Differentially Private Federated Knowledge Graphs Embedding

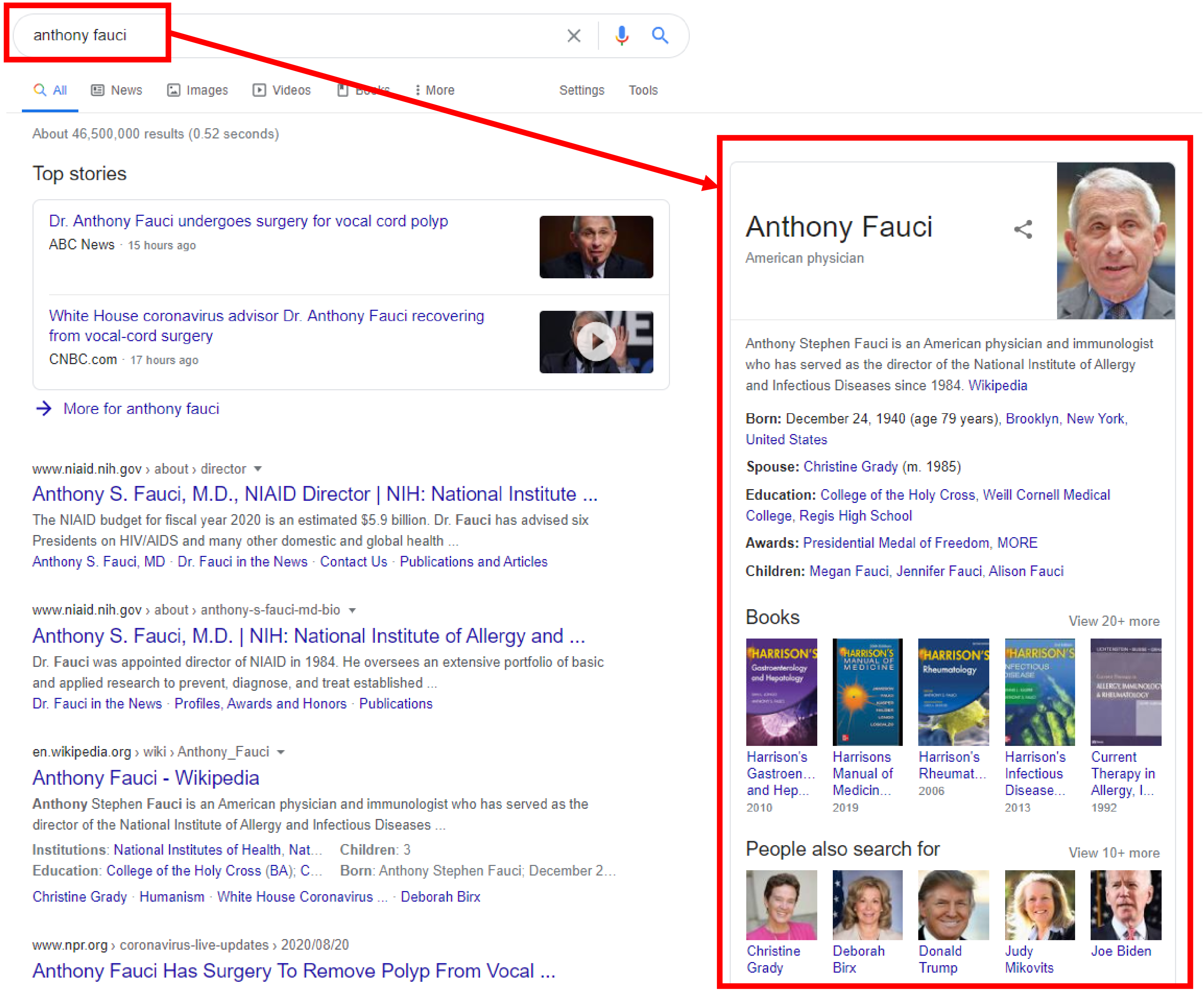
Hao Peng[1*], Haoran Li[2*], **Yangqiu Song**[2], Vincent Zheng[3], Jianxin Li[1]
[1]Beihang University
[2]Hong Kong University of Science and Technology
[3]AI Group, Webank Co., Ltd
*Equal contribution

Knowledge Graph

- A knowledge graph has many names in the history
  - Semantic networks, knowledge base, ontology, …

- In 2012, Google released its project "Google Knowledge Graph"
  - A graph-based knowledge representation connecting real-world entities to support search
    - Landmarks, celebrities, cities, sports teams, buildings, geographical features, movies, celestial objects, works of art and more
  - Get information instantly relevant to a query

https://googleblog.blogspot.com/2012/05/introducing-knowledge-graph-things-not.html


anthony fauci
All
News
Images
Videos
More
Settings
Tools
About 46,500,000 results (0.52 seconds)
Top stories
Dr. Anthony Fauci undergoes surgery for vocal cord polyp
ABC News · 15 hours ago
White House coronavirus advisor Dr. Anthony Fauci recovering from vocal-cord surgery
CNBC.com · 17 hours ago
More for anthony fauci
www.niaid.nih.gov › about › director
Anthony S. Fauci, M.D., NIAID Director | NIH: National Institute ...
The NIAID budget for fiscal year 2020 is an estimated $5.9 billion. Dr. Fauci has advised six Presidents on HIV/AIDS and many other domestic and global health ...
Anthony S. Fauci, MD · Dr. Fauci in the News · Contact Us · Publications and Articles
www.niaid.nih.gov › about › anthony-s-fauci-md-bio
Anthony S. Fauci, M.D. | NIH: National Institute of Allergy and ...
Dr. Fauci was appointed director of NIAID in 1984. He oversees an extensive portfolio of basic and applied research to prevent, diagnose, and treat established ...
Dr. Fauci in the News · Profiles, Awards and Honors · Publications
en.wikipedia.org › wiki › Anthony_Fauci
Anthony Fauci - Wikipedia
Anthony Stephen Fauci is an American physician and immunologist who has served as the director of the National Institute of Allergy and Infectious Diseases ...
Institutions: National Institutes of Health, Nat...
Children: 3
Education: College of the Holy Cross (BA); C...
Born: Anthony Stephen Fauci; December 2...
Christine Grady · Humanism · White House Coronavirus ... · Deborah Birx
www.npr.org › coronavirus-live-updates › 2020/08/20
Anthony Fauci Has Surgery To Remove Polyp From Vocal ...
Anthony Fauci
American physician
Anthony Stephen Fauci is an American physician and immunologist who has served as the director of the National Institute of Allergy and Infectious Diseases since 1984. Wikipedia
Born: December 24, 1940 (age 79 years), Brooklyn, New York, United States
Spouse: Christine Grady (m. 1985)
Education: College of the Holy Cross, Weill Cornell Medical College, Regis High School
Awards: Presidential Medal of Freedom, MORE
Children: Megan Fauci, Jennifer Fauci, Alison Fauci
Books
View 20+ more
Harrison's Gastroen... and Hep... 2010
Harrisons Manual of Medicin... 2019
Harrison's Rheumat... 2006
Harrison's Infectious Disease... 2013
Current Therapy in Allergy, I... 1992
People also search for
View 10+ more
Christine Grady
Deborah Birx
Donald Trump
Judy Mikovits
Joe Biden

COVID-19

All | News | Videos | Images | Shopping | More | Settings | Tools

About 5,790,000,000 results (0.76 seconds)

Top stories >

Hong Kong Free Press
Covid-19: New infections hit six-week low in Hong Kong, more public services to resume
4 hours ago

Mondaq News Alerts
Meeting Lockdown Challenges With Communication, Patience And Resolve...
17 hours ago

South China Morning Post
Expert warns virus transmission rate on rise as Hong Kong extends rules
4 days ago

Health information

Symptoms | Prevention | Treatments

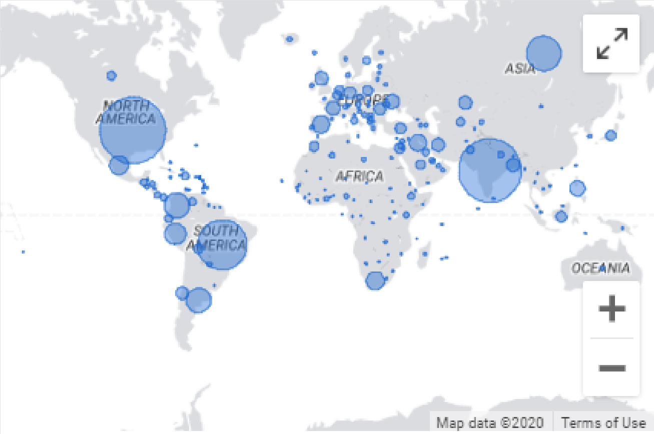
COVID-19 affects different people in different ways. Most infected people will develop mild to moderate illness and recover without hospitalization.

Most common symptoms:
- fever
- dry cough
- tiredness

Less common symptoms:
- aches and pains
- sore throat
- diarrhoea
- conjunctivitis
- headache
- loss of taste or smell
- a rash on skin, or discolouration of fingers or toes

Learn more on who.int

Map of cases (last 14 days)



Sources: Wikipedia and The New York Times. · About this data

Cases overview

Worldwide

Total cases	Recovered	Deaths
22.6M	14.5M	792K

More locations and statistics

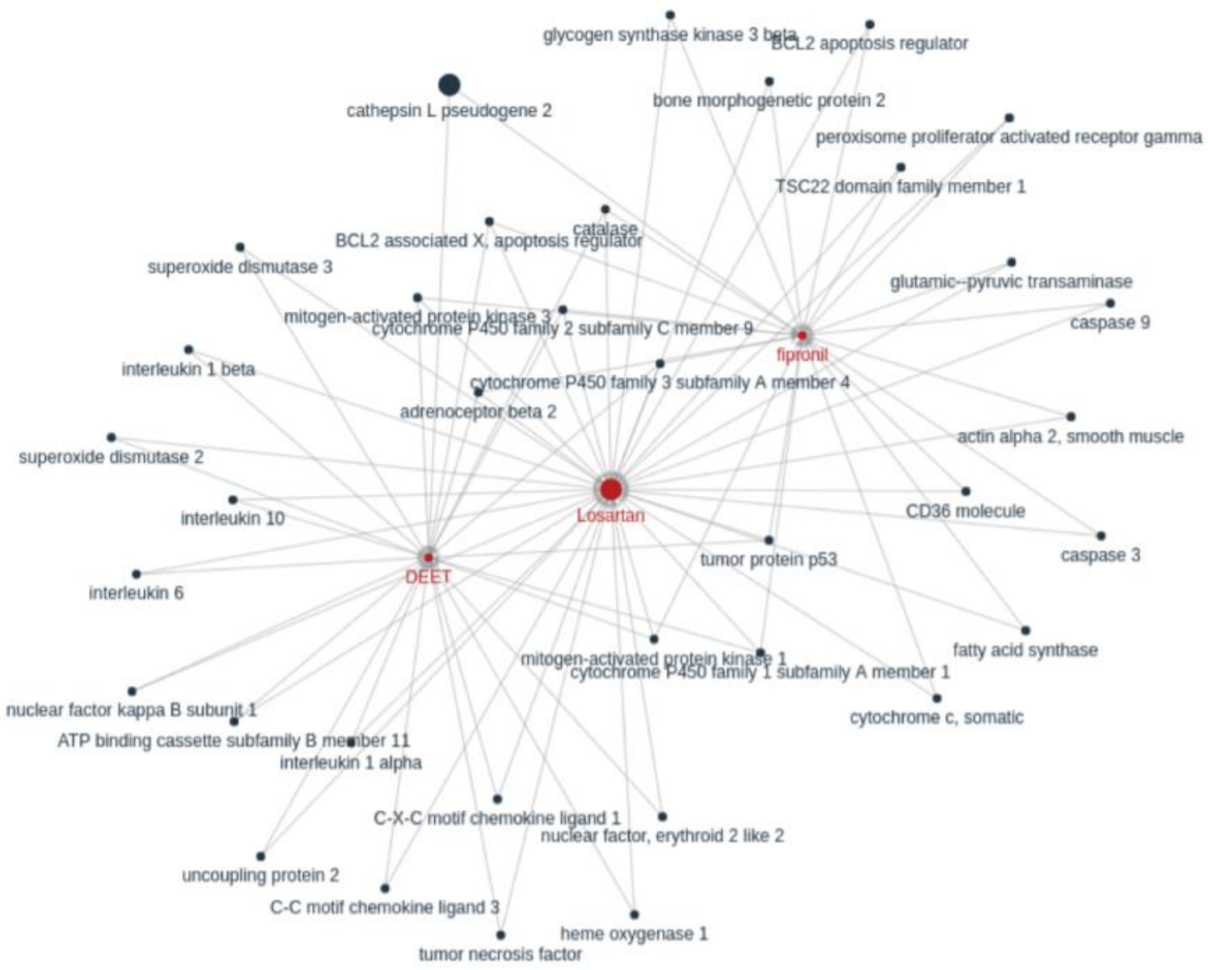
'+' shows new cases reported yesterday · Updated less than 4 hours ago ·
Source: Wikipedia · About this data

Coronavirus disease (COVID-19) is an infectious disease caused by a newly discovered coronavirus.

Most people who fall sick with COVID-19 will experience mild to

UIUC COVID-19 Literature Knowledge Graph

- http://blender.cs.illinois.edu/covid19/
- Extract entities, relations and events from text
  - 50,752 Gene nodes
  - 10,781 Disease nodes
  - 5,738 Chemical nodes
  - 535 Organism nodes
  - 133 relation types
  - 13 Event types
- Knowledge extraction from images, and do cross-media fusion and inference with entities and events

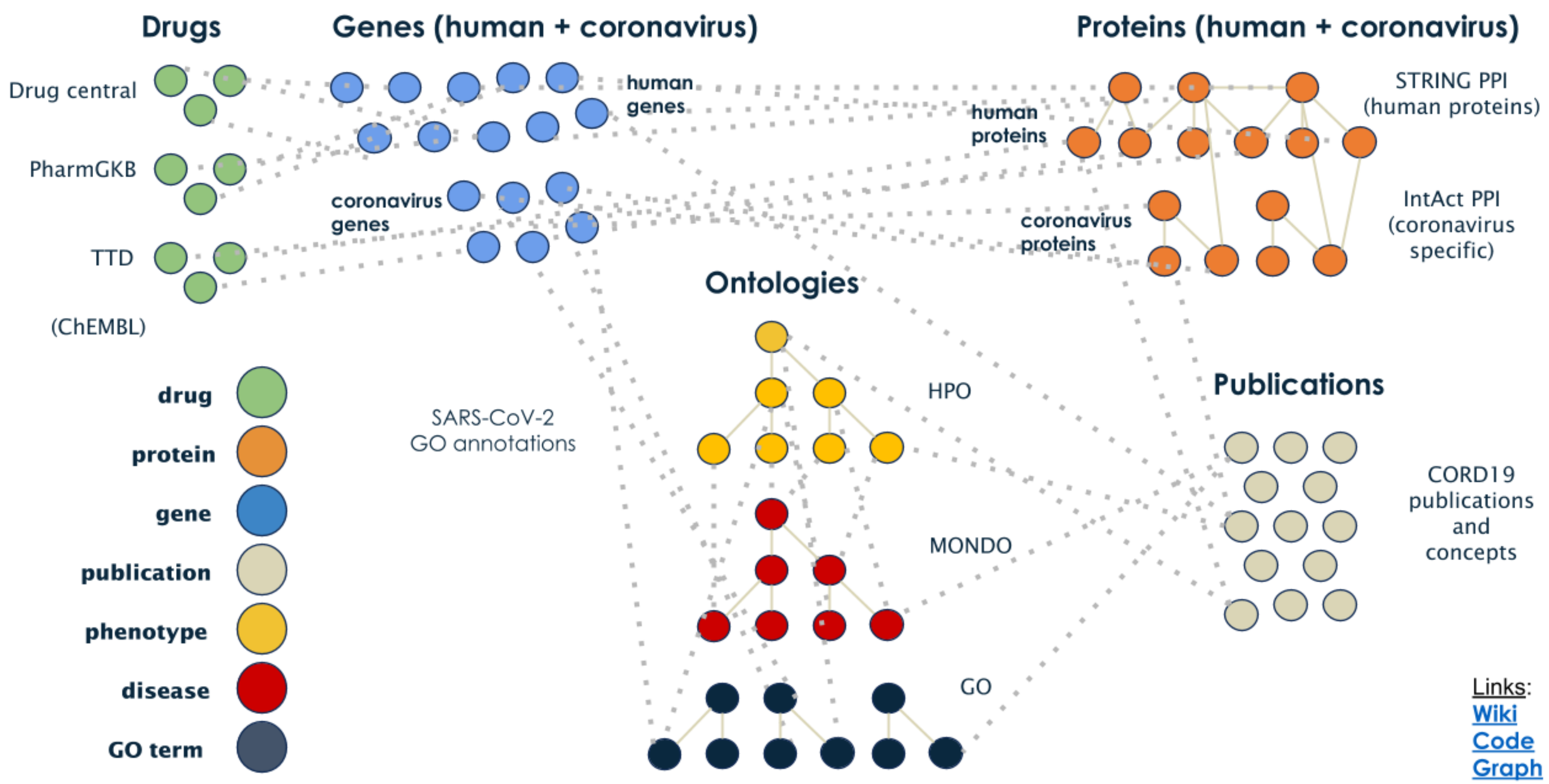


Q. Wang et al., COVID-19 Literature Knowledge Graph Construction and Drug Repurposing Report Generation. Arxiv, 2020

Berkeley Lab COVID-19 Knowledge Graph

32,000 drugs
21,000 human
272 viral proteins plus roughly the same number of genes
more than 50,000 scientific studies and clinical trials.

[KG-COVID-19 Knowledge Graph (Apr 2020)](https://github.com/Knowledge-Graph-Hub/kg-covid-19/wiki)

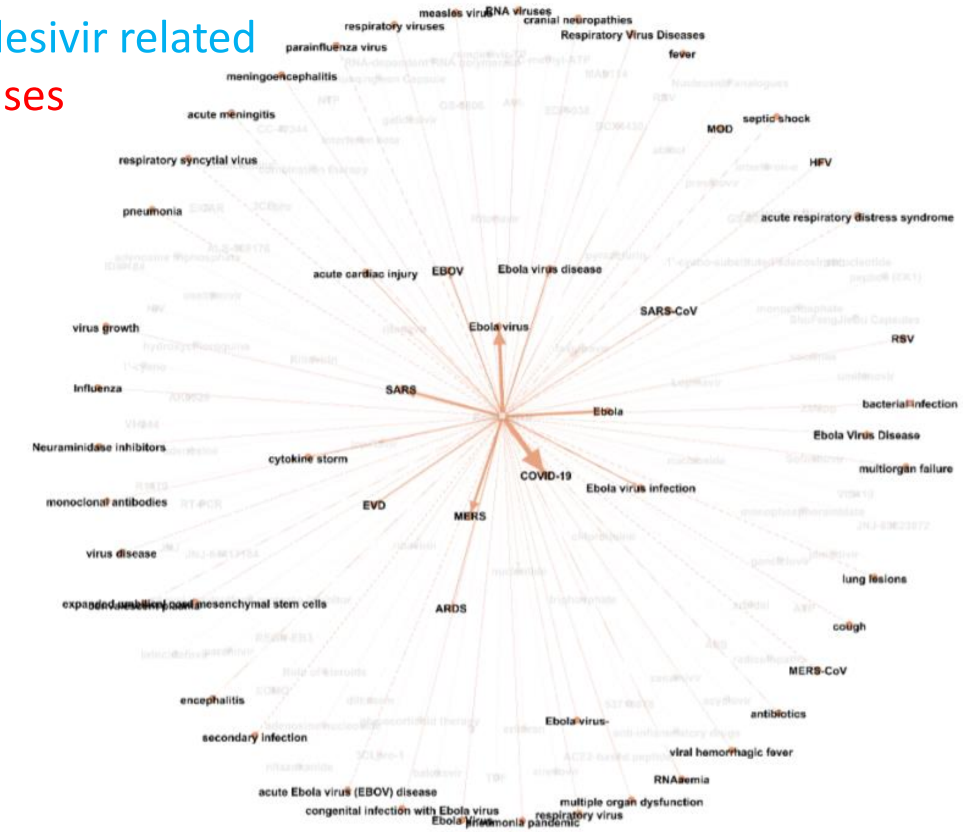


https://federallabs.org/news/berkeley-lab-creates-knowledge-graph-to-make-covid-19-drug-predictions
Image from: https://github.com/Knowledge-Graph-Hub/kg-covid-19/wiki

