Decentralised Graph Neural Networks: Adaptation, Training and Interpretability in Federated Environments

S M Asiful Huda

DC9 — TUAI Doctoral Network
University of Naples Federico II

Main Supervisor: Francesco Piccialli (UNINA)
Co-Supervisors: David Camacho (UPM), Shen Yin (NTNU), Dariusz Mrozek (SUT)

July 23, 2026

DC9 — Scope and Motivation

Objective

Advance GNNs within federated and decentralised learning frameworks: training across decentralised nodes without central data aggregation, with a major focus on explainability.

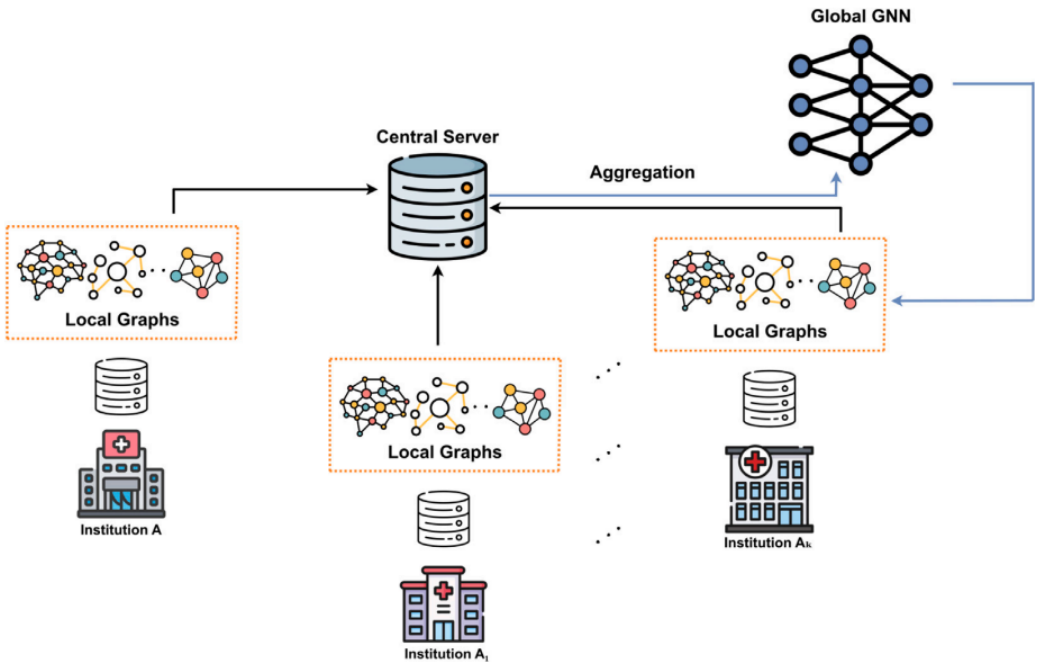


Figure: Decentralized graph data across multiple clients.

Outline

1. Motivation — decentralised graphs and the right to erasure
2. **Track A** — GFGU: isolating client influence for unlearning
3. **Track B** — FedConcept: structural concepts for federated foundation models
4. Research contributions
5. Evolution of scope and next steps

Track A: Machine Unlearning in Federated Graphs: Motivation

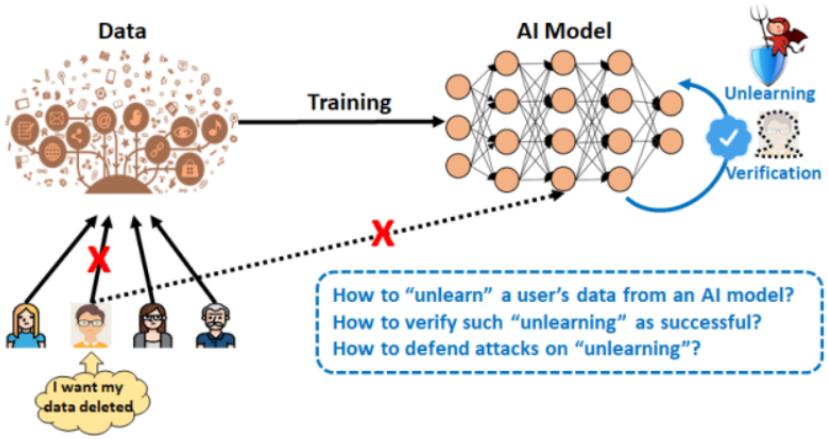


Figure: Illustration of a machine unlearning request.

- **EU GDPR (Article 17):** Grants individuals the *Right to Erasure* ("Right to be Forgotten"), allowing them to request the removal of their personal data.
- **Machine Unlearning:** Removes the influence of specific training samples from a trained model without retraining the entire model.

Can we isolate influence of the training data in Federated Graphs?

