Private Federated Learning of Knowledge Graph Representation

Tu Hoang

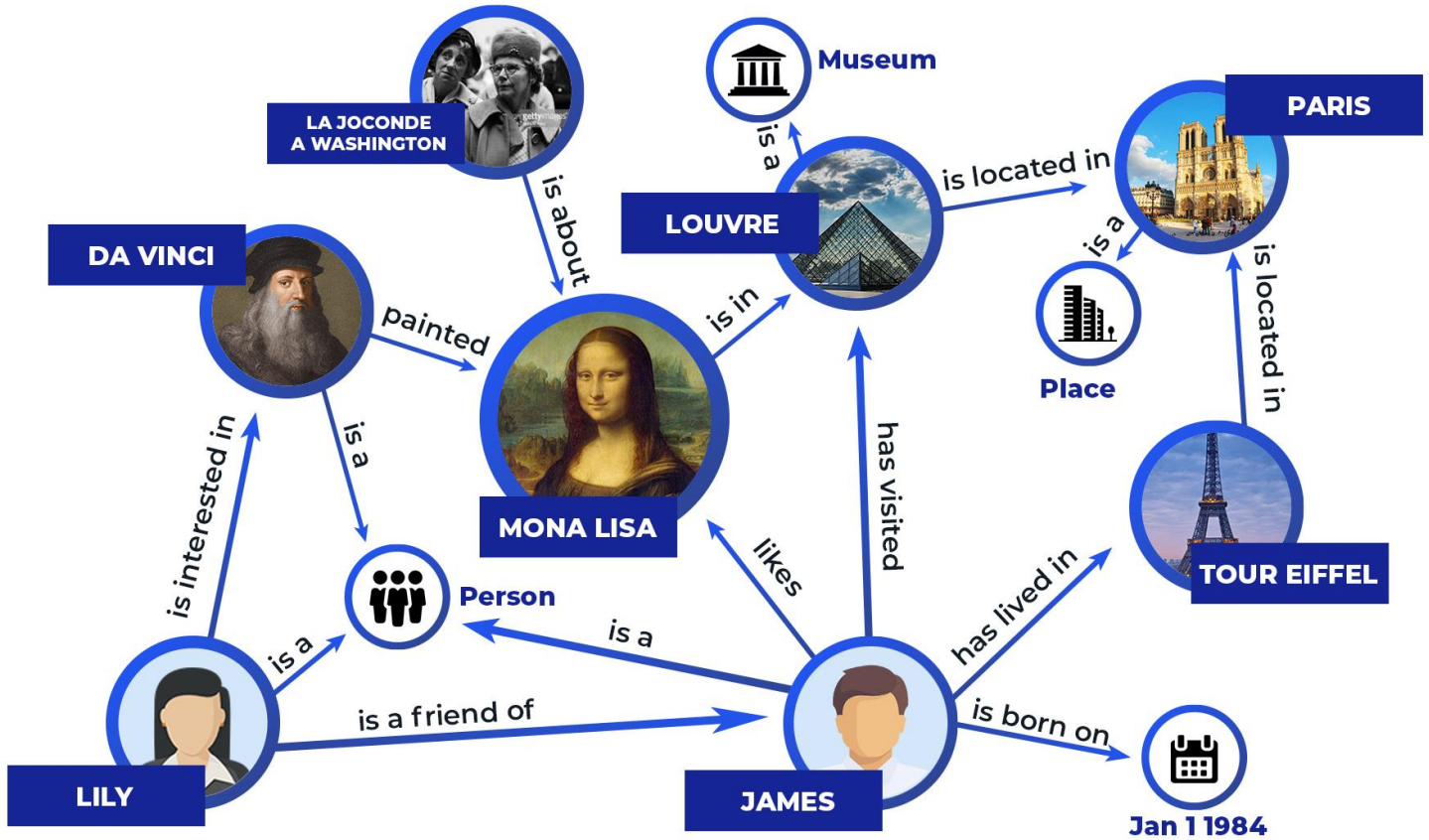
anhtu.hoang@uninsubria.it

DiSTA, University of Insubria, Italy

Content

- Introduction
- Knowledge Graph Representation
- Federated Learning of Knowledge Graphs
- Differential Privacy Techniques for Federated Learning of Knowledge Graphs
- Conclusion & Research Directions

What is a knowledge graph?



Knowledge Graph Sample[1]

Search engine Google Microsoft

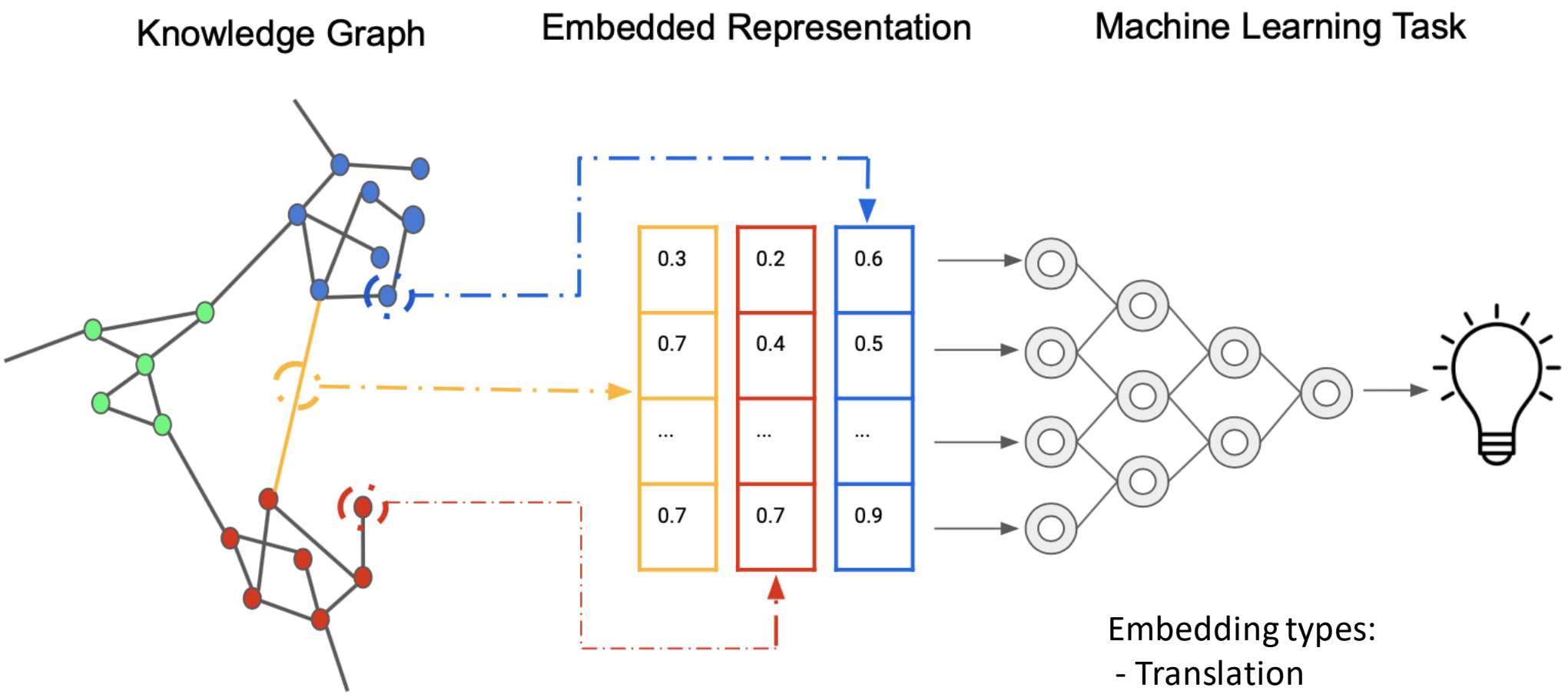
Product Knowledge Graph amazon ebay

Drug Discovery AstraZeneca

Others IBM

1. https://thetechbrook.com/inside-the-black-box-of-ai/

Knowledge Graph Representation



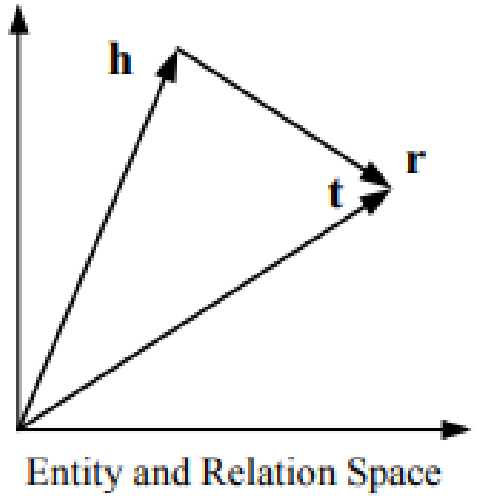
Knowledge Graph Embedding[2]