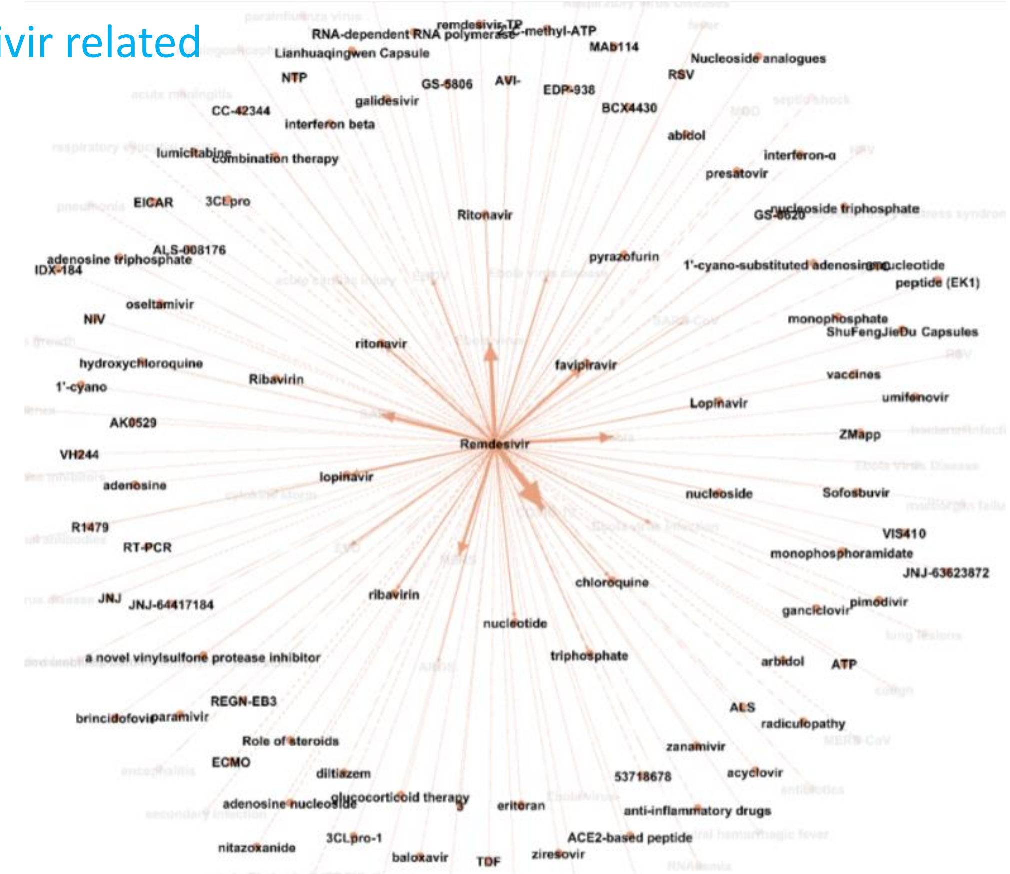
UT Austin COVID-19 Knowledge Graph

- 53,523 Drugs, 12,077 Diseases, 15,519 Species, 18,678 Genes, Gene mutations extracted from CORD-19 dataset

remdesivir related diseases

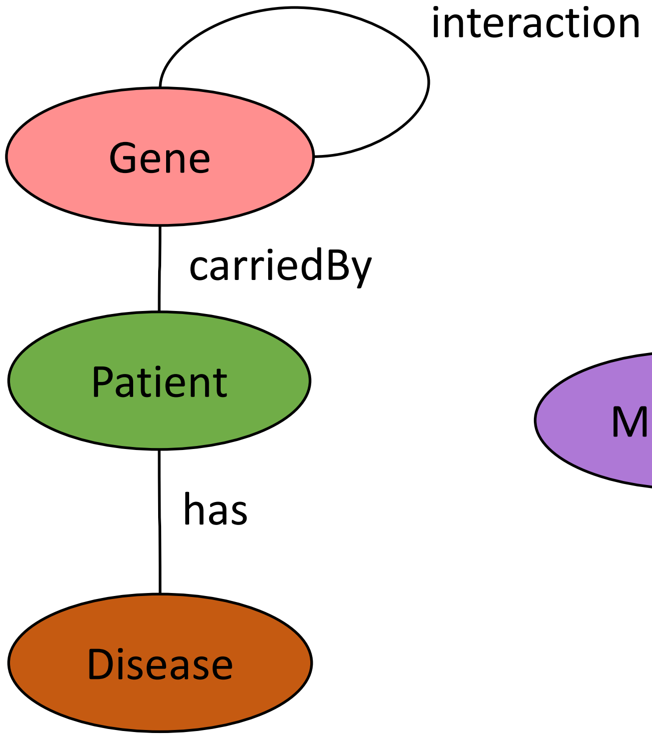


remdesivir related drugs

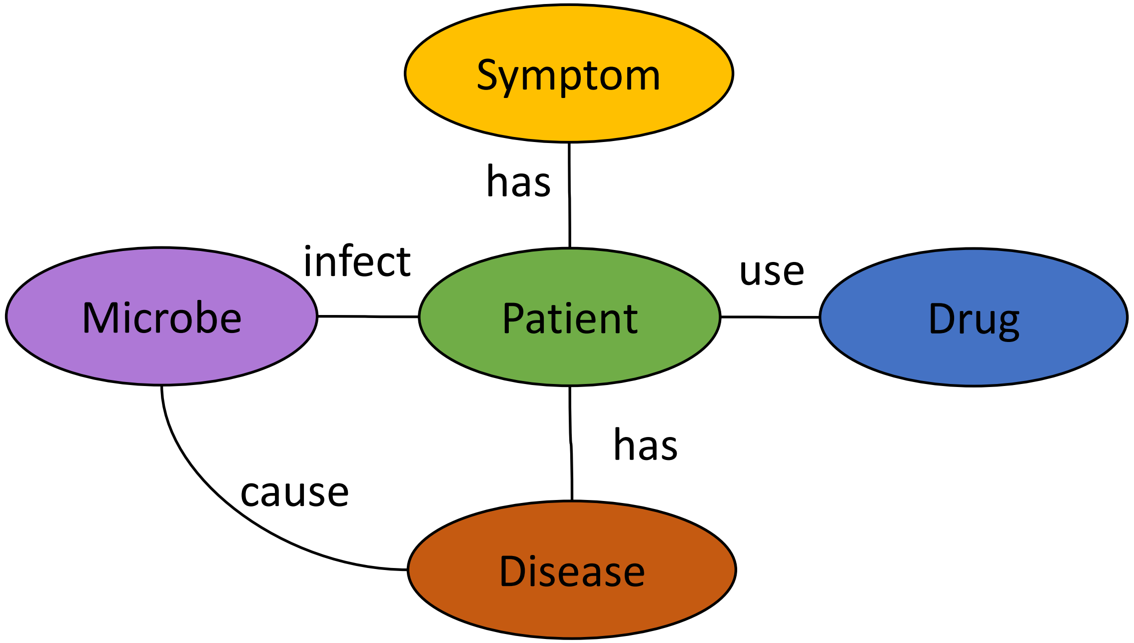


Chen, C., Ebeid, I.A., Bu, Y., & Ding, Y. (2020). Coronavirus knowledge graph: A case study. KDD Workshop on Knowledge Graph, 2020.
https://www.semanticscholar.org/cord19

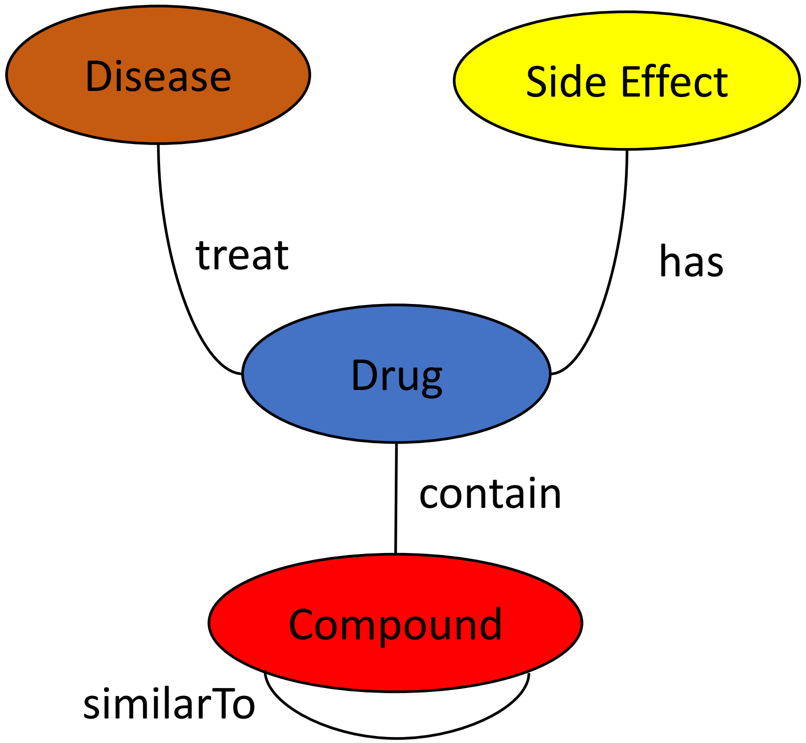
Knowledge Sharing



KG 1 from a gene engineering company



KG 2 from a hospital



KG 3 from a pharmaceutical company

Existing Approaches

- Federated database systems
  - Support unified query language over heterogeneous databases without doing actual data integration
  - Do not help improve individual KG's quality or service with private data preserved
- Learning based methods: aligned knowledge base embedding
  - Powerful for knowledge representation, reasoning, and many downstream applications
  - However, revealing vector representations to other parties can also leak private information
    - Reverse engineering individuals' properties and identities

Knowledge Sharing

- Each party has its private part of data, which cannot be disclosed to others
  - Patient information
  - Drag chemical compound
  - Personal gene expressions
- Even if privacy is not a concern, they would not expose their knowledge to other companies except they can also benefit from others
  - Existing drug repurposing failure cases
- Integrating knowledge itself is not trivial or easy
  - A lot of ambiguities
  - For example, amyotrophic lateral sclerosis, motor neurone disease, and Lou Gehrig's Disease refer to the same disease

Federated Machine Learning

- Horizontal federated learning
  - Node embeddings should be aligned
    - Very unlikely
- Vertical federated learning
  - Samples (nodes) should be partially aligned
    - Possible but sometimes unlikely
  - Aligned nodes are in different embedding space but features are not complementary
- Federated transfer learning
  - Nodes and their embeddings are aligned
    - Possible
  - Nodes and their embeddings are not aligned
    - Likely

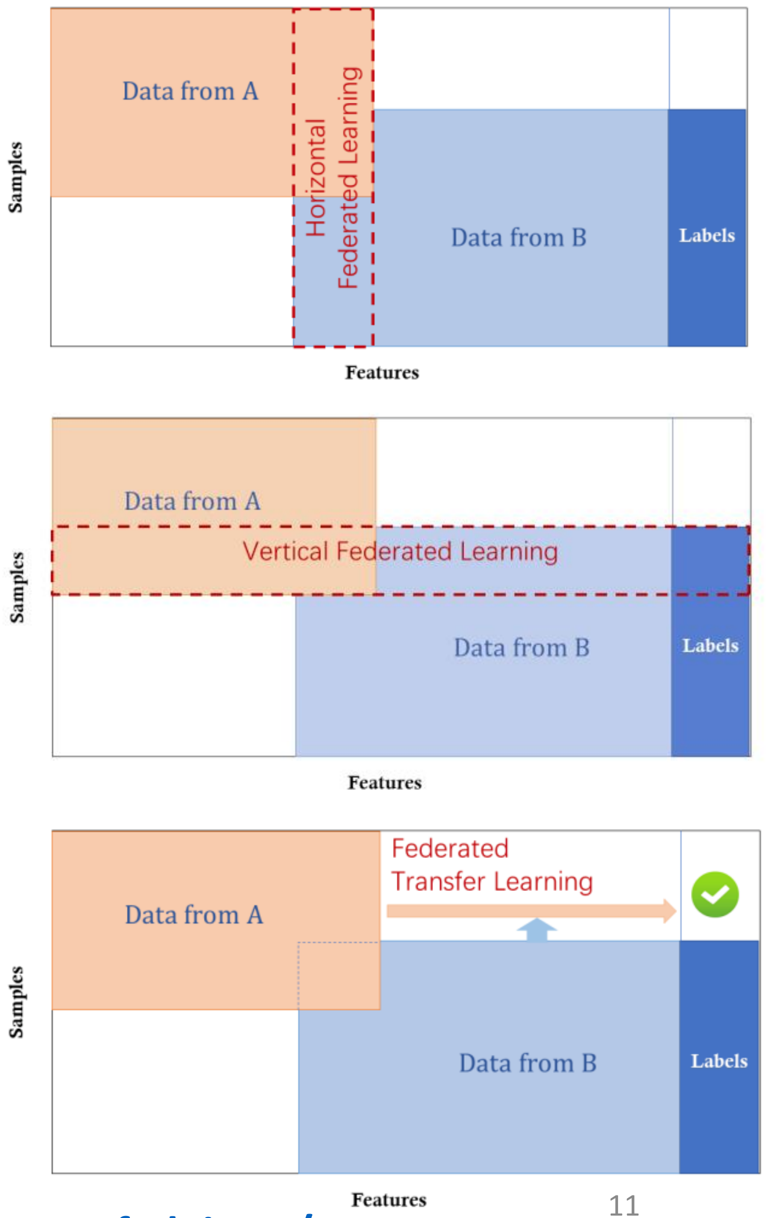
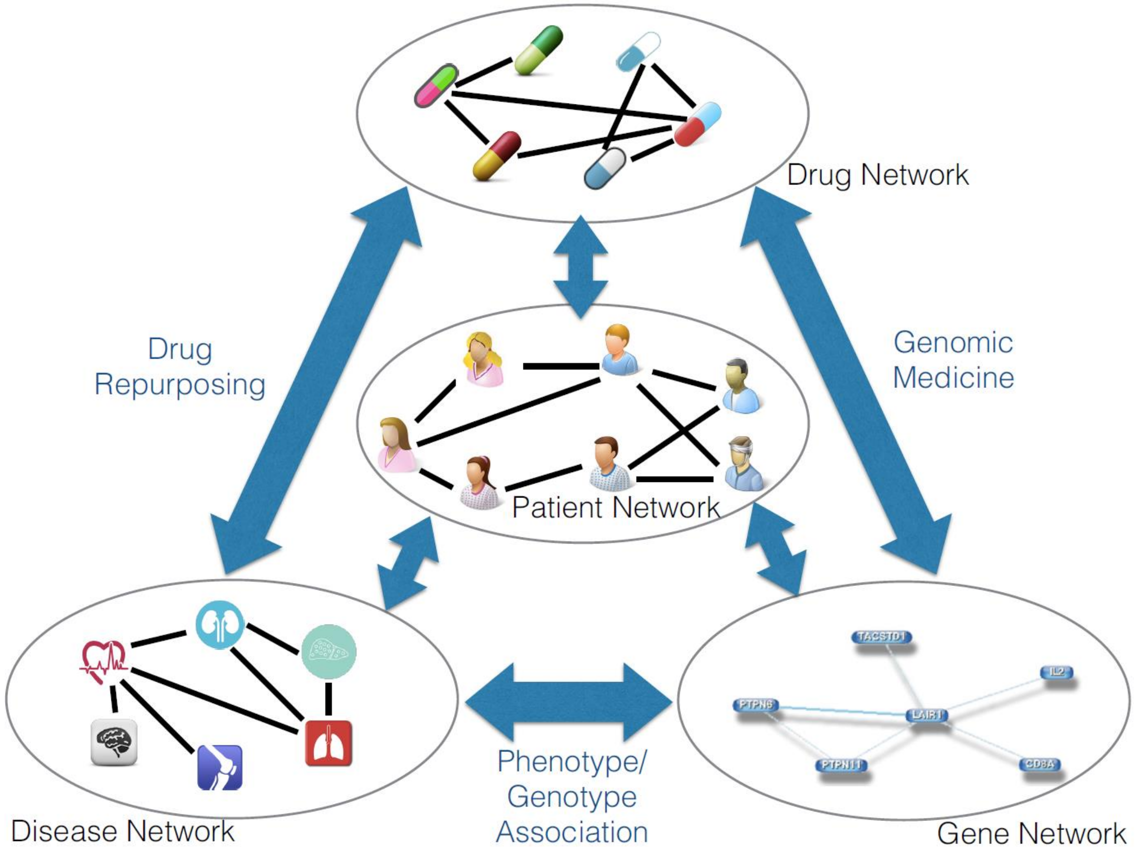


Figure credit: WeBank Tutorial, Chapter 1 - Introduction to Federated Learning. https://www.fedai.org/

Our Approach: Federated Knowledge Graphs Embedding (FKGE)

- Asynchronous and decentralized
  - Pairs up KGs from different domains
- Scalable and compatible with many base embedding models
  - A meta-algorithm for existing KG embedding methods through a handshake protocol
- FKGE is privacy-preserving and guarantees no raw data leakage
  - No raw data transmission between collaborators, and transmitted generated embeddings are differentially private

Background: Knowledge Graph Embedding



- Typical translational embedding
  - Nodes are treated as the same type
  - Relations are distinguished in triplets (head, relation, tail)

$$\mathrm{Score}\left(\boldsymbol{h}_i + \boldsymbol{r}_k - \boldsymbol{t}_j\right) = \left\|\boldsymbol{h}_i + \boldsymbol{r}_k - \boldsymbol{t}_j\right\|$$

Head entity ID i
e.g., Crohn's disease

Tail entity ID j
e.g., Diarrhea

Relation type k
e.g., SymptomOfDisease

$$\mathrm{Score}\left(\boldsymbol{h}_i + \boldsymbol{r}_k - \boldsymbol{t}_j\right) > \mathrm{Score}\left(\boldsymbol{h}_i + \boldsymbol{r}_k - \boldsymbol{t}_l\right) + \delta$$

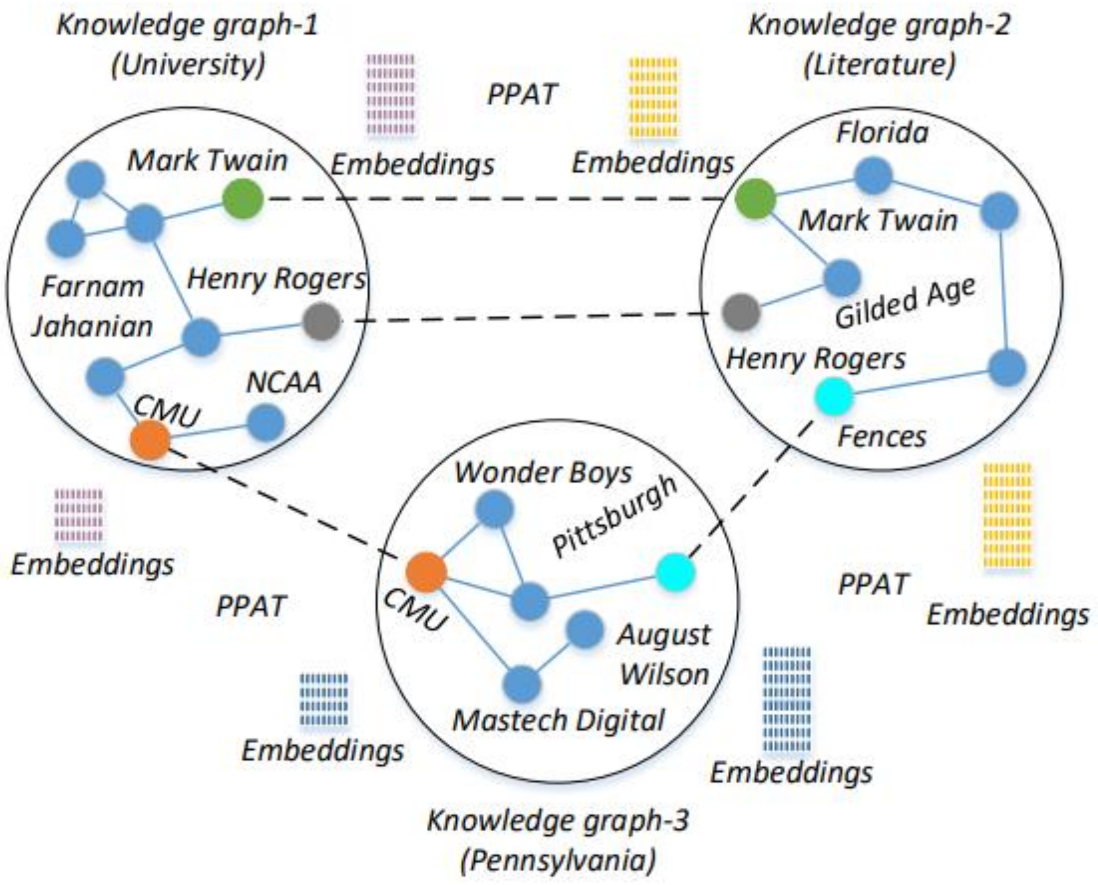
Figure Credit: Fei Wang

Antoine Bordes, Nicolas Usunier, Alberto García-Durán, Jason Weston, Oksana Yakhnenko: Translating Embeddings for Modeling Multi-relational Data. NIPS 2013

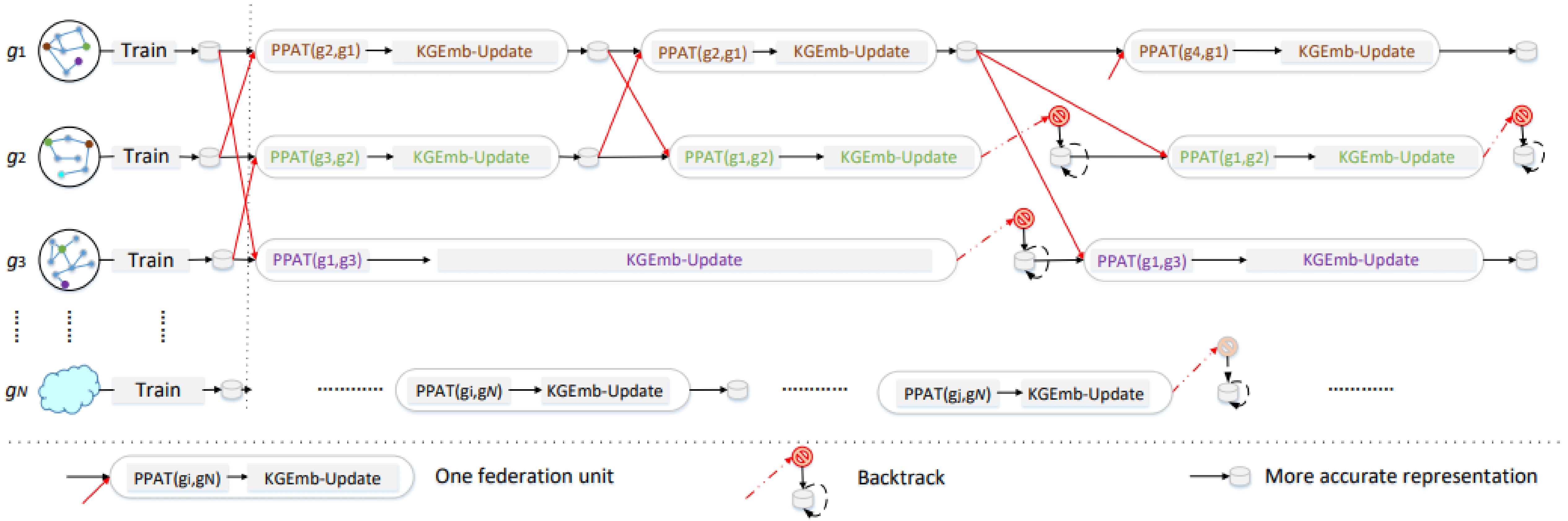
KG Embedding from Different Owners

- Existing knowledge graph embedding performs well on individual KG,
  - But may not be applied directly on multiple KGs

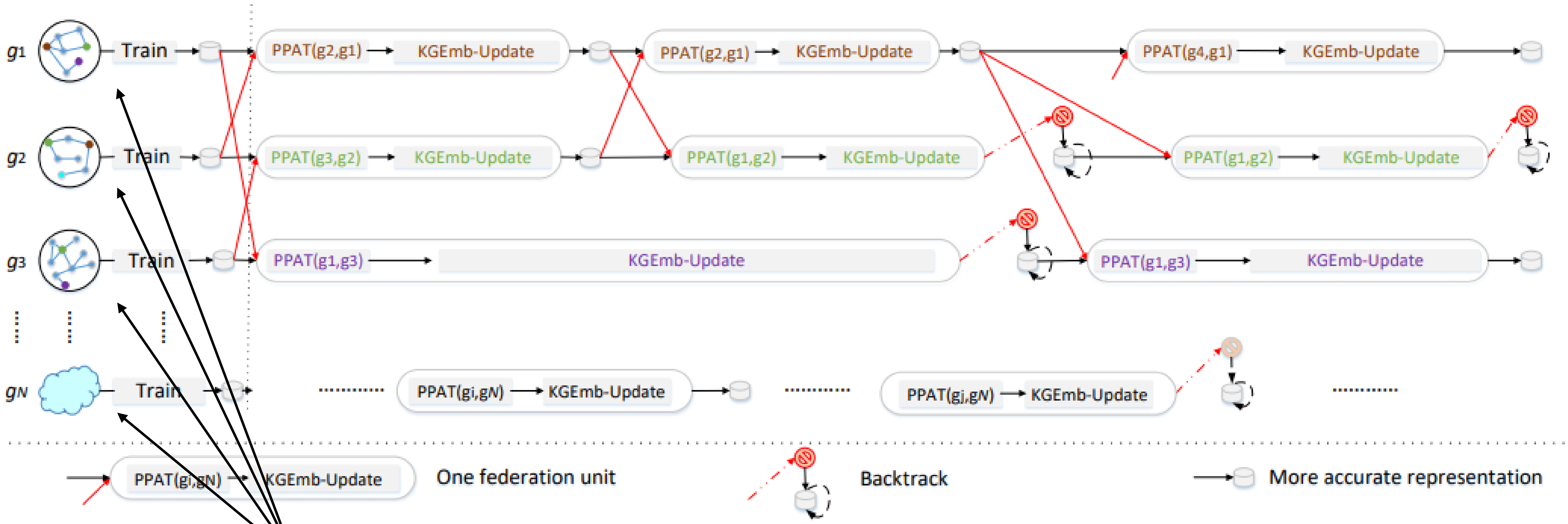
- They do have incentives to share KGs if they can:
  - Benefit from sharing
  - Improve their own services without revealing sensitive records



The FKGE Framework



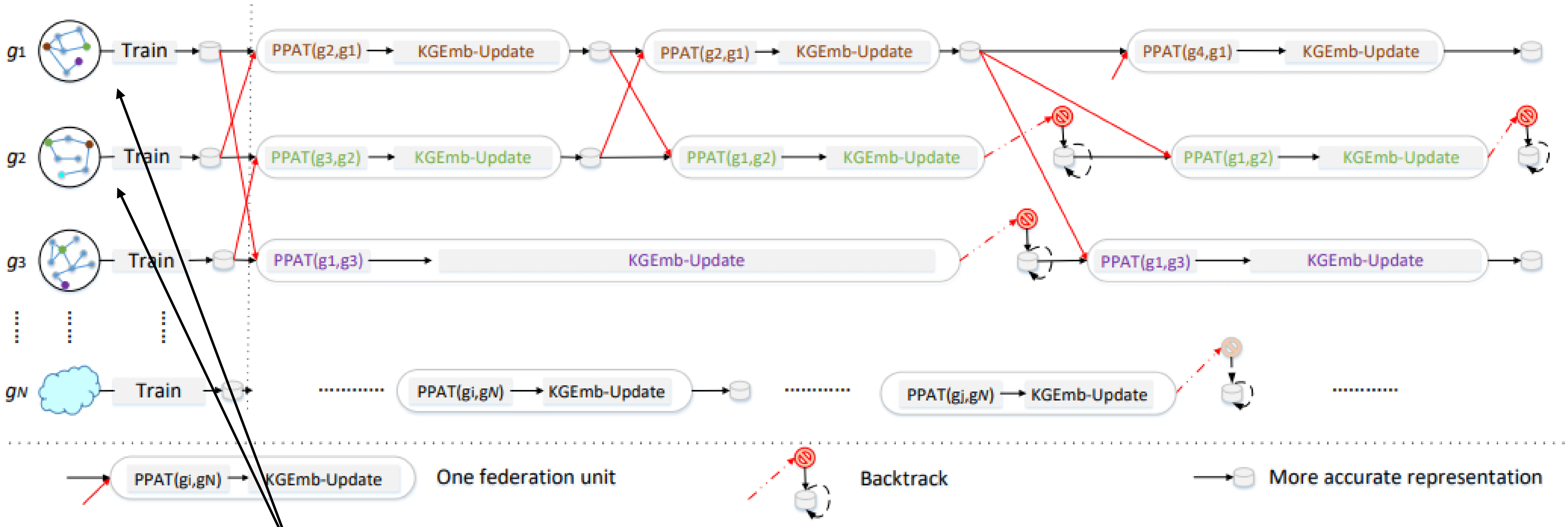
The FKGE Framework



Knowledge graphs g_i = {E_i, R_i, T_i} for entities, relations, and triples.