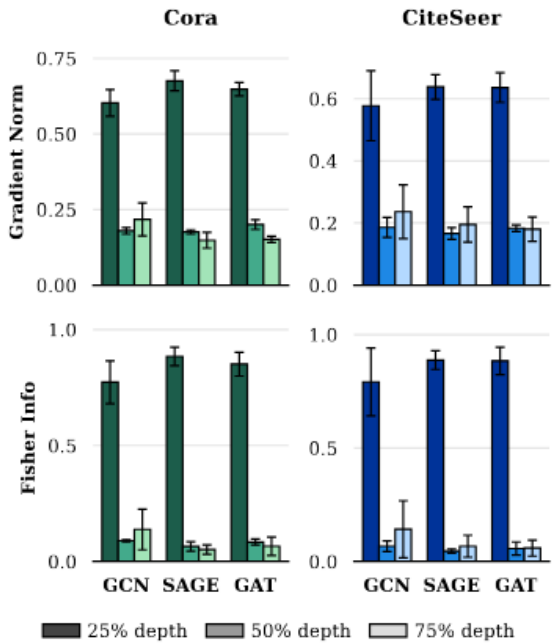


Figure: Early layers consistently exhibit higher sensitivity to client-level structural dependencies than deeper representations.

GFGU — Gated Federated Graph Unlearning Framework

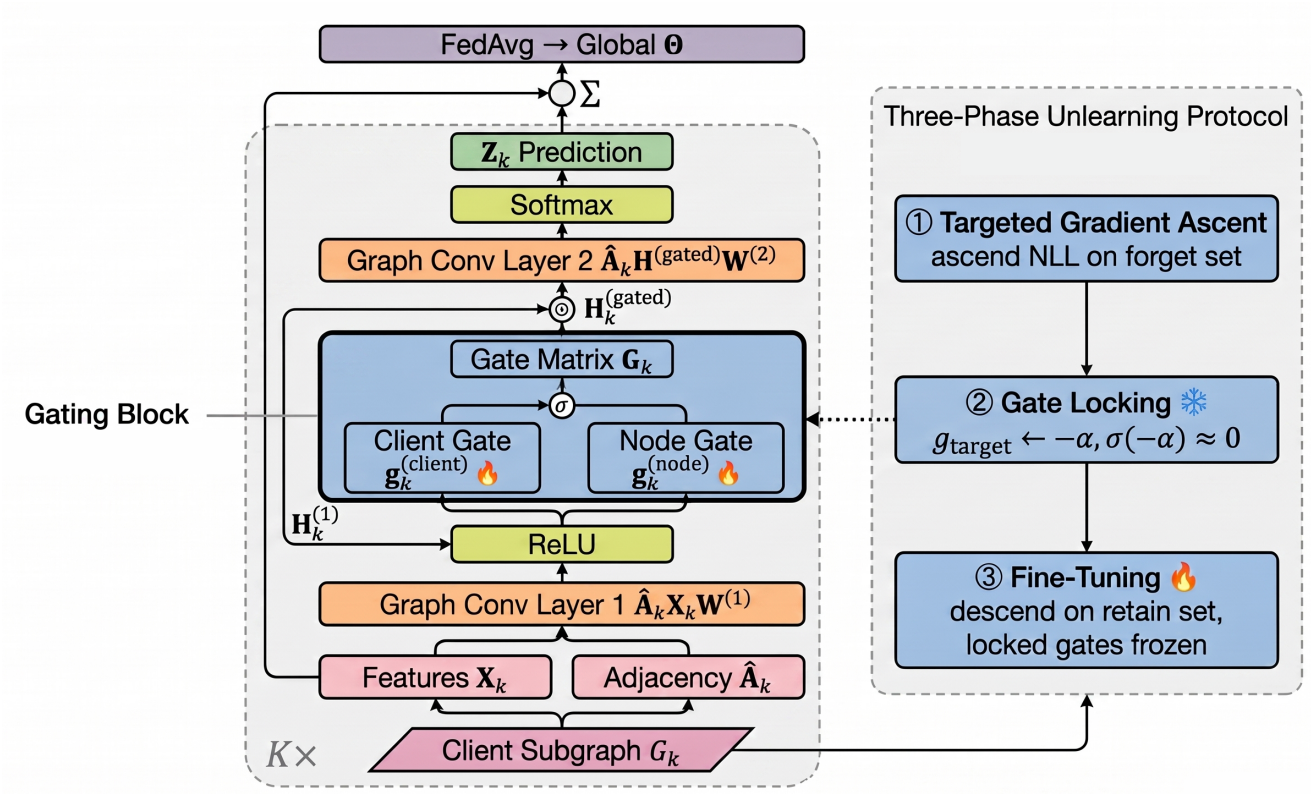


Figure: Overview of the GFGU Framework.

A learnable **multiplicative gate** sits after the first graph convolutional layer, at client and node granularity.

Benchmark Datasets

- **Citation Networks**
  - **Cora**, **PubMed**
  - Node classification on scientific publication graphs.
- **Co-authorship Network**
  - **CS**
  - Collaboration graph of computer science authors.
- **E-commerce Network**
  - **Photo**
  - Amazon product co-purchase graph for node classification.
- **Crowdsourcing Graphs**
  - **Tolokers**, **Minesweeper**
  - Graphs derived from crowdsourcing platforms for worker/task prediction.

