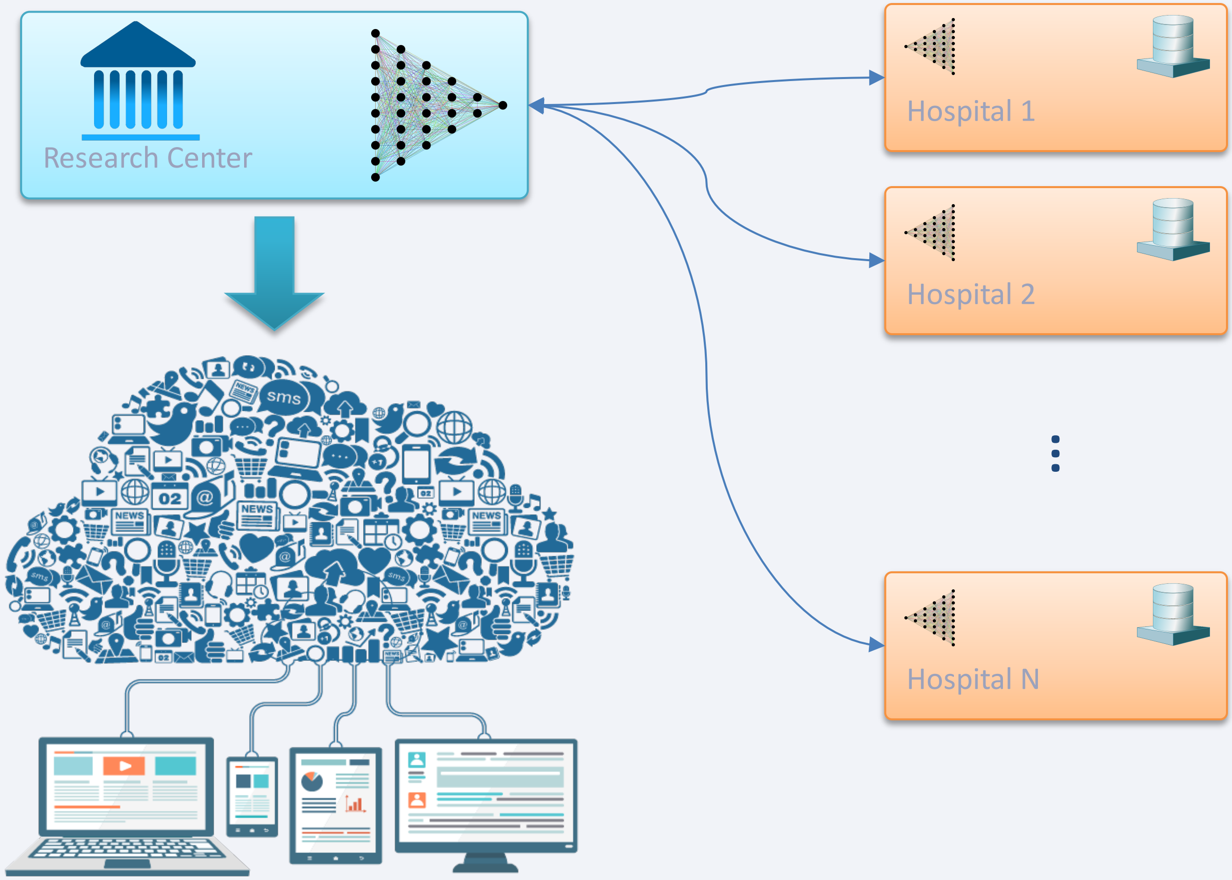
Promoters: Sander Klous, Leon Gommans

Multiscale Networked Systems Group

23 April 2021

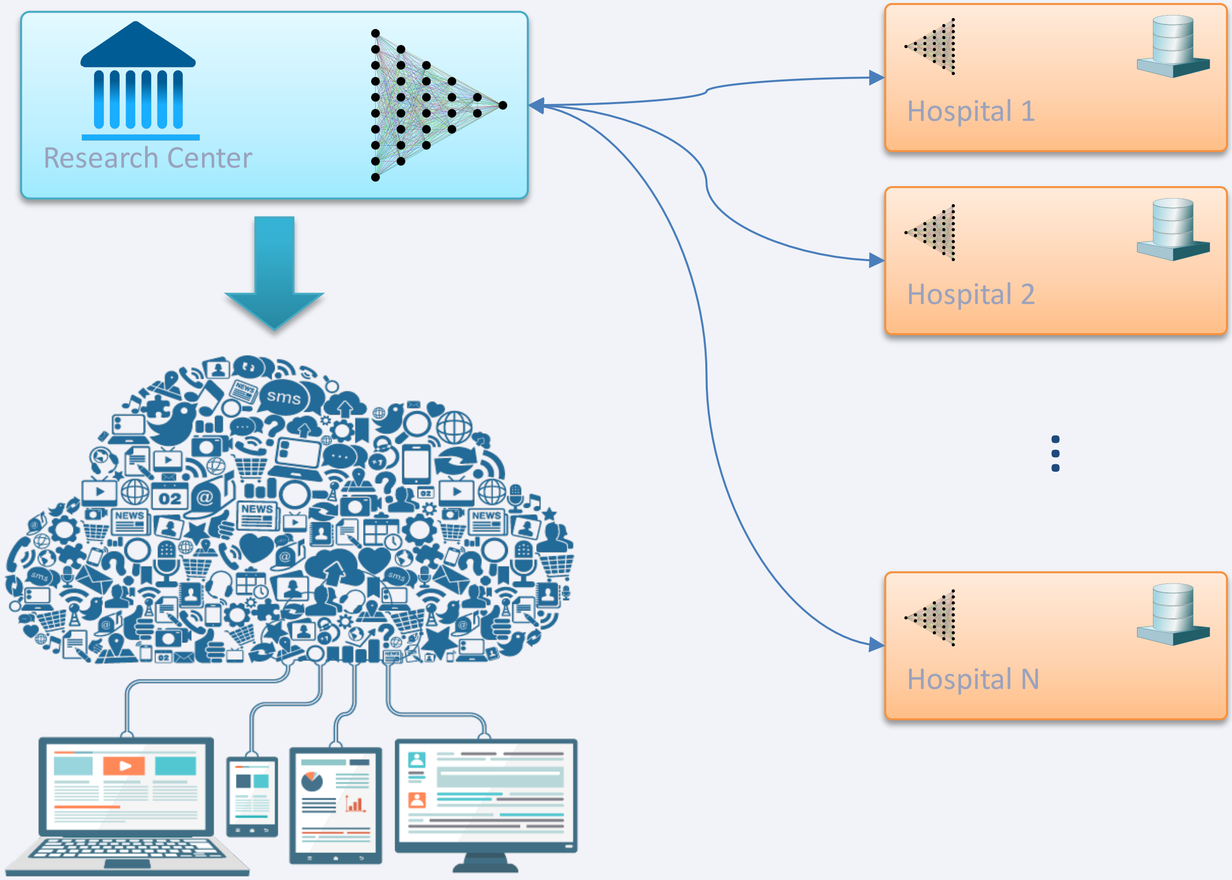
Research Domain

- Digital Health Twin



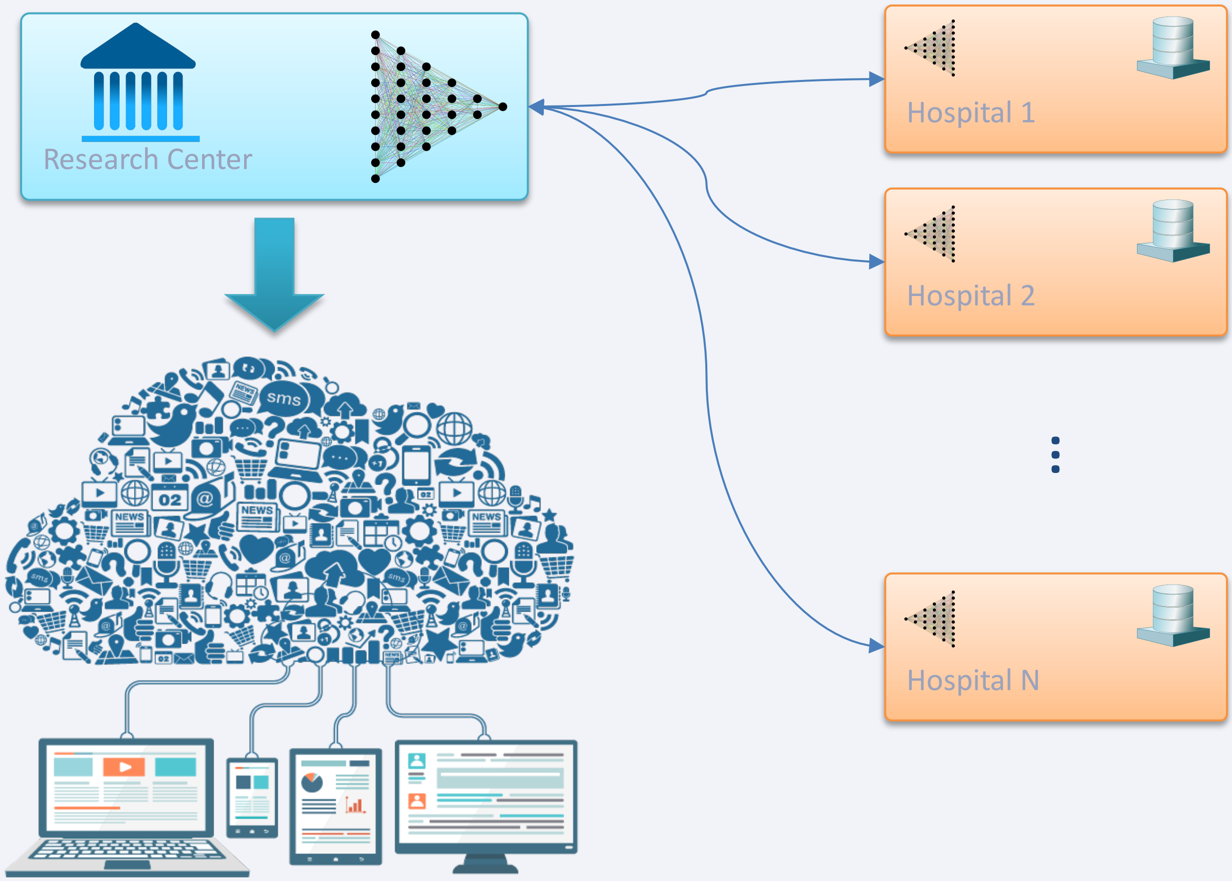