Embedding types:
- Translation
- Semantic Matching
- Matching with Neural Network

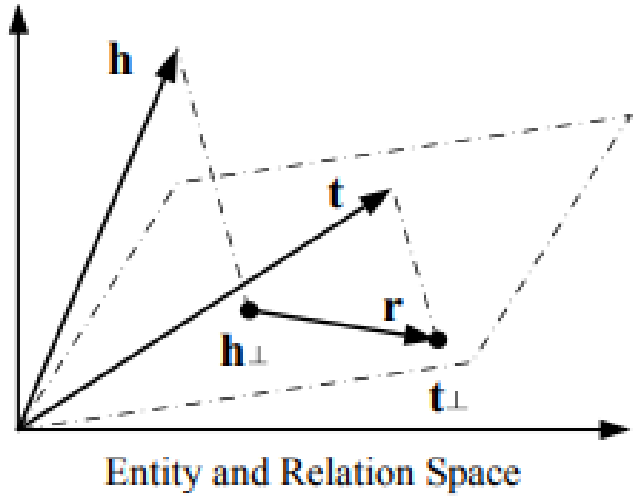
2. https://en.wikipedia.org/wiki/Knowledge_graph_embedding

Translation Models

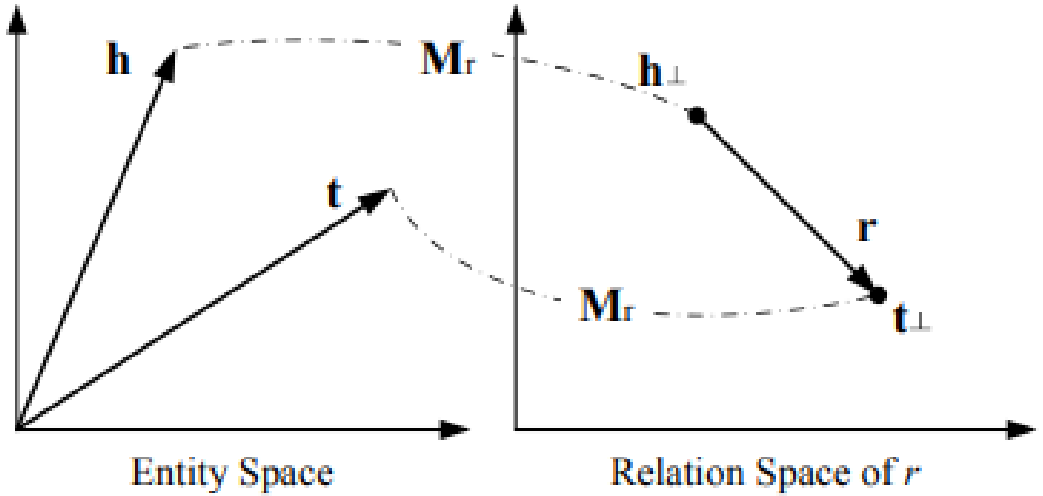
Each edge in a KG is a triplet (h, r, t), where h is the head, r is the relation, and t is the tail.
Example: (Tu, lives_in, Varese)



TransE[1]



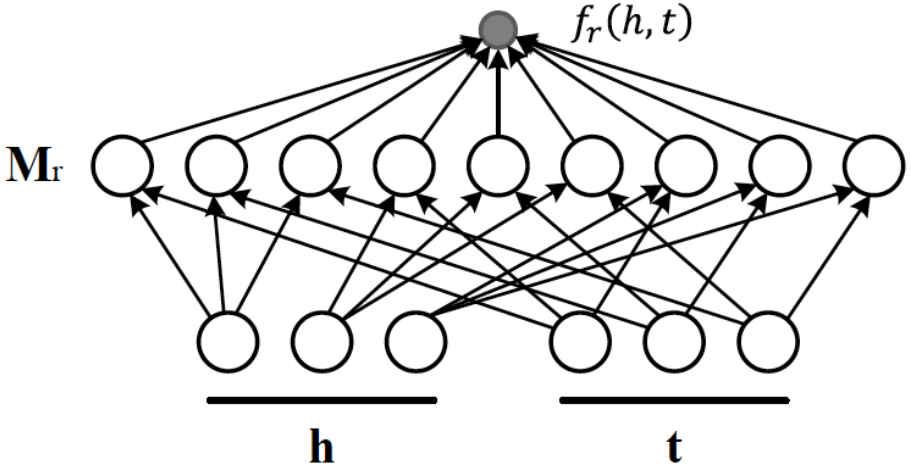
TransH[2]



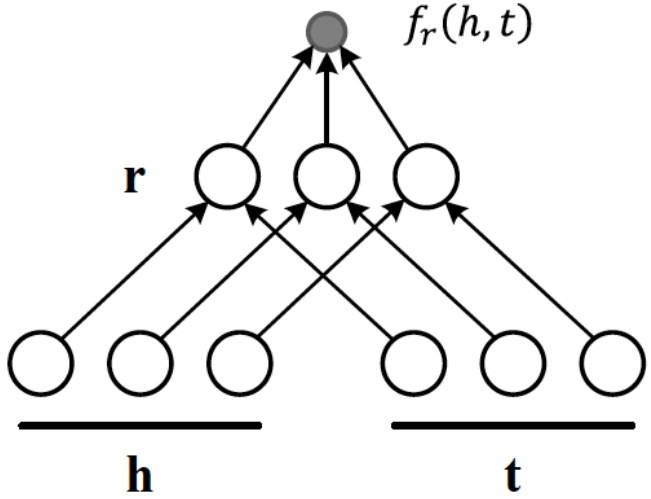
TransR[3]

[1] Antoine Bordes et al.: Translating Embeddings for Modeling Multi-relational Data. NIPS 2013: 2787-2795
[2] Zhen Wang et al.: Knowledge Graph Embedding by Translating on Hyperplanes. AAAI 2014: 1112-1119
[3] Yankai Lin et al. : Learning Entity and Relation Embeddings for Knowledge Graph Completion. AAAI 2015: 2181-2187

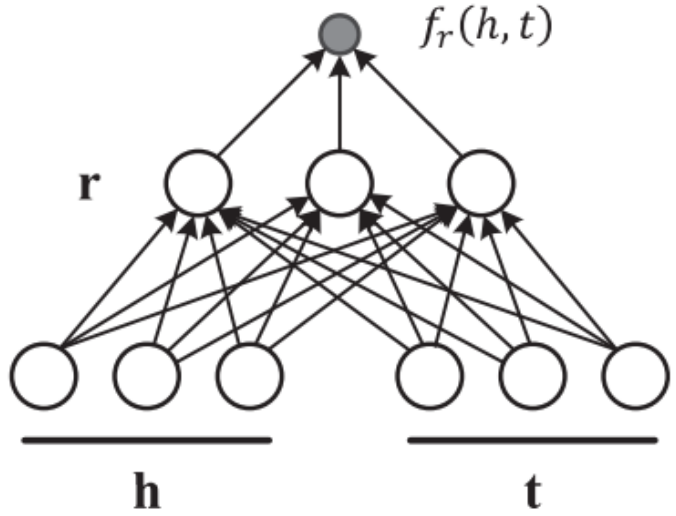
Semantic Matching Models



Rescal[4]



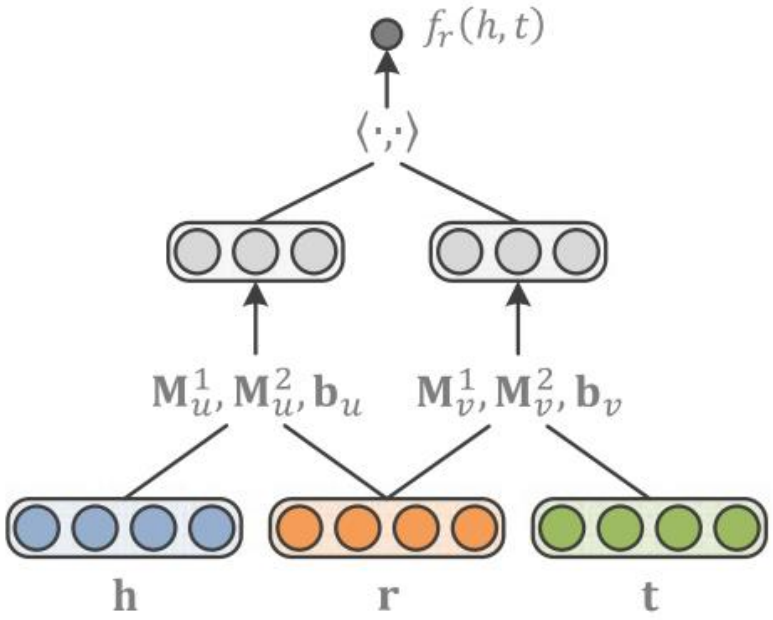
DistMult[5]