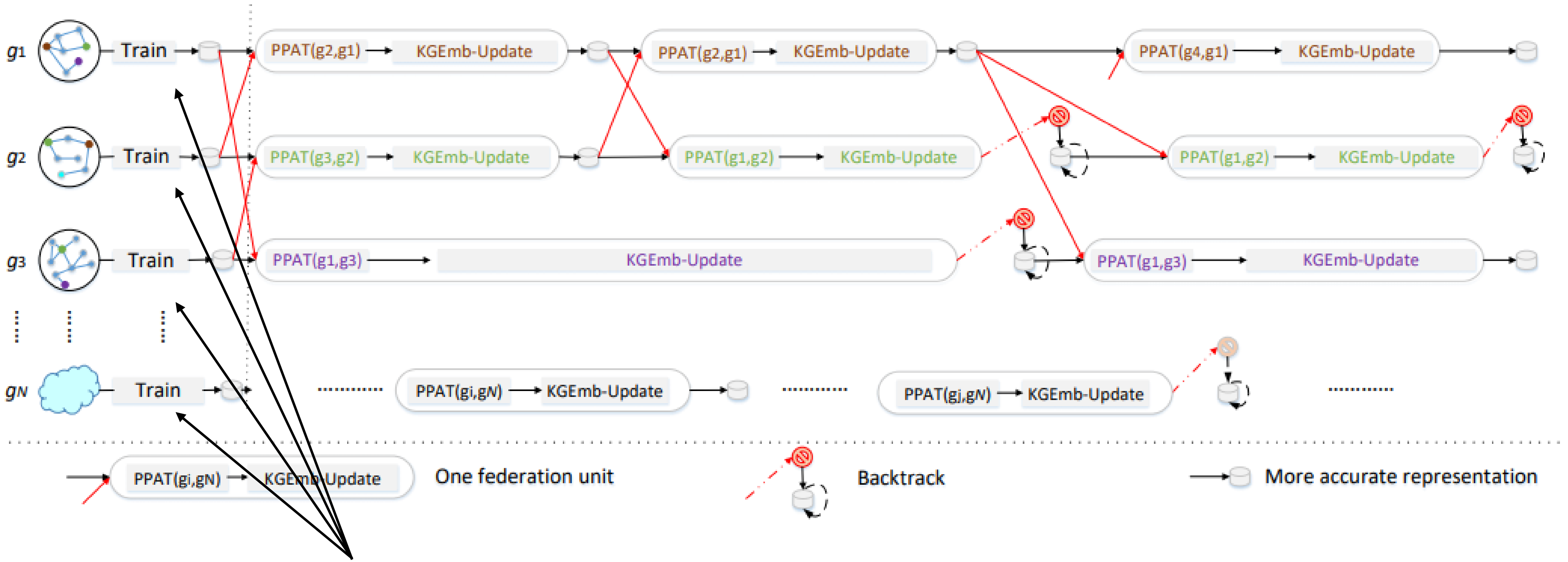
Every element in KG locates in different databases and cannot access other KGs' databases

The FKGE Framework



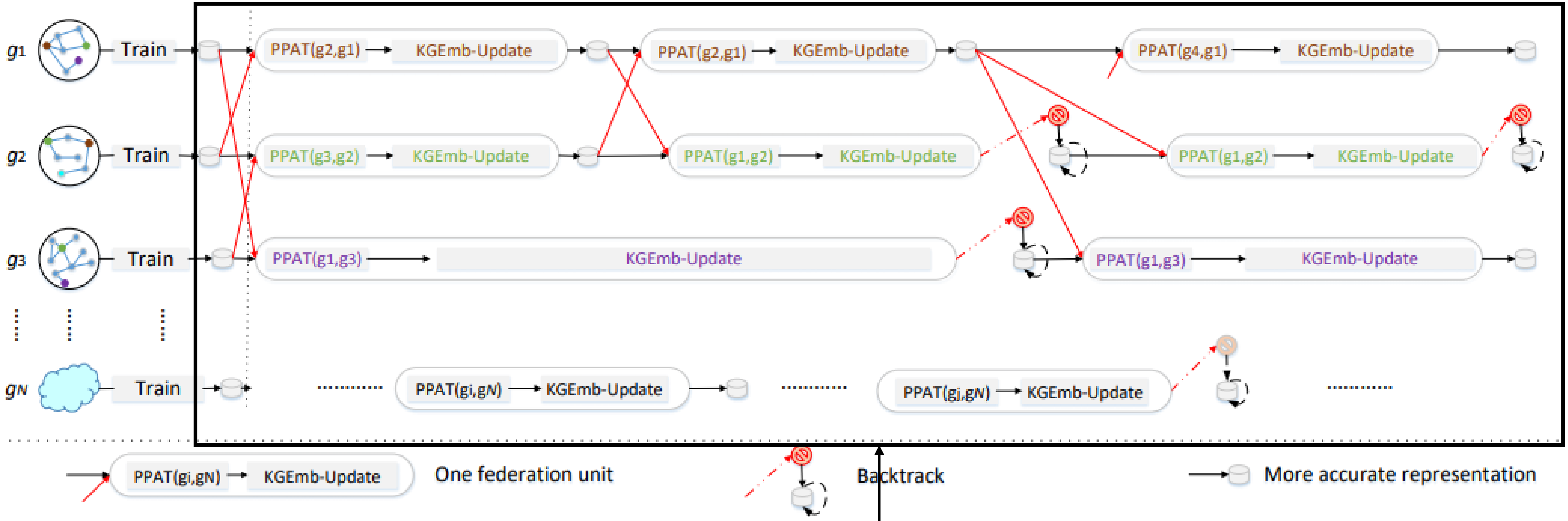
A subset of entities $E_i \cap E_j$ and relations $R_i \cap R_j$ in each pair of KGs is known to be the same.

The FKGE Framework



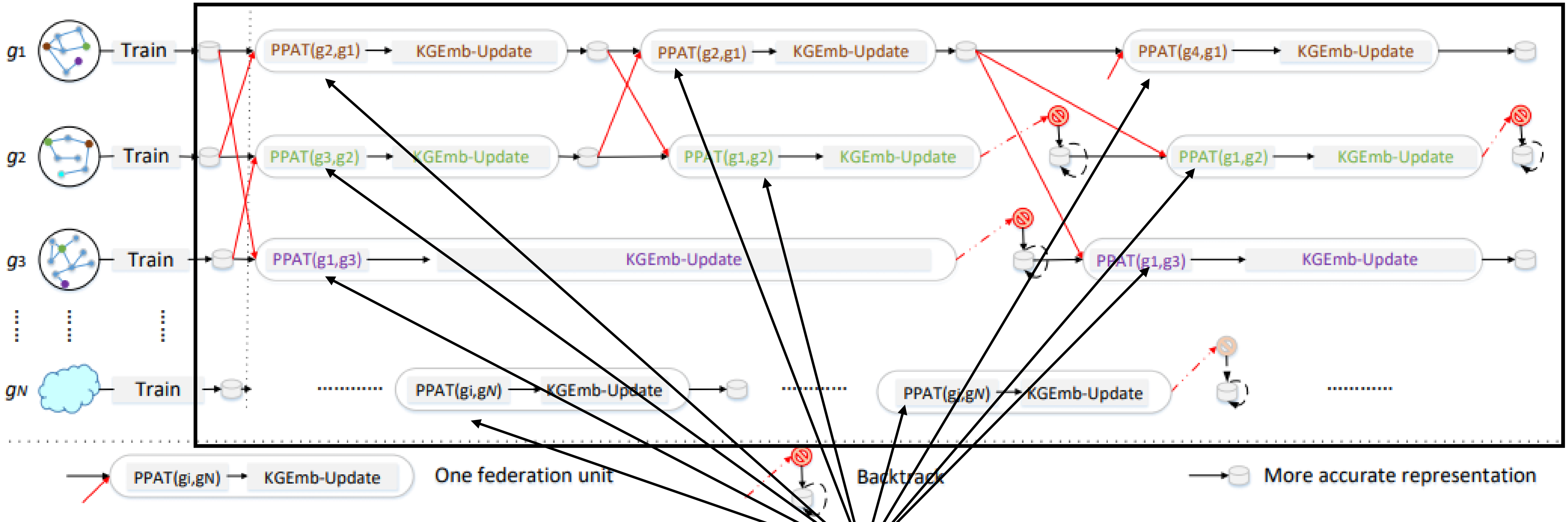
Each KG owner trains its own embeddings of entities and relations locally.

The FKGE Framework



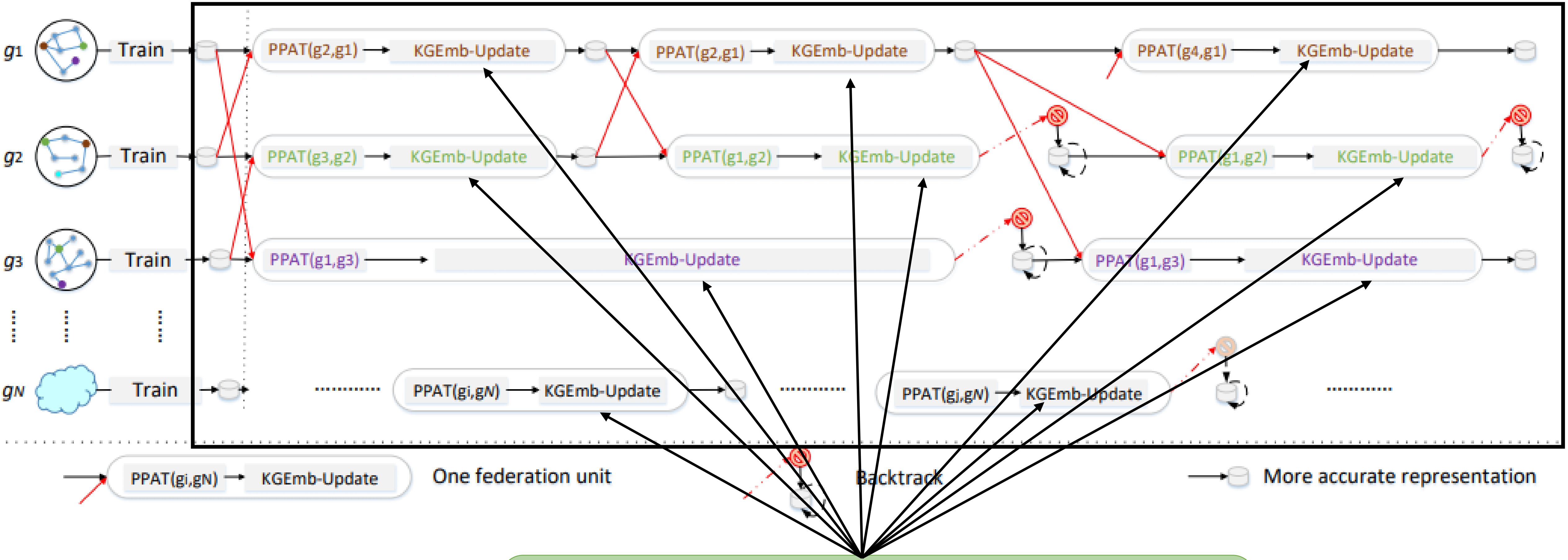
Based on the trained embeddings, FKGE aggregates the embeddings of both aligned entities and relations from paired KGs, and then updates all embeddings in a federated manner.

The FKGE Framework



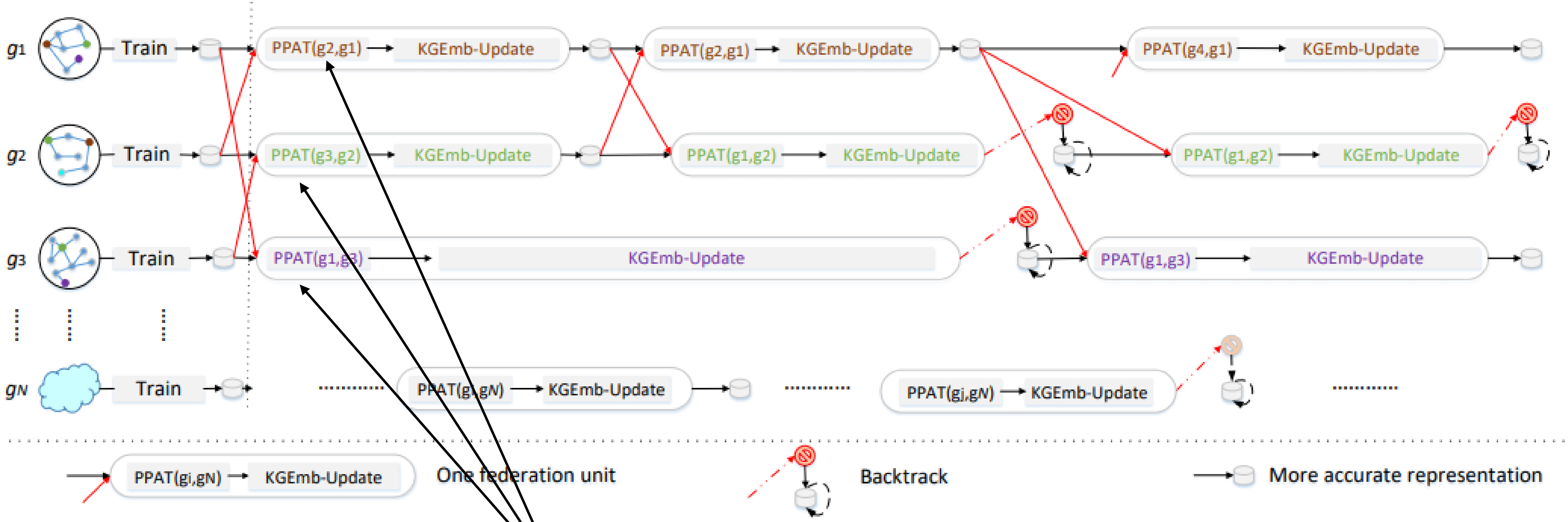
FKGE includes a secure pipeline that can refine the embeddings of $E_i \cap E_j$ and $R_i \cap R_j$ and further improve embeddings.

The FKGE Framework



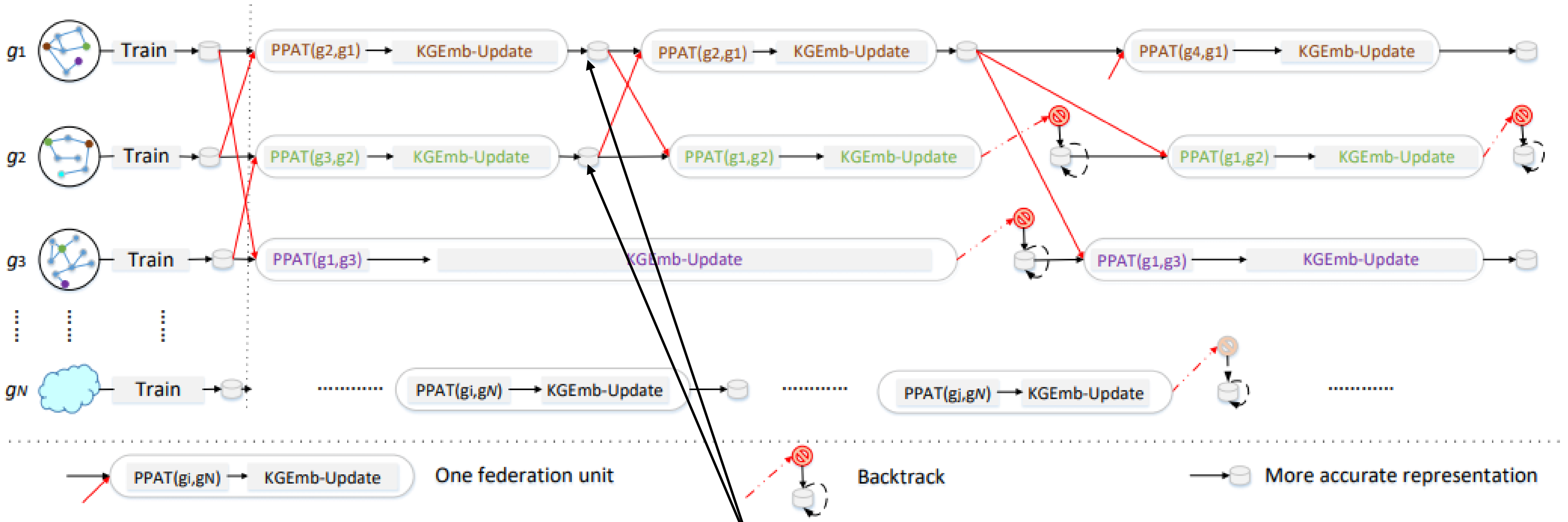
FKGE further improve embeddings of $E_i \cup R_i$ and $E_j \cup R_j$ individually.

The FKGE Framework: A Running Example



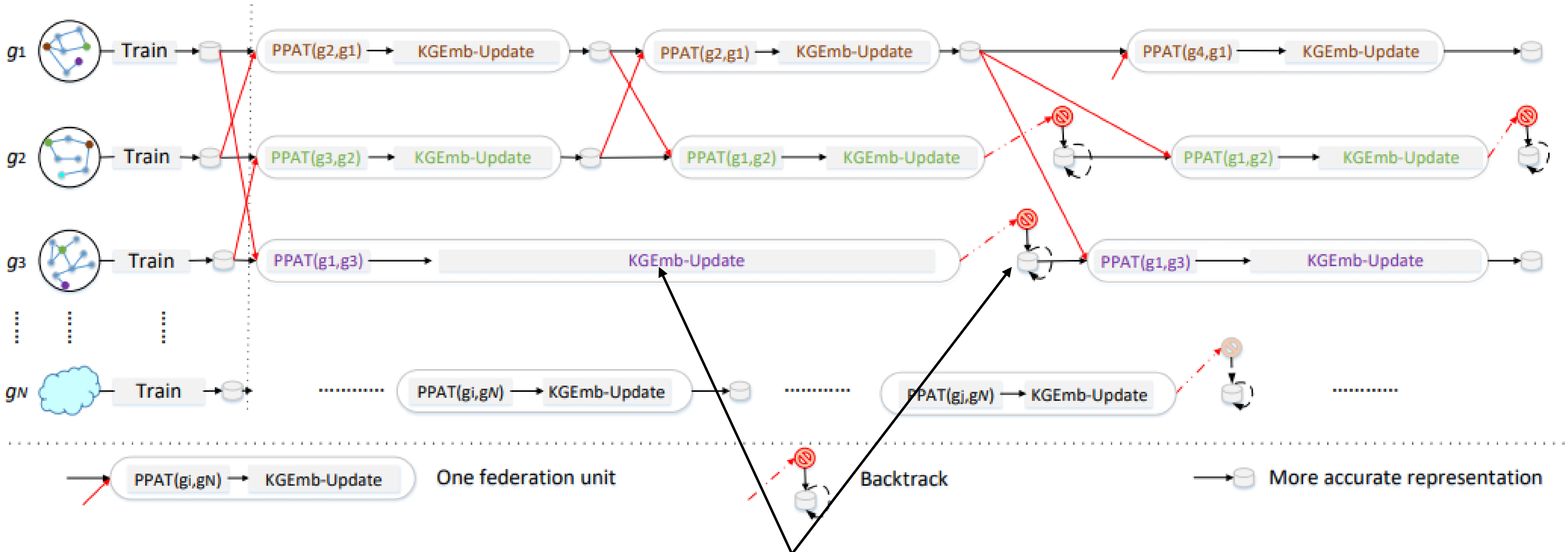
During first federation, they form 3 pairs of KGs: (g_1, g_3), (g_2, g_1), and (g_3, g_2).