Research Domain

- Digital Health Twin
- Distributed Learning



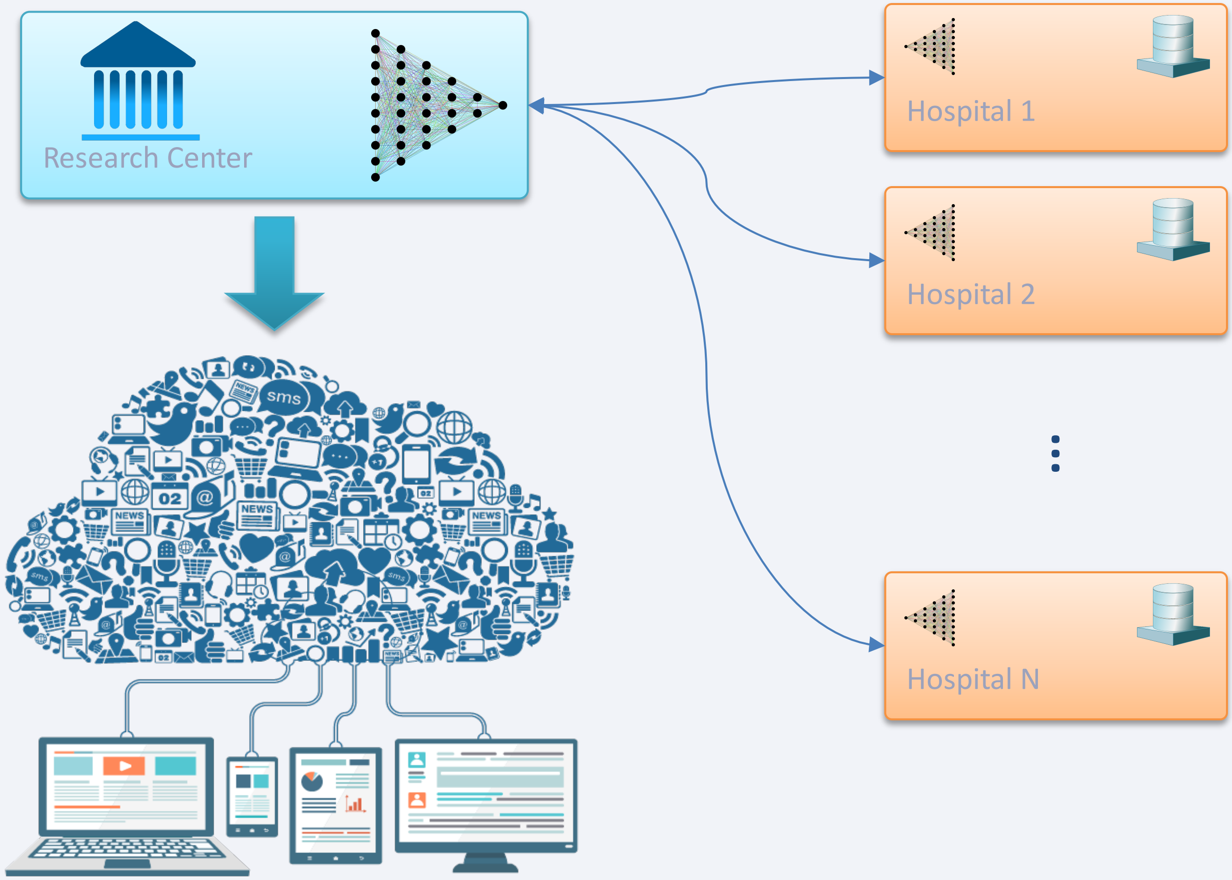
Research Domain

- Digital Health Twin
- Distributed Learning
- Privacy Preservation



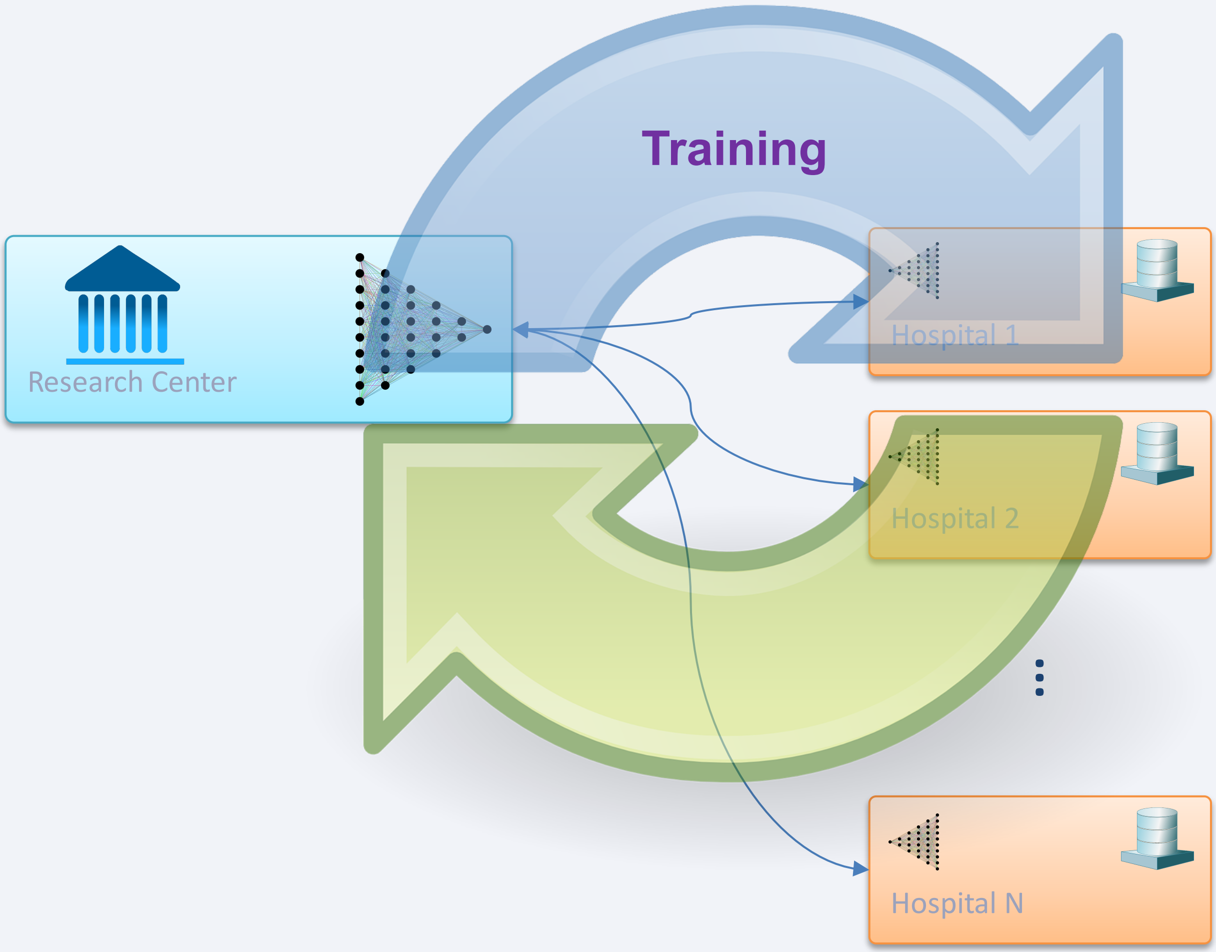
Definition of Privacy

- Digital Health Twin
- Distributed Learning
- Privacy Preservation
  - Definition: Providing patient/record level protection to every member of the training set while gaining useful insights about the populations as a whole