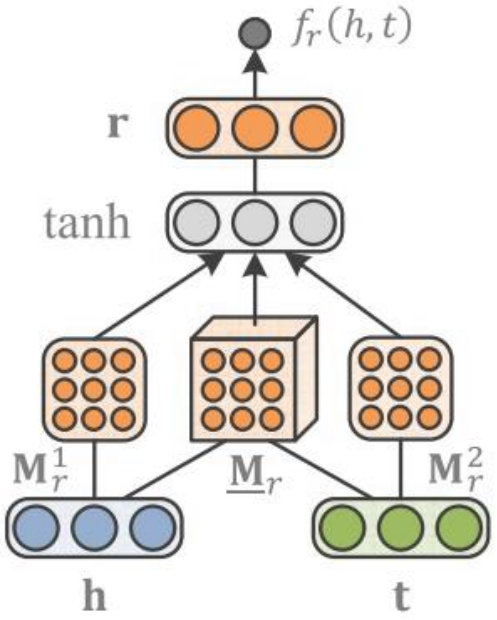
HolE[6]

[4] Maximilian Nickel et al.: A Three-Way Model for Collective Learning on Multi-Relational Data. ICML 2011: 809-816
[5] Bishan Yang et al.: Embedding Entities and Relations for Learning and Inference in Knowledge Bases. ICLR (Poster) 2015
[6] Maximilian Nickel, Lorenzo Rosasco, Tomaso A. Poggio: Holographic Embeddings of Knowledge Graphs. AAAI 2016: 1955-1961

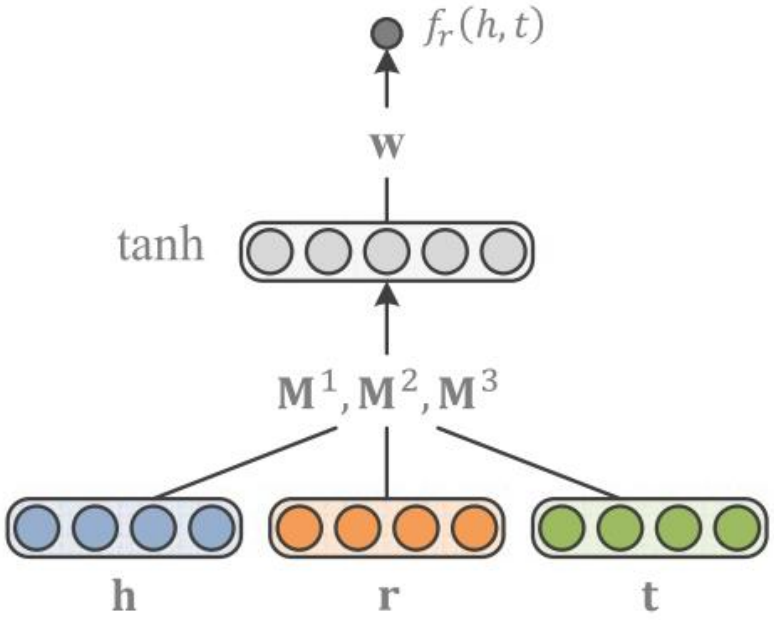
Matching with Neural Networks



SME[7]



NTN[8]



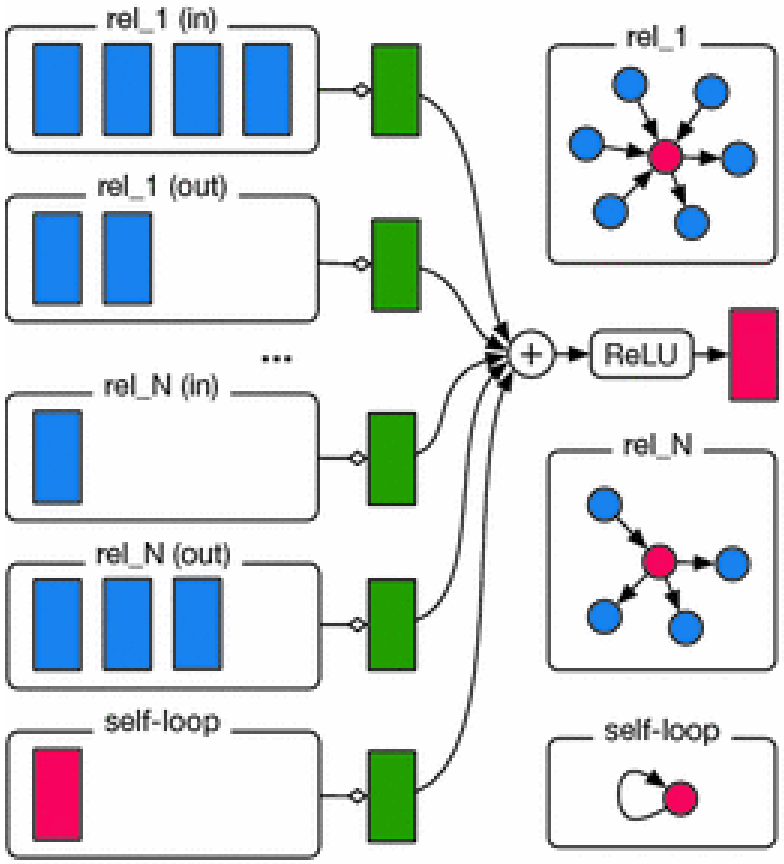
MLP[9]

[7] Xavier Glorot et al.: A Semantic Matching Energy Function for Learning with Multi-relational Data. ICLR (Workshop Poster) 2013

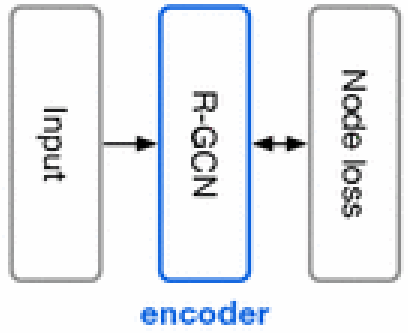
[8] Richard Socher et al.: Reasoning With Neural Tensor Networks for Knowledge Base Completion. NIPS 2013: 926-934

[9] Quan Liu et al. : Probabilistic Reasoning via Deep Learning: Neural Association Models. CoRR abs/1603.07704 (2016)

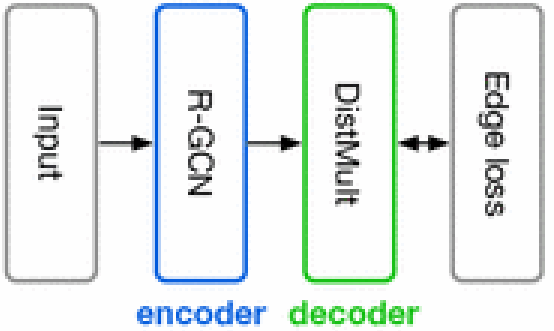
Relational Graph Convolution Network [10]



