FedScholar: Privacy-Preserving Academic Report Generation via Publication Clustered Federated Learning

Zhihan Guo
The Chinese University of Hong Kong
Hong Kong SAR, China
zhguo22@cse.cuhk.edu.hk

Yukun Zhang
Harbin Institute of Technology
Shenzhen, China
23s051048@stu.hit.edu.cn

Zenglin Xu
Fudan University
Shanghai, China
zenglinxu@fudan.edu.cn

Irwin King
The Chinese University of Hong Kong
Hong Kong SAR, China
king@cse.cuhk.edu.hk

Abstract

The explosive growth of academic literature necessitates utilizing Large Language Models (LLMs) to generate comprehensive scientific reports. However, current general LLMs often suffer from insufficient disciplinary knowledge, leading to superficial or hallucinatory outputs when addressing specialized domains. The limitation stems from three obstacles: (1) high-quality literature fragmented across distinct publishers under strict privacy constraints, preventing the centralized training; (2) the scarcity of training data tailored for long-sequence academic writing; (3) the difficulty in designing objective reward signals for open-ended generation. To address these challenges, we introduce **FedScholar**, a **Fed**erated reinforcement learning framework tailored for **Scholar**ly report generation. First, we propose a Publication Clustered Federated Learning (PCFL) to bridge the knowledge gap with data privacy. Second, we implement a prompt-driven annotation pipeline to resolve data scarcity. Third, we design a novel reward signal of objective evaluation. Extensive experiments across 10 academic disciplines on 23 top-tier publications demonstrate that FedScholar significantly outperforms state-of-the-art baselines. Notably, FedScholar outperforms all baselines, achieving state-of-the-art scores of 63.79% in local retention and 62.09% in global generalization.

CCS Concepts

• **Computing methodologies** → **Natural language processing**.

Keywords

federated learning, long text generation, reinforment learning

ACM Reference Format:
Zhihan Guo, Yukun Zhang, Zenglin Xu, and Irwin King. 2026. FedScholar: Privacy-Preserving Academic Report Generation via Publication Clustered Federated Learning. In *Companion Proceedings of the ACM Web Conference 2026 (WWW '26 Companion), April 13–17, 2026, Dubai, UAE.* ACM, New York, NY, USA, 10 pages. https://doi.org/XXXXXXX.XXXXXXX

WWW '26 Companion, Dubai, UAE

ACM ISBN 978-1-4503-XXXX-X/2018/06
https://doi.org/XXXXXXX.XXXXXXX

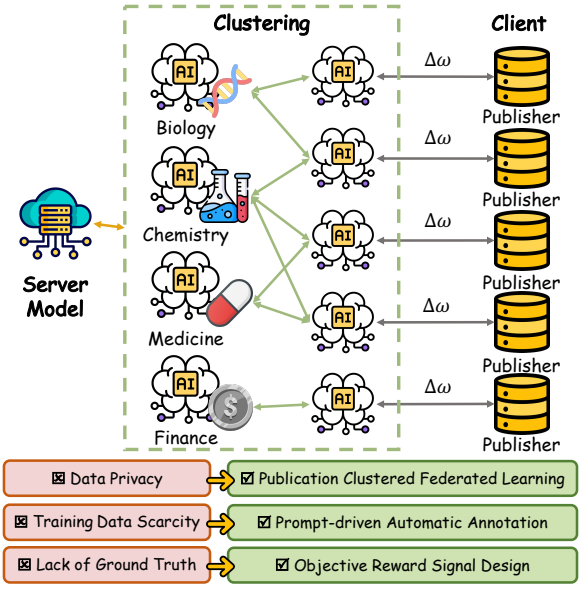

Figure 1: Illustrations of primary challenges and solutions in current academic report generation.

1 Introduction

Academic report writing demands the synthesis of extensive knowledge to answer complex research inquiries. However, as the volume of scientific literature grows exponentially [19], relying on traditional manual methods to curate and summarize such vast information has become increasingly impractical [23]. Consequently, there is an urgent need for automated approaches, particularly Large Language Models (LLMs), to assist in this intensive knowledge work [26]. Nevertheless, the performance of these automated systems is fundamentally constrained by data accessibility. Training high-quality academic models requires access to premium, authoritative literature. However, a significant portion of this high-value data is non-public, residing within proprietary databases or protected by strict copyright laws. Centralizing such sensitive data for

training inevitably raises severe privacy concerns and risks infringing upon publisher rights. Therefore, exploring privacy-preserving mechanisms that enable effective learning without compromising data ownership has become a critical priority.