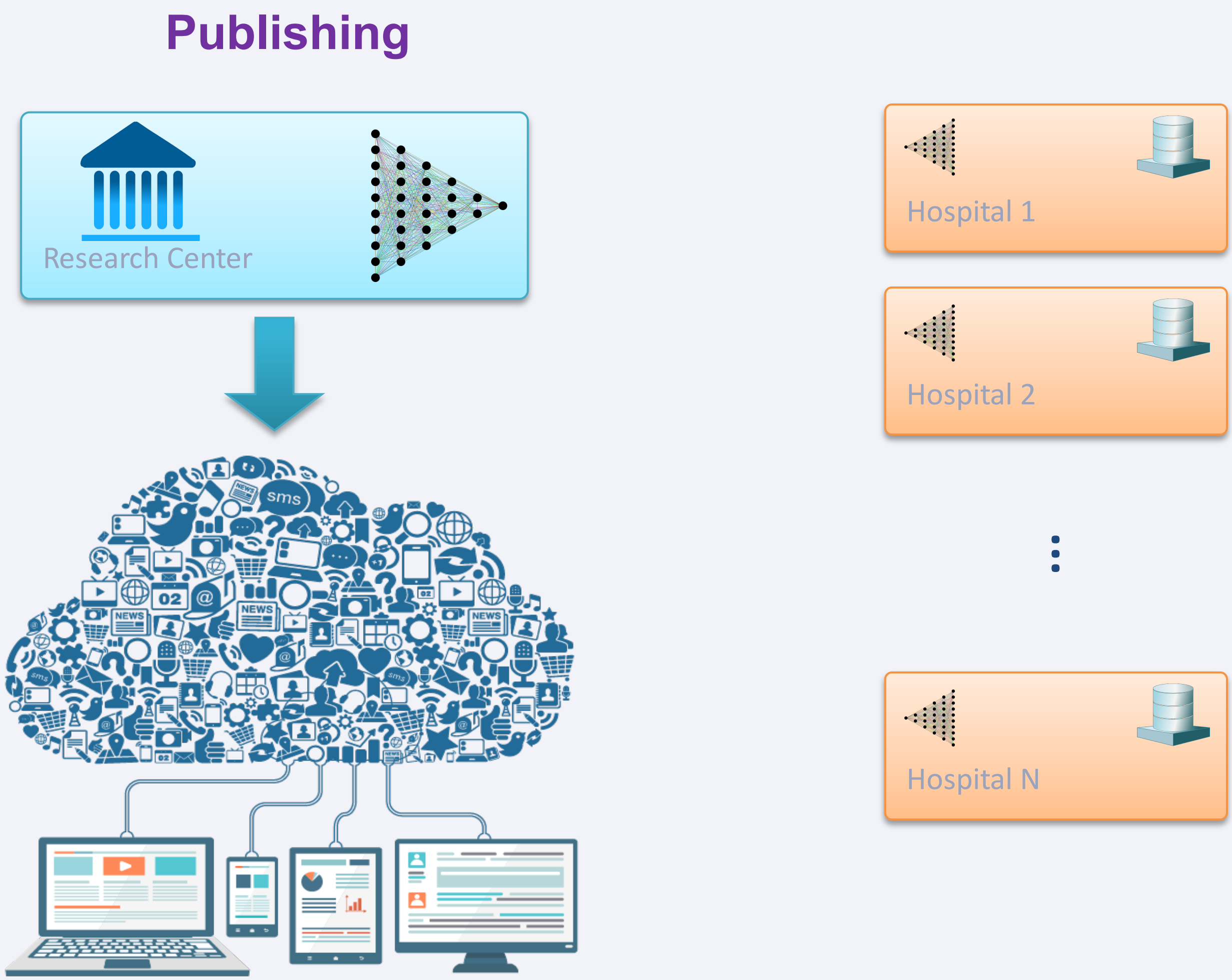
Typical Federated Learning Scenario

- Digital Health Twin
- Distributed Learning
- Privacy Preservation
  - Definition: Providing patient/record level protection to every member of the training set while gaining useful insights about the populations as a whole



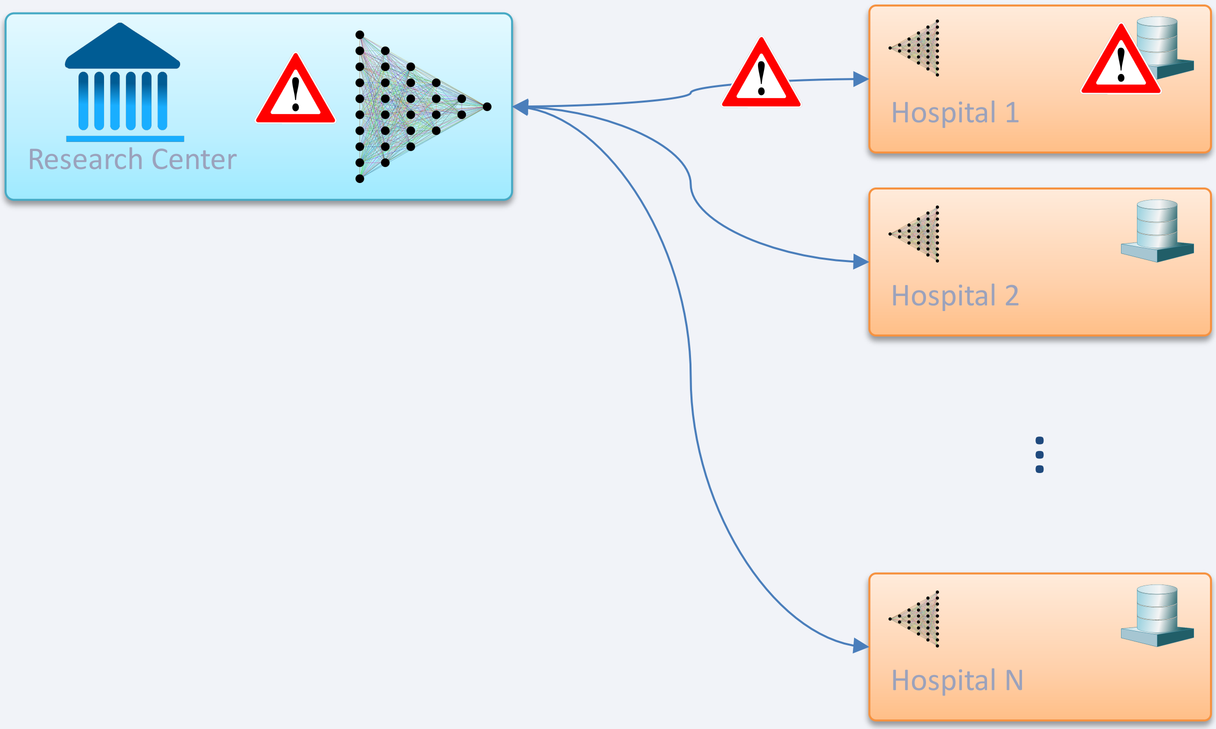
Typical Federated Learning Scenario - Publishing

- Digital Health Twin
- Distributed Learning
- Privacy Preservation
  - Definition: Providing patient/record level protection to every member of the training set while gaining useful insights about the populations as a whole



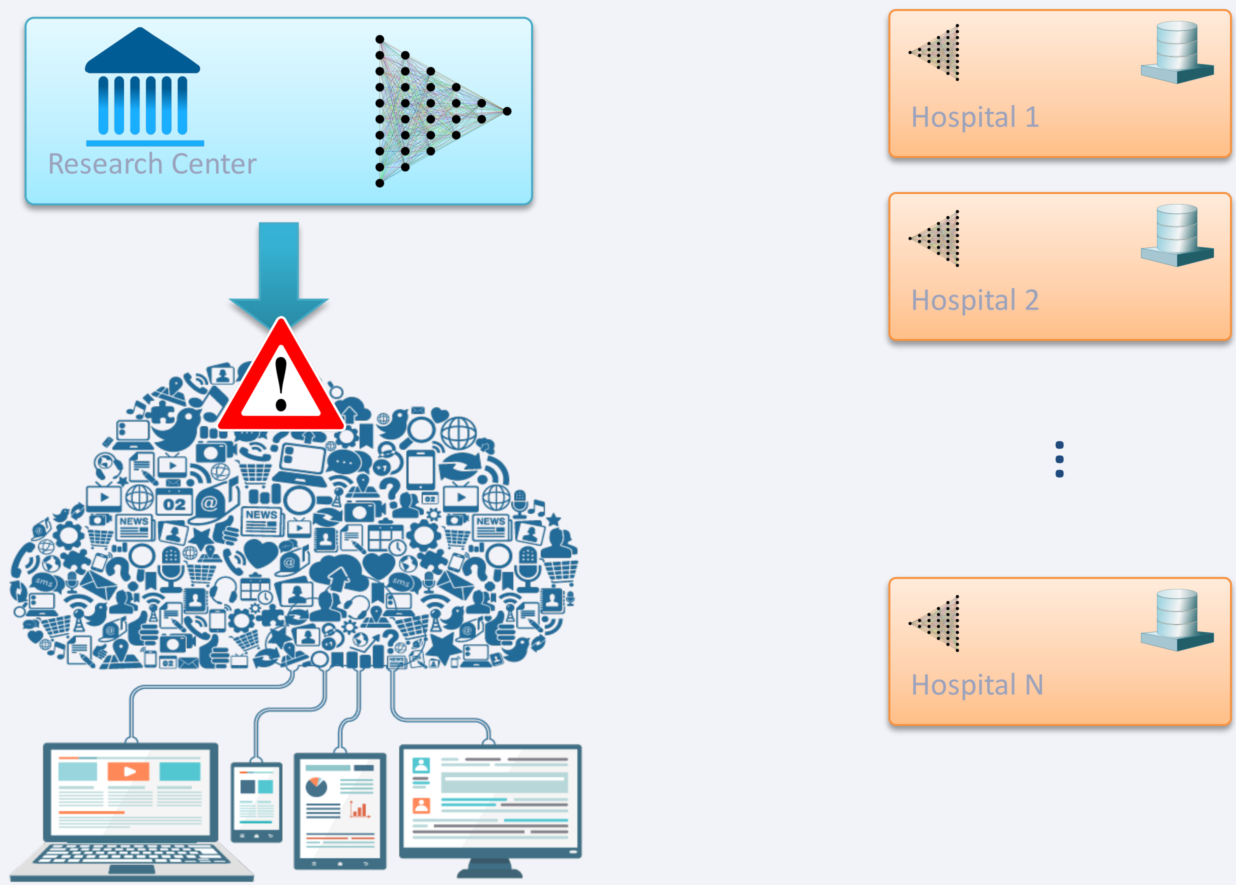
Typical Federated Learning Scenario – Training Risks