(a) Single R-GCN layer



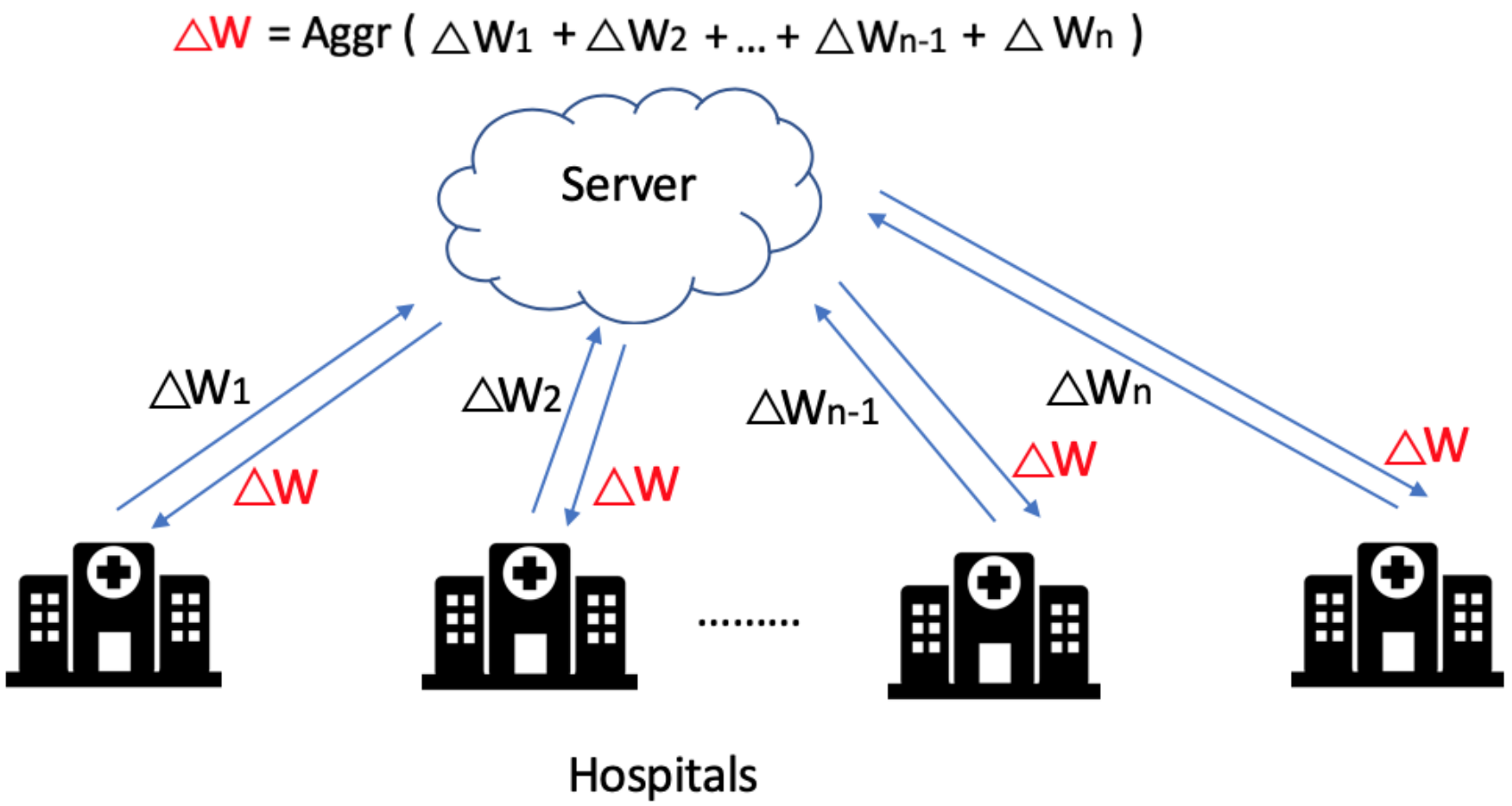
(b) Entity classification model



(c) Link prediction model

[10] Michael Sejr Schlichtkrull et al.: Modeling Relational Data with Graph Convolutional Networks. ESWC 2018: 593-607

Federated Learning

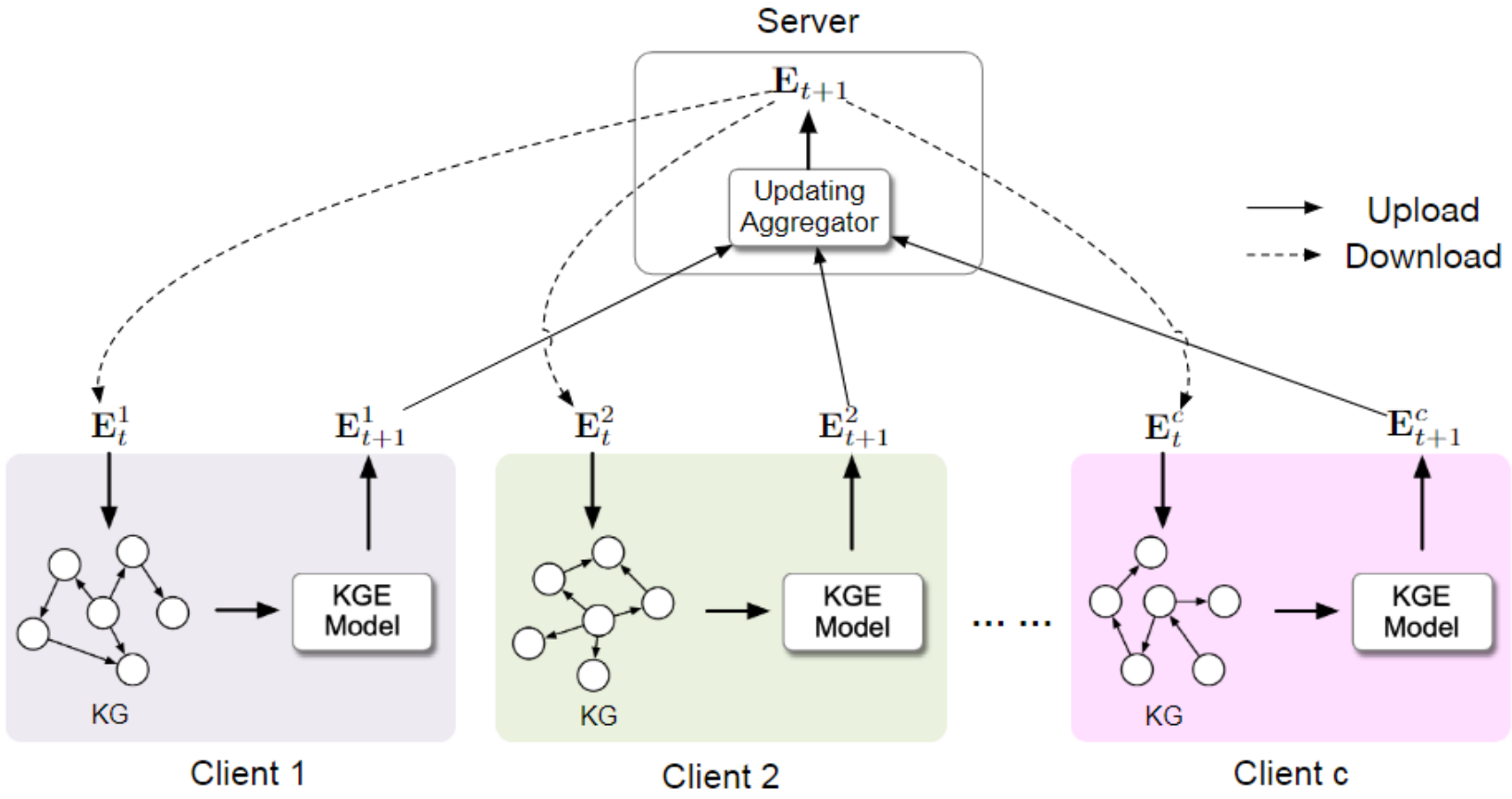


Federated Learning Workflow[3]

3. https://medium.com/@vaikkunthmugunthan/a-laymans-introduction-to-privacy-preserving-federated-learning-8ca0e6c73ad4

FedE [11]

1. Initialization: find all entities, initialize embeddings
2. Training: train/update client's models and aggregate global one
3. Model Fusion: combine local and federated learning model