The FKGE Framework: A Running Example



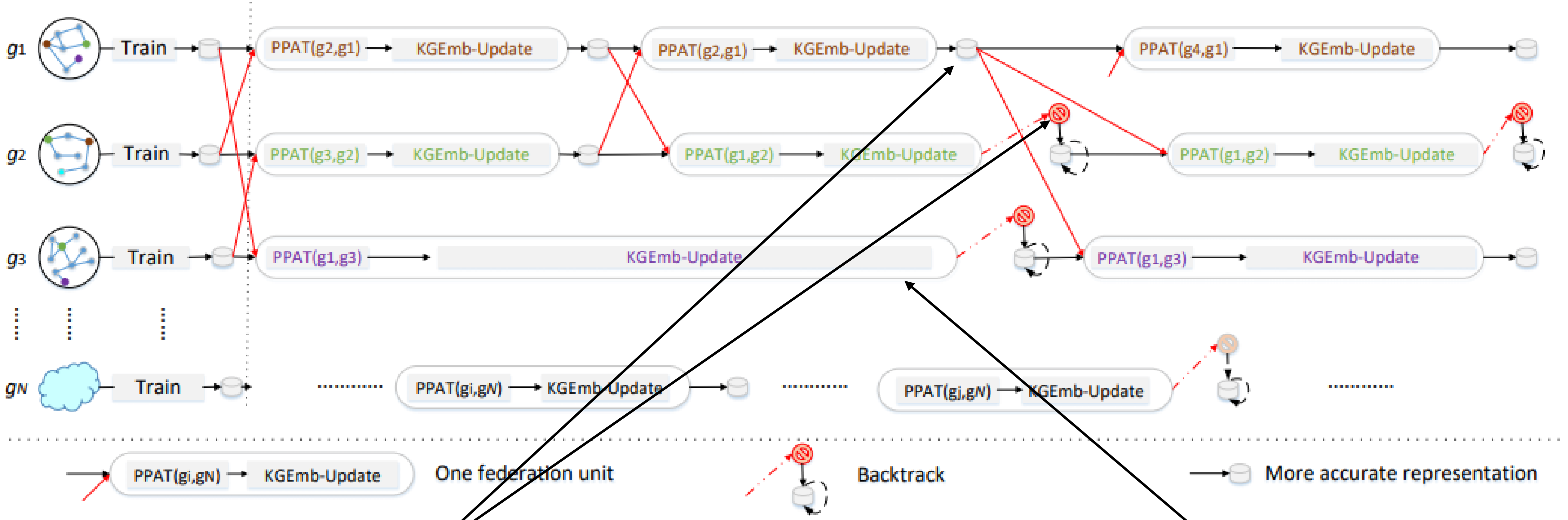
After the first federation, g_1 and g_2 gain improvement for overall embeddings.

The FKGE Framework: A Running Example



g_3's training takes longer time and fails to improve its embedding; therefore, g_3 backtracks to initial embedding.

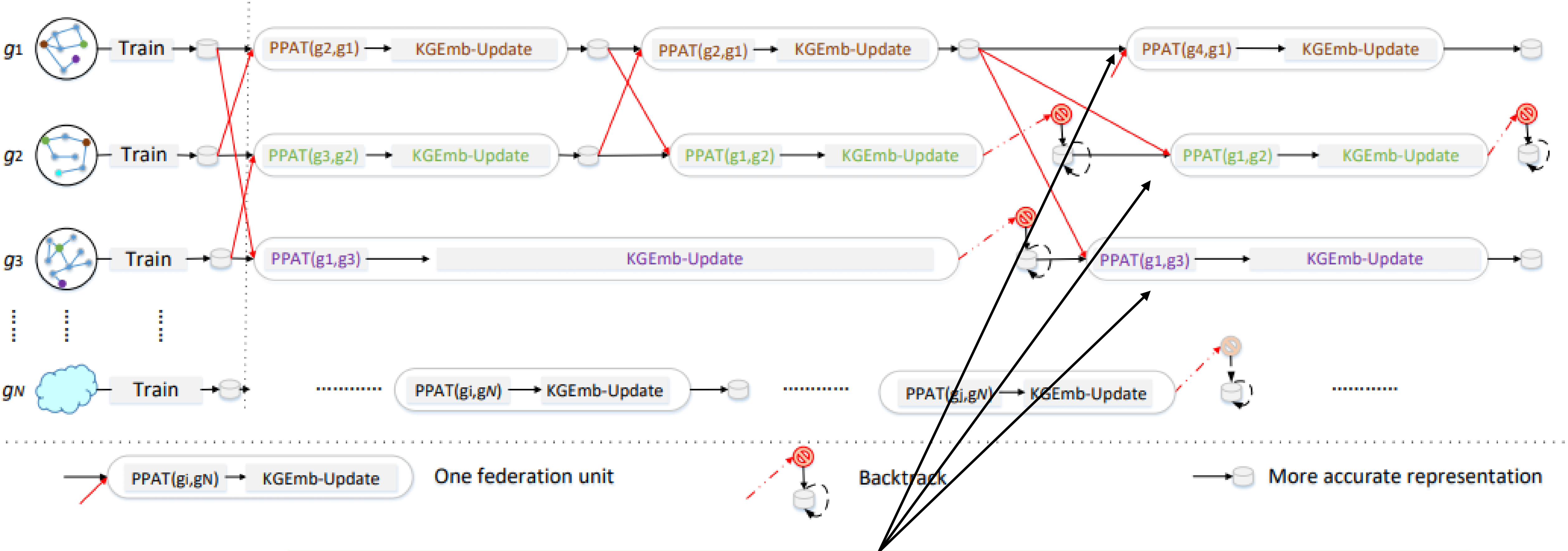
The FKGE Framework: A Running Example



During second federation, g_1 and g_2 pair up as (g_2, g_1) and (g_1, g_2) and only g_1 gains improvements. g_2 backtracks to previous embedding.

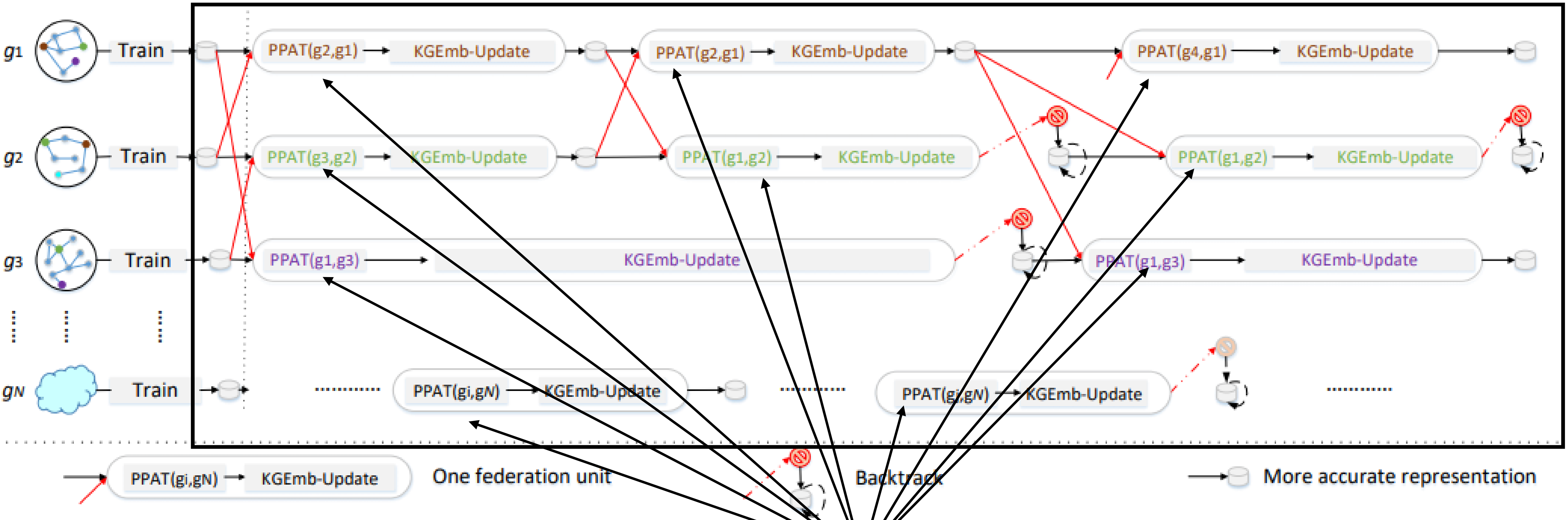
Since g_3 is still on the training process, it will not join second federation and will go to sleep state if no available KG exists.

The FKGE Framework: A Running Example



For third federation, g_1 finishes its training and broadcasts g_3 to wake up. Then they form (g_1, g_3), (g_1, g_2) and (g_4, g_1) pairs for federation based on each queue owned by each KG.

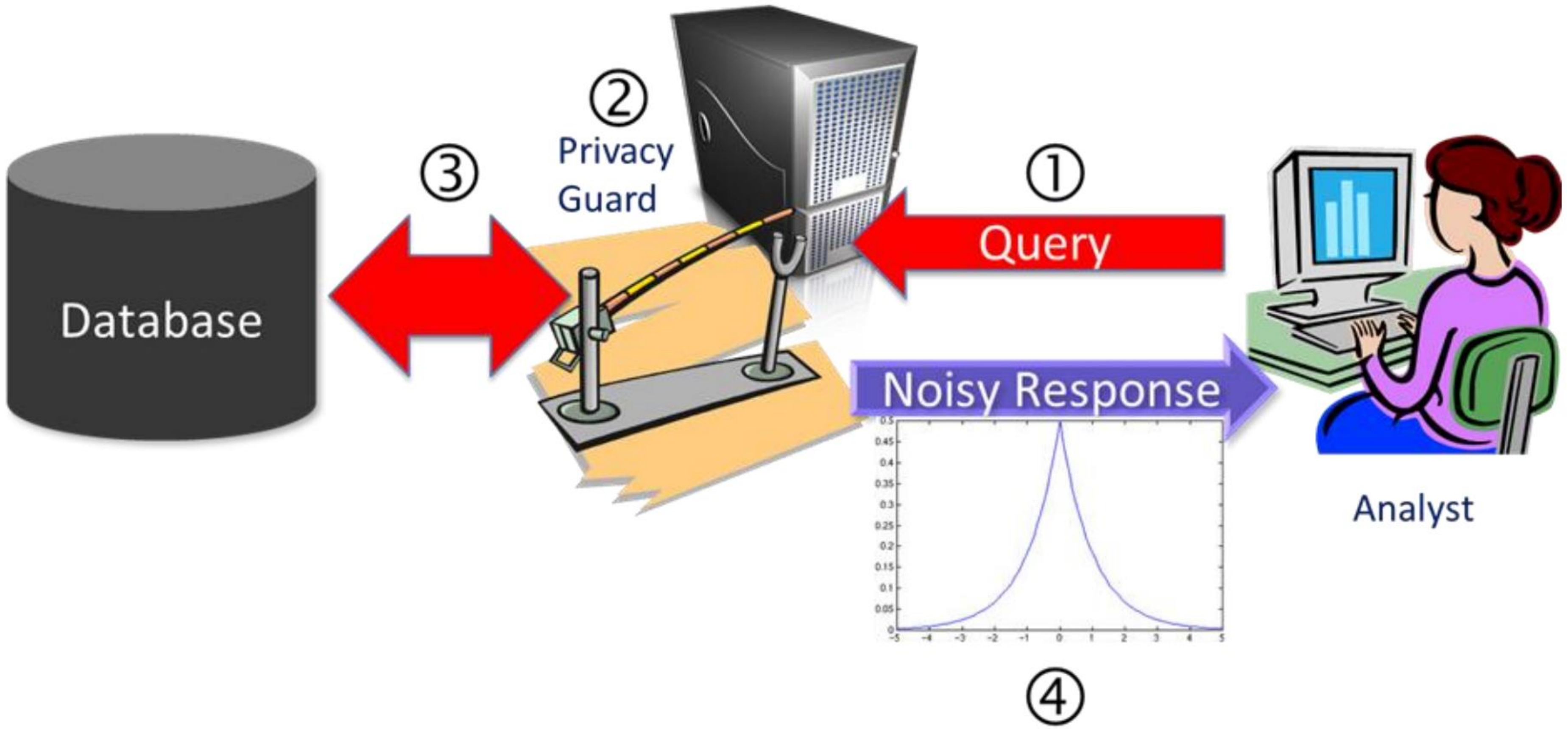
The FKGE Framework



Remaining problem: How to perform secure alignment of embeddings?

Background: Differential Privacy (DP)

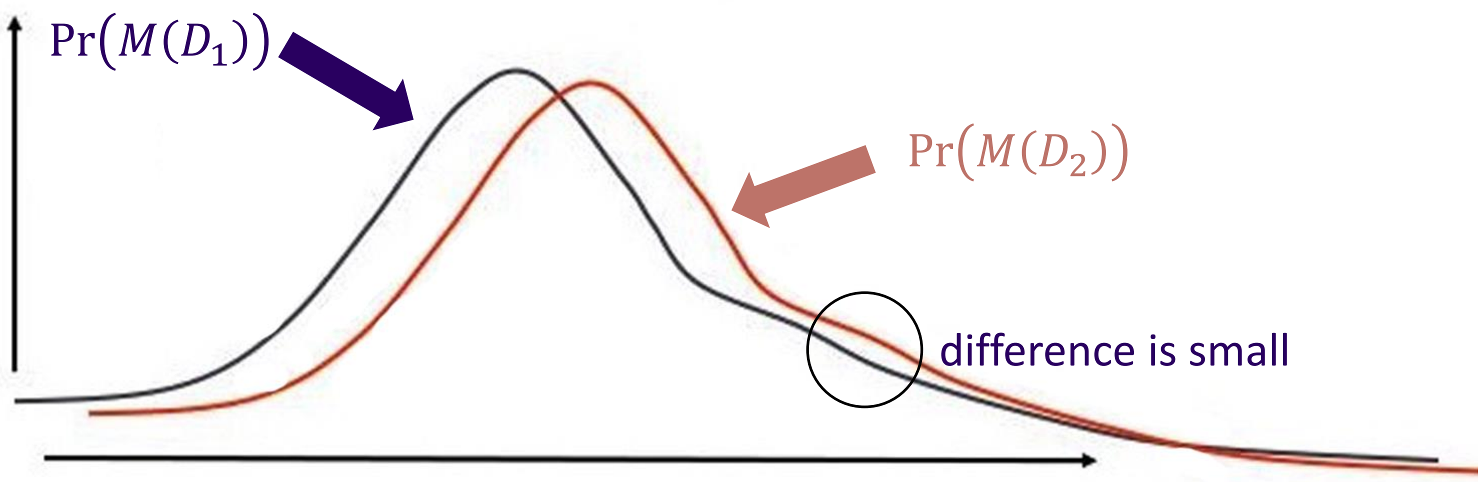
A lightweight privacy preserving solution



Slides credit: WeBank Tutorial, Chapter 2: Privacy-Preserving Techniques. https://www.fedai.org/

Background: Differential Privacy (DP)

- **Definition: Differential Privacy (DP)** [Dwork 2008]
- A randomized mechanism M is ϵ-differentially private, if for all output t of M, and for two databases D_1 and D_2 which differ by at most one element, we have
  - $\Pr(M(D_1) = t) = e^{\epsilon}\Pr(M(D_2) = t).$



Intuition: changes in the distribution are too small to be perceived with variations on a single element.

Cynthia Dwork, 2008. Differential privacy: a survey of results. Theory and Applications of Models of Computation.

Slides credit: WeBank Tutorial, Chapter 2: Privacy-Preserving Techniques. https://www.fedai.org/

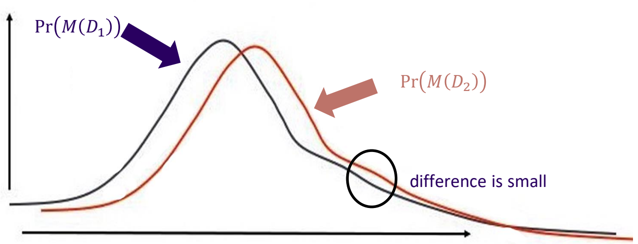
Background: DP in Machine Learning

- WISH: parameters of ML models to encode general patterns
  - ‘‘patients who smoke are more likely to have heart disease’’
- Rather than facts about specific training examples
  - “Jane Smith has heart disease”

- REALITY: ML algorithms do not learn to ignore specifics by default
  - So here the randomized mechanism M in machine learning is a learning algorithm that can satisfy the differential privacy
  - Differential privacy is in fact well aligned with the goals of machine learning
    - Reduce overfitting

Background: Private Aggregation of Teacher Ensembles (PATE)

- A framework of differential privacy requires that
  - the probability change (the *privacy budget*) of learning any particular set of parameters stays roughly the same
  - if we change a single training example in the training set
    - add a training example
    - remove a training example
    - change the values within one training example



- If a single patient (Jane Smith) does not affect the outcome of learning, then that patient's records cannot be memorized and her privacy is respected
- Smaller privacy budgets correspond to stronger privacy guarantees

Nicolas Papernot, Martín Abadi, Úlfar Erlingsson, Ian Goodfellow, Kunal Talwar. Semi-supervised Knowledge Transfer for Deep Learning from Private Training Data. ICLR 2027

Background: Private Aggregation of Teacher Ensembles (PATE)

- Assume that Jane Smith contributed to the training data of one of models only
  - If that model predicts that a patient like Jane has cancer
  - whereas the other model predicts the contrary,
  - this reveals private information about Jane.

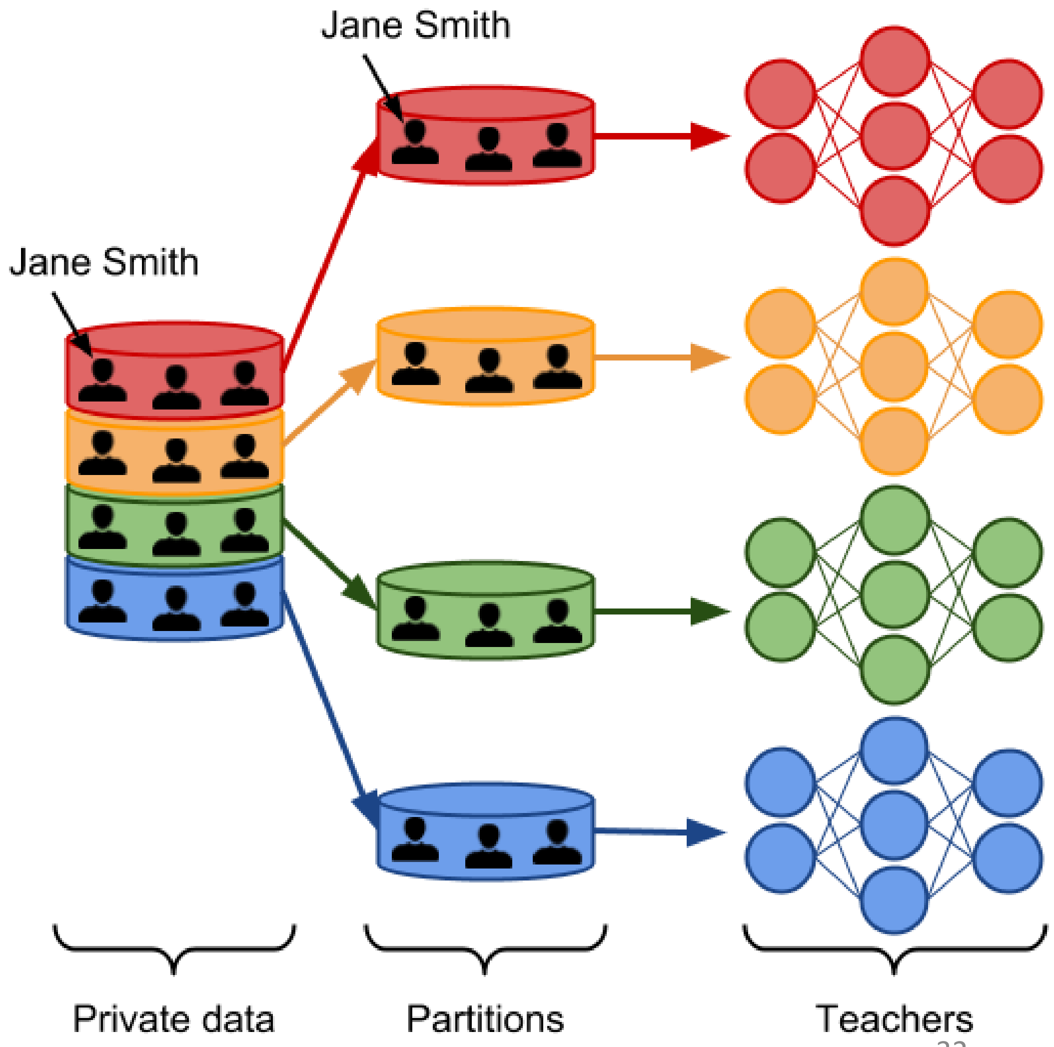


Figure from: https://www.cleverhans.io/privacy/2018/04/29/privacy-and-machine-learning.html

Background: PATE

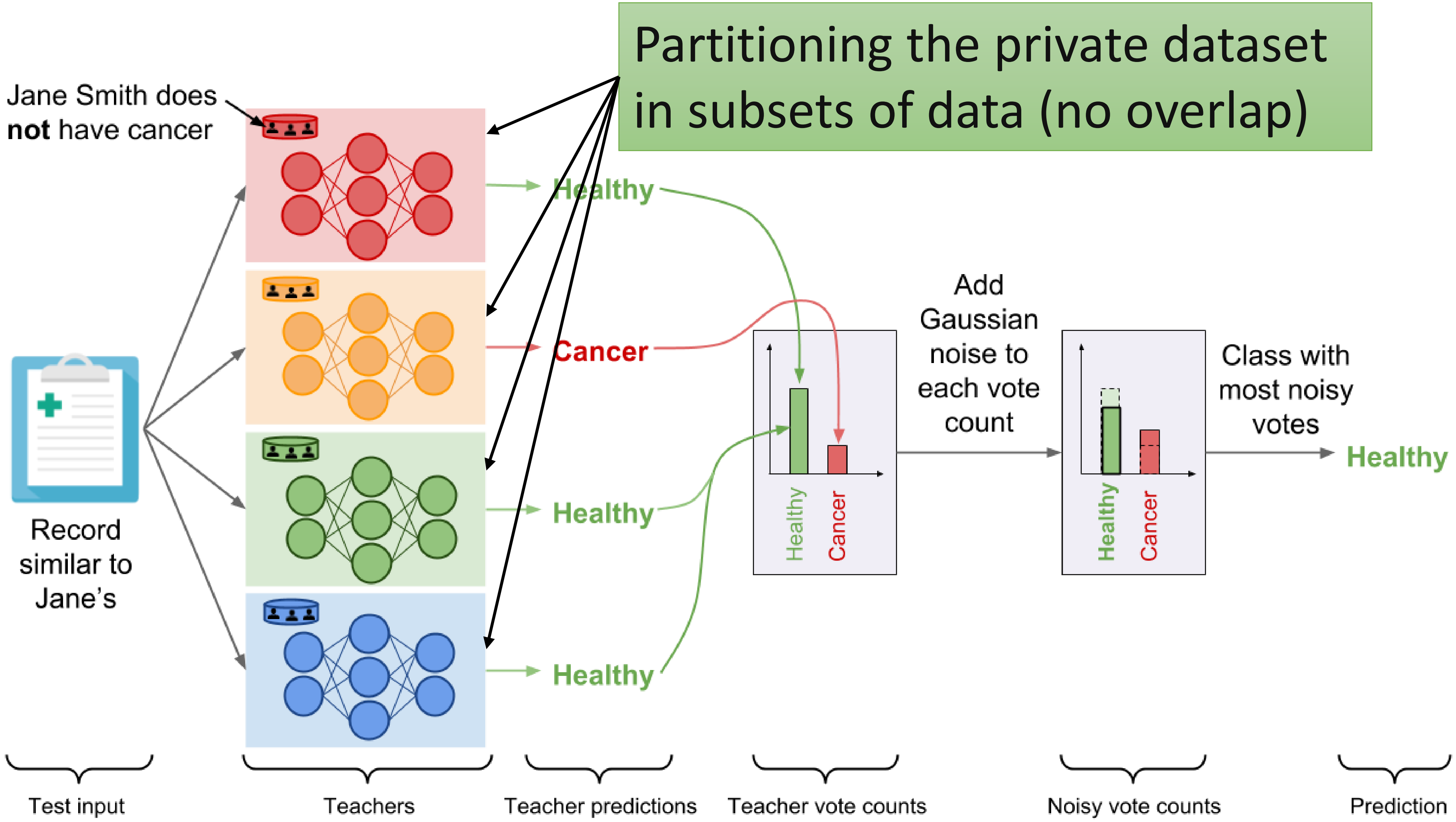


Figure from: https://www.cleverhans.io/privacy/2018/04/29/privacy-and-machine-learning.html