- Digital Health Twin
- Distributed Learning
- Privacy Preservation
  - Definition: Providing patient/record level protection to every member of the training set while gaining useful insights about the populations as a whole
  - What is not private?
    - ☐ Data
    - ☐ Communication
    - ☐ Infrastructure



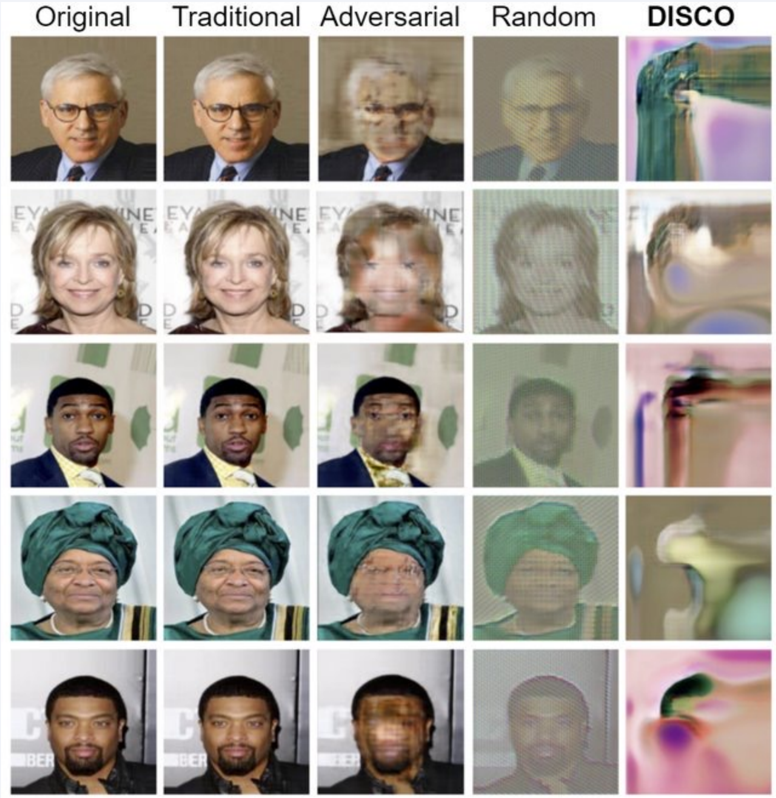
Typical Federated Learning Scenario – Publication Risks

- Digital Health Twin
- Distributed Learning
- Privacy Preservation
  - Definition: Providing patient/record level protection to every member of the training set while gaining useful insights about the populations as a whole
  - What is not private?
    - Data
    - Communication
    - Infrastructure
    - **Machine learning model output**



The Need for Privacy Preserving Machine Learning

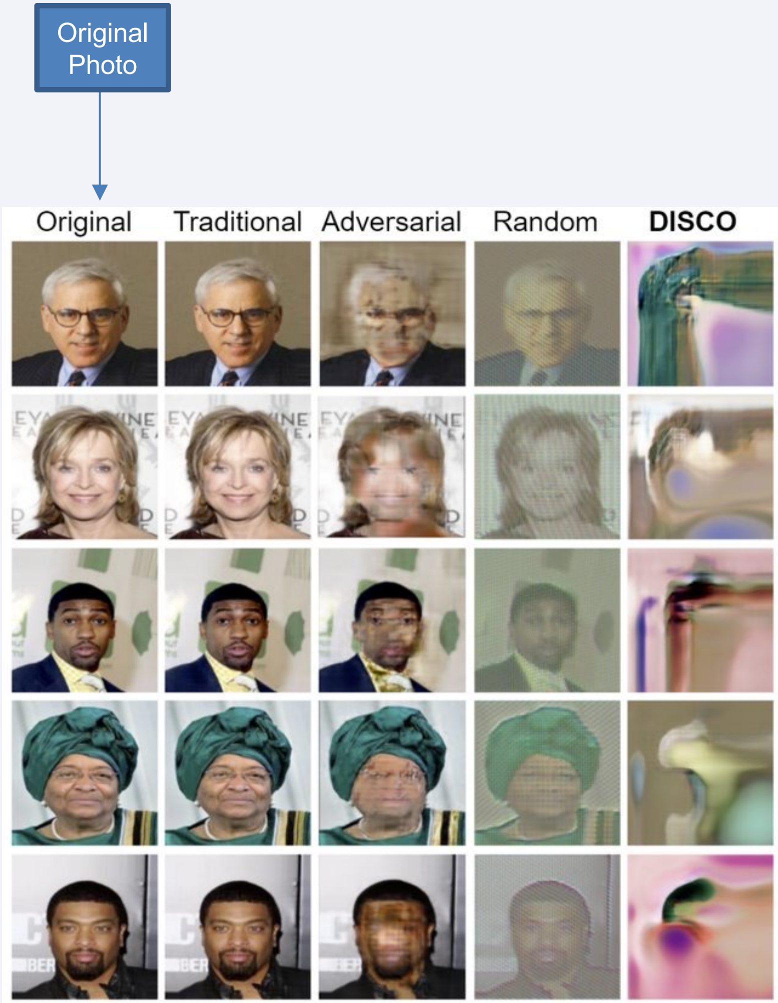
- Digital Health Twin
- Distributed Learning
- Privacy Preservation
  - Definition: Providing patient/record level protection to every member of the training set while gaining useful insights about the populations as a whole
  - What is not private?
    - Data
    - Communication
    - Infrastructure
    - **Machine learning model output**