Evaluation Metrics for Machine Unlearning

- **Node Classification Accuracy**
  - Measures the classification performance of the model on graph node prediction after unlearning.
- **Test Accuracy**
  - Evaluates the overall predictive performance on the test dataset after unlearning.
- **Forgetting Score (JSD)**
  - Divergence between the *unlearned model* and the *fully retrained model*.
  - Lower D indicates that the unlearned model closely approximates complete retraining.
- **Membership Inference Attack (MIA AUC)**
  - Measures the vulnerability of the model to membership inference attacks.
  - An AUC close to 0.5 indicates stronger privacy protection and more effective unlearning.

Track A — Predictive Performance

Node classification accuracy, ACC ± STD (%). Best per column in **bold**.

	5 Clients		10 Clients		15 Clients		20 Clients
	Cora	PubMed	CS	Photo	Tolokers	Minesweeper	Amazon-ratings
Client Unlearning (Unlearn Ratio = 0.2)							
Retrain	83.98 ± 1.52	83.90 ± 0.82	82.18 ± 0.43	78.41 ± 1.37	78.52 ± 1.51	79.88 ± 0.11	**41.09 ± 0.48**
FedEraser	83.85 ± 2.37	78.52 ± 3.21	78.22 ± 1.46	54.66 ± 4.72	78.22 ± 1.61	79.81 ± 0.26	36.89 ± 0.40
ReGeNUnlearn	83.83 ± 1.29	83.74 ± 0.95	78.27 ± 1.92	58.43 ± 9.78	78.54 ± 1.51	80.04 ± 0.21	40.20 ± 1.04
FUSED	49.43 ± 2.78	83.17 ± 0.40	72.68 ± 4.17	67.60 ± 5.34	78.47 ± 1.54	79.86 ± 0.10	39.08 ± 0.48
MoDe	80.33 ± 4.04	83.20 ± 1.13	73.90 ± 3.42	42.47 ± 8.23	78.47 ± 1.54	79.86 ± 0.10	36.63 ± 4.19
PAGE	83.91 ± 1.29	83.57 ± 0.54	78.66 ± 1.06	69.73 ± 2.16	78.52 ± 1.49	**80.06 ± 0.04**	40.75 ± 0.58
GFGU (Ours)	**84.23 ± 1.39**	**87.14 ± 0.75**	**88.03 ± 1.06**	**91.52 ± 0.92**	**79.82 ± 1.44**	79.81 ± 0.40	37.21 ± 0.79
Node Unlearning (Unlearn Ratio = 0.1)							
Retrain	78.65 ± 0.04	83.91 ± 0.17	87.74 ± 0.09	88.62 ± 0.13	78.52 ± 0.18	79.90 ± 0.10	**46.58 ± 0.06**
GIF	81.02 ± 0.38	84.21 ± 0.13	87.67 ± 0.05	80.44 ± 0.95	78.12 ± 0.08	79.69 ± 0.33	45.62 ± 0.13
D2DGN	75.83 ± 0.61	83.23 ± 0.15	86.98 ± 0.33	79.69 ± 3.53	77.49 ± 0.26	76.14 ± 1.16	45.74 ± 0.20
MEGU	77.96 ± 3.80	83.43 ± 0.05	86.94 ± 0.32	82.03 ± 3.03	78.34 ± 0.04	79.87 ± 0.00	42.63 ± 0.28
FedLU	71.96 ± 0.43	73.14 ± 0.17	80.44 ± 0.16	80.21 ± 0.51	78.32 ± 0.04	79.87 ± 0.00	44.11 ± 0.09
FedDM	80.11 ± 0.63	80.89 ± 0.20	80.67 ± 3.54	51.87 ± 2.57	76.82 ± 1.37	79.89 ± 0.08	32.09 ± 1.23
PAGE	81.29 ± 0.17	84.39 ± 0.09	**88.66 ± 0.07**	88.84 ± 0.39	**78.69 ± 0.19**	**79.92 ± 0.10**	46.21 ± 0.13
GFGU (Ours)	**84.34 ± 1.09**	**86.58 ± 0.33**	86.53 ± 1.41	**91.23 ± 0.45**	78.17 ± 0.50	79.87 ± 0.45	38.41 ± 0.62

- Best on 5 of 7 datasets for client unlearning; best on 3 of 7 for node unlearning. Competitive on others.