Background: PATE

ML model, called a *teacher*, on each of these partitions

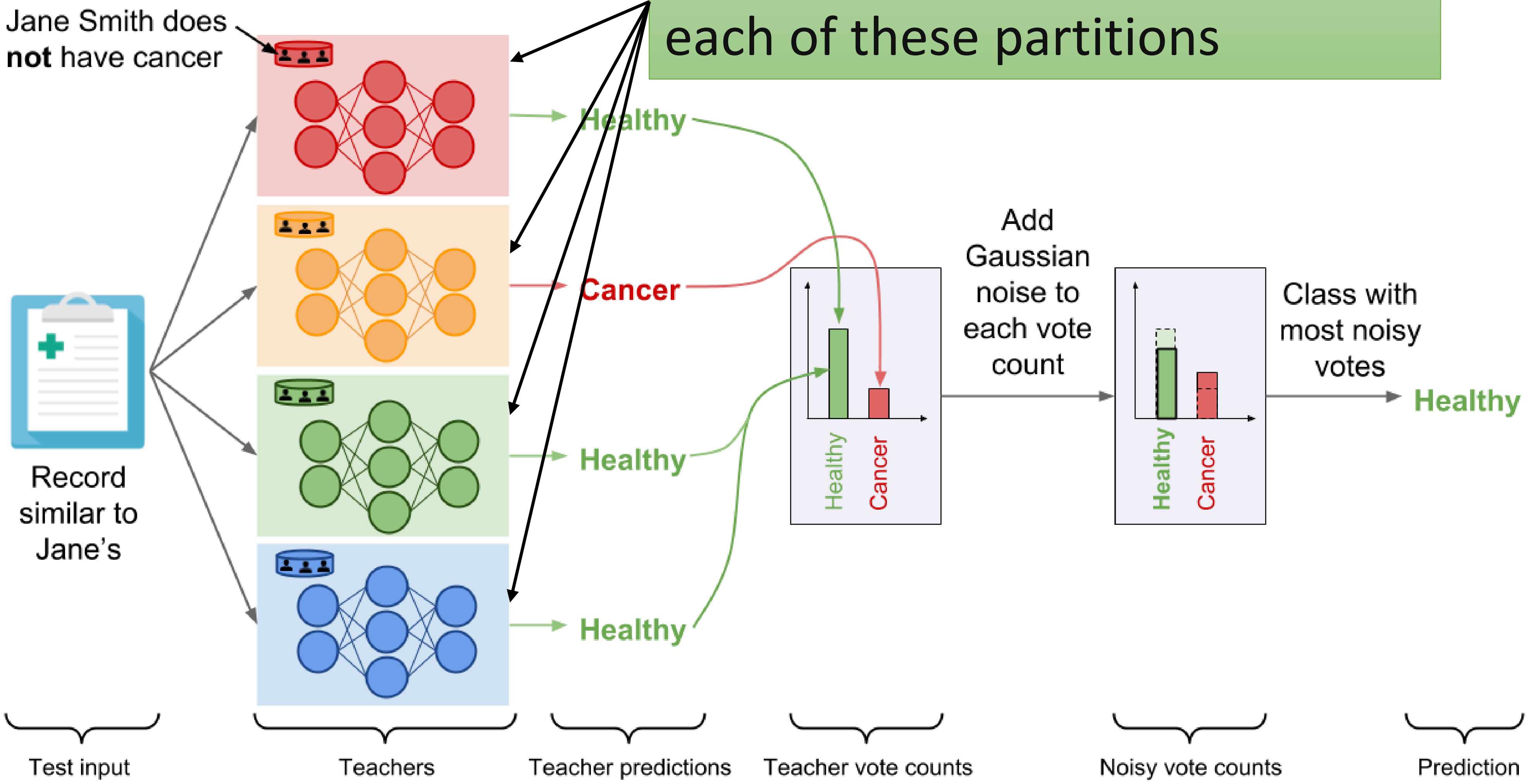


Figure from: https://www.cleverhans.io/privacy/2018/04/29/privacy-and-machine-learning.html

Background: PATE

Add noise while aggregating the predictions to form a single common prediction

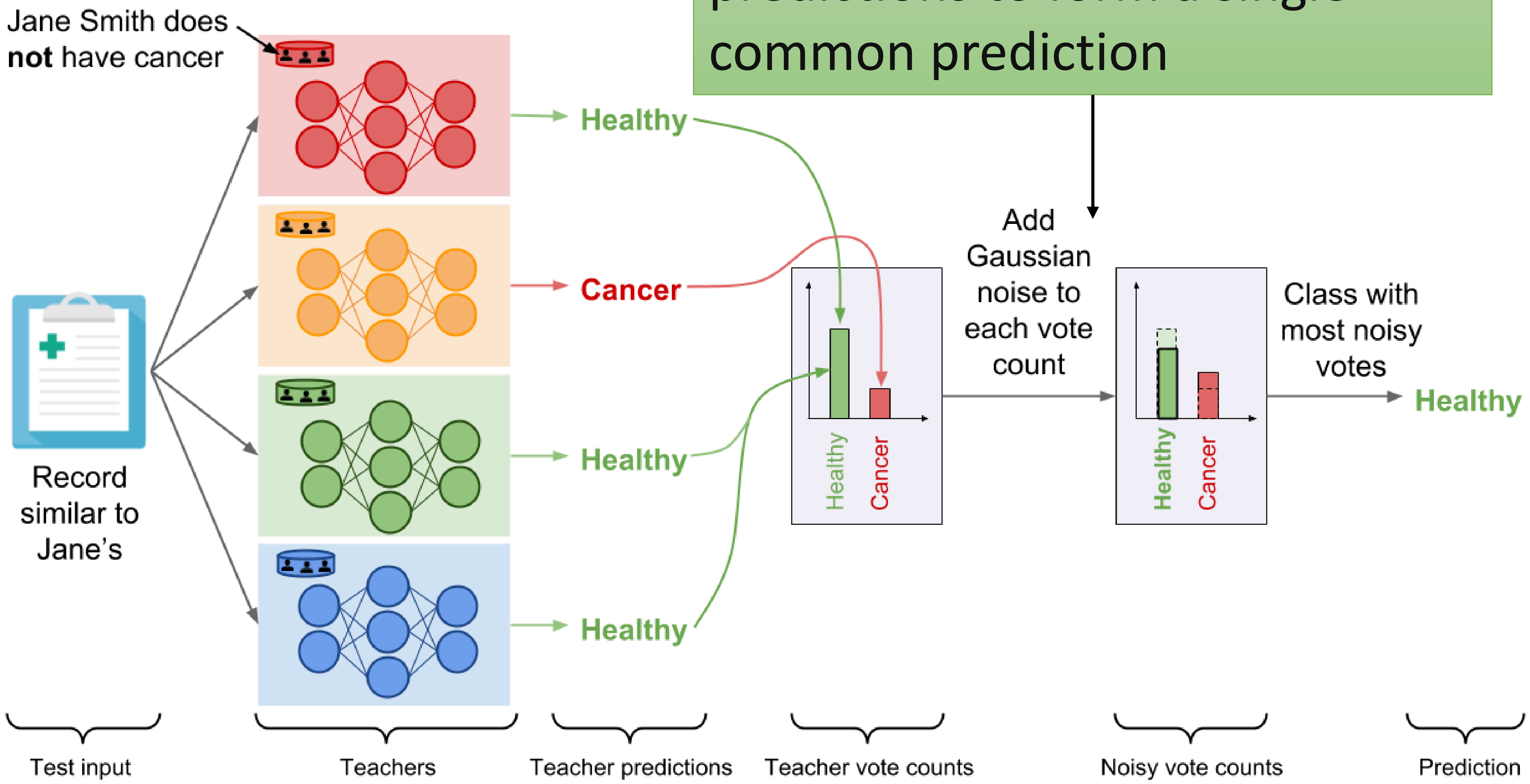


Figure from: https://www.cleverhans.io/privacy/2018/04/29/privacy-and-machine-learning.html

Background: PATE

Noisymax mechanism

$$PATE_\lambda(x) = \arg\max_{j\in\{0,1\}}(n_j(x) + V_j)$$

V_0, V_1 are i.i.d. $Lap(\frac{1}{\lambda})$ (or Gaussian noise)

$$n_j(x) = |\{T_i : T_i(x) = j\}| \quad \text{for } j = 0, 1$$

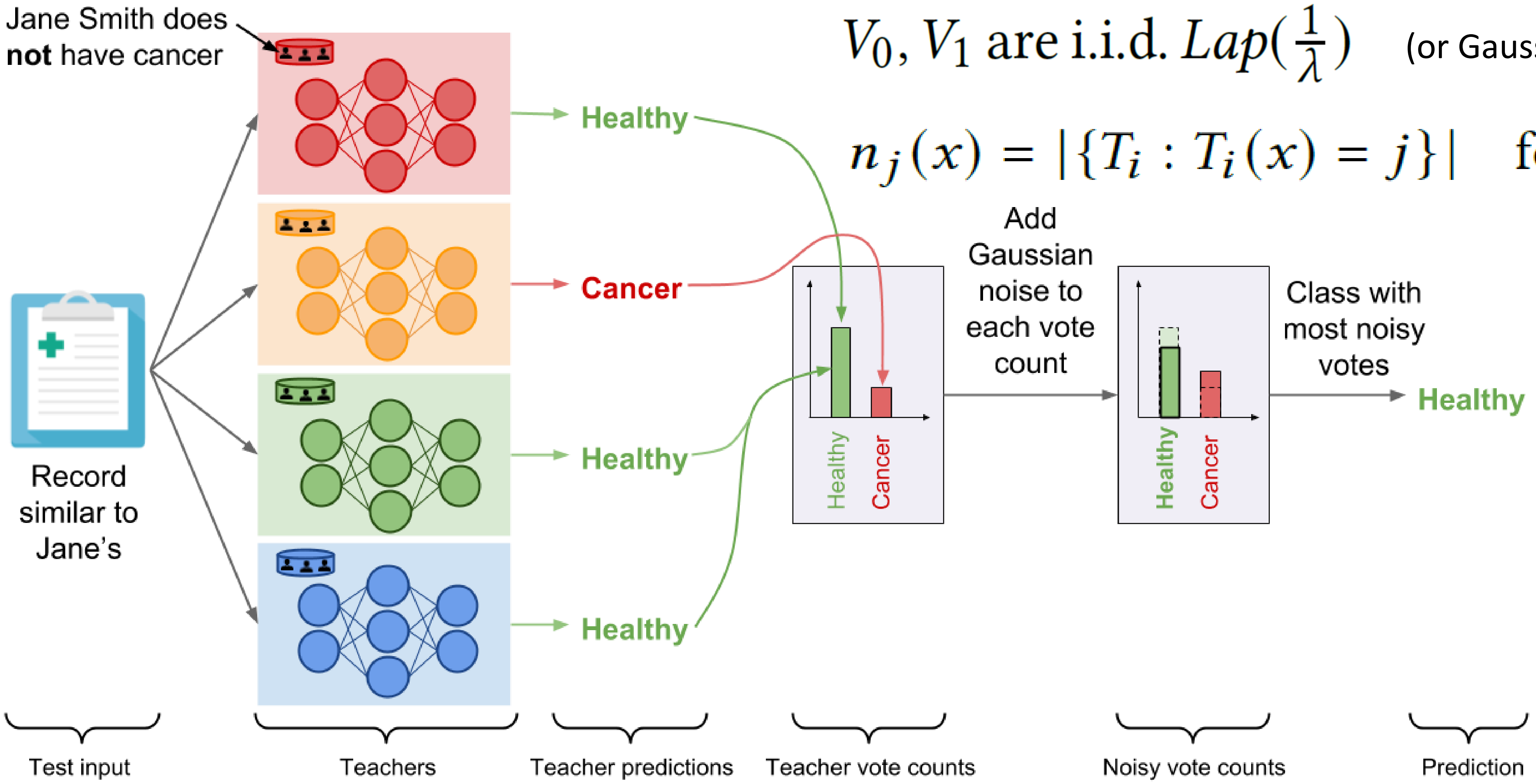


Figure from: https://www.cleverhans.io/privacy/2018/04/29/privacy-and-machine-learning.html

Background: PATE

If most of the teachers agreed on the same class, adding noise to the vote counts will not change the fact that this class received the most votes

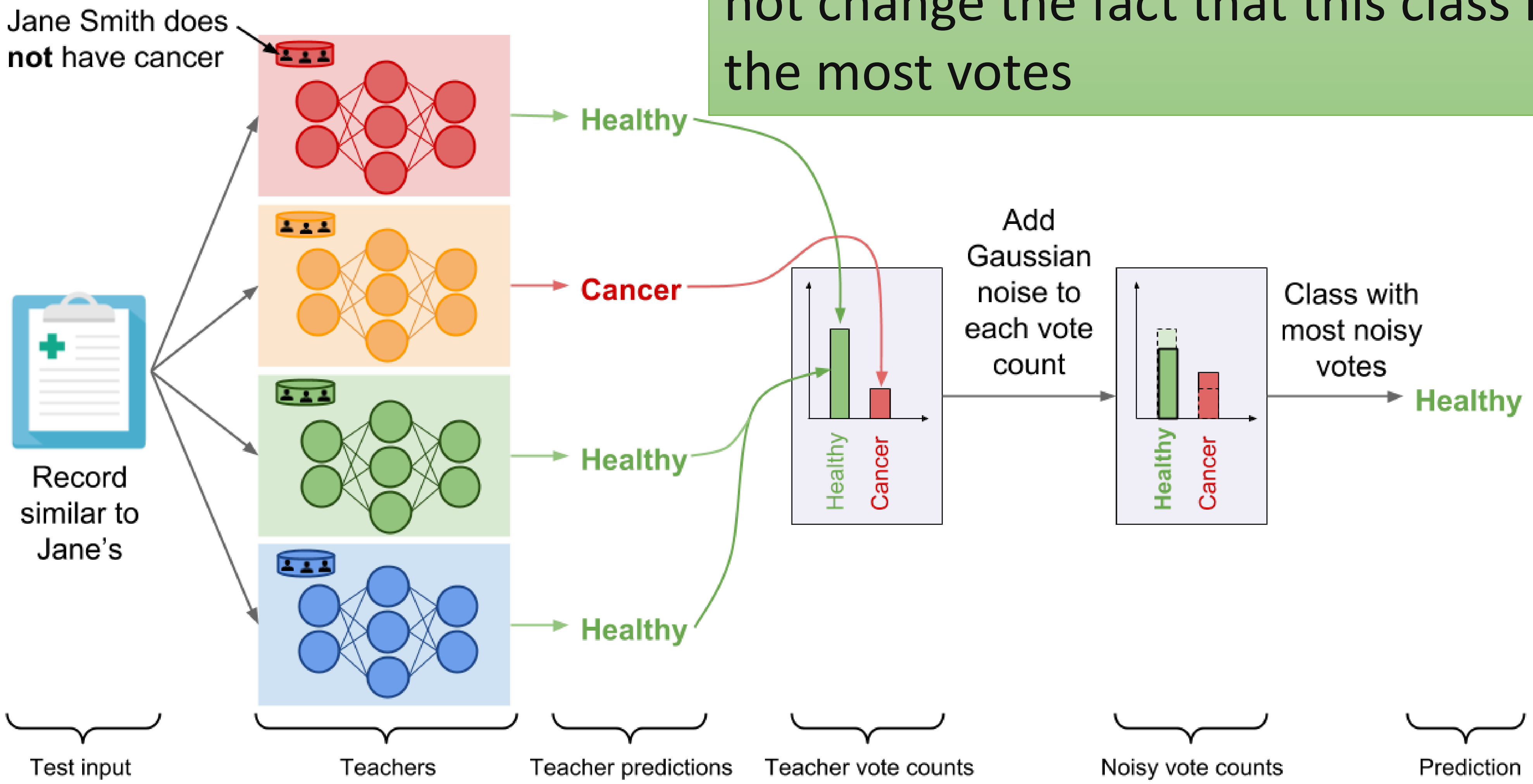


Figure from: https://www.cleverhans.io/privacy/2018/04/29/privacy-and-machine-learning.html

Background: PATE

When two teachers voting for the label "Cancer" while the two teachers vote for "Healthy":
The random noise prevents the outcome from reflecting any individual teachers to protect privacy: the noisy aggregation's outcome is equally likely to be "Healthy" or "Cancer".

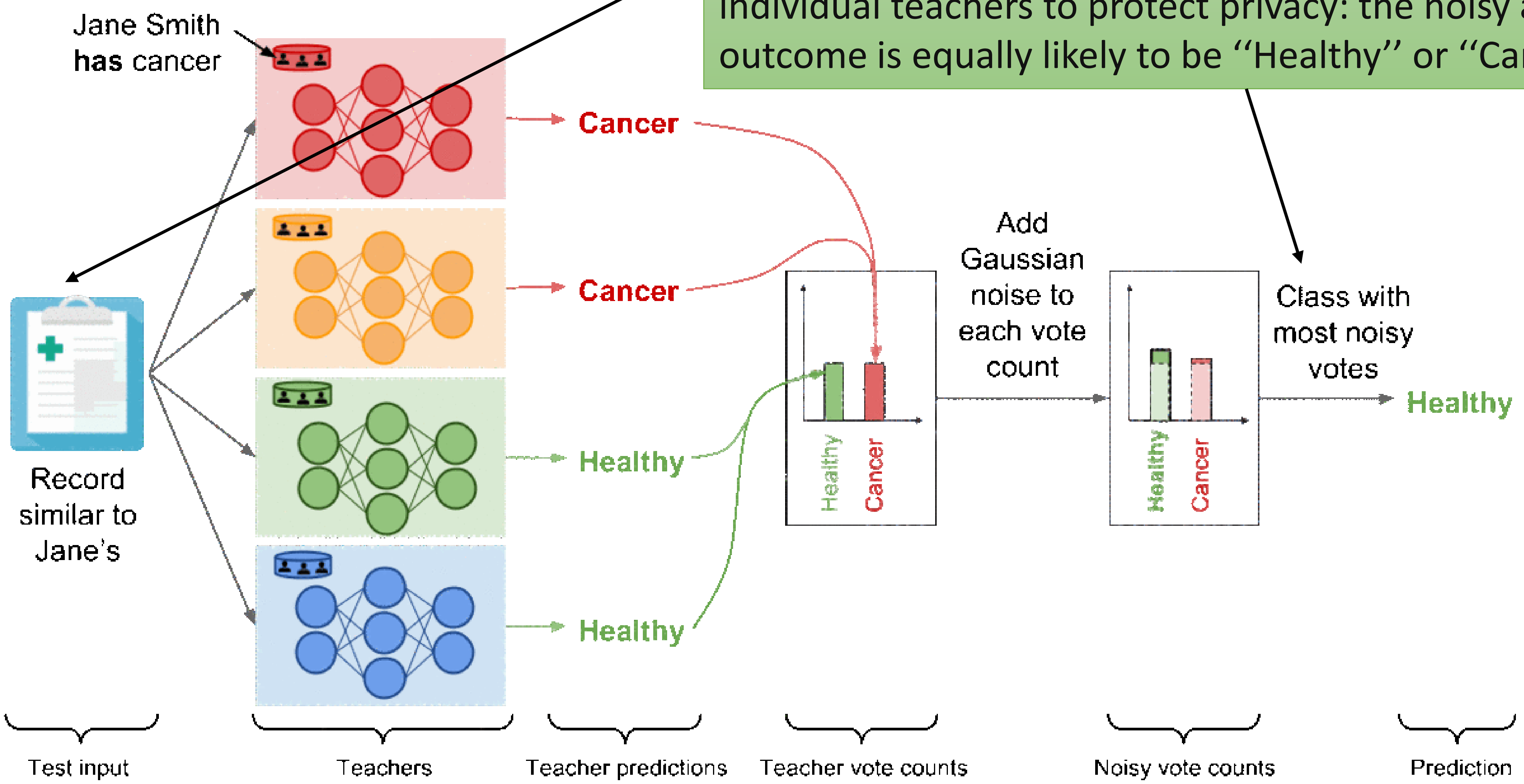


Figure from: https://www.cleverhans.io/privacy/2018/04/29/privacy-and-machine-learning.html

Background: Student Model in PATE

- Each prediction made by the aggregation mechanism increases the total privacy budget
  - The total privacy budget eventually becomes too large when many labels are predicted
- We can’t publicly publish the ensemble of teacher models

- One additional step in PATE: creating a student model

Background: PATE

- The student is trained by transferring knowledge acquired by the teacher ensemble in a privacy-preserving way.
- The student has access to a relatively large set of unlabeled inputs and it must learn with as little supervision from the teachers as possible.