Singh, Abhishek, et al. "DISCO: Dynamic and Invariant Sensitive Channel Obfuscation for deep neural networks." *arXiv preprint arXiv:2012.11025* (2020)

The Need for Privacy Preserving Machine Learning

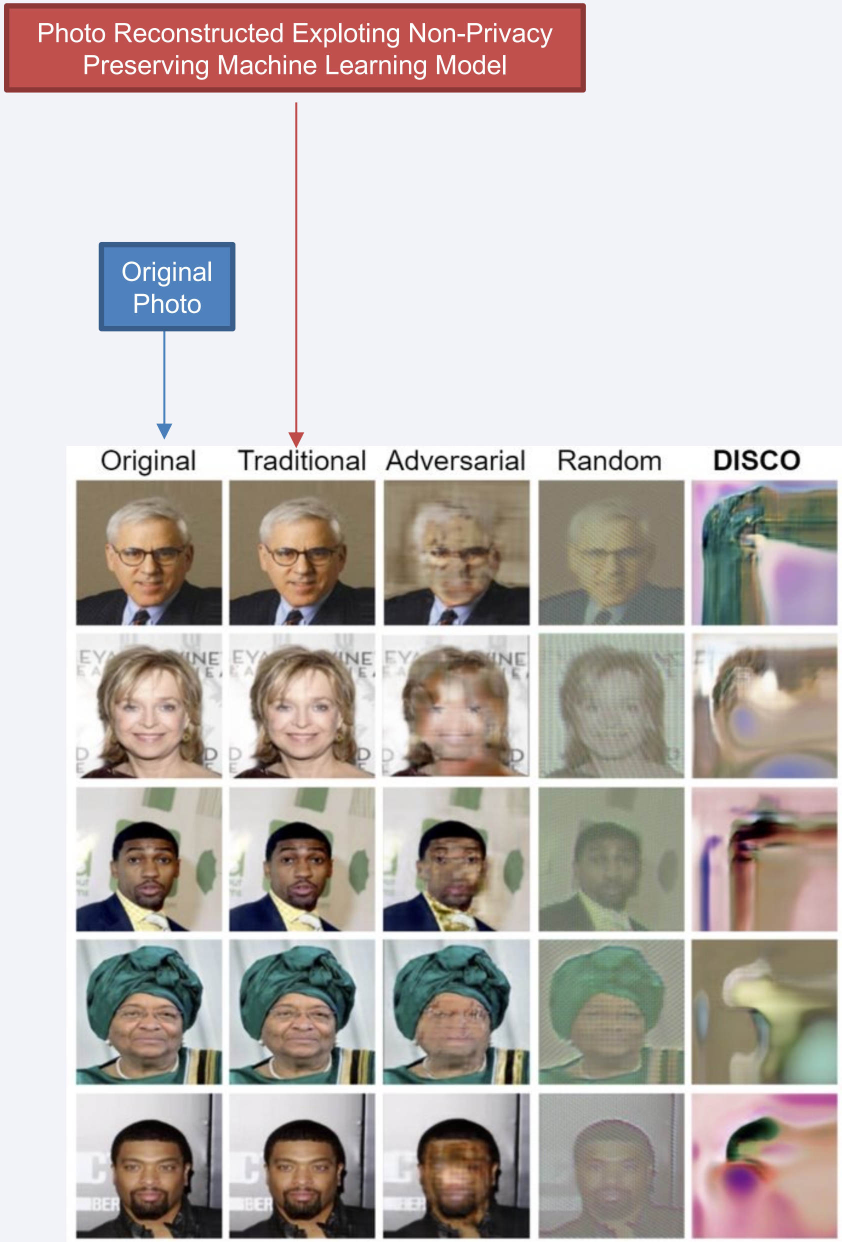
- Digital Health Twin
- Distributed Learning
- Privacy Preservation
  - Definition: Providing patient/record level protection to every member of the training set while gaining useful insights about the populations as a whole
  - What is not private?
    - Data
    - Communication
    - Infrastructure
    - **Machine learning model output**



Singh, Abhishek, et al. "DISCO: Dynamic and Invariant Sensitive Channel Obfuscation for deep neural networks." *arXiv preprint arXiv:2012.11025* (2020)

The Need for Privacy Preserving Machine Learning

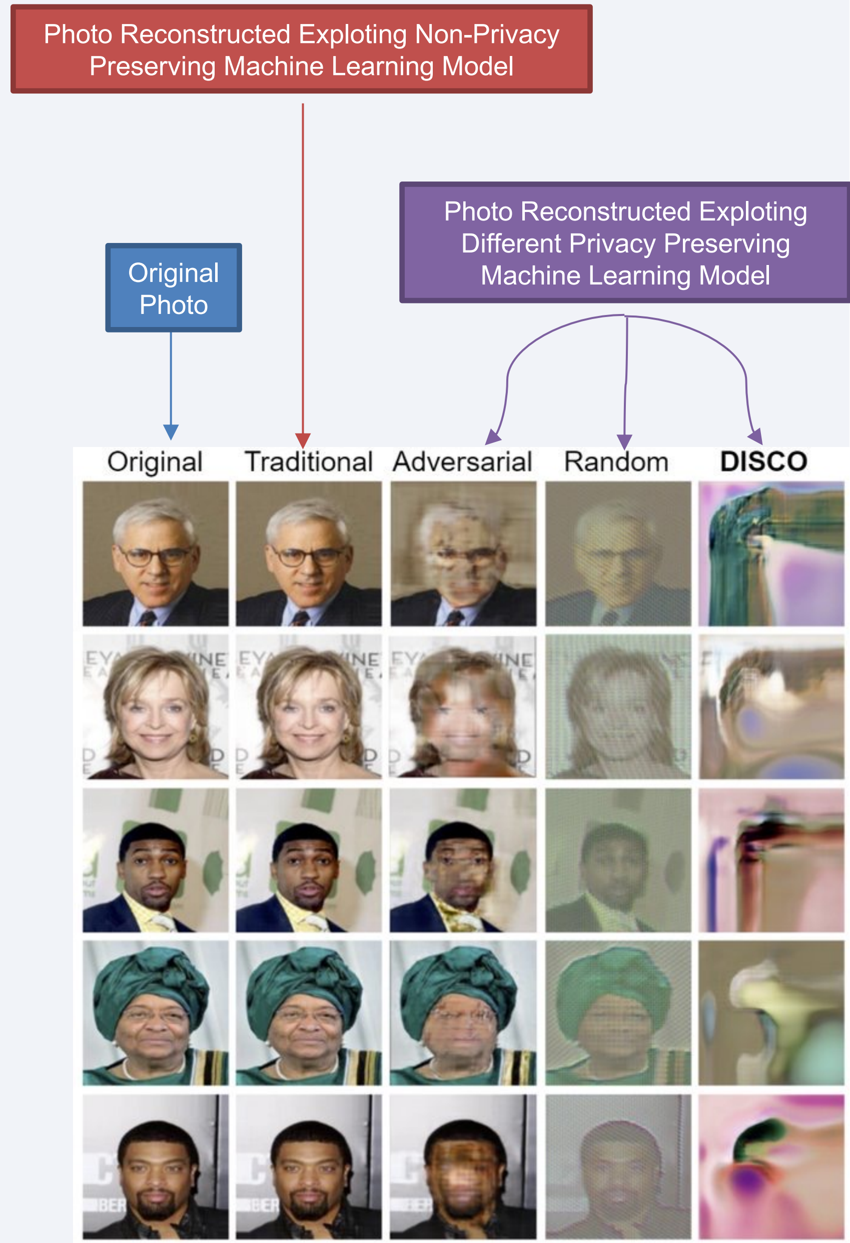
- Digital Health Twin
- Distributed Learning
- Privacy Preservation
  - Definition: Providing patient/record level protection to every member of the training set while gaining useful insights about the populations as a whole
  - What is not private?
    - Data
    - Communication
    - Infrastructure
    - **Machine learning model output**