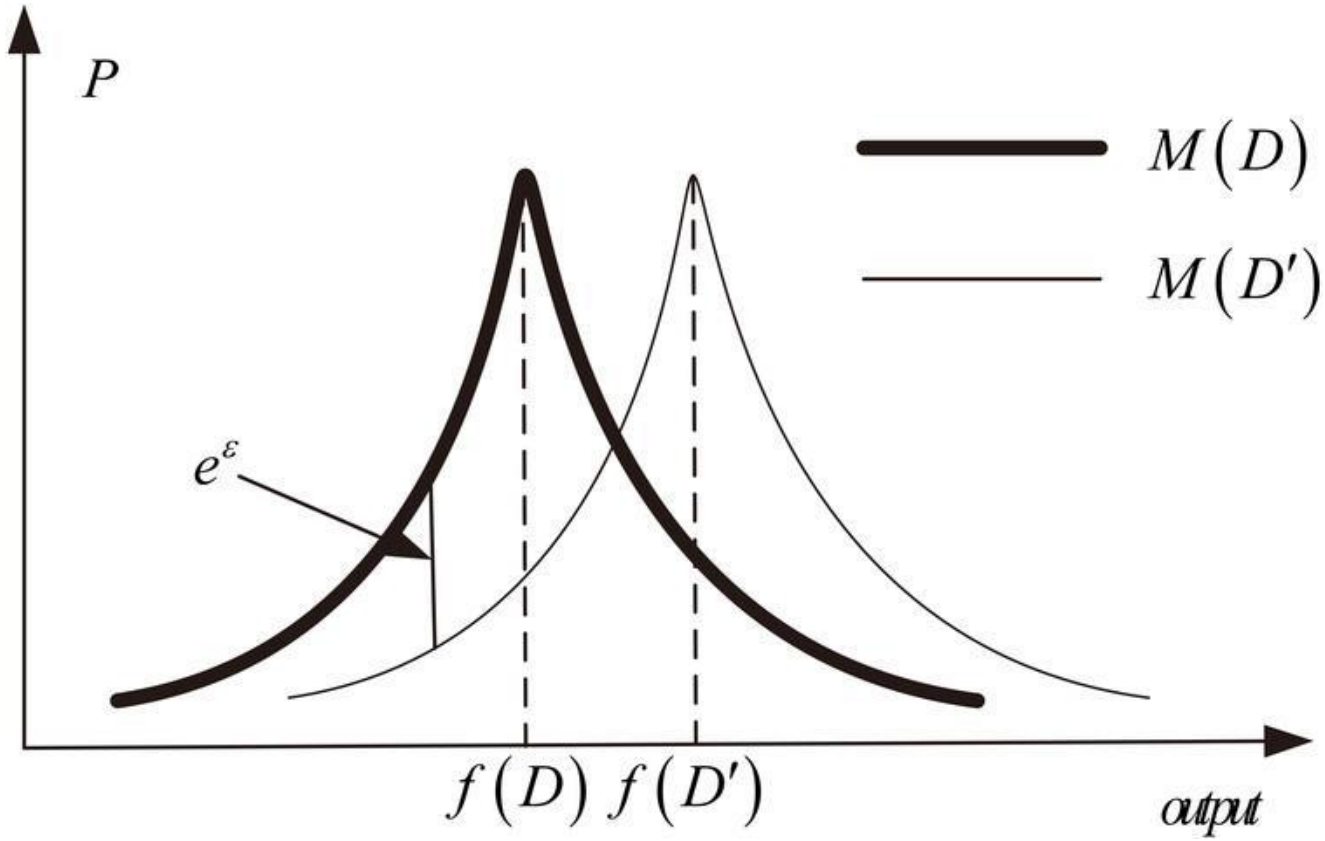
[11] Mingyang Chen et al.: FedE: Embedding Knowledge Graphs in Federated Setting. IJCKG 2021: 80-88

Differential Privacy (DP)

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

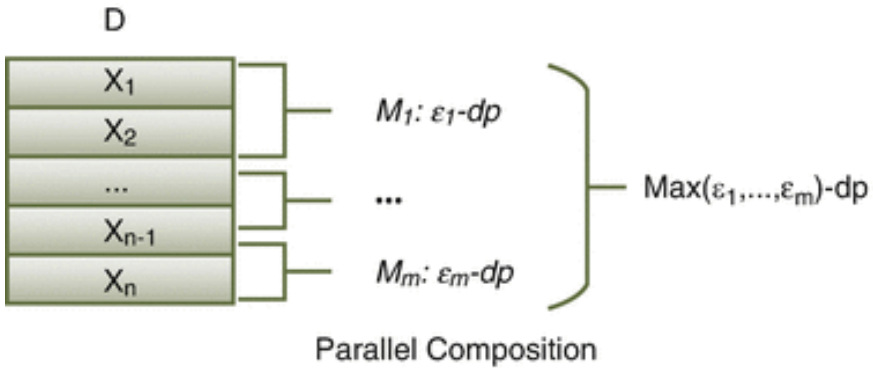
$$ln\left(\frac{Pr[\mathcal{M}(x) \in S]}{Pr[\mathcal{M}(x') \in S]}\right) \leq \epsilon$$

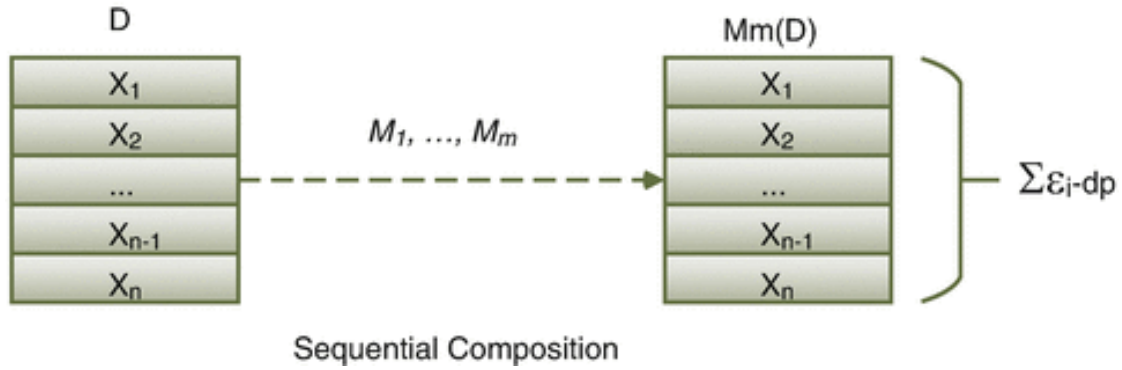
DP Simplified Definition[4]



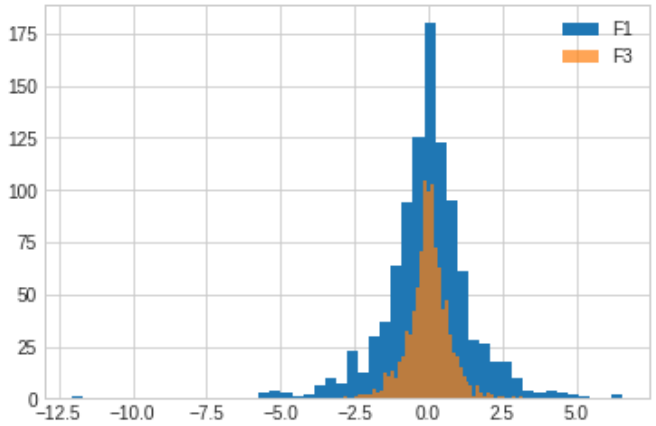
4. https://kc2853.medium.com/side-notes-on-differential-privacy-43ad18809972

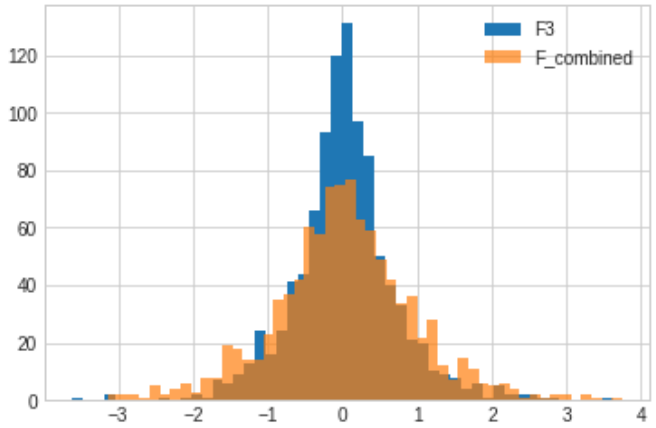
DP Compositions





DP Compositions[12]





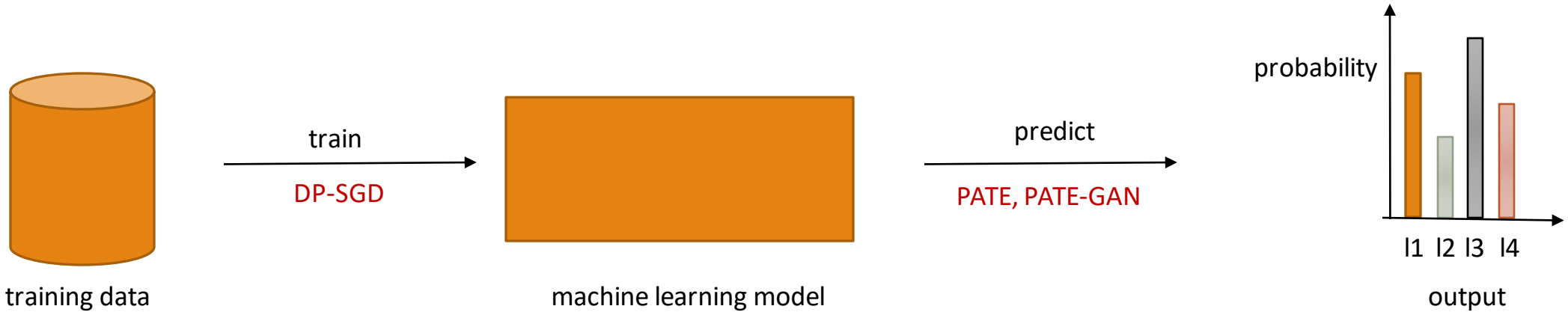
Estimated and Actual Distribution[13]

F1: e=1, F3: e=2,
F_combined: do F1 two times (estimated to have e=2)

The flexibility of DP:
1. Combine different DP-mechanisms to create a more complicated one,
2. Estimate the privacy budget of the combined mechanisms,
3. Monitor the estimated privacy budget to stop when it is too high.