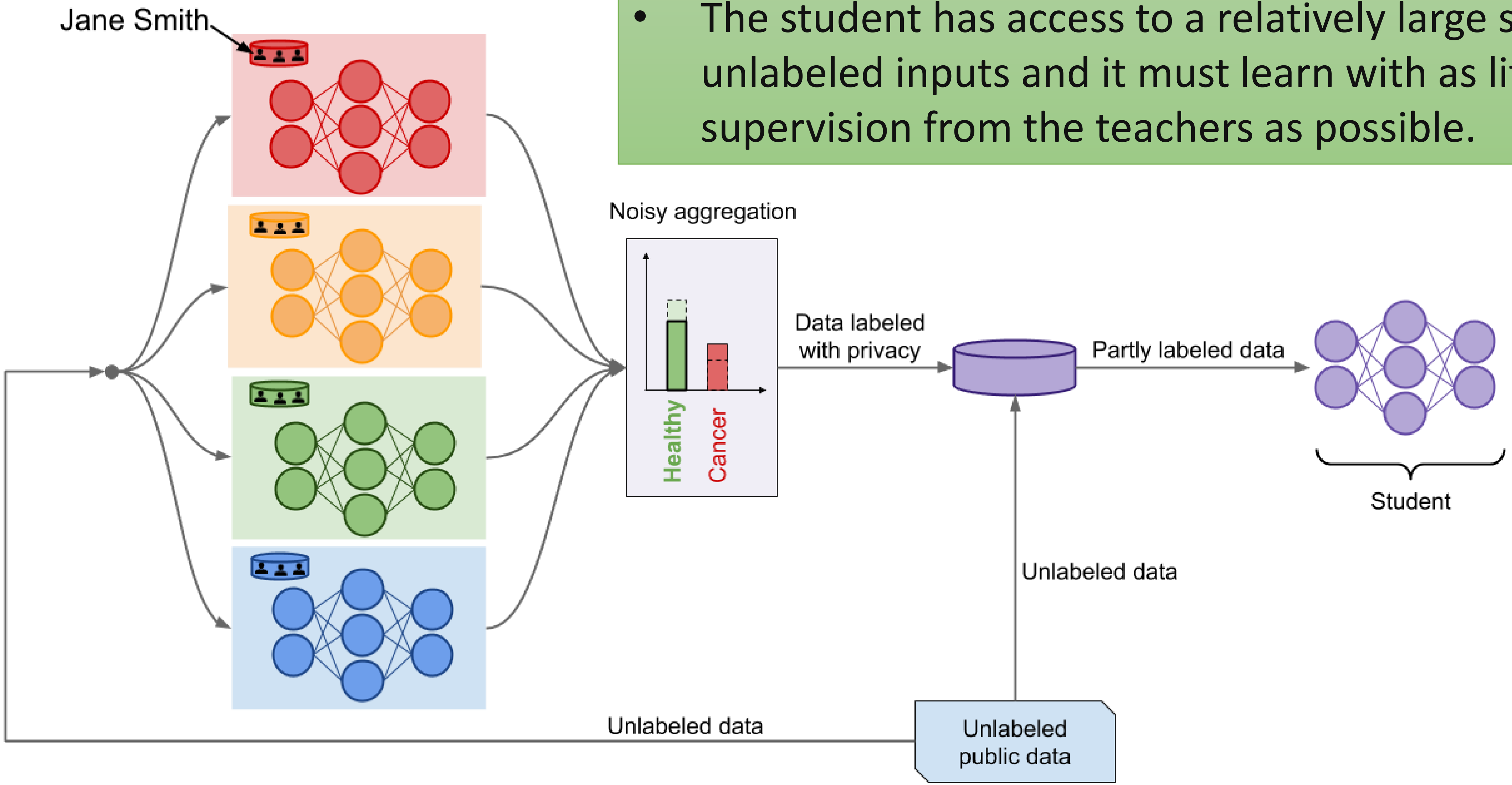
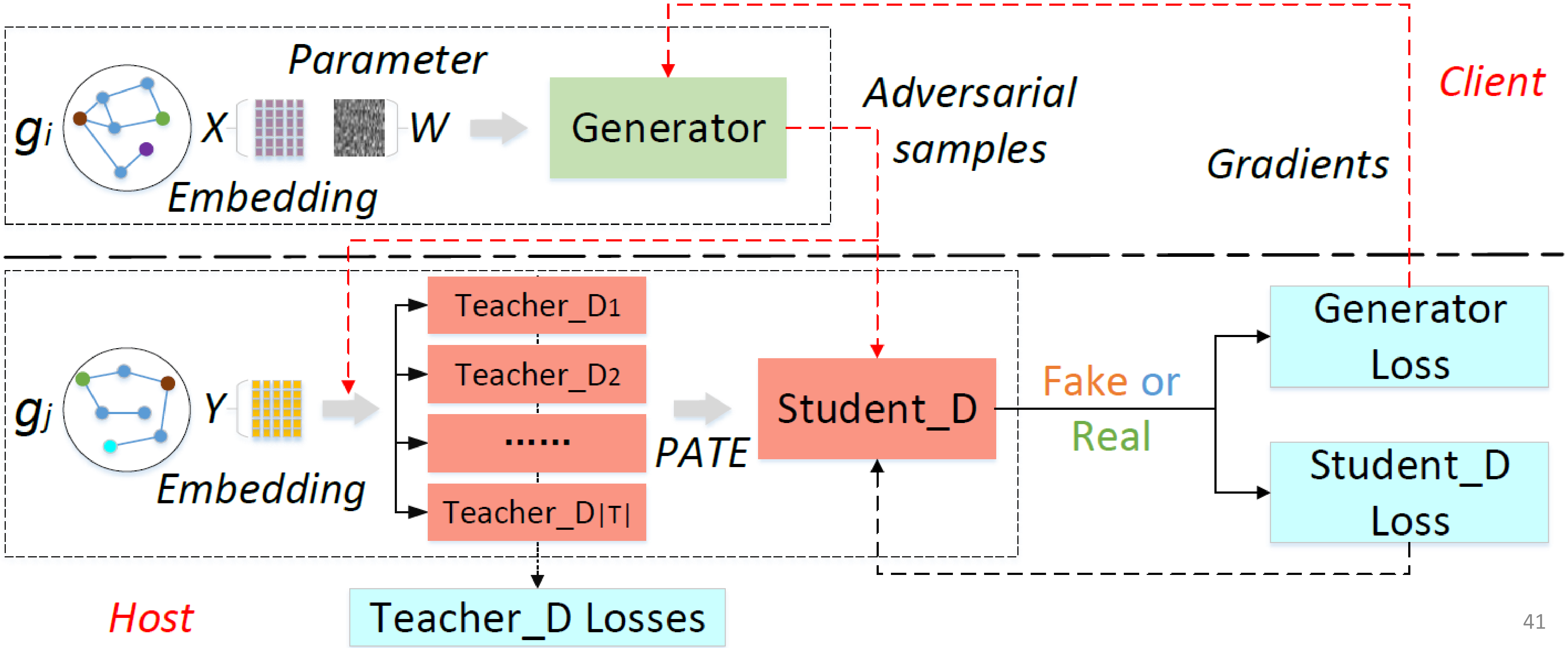


Figure from: https://www.cleverhans.io/privacy/2018/04/29/privacy-and-machine-learning.html

The Privacy-Preserving Adversarial Model

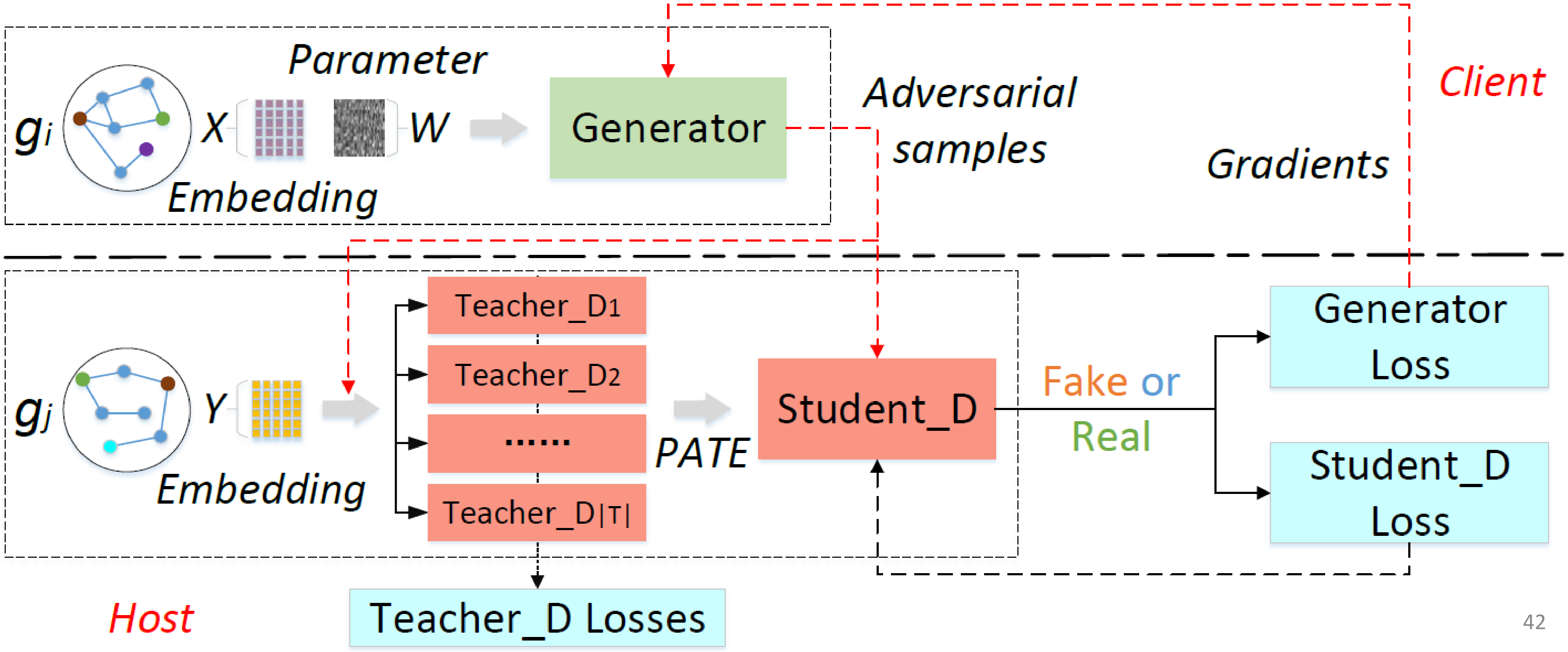
Privacy-Preserving Adversarial Translation (PPAT) network



PPAT Network

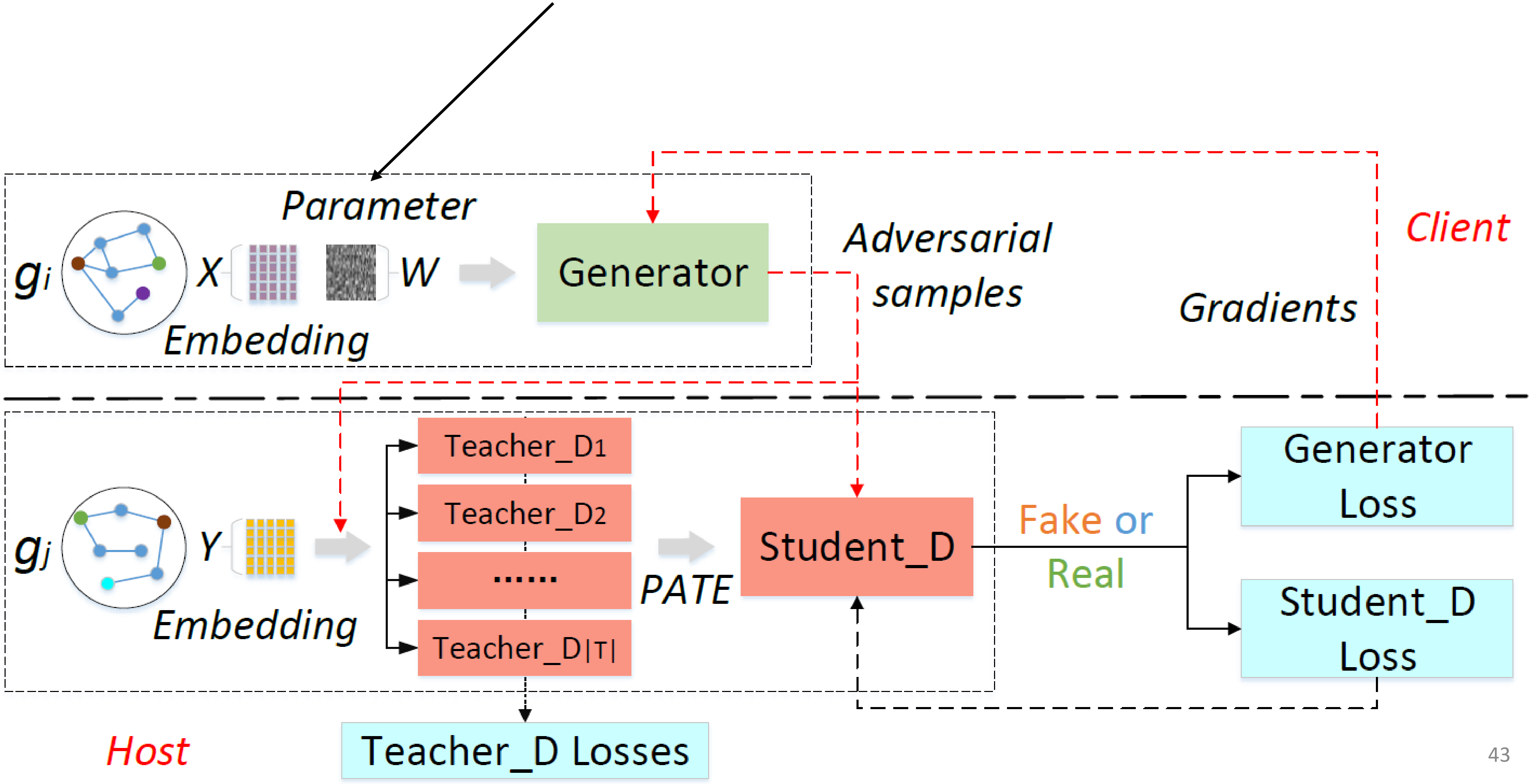
PPAT network exploits GAN structure to generate differentially private synthetic embedding with high utility

- We replace the original GAN discriminator with multiple teacher discriminators and
- One student discriminator to achieve differential privacy



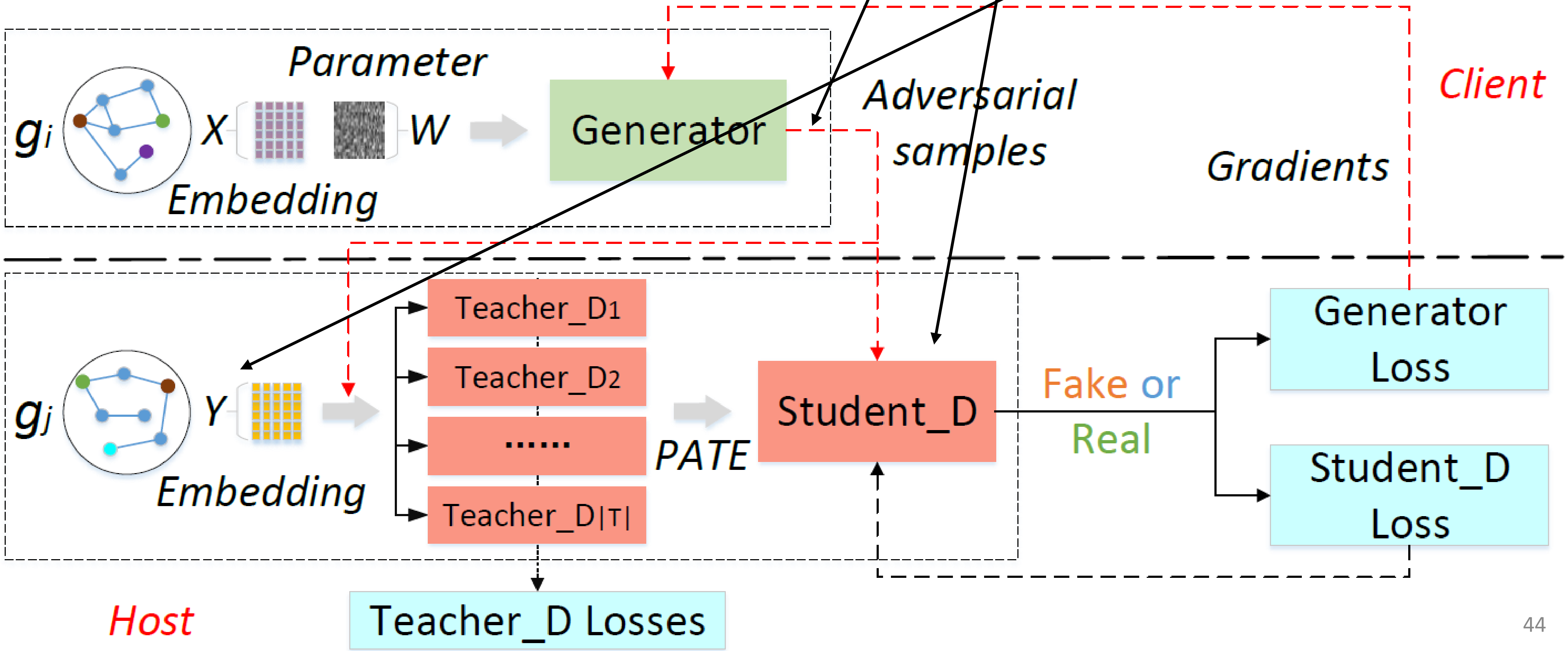
PPAT Network

The generator G is a translational mapping matrix:
$\theta_G = W$



PPAT Network

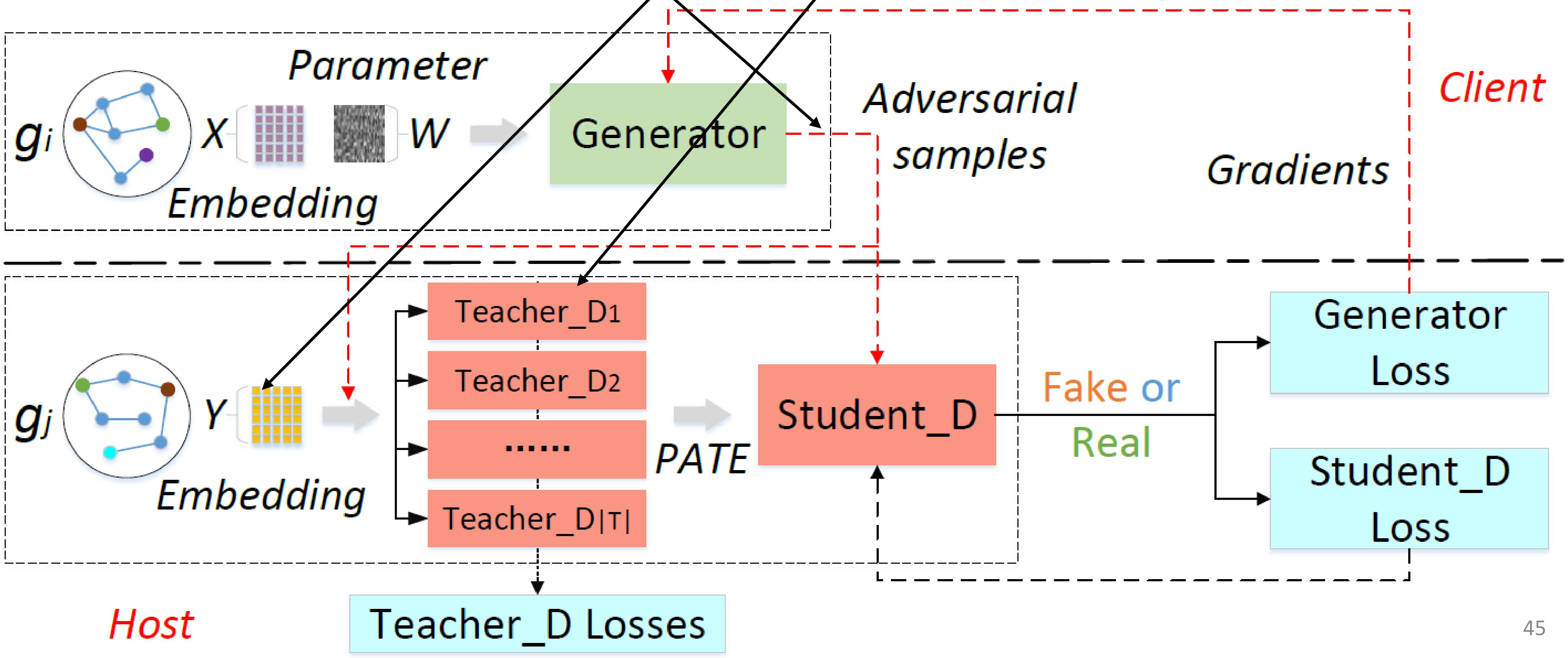
The objective of the generator G is to generate adversarial samples by making $G(X)=WX$ and Y similar so that the student discriminator S cannot distinguish them



PPAT Network

The learning objective of teacher discriminators is the same as the original discriminator that distinguishes between fake samples $G(X)$ and real samples Y, trained on disjointly partitioned data

$$L_T^i(\theta_T^i; G) = -[\sum_{m=1}^{n} \log(1 - T_i(\boxed{G(x_m)}; \theta_T^i)) + \sum_{y_k \in D_i} \log(T_i(\boxed{y_k}; \theta_T^i))]$$

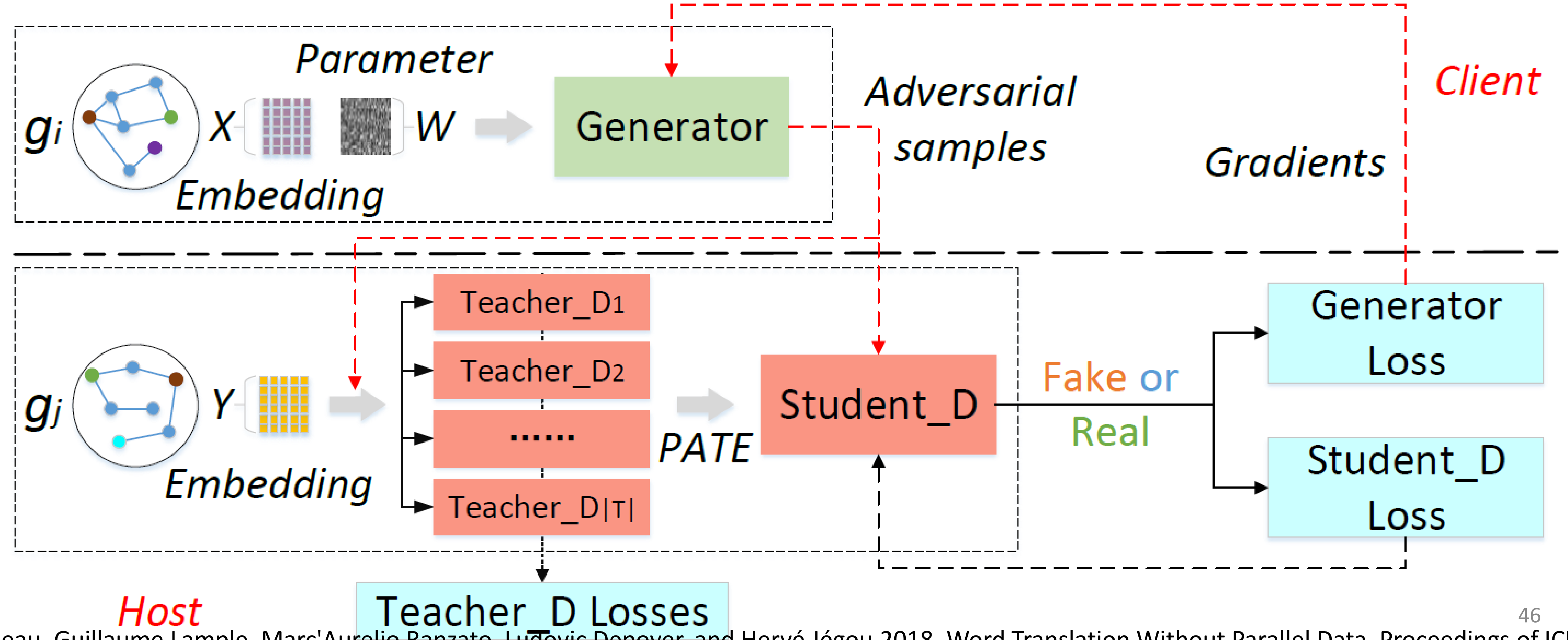


PPAT Network

The discriminator is parameterized by θ_S, which takes embeddings of both $G(X)$ and Y as an input under the CSLS metric used by MUSE

$$\mathrm{CSLS}(Wx_s, y_t) = 2\cos(Wx_s, y_t) - r_{\mathrm{T}}(Wx_s) - r_{\mathrm{S}}(y_t)$$

$$r_{\mathrm{T}}(Wx_s) = \frac{1}{K} \sum_{y_t \in \mathcal{N}_{\mathrm{T}}(Wx_s)} \cos(Wx_s, y_t)$$

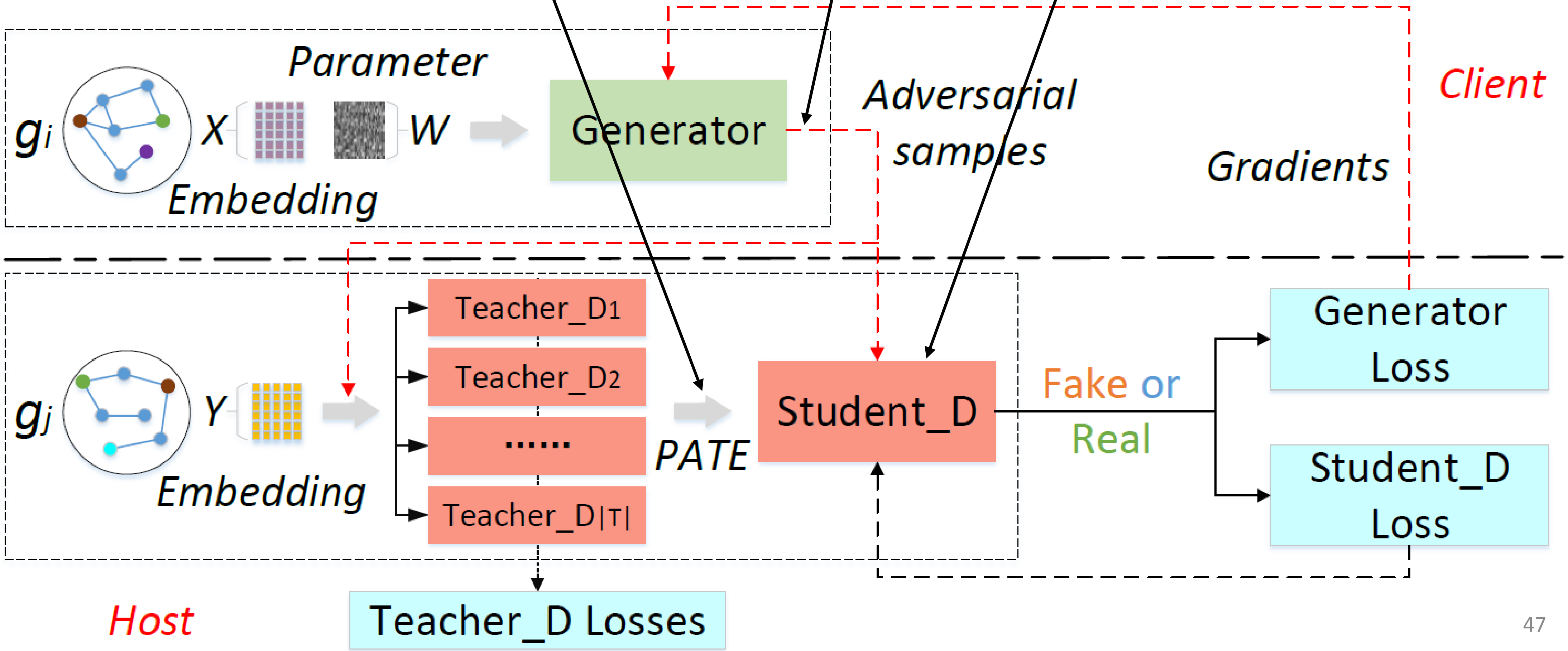


Alexis Conneau, Guillaume Lample, Marc'Aurelio Ranzato, Ludovic Denoyer, and Hervé Jégou.2018. Word Translation Without Parallel Data. Proceedings of ICLR.