Singh, Abhishek, et al. "DISCO: Dynamic and Invariant Sensitive Channel Obfuscation for deep neural networks." *arXiv preprint arXiv:2012.11025* (2020)

The Need for Privacy Preserving Machine Learning

- Digital Health Twin
- Distributed Learning
- Privacy Preservation
  - Definition: Providing patient/record level protection to every member of the training set while gaining useful insights about the populations as a whole
  - What is not private?
    - Data
    - Communication
    - Infrastructure
    - **Machine learning model output**



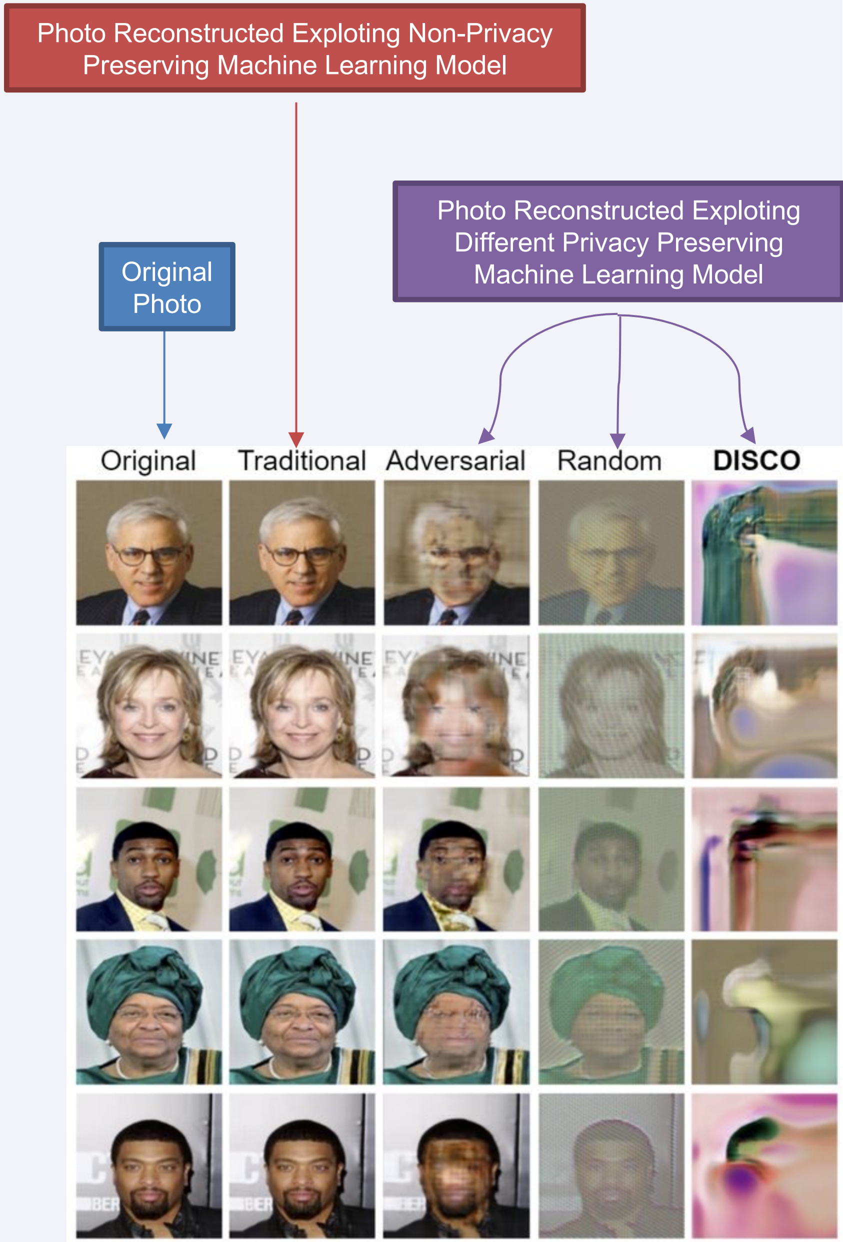
Singh, Abhishek, et al. "DISCO: Dynamic and Invariant Sensitive Channel Obfuscation for deep neural networks." *arXiv preprint arXiv:2012.11025* (2020)

The Need for Privacy Preserving Machine Learning

- Digital Health Twin
- Distributed Learning
- Privacy Preservation
  - Definition: Providing patient/record level protection to every member of the training set while gaining useful insights about the populations as a whole
  - What is not private?
    - Data
    - Communication
    - Infrastructure
    - **Machine learning model output**
  - Solution
    - **Privacy Preserving Machine Learning**
  - Mechanism
    - **Differential Privacy**

Satisfies our privacy definition

Dwork, Cynthia, and Aaron Roth. "The algorithmic foundations of differential privacy." *Foundations and Trends in Theoretical Computer Science* 9.3-4 (2014): 211-407.