Federated Learning (FL) presents a robust solution to this privacy challenge. By allowing models to train locally on decentralized data sources and aggregating only the model updates—rather than the raw data—FL effectively safeguards privacy while leveraging distributed knowledge [9, 21]. This paradigm has already demonstrated profound value across various high-stakes industries. For instance, in healthcare, FL enables hospitals to collaboratively train diagnostic models without sharing sensitive patient records [5, 14, 20]; similarly, in finance, it allows institutions to improve fraud detection systems without exposing proprietary transaction data [2, 7]. Despite its proven success in these domains, the application of Federated Learning specifically for academic research automation remains surprisingly limited.

Despite these needs, current approaches face two primary challenges in this specific context. **First, the distributed nature of high-quality academic data creates a significant barrier to traditional model training.** Valuable literature is fragmented across distinct publishers and proprietary repositories, rather than being centrally available. Consequently, centralized aggregation of this data for training purposes is often infeasible, as it risks severe privacy leakage and violates the data sovereignty of different content owners. **Second, high-quality training datasets for generating long academic reports are extremely scarce.** Unlike general post-tuning, constructing datasets for this domain demands expert-level comprehension to ensure accuracy and structural integrity. Relying on manual annotation by human experts is prohibitively expensive and unscalable, thereby limiting the availability of large-scale supervision. **Third, optimizing the quality of generated reports via Reinforcement Learning (RL) presents a unique difficulty in reward signal design.** Unlike short-form tasks with unique answers, academic report writing is long-form and open-ended, lacking definitive ground-truth labels. This subjectivity makes it extremely challenging to formulate a precise reward mechanism that can automatically and objectively evaluate the depth and logical coherence of the model's output.

To address these challenges, we propose FedScholar, a novel framework for privacy-preserving and high-quality academic report generation, as shown in Figure 1. Our main contributions are summarized as follows:

- Discipline-Aware Federated Architecture: We implement a clustered federated learning strategy where individual publishers serve as distributed clients. To handle the heterogeneity of scientific knowledge, model updates are aggregated based on academic disciplines to train the global server model, thereby utilizing multi-source proprietary data without compromising privacy.
- Prompt-Driven Automated Annotation: We devise a cost-effective data construction method that eliminates the need for heavy manual labor. By utilizing a single, concise prompt, we empower LLMs to autonomously annotate and generate high-quality training datasets, significantly reducing the costs associated with human supervision.
- Reference-Free Reward Mechanism: We design a novel reinforcement learning reward signal tailored for open-ended tasks. This mechanism leverages the long-context understanding capabilities of LLMs to generate informative reward signals that guide model optimization, effectively overcoming the dependency on traditional gold references.

2 Related work

2.1 LLMs for academic long text generation

The application of Large Language Models (LLMs) in the academic domain has increasingly focused on the challenge of generating coherent, long-form scientific text. Unlike short-form dialogue, academic report writing requires models to maintain logical consistency and contextual integrity over extremely long sequences [4]. Specialized models, such as Galactica [16] and InteractiveSurvey [18], have been developed to master scientific nomenclature and complex reasoning. These works demonstrate that scaling model capacity and context window size is crucial for synthesizing extensive literature into structured reports. However, existing research predominantly relies on centralized training strategies utilizing public repositories like arXiv or PubMed. This centralized paradigm overlooks the vast landscape of high-quality, proprietary academic literature restricted by copyright. Furthermore, standard training objectives are often insufficient for optimizing the global coherence of long open-ended reports, necessitating more sophisticated guidance mechanisms.