PPAT Network

The learning objective of the student discriminator S is to classify generated samples given aggregated PATE noisy labels

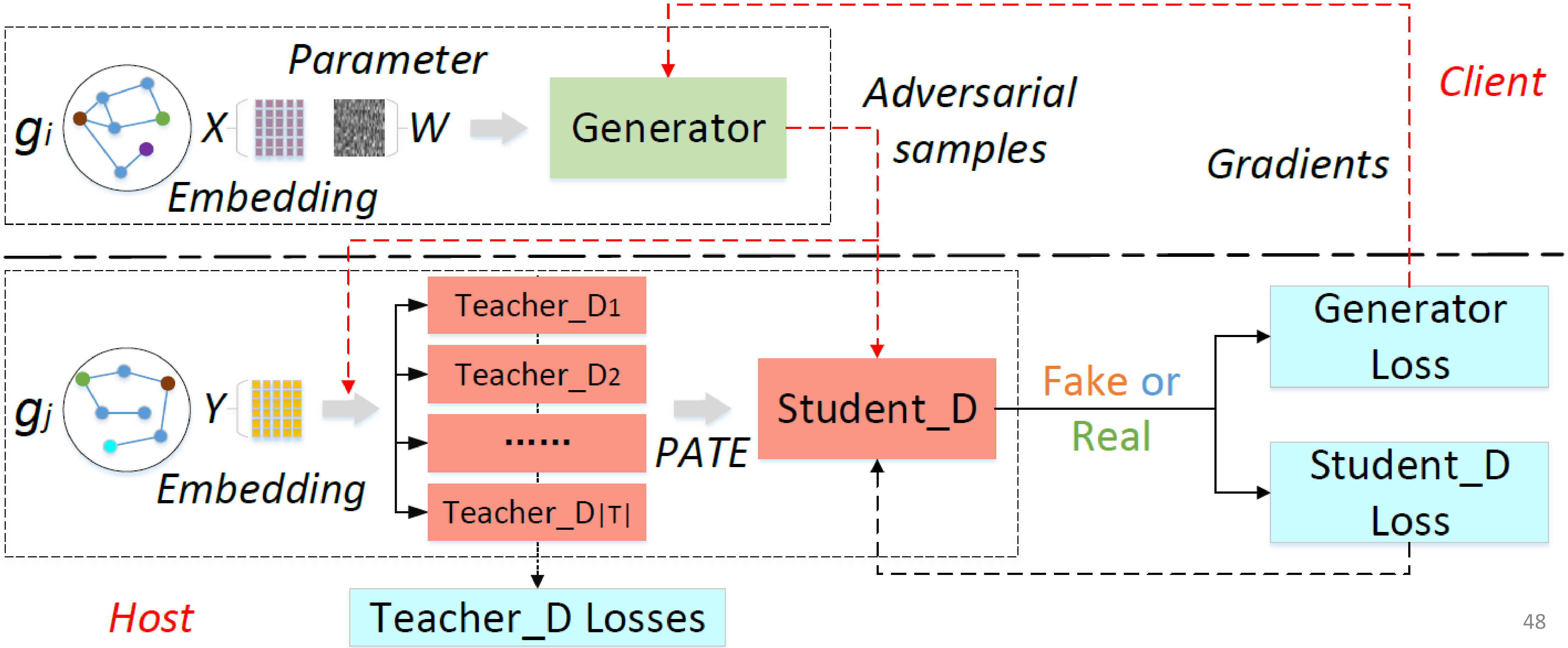
$$L_S(\theta_S; T, G) = \frac{1}{n}\sum_{i=1}^{n}[\gamma_i \log S(G(x_i); \theta_S) + (1-\gamma_i)\log(1-S(G(x_i); \theta_S))]$$



PPAT Network

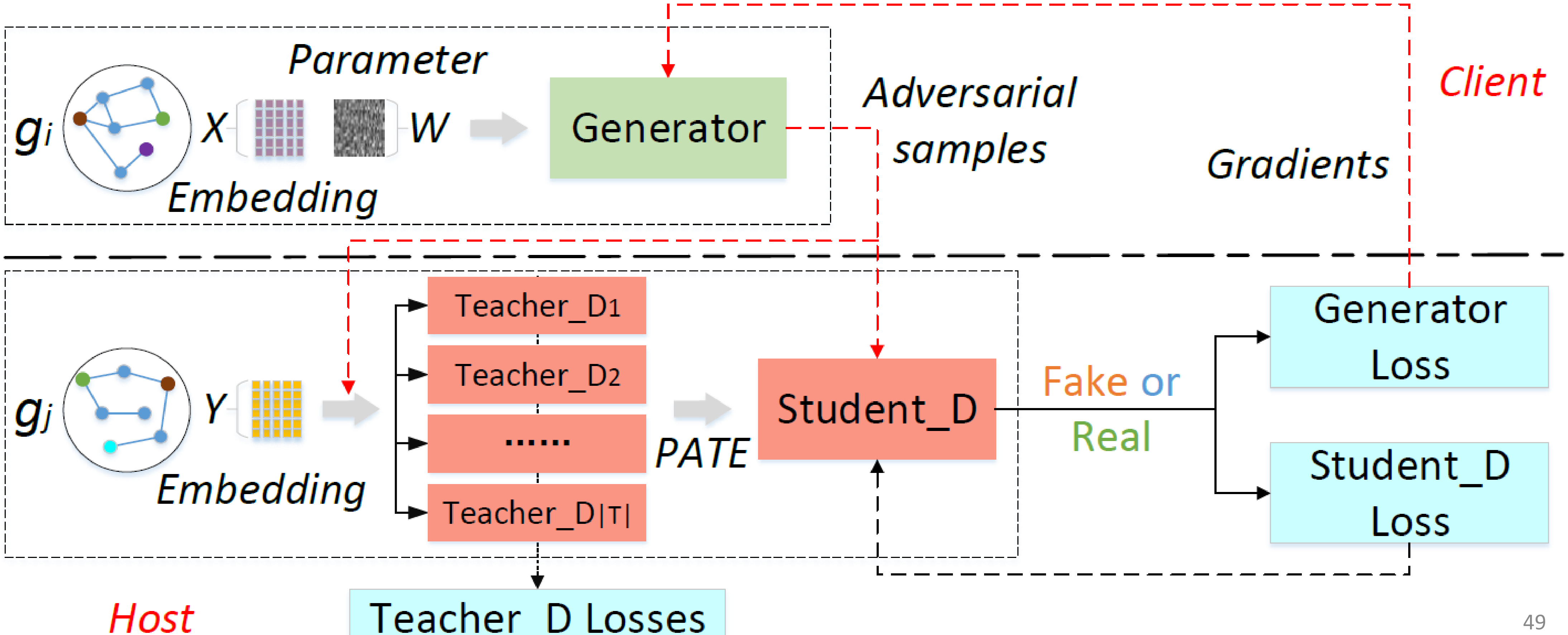
For student discriminator S, no data is publicly available. The training is solely based the generated samples: uniformly generated using Xavier initialization

$$L_S(\theta_S; T, G) = \frac{1}{n}\sum_{i=1}^{n}[\boxed{y_i}\log S(\boxed{G(x_i)};\theta_S)+\boxed{(1-y_i)}\log(1-S(\boxed{G(x_i)};\theta_S))]$$



PPAT Network

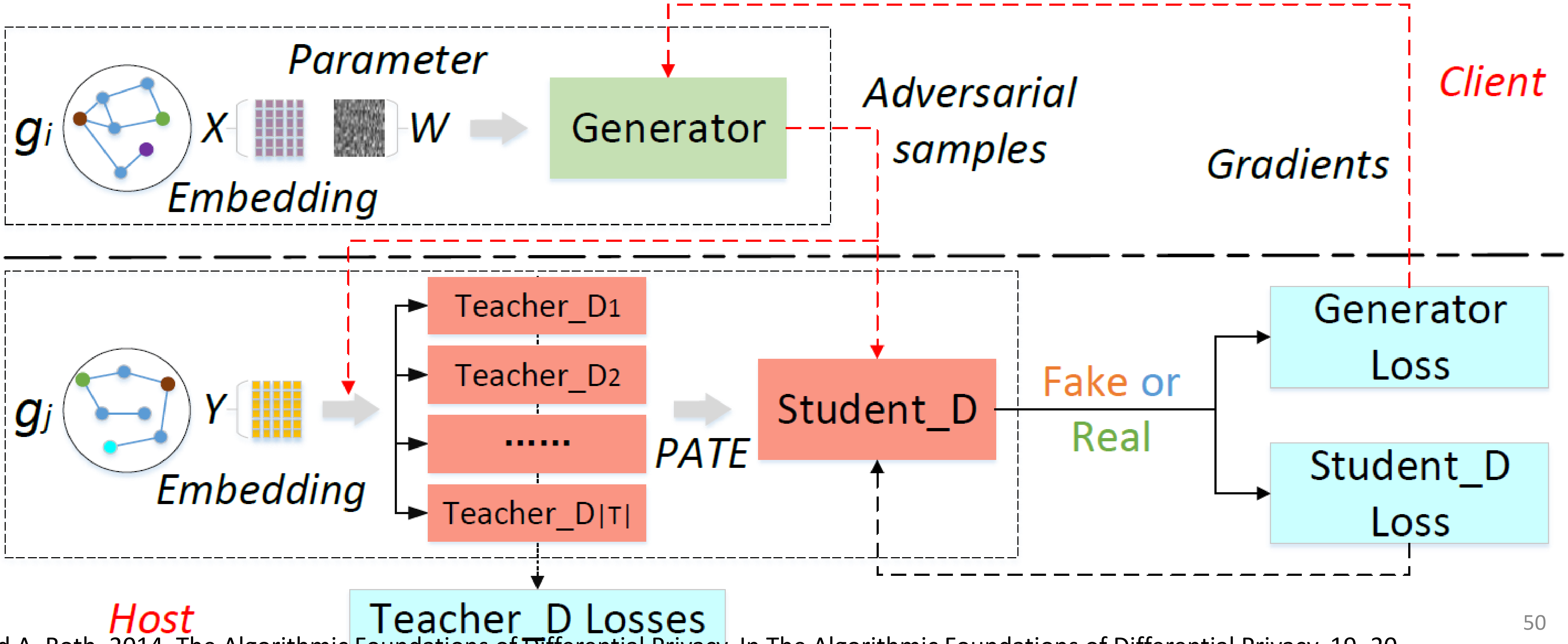
By the Post-Processing Theorem, the student discriminator S is differentially private since it is trained by differentially private labels.
The generator G is differentially private since G is trained by student discriminator S.



C. Dwork and A. Roth. 2014. The Algorithmic Foundations of Differential Privacy. In The Algorithmic Foundations of Differential Privacy. 19–20.

PPAT Network

The host calculates the generator's and all discriminators' loss functions locally; Gradients of generator loss are sent back to the generator to update its parameters.



C. Dwork and A. Roth. 2014. The Algorithmic Foundations of Differential Privacy. In The Algorithmic Foundations of Differential Privacy. 19–20.

Privacy Budget

- Smaller privacy budgets correspond to stronger privacy guarantees
- Similar to PATE and PATE-GAN, In practice, the privacy budget primarily depends on the how much noise is added and consensus between teachers

$$P[M(D) \in S] \leq e^{\epsilon} P[M(D') \in S] + \delta.$$

l: the new parameter introduced by moments accountant method for iterating DP bound based on α

$$\hat{\epsilon} = \min_{l} \frac{\alpha(l) + \log(\frac{1}{\delta})}{l}$$

$$\alpha(l) = \alpha(l) + \min\left\{2\lambda^2 l(l+1), \log\left((1-q)\left(\frac{1-q}{1-e^{2\lambda}q}\right)^l + qe^{2\lambda l}\right)\right\}$$

How much noise is added

Consensus between teachers

$$q = \frac{2 + \lambda|n_0 - n_1|}{4\exp(\lambda|n_0 - n_1|)}$$

larger amounts of noise→smaller privacy budget (smaller lambda: larger scale parameter $\sigma^2 = 2b^2$ where $b = 1/\lambda$)

Higher consensus→smaller privacy budgets

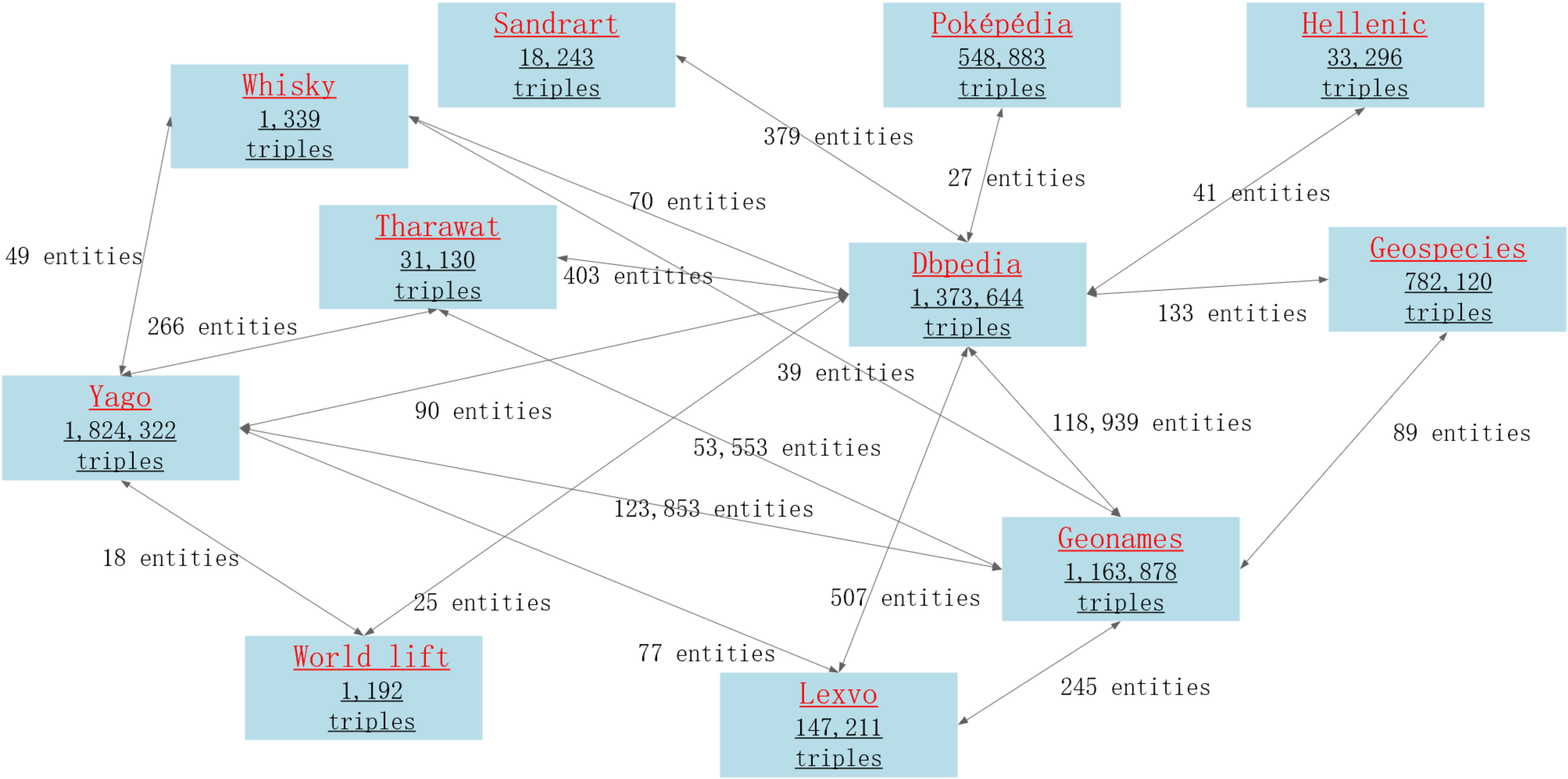
Experiments

- 11 KGs at different scales from the Linked Data community
- In total, there are more than 1-million nodes and 5-million edges
- Train:dev:test=90:5:5

KGs	#Relation	#Entity	#Triple
Dbpedia	14,085	49,1078	1,373,644
Geonames	6	300,000	1,163,878
Yago	37	286,389	1,824,322
Geospecies	38	41,943	782,120
Poképédia	28	238,008	548,883
Sandrart	20	14,765	18,243
Hellenic	4	11,145	33,296
Lexvo	6	9,810	147,211
Tharawat	12	4,693	31,130
Whisky	11	642	1,339
World lift	10	357	1,192
Summation	14,257	1,398,830	5,915,596

Experiments

- Number of AEs (Aligned Entities): Ranging from tens to >100K



Privacy Setting

According to PATE, (ϵ, δ)=(2, 10E−5) satisfies normal privacy budget while (ϵ, δ)=(8, 10E−5) is a relatively looser bound.

- λ = 0.05
- δ = 10E−5
- α (l) = 0.29
- δ = 1/11.5
- l = 9
- ϵ = 2.73

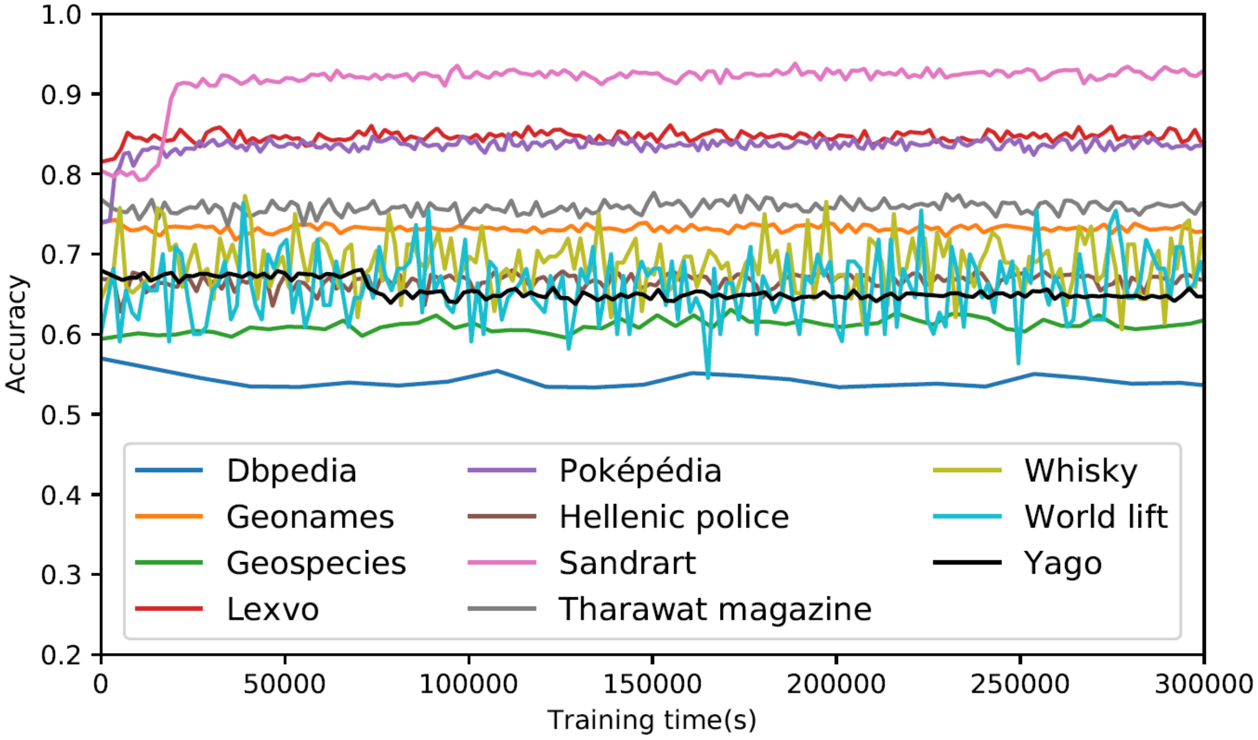
$$\hat{\epsilon} = \min_{l} \frac{\alpha(l) + \log(\frac{1}{\delta})}{l}$$