Track A — Forgetting Quality and Efficiency

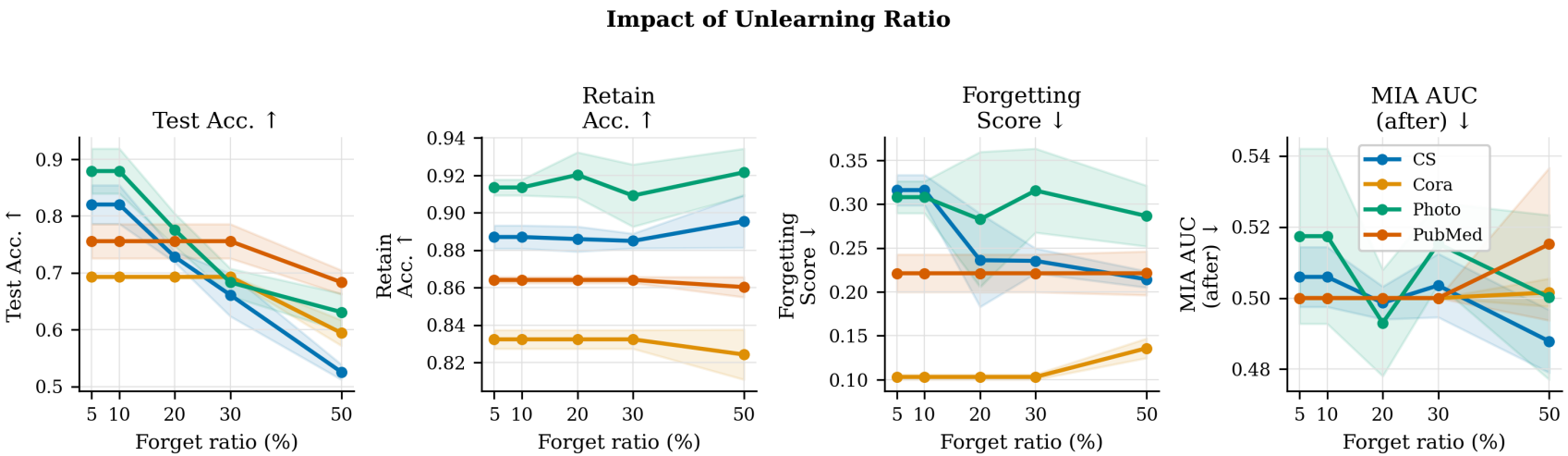


Figure: Early layers consistently exhibit higher sensitivity to client-level structural dependencies than deeper representations.

- **MIA AUC near** 0.5: a membership inference adversary cannot distinguish erased members from non-members.
- **Low Forgetting Score** (JSD against the retrained model): outputs stay close to the gold standard.
- Achieved at a **fraction of the cost** of full retraining.

Graph Foundation Models

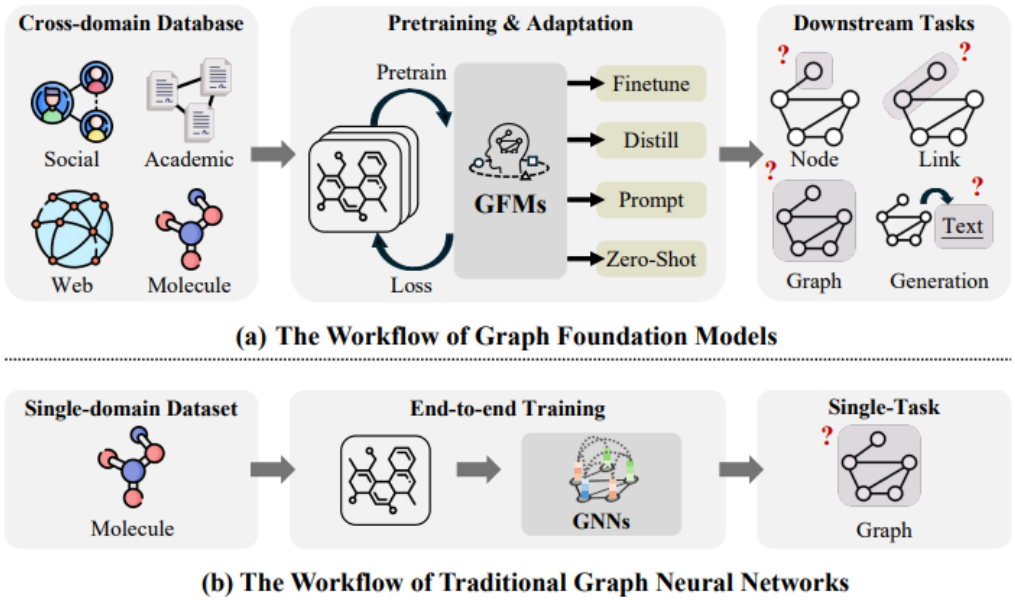


Figure: Illustration of a Graph Foundation Model.

- **Graph Foundation Models (GFMs):** Large pre-trained graph models that learn general-purpose representations from diverse graph data.
- **Pre-training:** Trained on large-scale graph datasets using self-supervised or contrastive learning objectives.
- **Adaptability:** Can be fine-tuned for downstream graph tasks such as node classification, link prediction, and graph classification.