2.2 Clustered federated learning

Federated Learning (FL) enables collaborative model training across distributed clients without sharing raw data, addressing the aforementioned privacy concerns [10]. Conventional algorithms aim to train a single global model that generalizes across all clients, often struggling when data distributions vary significantly [1, 8]. However, a critical challenge in FL is the statistical heterogeneity (Non-IID data) among clients, which can degrade the performance of the global model. Clustered Federated Learning (CFL) emerges as a strategic middle ground between these extremes, specifically designed to mitigate the challenges of Non-IID (Non-Independent and Identically Distributed) data [15]. The fundamental premise of CFL is that clients exhibiting comparable data-generating distributions should be aggregated into cohesive clusters. By doing so, the decentralized datasets within each cluster approximate IID conditions, satisfying conventional machine learning assumptions that are otherwise violated in standard FL [3]. Consequently, CFL involves training a shared, specialized model for each cluster, which improves performance through intra-cluster collaboration. In the context of academic publishing, this framework provides a theoretical basis for grouping publishers by discipline, allowing for the training of specialized models that are robust to domain heterogeneity.

3 Preliminaries

3.1 Preference alignment

Let $\mathcal{D} = \{(x, y^+, y^-)\}$ denote a dataset of preferences, where x is an input prompt, y^+, y^- are the responses labeled as preferred and

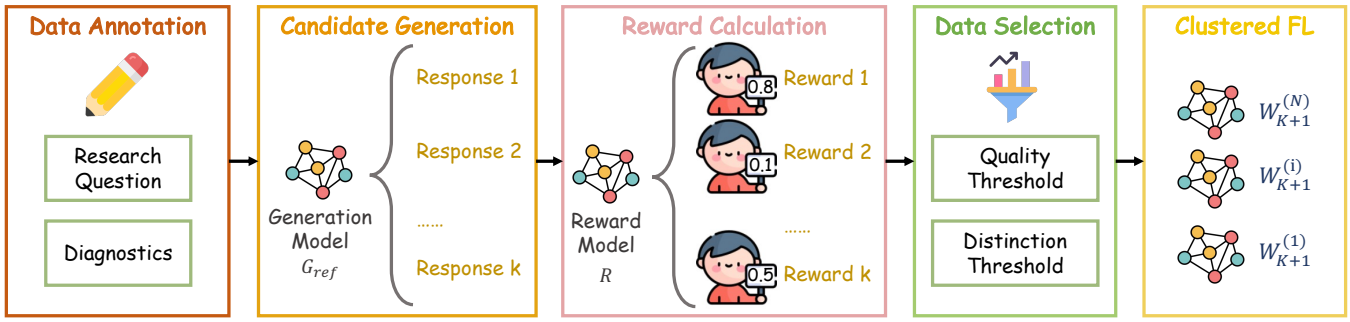

Figure 2: FedScholar framework overview.

dis-preferred, respectively. The purpose of preference alignment is to designing a policy π that maps prompts to responses, maximizing a reward that reflects human preferences using the Bradley–Terry (BT) model:

$$p(y_1 \succ y_2|x) = \sigma\left(r^*(x, y_1) - r^*(x, y_2)\right), \tag{1}$$

where $r^*(x, y_1)$ represents the oracle reward of a response given a prompt, and $\sigma(z) = \{1 + exp(-z)\}^{-1}$ is the sigmoid function, mapping differences in rewards to probabilities. A parameterized reward model r_θ is estimated by solving a maximum likelihood estimation (MLE) objective:

$$\mathcal{L}(\theta) = -\mathbb{E}_{(x,y_w,y_l)\sim\mathcal{D}}[\log \sigma(r^*_\theta(x, y_w) - r^*_\theta(x, y_l)], \tag{2}$$

where y_w, y_l are preferred and dis-preferred sample respectively. The direct preference optimization (DPO) [**?**] we used in this work chose the reward as:

$$r_\theta(x, y) = \beta \log \frac{\pi_\theta(y \mid x)}{\pi_{\text{ref}}(y \mid x)}, \tag{3}$$

to directly optimize the policy π_θ using the loss $L(\theta)$ in Equation 2, with the reward function r specified by Equation 3. As the reward is implicitly defined by the policy itself, the objective becomes fully dependent on θ eliminating the need for a separately trained reward model. Consequently, this reformulation significantly improves the computational efficiency of the alignment process.