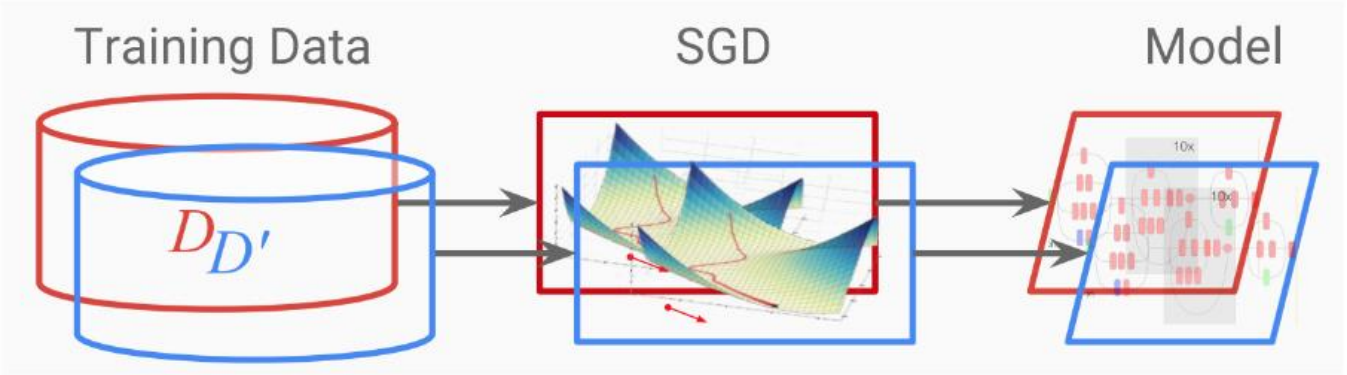
[12] Tianqing Zhu et al.: Differential Privacy and Applications. Advances in Information Security 69, Springer 2017, ISBN 978-3-319-62002-2, pp. 1-222
[13] Near, J.P. and Abuah, C.: Programming Differential Privacy. https://programming-dp.com/book.pdf.

Differential Privacy for Machine Learning



DP-SGD[14]

1. Add noise to the gradient
2. Use the noisy gradient to update the models



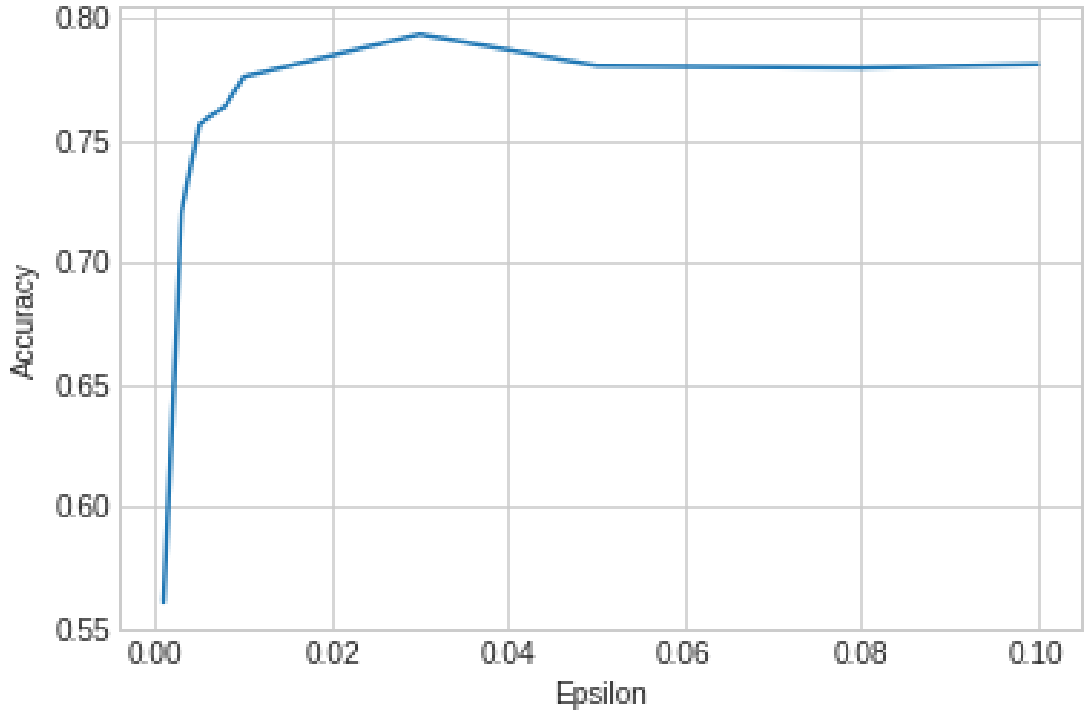
Sensitivity:
1. Clip the gradients with parameter b
2. The sensitivity is b

Noise Estimation:
1. In one epoch, the privacy budget is e.
2. In all epochs, the privacy budget is e * n_epochs.

Noises depending on:
1/ n_epochs
2/ b
3/ e

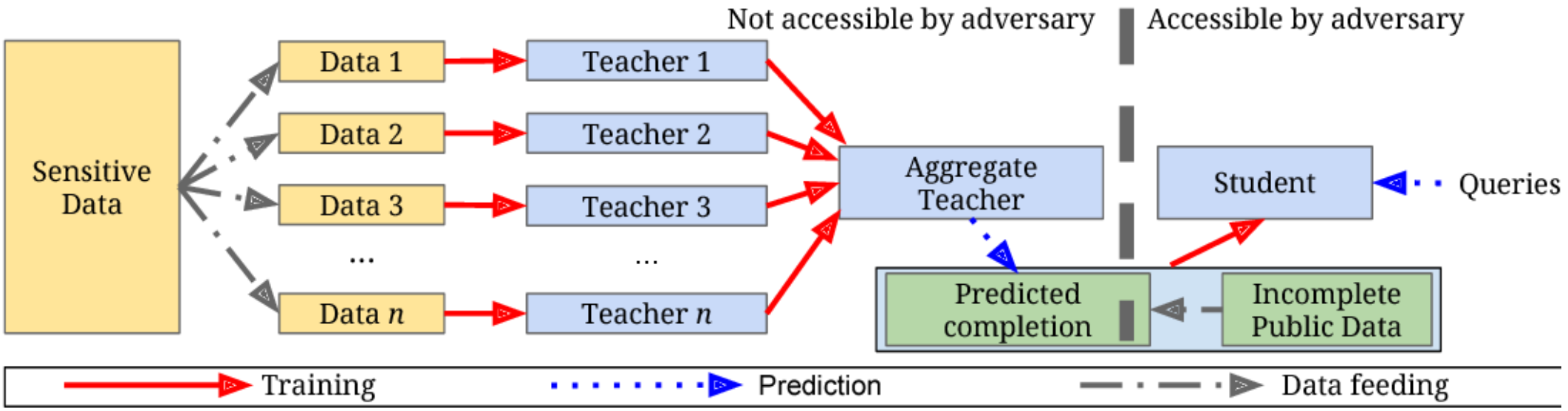
[14] Martín Abadi et al.: Deep Learning with Differential Privacy. CCS 2016: 308-318

Impact of privacy budgets on models' quality [13]