Interpretable Foundation Model via Concept Bottolneck Model

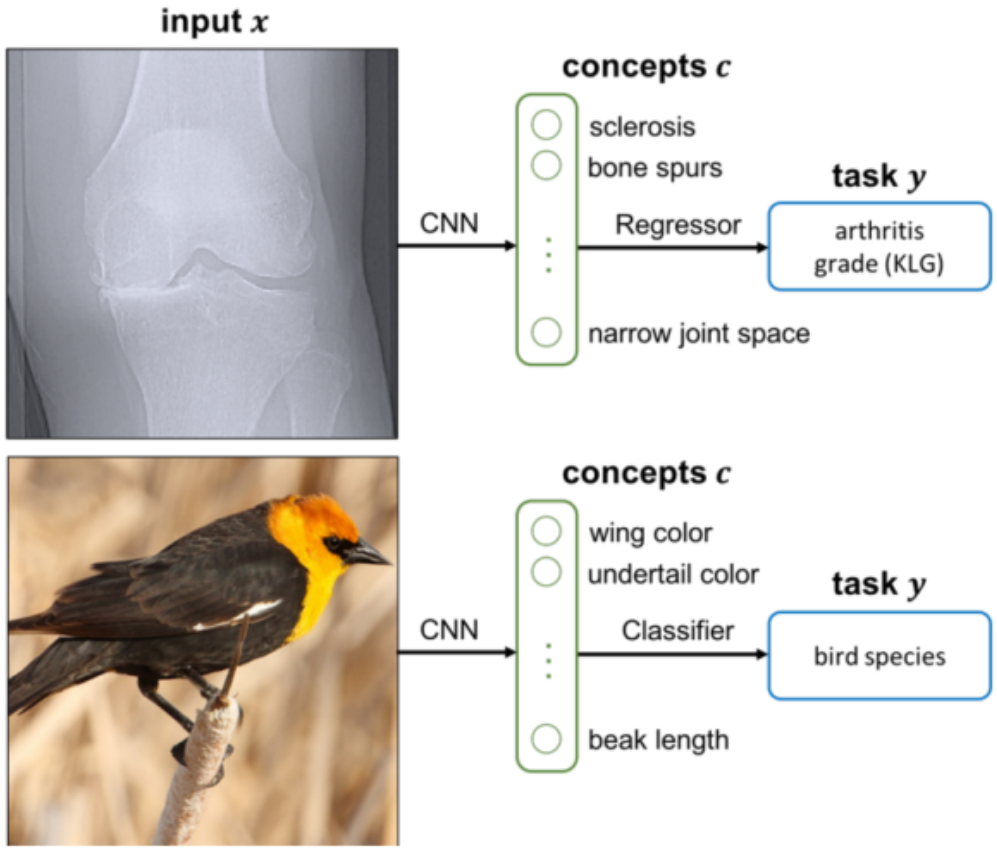


Figure: Main idea of a Concept Bottolneck Model.

Track B — Can we find any structural concepts from Graph data?

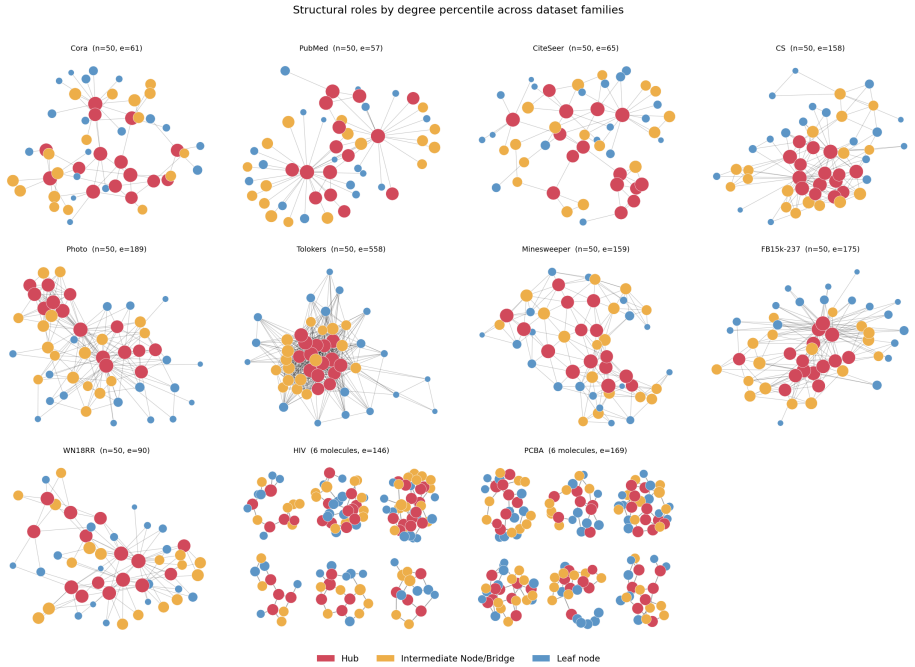


Figure: Visualization of degree percentiles among graphs from different domains.

- Degree percentile captures a node's **relative connectivity position** within its graph domain, rather than its absolute degree value.
- Instead of comparing raw degrees across domains, map every node to its corresponding **degree percentile rank** to obtain a normalized structural measure.