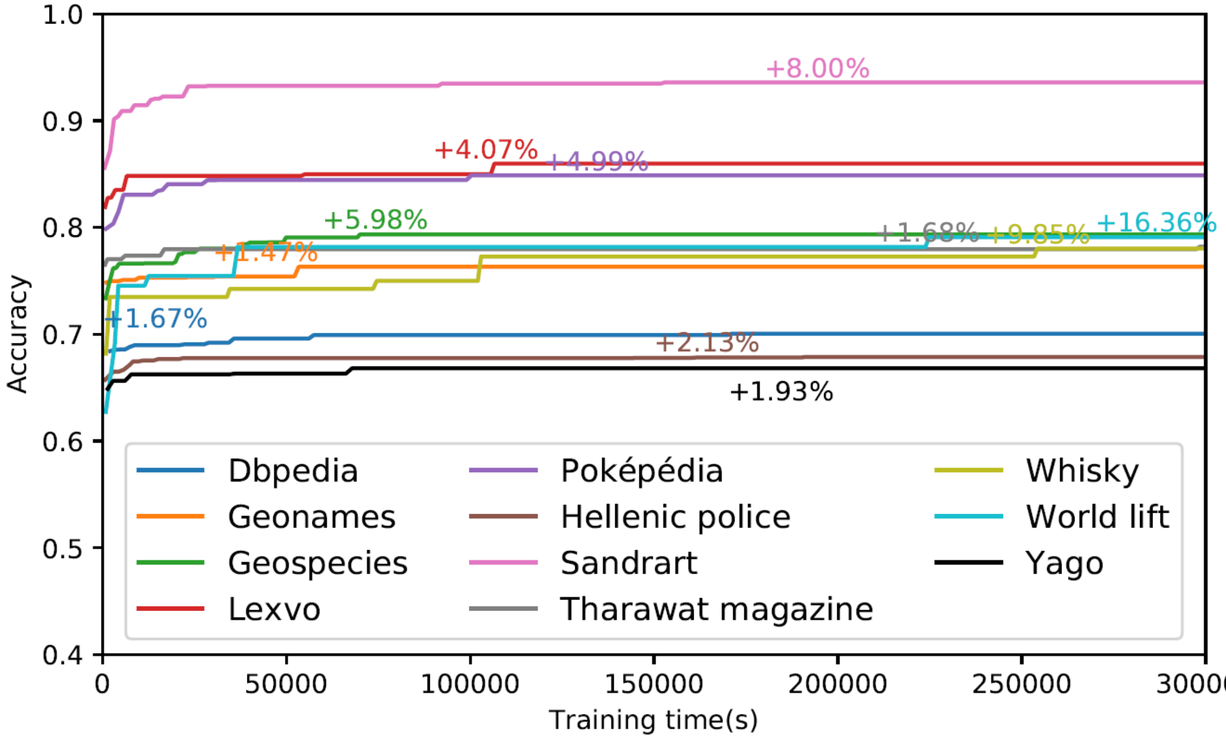
$$\alpha(l) = \alpha(l) + \min\left\{2\lambda^2 l(l+1), \log\left((1-q)\left(\frac{1-q}{1-e^{2\lambda}q}\right)^l + qe^{2\lambda l}\right)\right\}$$

$$q = \frac{2 + \lambda|n_0 - n_1|}{4\exp(\lambda|n_0 - n_1|)}$$

Performance on Triple Classification

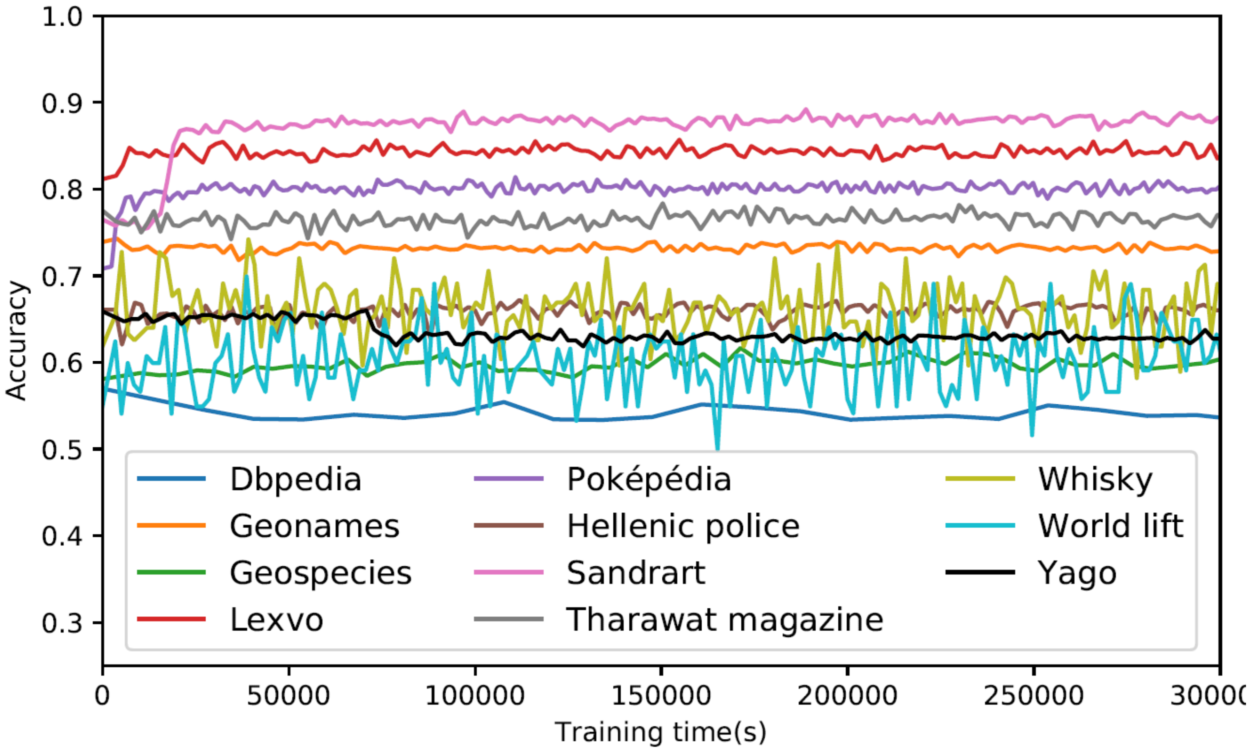
- Comparison based on TransE
- All KGs are improved ranging from 1.47% to 16.36%

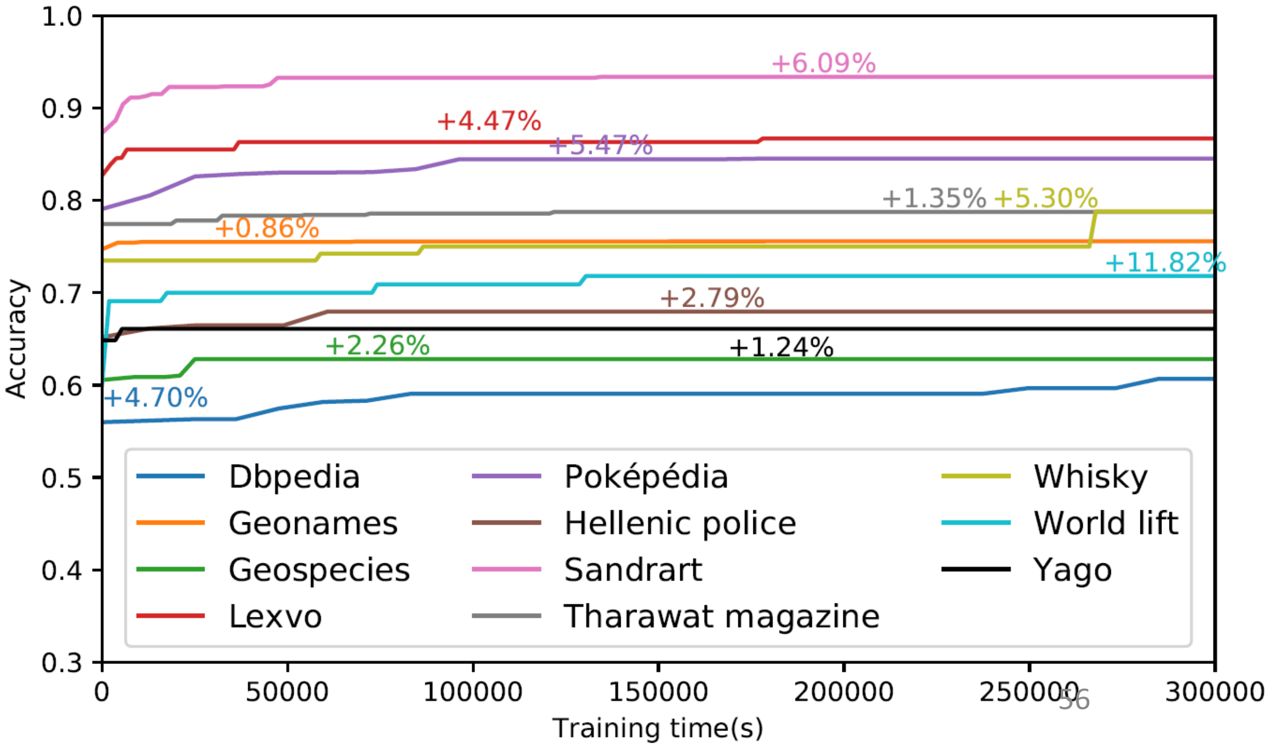




Performance on Triple Classification

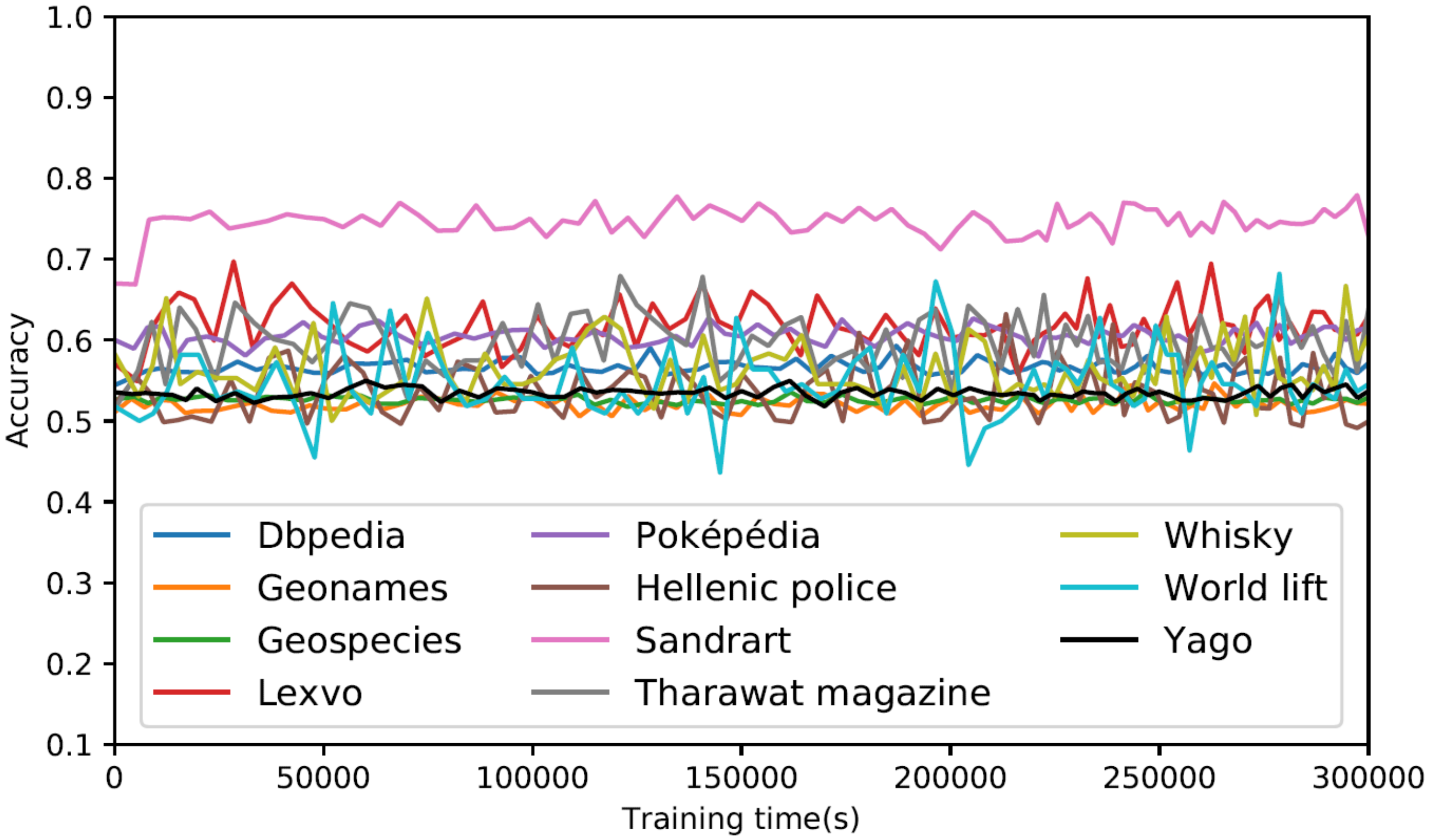
- Comparison based on different embedding methods:
  - Dbpedia (TransR), Geonames (TransD), Yago (TransE), Geospecies (TransR), Poképédia (TransE), Sandrart (TransD), Hellenic (TransD), Lexvo (TransD), Tharawat (TransD), Whisky (TransH), and World lift (TransR)
- All KGs are improved ranging from 0.86% to 11.82%





Performance on Triple Classification

- TransE trained based on a unified KG:



Performance on Link Prediction

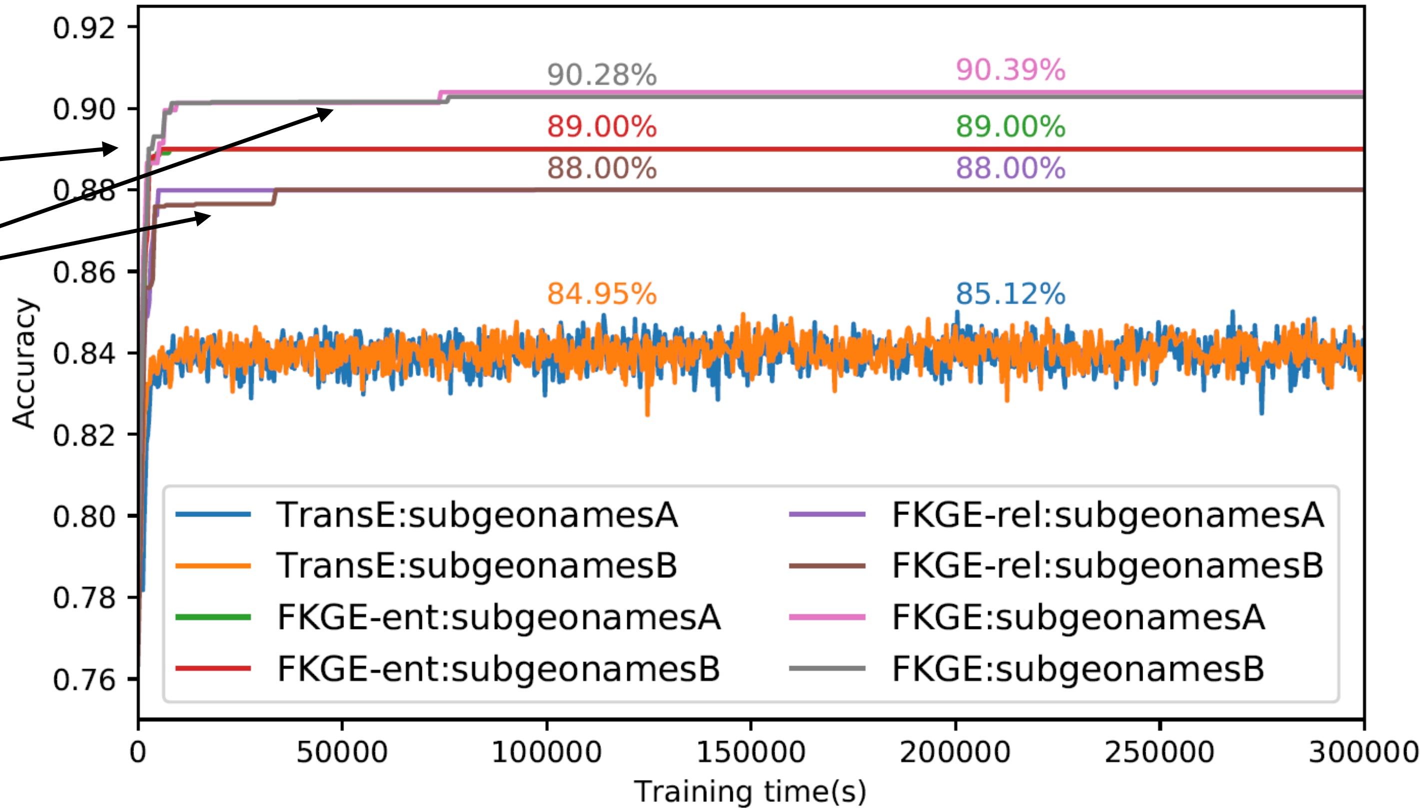
- We observe similar improvements on same settings

Methods	Independent-TransE			FKGE			Random-Independent-KGE			Multi-FKGE		
Metric	Hit@10	Hit@3	Hit@1	Hit@10	Hit@3	Hit@1	Hit@10	Hit@3	Hit@1	Hit@10	Hit@3	Hit@1
Dbpedia	23.29	12.88	5.12	**25.07**	**14.41**	**6.37**	5.46	2.51	1.10	**6.67**	**3.20**	**1.24**
Geonames	8.82	3.69	1.93	**9.65**	**4.88**	**2.12**	8.45	4.53	1.90	**8.85**	**4.97**	**2.14**
Yago	2.05	0.76	0.25	**2.59**	**0.88**	**0.29**	2.03	**0.75**	**0.24**	**2.36**	**0.75**	**0.24**
Geospecies	58.49	45.81	34.01	**60.97**	**46.95**	**35.03**	38.68	26.43	13.12	**40.92**	**28.04**	**14.38**
Poképédia	38.14	29.04	19.31	**45.58**	**35.48**	**24.90**	34.22	25.13	16.43	**42.12**	**32.14**	**22.65**
Sandrart	87.39	83.16	67.18	**88.65**	**84.97**	**72.14**	87.71	83.71	68.91	**87.99**	**84.22**	**69.69**
Hellenic	32.18	21.87	18.96	**33.00**	**22.87**	**19.35**	32.21	22.23	18.59	**32.82**	**22.59**	**19.44**
Lexvo	85.67	76.07	58.29	**87.35**	**77.74**	**62.90**	84.21	75.82	58.09	**85.72**	**76.99**	**59.76**
Tharawat	12.48	4.56	1.67	**13.45**	**5.26**	**2.19**	12.30	4.38	1.39	**12.55**	**5.21**	**1.77**
Whisky	28.78	15.15	9.84	**35.60**	**18.93**	**10.60**	28.78	18.93	12.87	**30.12**	**19.45**	**12.92**
World lift	45.76	24.57	7.62	**51.69**	**28.88**	**11.17**	18.64	8.47	1.69	**18.85**	**9.32**	**2.54**

Effects of Aligned Entities and Relations

- Geonames KG is split into two subsets
  - FKGE-ent: only align entities
  - FKGE-rel: only align relations
  - FKGE: align both

- Evaluated on triple classification



Noise Scales

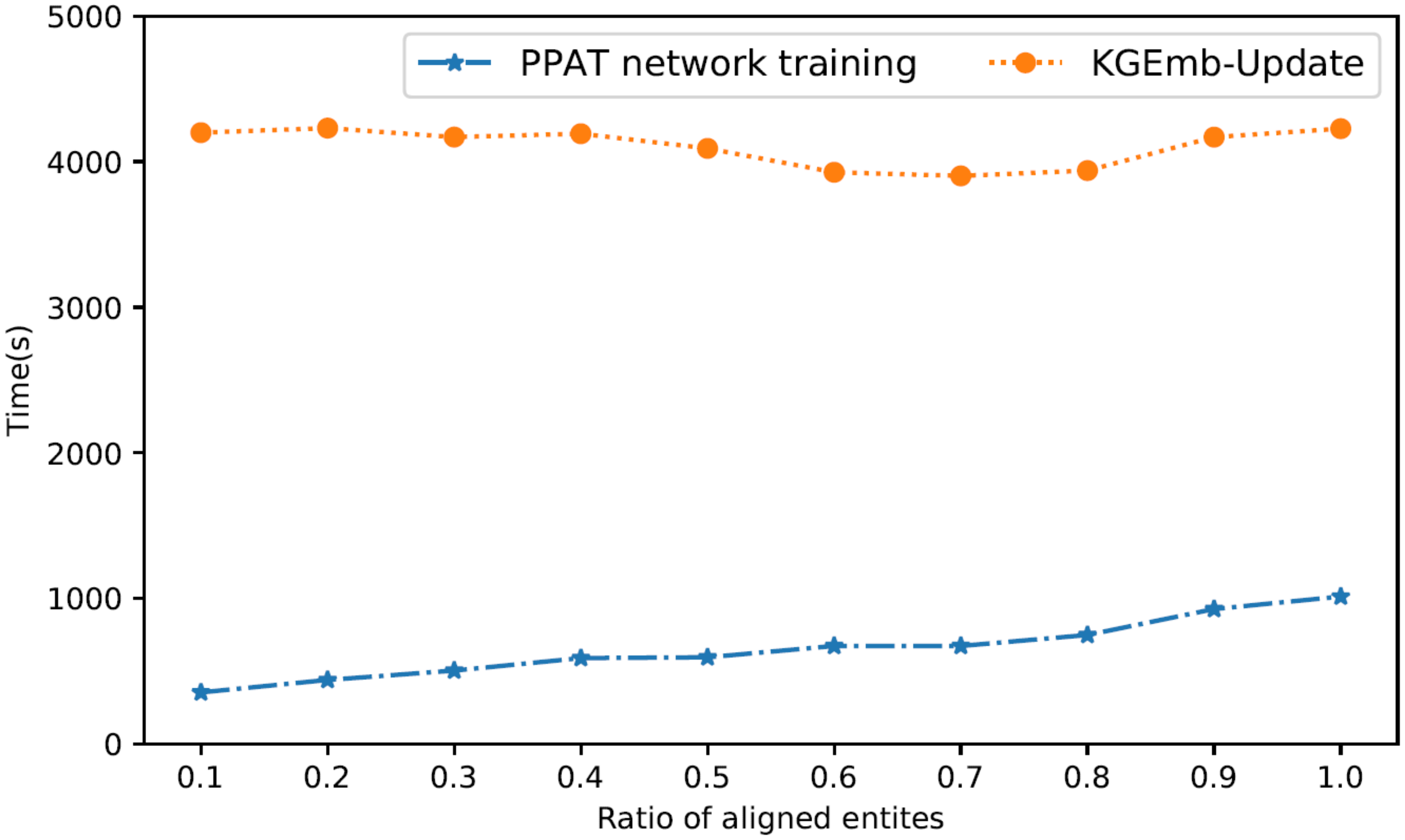
- Variance of Laplace distribution: $\sigma^2 = 2b^2$ where $b = 1/\lambda$
  - Larger λ means smaller variance (add less noise to teachers) and larger privacy budget

- All accuracies are similar with no difference greater than 0.6%
- PPAT network tends to be robust by introducing acceptable randomness

Noise λ	No noise	0.05	1	2	5
Dbpedia	68.51%	67.94%	68.29%	68.54%	68.11%
Geonames	74.29%	74.21%	74.28%	74.34%	74.02%

Execution Time

- KGEmb-Update usually costs much more time than PPAT network
- The cost for PPAT training increases roughly linearly from 350s to 1,000s as number of aligned entities increases
- With batch size = 32, d = 100, and 64 bit for double precision, total communication cost for a batch training of the PPAT network is at most 0.845 Mb



Conclusions

- We proposed a new differentially private knowledge graph embedding framework FKGE:
  - Asynchronous and decentralized
  - Scalable and compatible with many base embedding models
  - Privacy-preserving and guaranteeing no raw data leakage

- Code is available at: https://github.com/HKUST-KnowComp/FKGE

Thank You! ☺