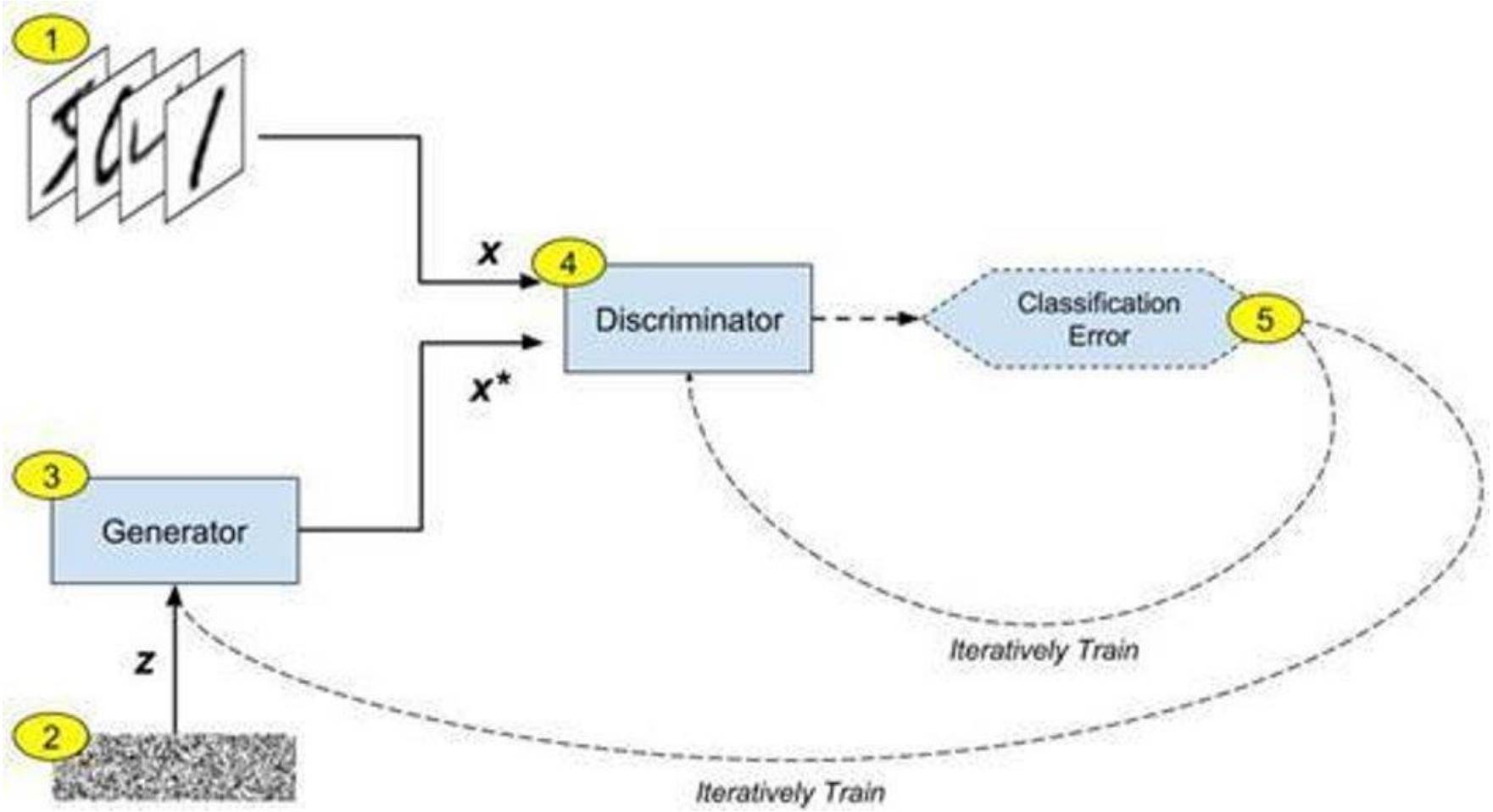
Decreasing e increases privacy protection but decreases the models' quality

PATE [15]



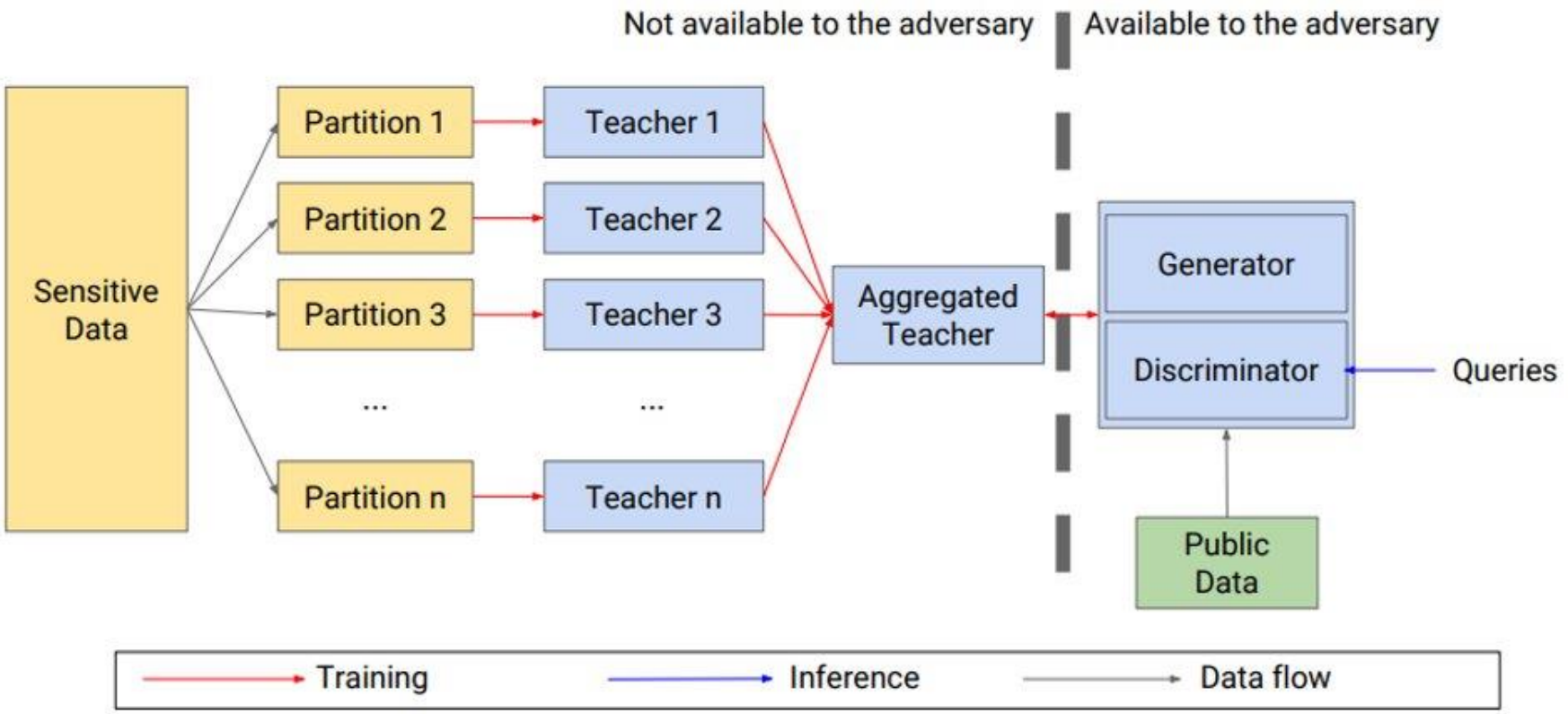
- Aggregation satisfies DP
- Add less noise than DP-SGD

Generative Adversarial Network (GAN) [16]