Track B — FedConcept Framework

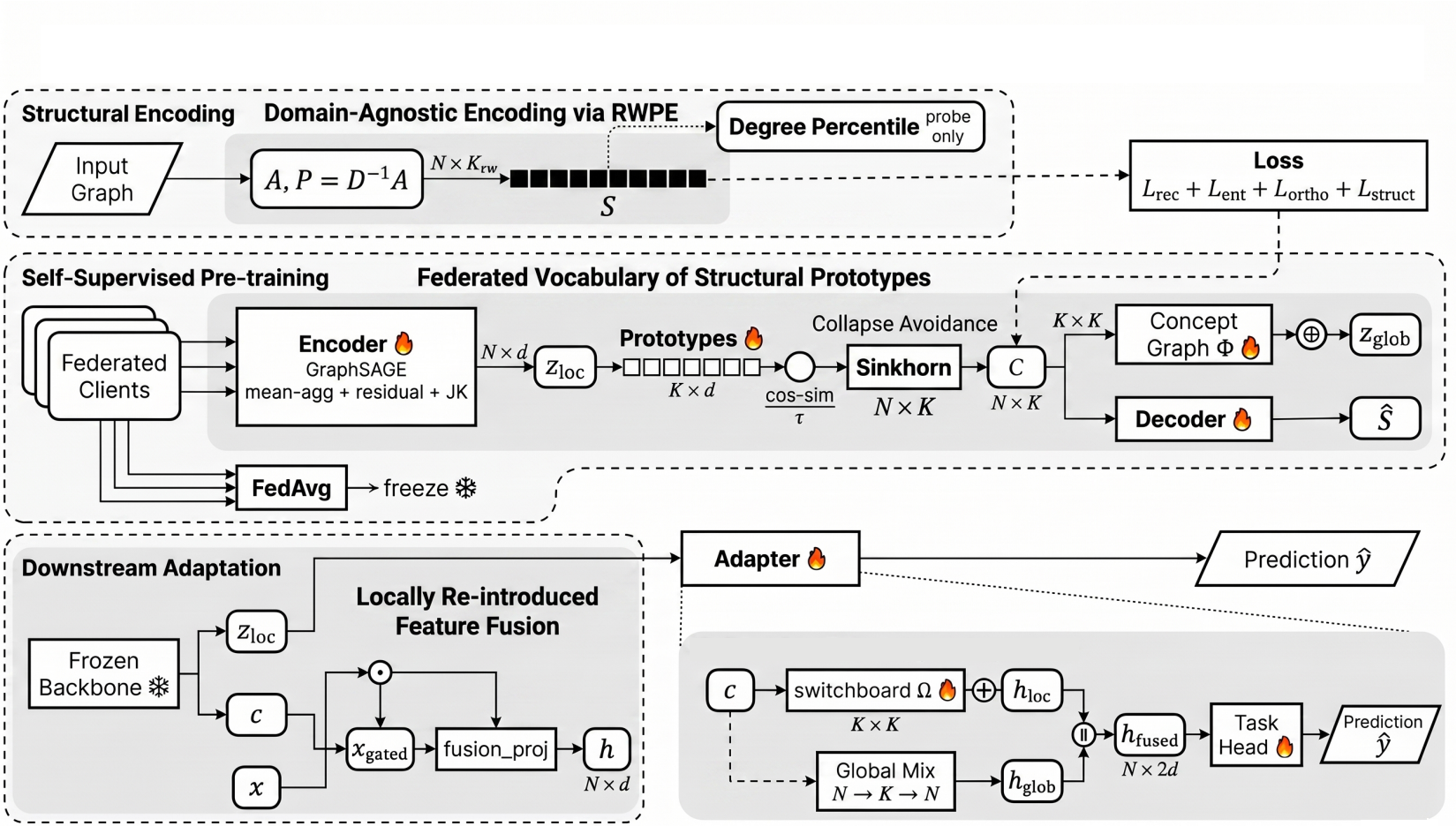
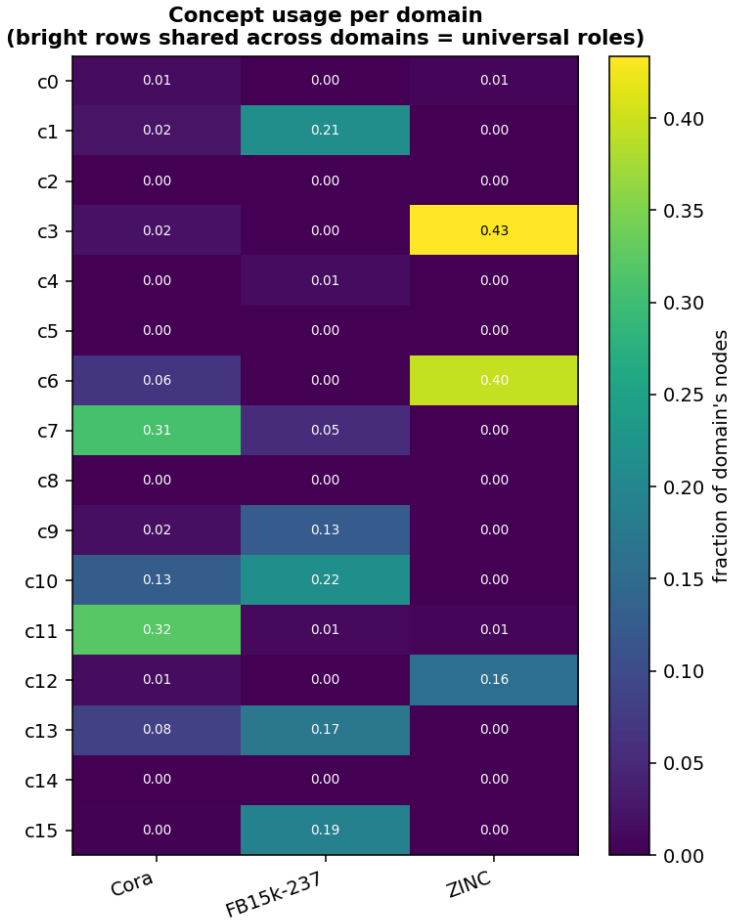