Singh, Abhishek, et al. "DISCO: Dynamic and Invariant Sensitive Channel Obfuscation for deep neural networks." *arXiv preprint arXiv:2012.11025* (2020)

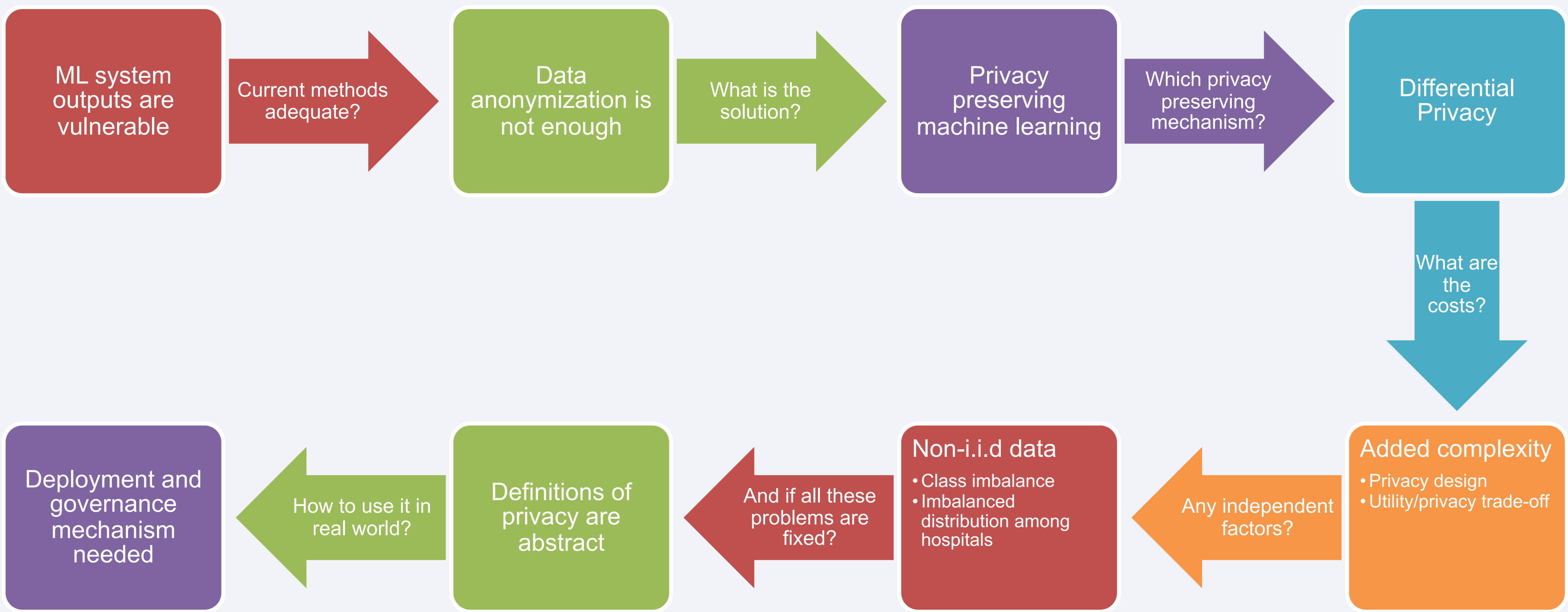
Past and Present Activities

- Supervision of 3 B.Sc. AI Theses (concluded)[1]
- Supervision of 3 M.Sc. computer science literature reviews (concluded) [1]
- Short paper on *local differentially private federated learning through compression* (PPAI@AAAI-21) [2]
- Research on *local and global differentially private federated learning through compression* (experiments underway, paper being prepared)
- Review paper on *differentially private synthetic data generation* submitted (pending editorial decision) [1]
- Review paper on *privacy attacks against machine learning* systems (receiving internal feedback) [1]
- Review paper on *privacy preserving distributed machine learning* w/ Corinne (being prepared)
- General paper on *EPI project* (being prepared)
- Supervision of 4 M.Sc. computer science and data science theses (underway)
- Research on DP distributed synthetic data generation (underway)

[1] Reports and paper available upon request; Code will be published by August 2021 depending on permission from consortium

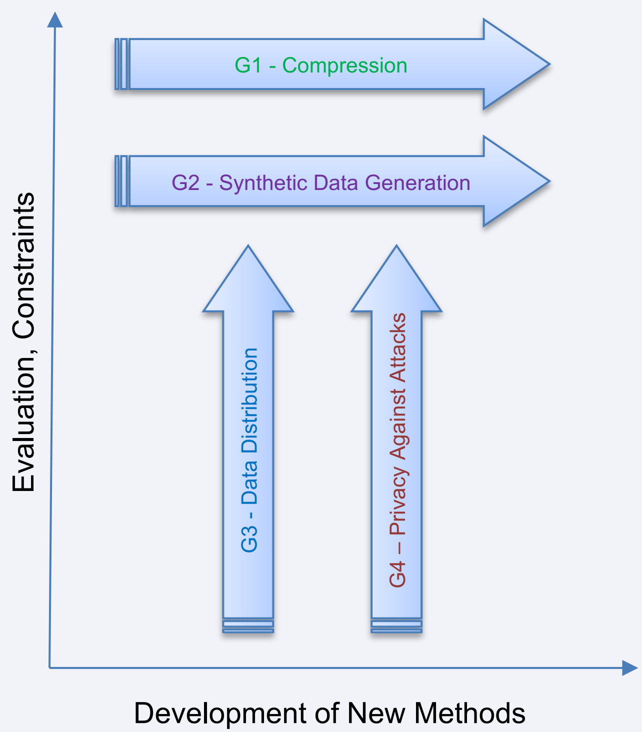
[2] https://ppai21.github.io/files/29-paper.pdf

Lessons learned



Research Goals

- G1 - Achieve Differential Privacy Through Compression
- G2 - Generate differentially-private synthetic tabular data in a distributed setting
- G3 - Analyze the effect of non-i.i.d data on the performance of differentially private machine learning models
- G4 - Measure the privacy level of DP machine learning methods from the perspective of privacy attacks



Future Works

- [1] July 2021
  - Research on local/global compressive differentially private federated learning
  - Research on comparison of JAX framework against Pytorch for privacy preserving federated learning [1]
  - Research on privacy preserving federated learning on Vantage6 framework [2]
  - Output: paper; Code + experiments
- [2] September 2021
  - Research on distributed DP synthetic data generation using VAEs
  - Research on distributed DP synthetic data generation using GANs
  - Output: paper; Code + experiments
- [3] October 2021
  - Research on effect of non-i.i.d data in privacy preserving and non-privacy preserving federated learning
  - Output: paper; Code + experiments **(repo ready)**
- [4] 2021 Q4, 2022 Q1
  - Research on extension of [2]
- [5] 2022 Q2
  - Utilization of the results of [3] in [2], [4]
- [6] 2022 Q2, Q3
  - Research on privacy analysis by measuring resiliency against privacy attacks

[1]JAX developed by Google (See https://opensource.google/projects/jax), Pytorch developed by Facebook (See https://pytorch.org/)
[2] Developed by Maastricht University Medical Centre+, Netherlands Comprehensive Cancer Organization (See https://vantage6.ai/vantage6)

Thank you!