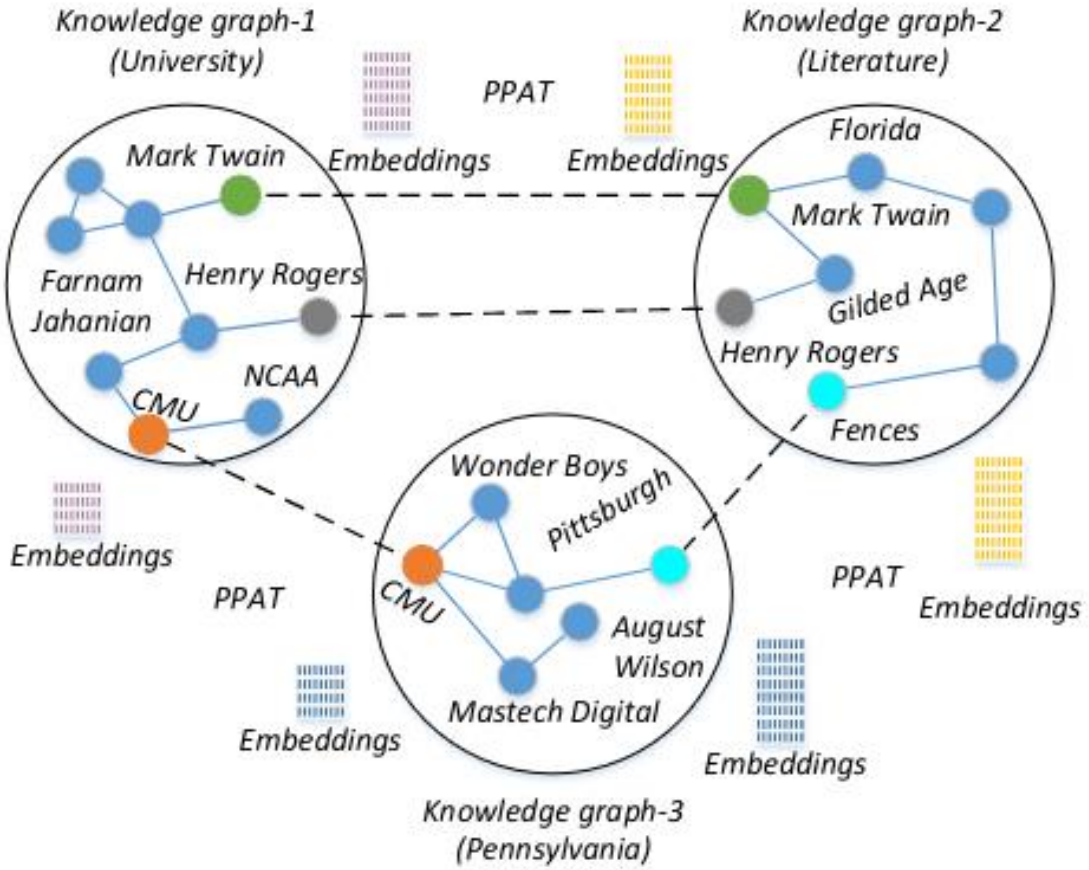


[16] Ian J. Goodfellow et al.: Generative adversarial nets. In <i>Proceedings of the 27th International Conference on Neural Information Processing Systems - Volume 2. 2014. MIT Press, Cambridge, MA, USA, 2672–2680.

PATE-GAN[17]


Not available to the adversary
Available to the adversary
Sensitive Data
Partition 1
Partition 2
Partition 3
...
Partition n
Teacher 1
Teacher 2
Teacher 3
...
Teacher n
Aggregated Teacher
Generator
Discriminator
Queries
Public Data
Training
Inference
Data flow

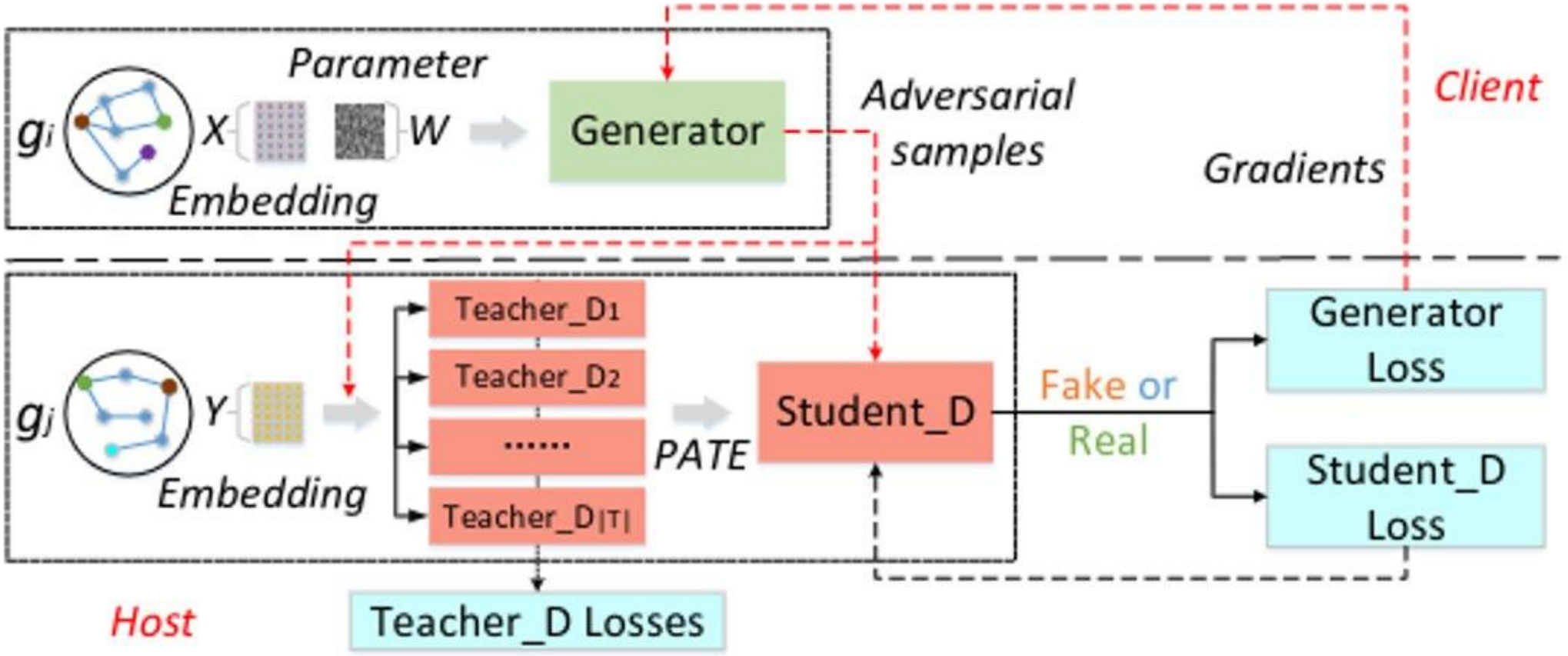
Federated learning KG Embeddings (FKGE)[18]



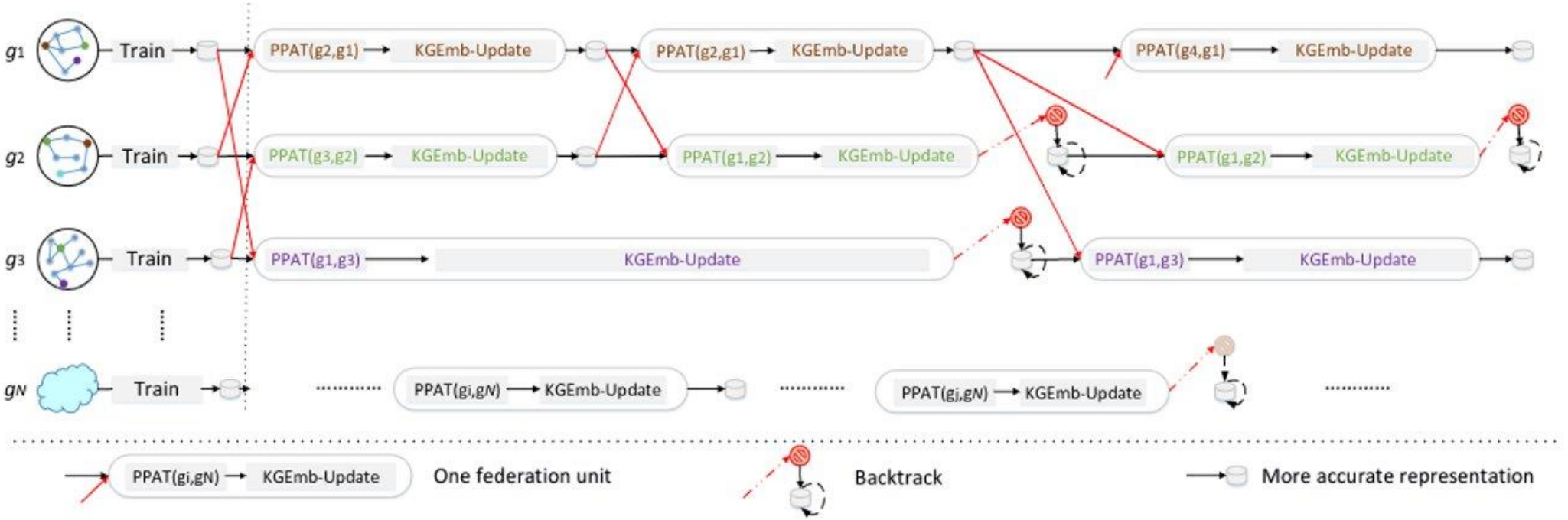
Privacy:
Protecting the existence of entities in KGs of all data providers

[18] Hao Peng, Haoran Li, Yangqiu Song, Vincent W. Zheng, Jianxin Li: Differentially Private Federated Knowledge Graphs Embedding. CIKM 2021: 1416-1425

FKGE architecture



Training Procedure



Issues:

1. Training time: each node contacts with only one node at a time.
2. Quality: only aggregate with models of one node at a time.

Conclusion

- Combining various federated learning and differential privacy approaches for training KGs' embeddings.

- Research Directions:
  - Training time: can we have a more efficient training procedure?
  - Quality: can we restrict peers from sharing high noisy models?

Thank you for your attention