3.2 Synthetic preference alignment pipeline

Given the generation policy $\mathcal{G}$ parameterized by θ and an LLM-based reward model $\mathcal{R}$. The synthetic preference alignment pipeline typically consists of the following stages:

Response Generation. Given a dataset of prompts $\mathcal{X} = \{x_1, \cdots, x_n\}$, the policy $\mathcal{G}_\theta$ generate a set of responses $\{y_i^1, y_i^2, \cdots\}$ which are intended to cover diverse output patterns for each prompt x_i.

AI-based Reward Assignment. For each response y_i^j, reward score $r(x_i, y_i^j)$ is calculated by reward model $\mathcal{R}$, which acts as an automatic evaluator.

Policy Optimization. The policy $\mathcal{G}_\theta$ is then fine-tuned using the synthetic reward signal. DPO are commonly used to align the policy with the feedback provided by reward model $\mathcal{R}$.

4 Methodology

In this section, we introduce FedScholar, a novel reinforcement learning framework for multi-domain academic report generation. As shown in Figure 2, our approach comprises three key components. First, an automatically constructed **FedScholar Dataset** (Section 4.1), an automatically constructed corpus that leverages LLMs to generate diverse research inquiries and corresponding checklist pairs, thereby avoiding the rigidity of fixed patterns. Second, we propose the **FedScholar Reward** (Section 4.2), a reference-free mechanism that provides targeted signals to guide the reinforcement learning optimization process. Third, we implement a **Publisher Clustered Federated Learning (PCFL)** system (Section 4.3), a clustered federated learning strategy that mitigates data heterogeneity among publishers while strictly ensuring data privacy.

4.1 FedScholar dataset collection

A fundamental challenge in academic report generation is the open-ended nature of the task. Unlike factoid QA, long-form reports do not adhere to fixed patterns, making it inherently difficult to establish standard ground-truth references for training. To address this, we construct the FedScholar Dataset, a comprehensive corpus designed to facilitate training across 10 distinct disciplines: Computer Science (CS), Finance, Medicine, Biology, Chemistry, Environmental Science (Env), Energy, Building & Construction (B&C), Earth Science (Earth), and Materials. Formally, the dataset comprises a set of research questions $M = \{m_1, m_2, ..., m_n\}$, where each research question m_i a lengthy, context-rich response. For every research question m_i, we derive a corresponding set of diagnostic question-answer pairs, denoted as $P_i = \{(q_{i1}, a_{i1}), (q_{i2}, a_{i2}), ..., (q_{il}, a_{il})\}$.

Our core methodological innovation lies in transforming the subjective expert evaluation of long-form text quality into a set of objective, granular reading comprehension tasks. These diagnostic pairs (q, a) function as an objective "checklist" to quantify the information density of a generated report. Each diagnostic question q is designed to query a specific key information point that should be present in a high-quality response to m. The corresponding answer a is formulated as a Boolean value. Crucially, we ensure that the majority of diagnostic questions expect "True" as the answer. This design choice effectively treats the evaluation as a recall-oriented

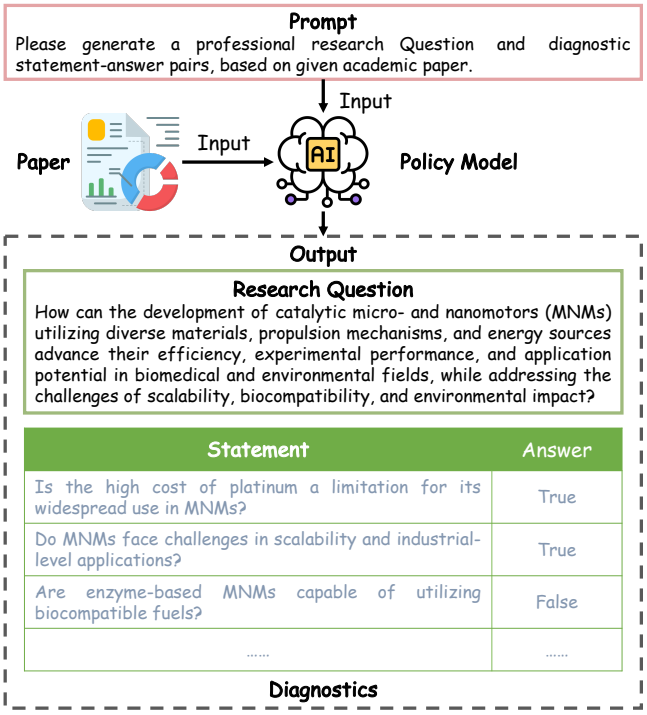

Figure 3: The pipeline of research questions and diagnostic Question-Answer pairs generation in FedScholar.