Figure: Framework of FedConcept based on Degree Percentile based Concept Prototype.

Concept Usage Heatmap



(a) Concept Usage Heatmap.

Embedding Visualization of FedConcept

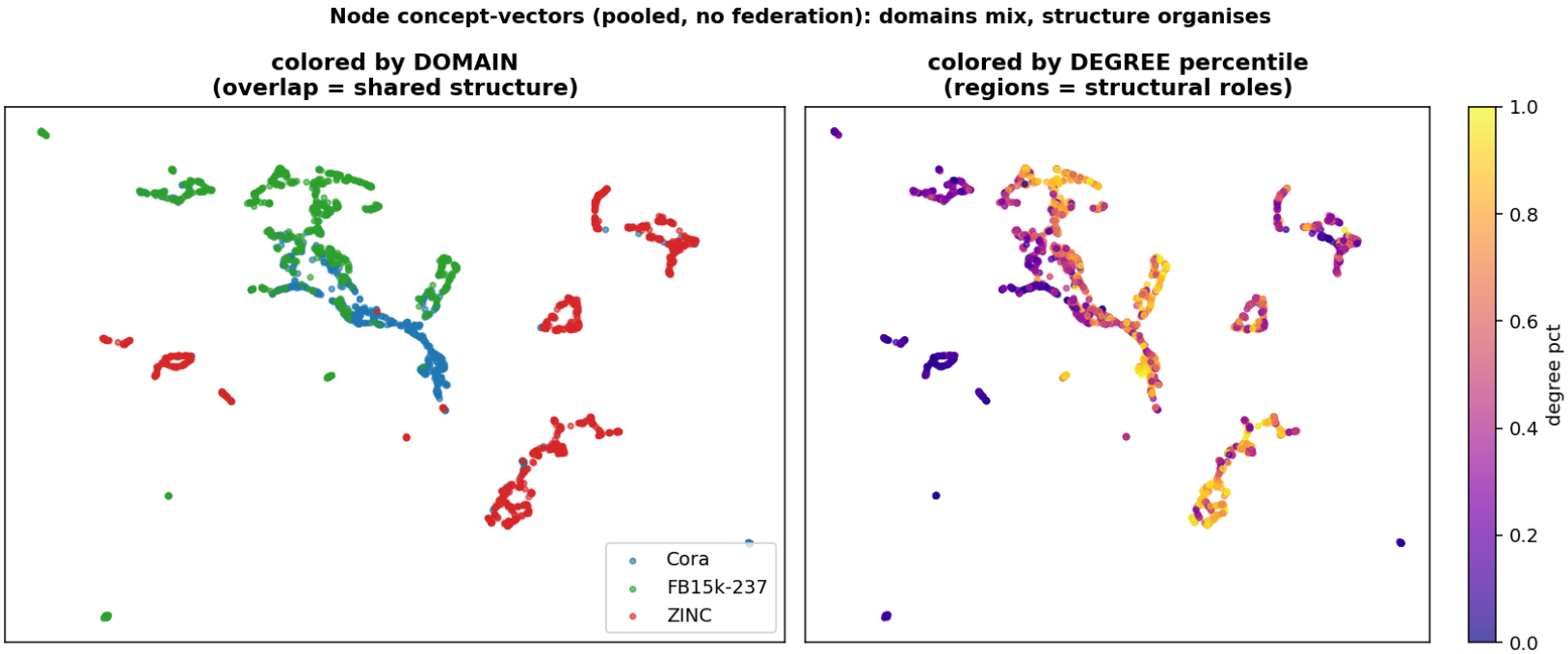


Figure: VIsualization of the FedConcept Embedding in three different graph domains.

Benchmark Datasets

- **Citation Networks**
  - **Cora**, **PubMed**, **OGB-Arxiv**
  - Scientific publication graphs used for node classification.
- **Wikipedia Network**
  - **WikiCS**
  - Hyperlink graph of Wikipedia computer science articles for node classification.
- **Knowledge Graphs**
  - **FB15k-237**, **WN18RR**
  - Entity-relation graphs for knowledge graph completion and link prediction.
- **Molecular Graphs**
  - **HIV**, **PCBA**
  - Molecular property prediction benchmarks for drug discovery and bioinformatics.

Track B — Preliminary Results

Accuracy for node/edge tasks, AUC for molecular graph tasks.

Method	Cora	PubMed	OGB-arxiv	WikiCS	FB15K	WN18RR	HIV	PCBA
GCN	75.22	82.07	70.68	75.54	66.95	79.74	60.30	63.15
GAT	76.13	82.75	68.02	74.40	68.59	81.35	60.12	66.41
GraphSAGE	74.61	83.53	68.31	74.87	69.94	80.33	61.02	66.50
FedAvg	75.76	82.84	70.35	75.31	68.43	80.07	63.95	68.62
MOON	76.43	85.44	71.19	76.37	68.44	80.62	65.96	69.42
FedBook	81.32	**88.19**	**79.87**	76.21	74.33	86.21	70.10	75.41
FedGFM	80.59	87.17	78.69	74.51	73.23	84.31	69.70	74.81
FedConcept	**83.95**	85.61	66.90	**84.84**	**82.14**	**86.70**	**74.13**	**78.79**

Best per column in **teal**.

Track B — 2-Shot Results

Accuracy (%); FedConcept reported as mean over 10 seeds. Best per column in **bold**.

Method	Cora	PubMed	OGB-arxiv	WikiCS	FB15K237	WN18RR
OFA	50.23	46.26	17.11	38.06	19.44	30.42
GFT	54.89	47.93	18.76	40.48	18.99	29.14
UniGraph	49.80	46.47	18.74	37.51	20.34	28.59
GQT	53.11	48.35	21.50	38.13	20.16	30.66
GraphCLIP	51.11	50.31	20.84	38.37	19.74	29.51
FedGFM+	55.24	52.10	21.76	42.27	22.45	32.32
FedBook	**56.34**	55.23	**22.46**	**45.05**	**25.13**	**33.29**
FedConcept	37.66	**65.10**	21.58	43.24	23.02	31.57

Results in progress.

Seletcted Research Contributions

- **Published**
  - Graph-based Federated Learning for Smart Healthcare: A Comprehensive Survey
- **Submitted**
  - Mitigating Structural Entanglement via Architecture-Aware Unlearning in Federated Graphs
- **In Progress**
  - Graph of Concepts: Towards Federated Graph Foundation Models via Structural Bottlenecks

Evolution of Scope

The DoA anticipated partitioning, aggregation and interpretability for federated GNN *training*. The work has moved toward two questions that emerged from the interpretability objective itself.

- Asking *where* a client's influence lives is an interpretability question.

Evolution of Scope

The DoA anticipated partitioning, aggregation and interpretability for federated GNN *training*. The work has moved toward two questions that emerged from the interpretability objective itself.

- Asking *where* a client's influence lives is an interpretability question.
- Asking whether decentralised representations can be *named* led to a shared structural vocabulary, and hence to a **foundation-model** formulation.

Evolution of Scope

The DoA anticipated partitioning, aggregation and interpretability for federated GNN *training*. The work has moved toward two questions that emerged from the interpretability objective itself.

- Asking *where* a client's influence lives is an interpretability question.
- Asking whether decentralised representations can be *named* led to a shared structural vocabulary, and hence to a **foundation-model** formulation.
- Current focus: decentralised training without raw data exchange, non-IID structured data. Robustness, Data partition, and interpretability are the main focus in the presented research

Summary and Next Steps

1. **Isolating the Influence** — removal of client- or node-level influence at a fraction of retraining cost, with MIA AUC near chance.

Summary and Next Steps

1. **Isolating the Influence** — removal of client- or node-level influence at a fraction of retraining cost, with MIA AUC near chance.
2. **Cross-domain generalizability** — a shared vocabulary of structural concepts gives cross-domain federated learning an interpretable intermediate layer.

Summary and Next Steps

1. **Isolating the Influence** — removal of client- or node-level influence at a fraction of retraining cost, with MIA AUC near chance.
2. **Cross-domain generalizability** — a shared vocabulary of structural concepts gives cross-domain federated learning an interpretable intermediate layer.

Summary and Next Steps

1. **Isolating the Influence** — removal of client- or node-level influence at a fraction of retraining cost, with MIA AUC near chance.
2. **Cross-domain generalizability** — a shared vocabulary of structural concepts gives cross-domain federated learning an interpretable intermediate layer.

Next steps

What exactly drives the uncertainity and interpretability in federated graph neural network?

Concluding Remarks

Where the project stands

Two federated GNN frameworks are implemented and evaluated, both built on the same idea: that influence in a decentralised graph model can be **located** and **named** rather than treated as opaque.

- Interpretability remains the organising objective of DC9, pursued through model structure rather than post-hoc attribution.

Thank you, questions welcome.

S M Asiful Huda · DC9 · UNINA