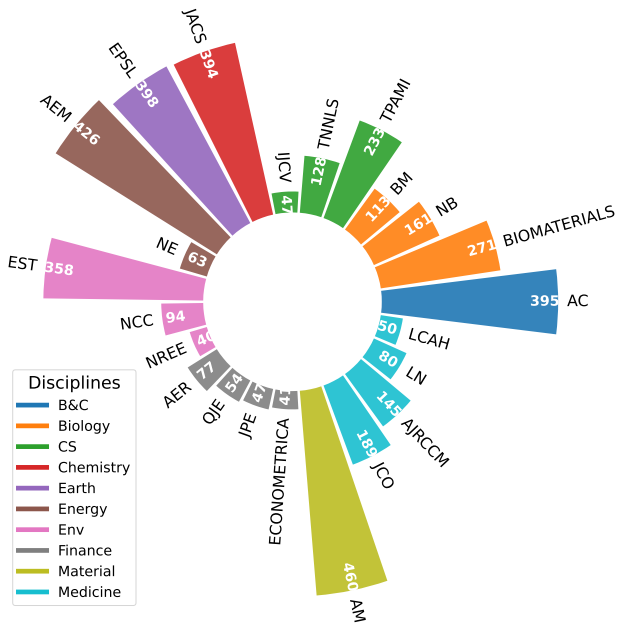

Figure 4: Publications distribution in FedScholar dataset across 10 disciplines: CS (Computer Science), Finance, Medicine, Biology, Chemistry, Env (Environmental Science), Energy, B&C (Building and Construction), Earth (Earth Science), and Materials. The correspondence between publication abbreviations and full titles is provided in Appendix B.

verification task, where a False" value would imply that the information queried by q is irrelevant to the core research topic m, thus rendering it unnecessary for the checklist.

The construction pipeline for the FedScholar Dataset prioritizes scalability and cost-efficiency. Instead of relying on expensive manual annotation, we employ a prompt-driven approach via LLM APIs. We initially selected 20 publications (mapped to the 10 major disciplines) that require complex reasoning, such as Nature Biotechnology, IEEE Transactions On Pattern Analysis And Machine Intelligence, and Lancet Neurology. As illustrated in Figure 3, the automated generation process proceeds in two steps. First, we prompt the LLM to generate a large-scale set of 10,557 research questions ($\mathcal{M}$) requiring deep contextual answers. Second, for each research question m_i, we instruct the LLM to generate approximately 15 Boolean diagnostic pairs ($\mathcal{P}_i$). The specific prompts used are detailed in Figure A, and the statistical distribution of the dataset is presented in Figure 4.

4.2 FedScholar reward mechanism

A significant bottleneck in RLHF (Reinforcement Learning from Human Feedback) for academic domains is the scarcity of large-scale, high-quality preference data. To address this, we implement an Active Exploration strategy driven by a reference-free reward mechanism. This process involves a generator model $\mathcal{G}$ and an evaluator (reward) model $\mathcal{R}$, proceeding in three distinct phases: Candidate Generation, Reward Calculation, and Preference Pair Construction.

Candidate Generation. Given a research question m from the FedScholar Dataset, we first employ the policy model $\mathcal{G}$ to generate k diverse candidate reports, denoted as $\mathcal{T} = \{r_1, r_2, ..., r_k\}$. To ensure diversity in the reasoning paths and structural organization of the reports, we apply temperature sampling during this inference phase.

Reference-Free Reward Calculation. To evaluate these open-ended reports without relying on expensive human references, we hypothesize that the quality of a long-form response r_i is directly correlated with its information density—specifically, its ability to answer the pre-generated diagnostic checklist $\mathcal{P} = \{(q_j, a_j)\}_{j=1}^{l}$ (from Section 4.1). We feed the generated report r_i and the diagnostic questions into the reward model $\mathcal{R}$. The model predicts an answer a'_{ij} for each question q_j based solely on the context provided by r_i. The reward score $\mathcal{S}(r_i)$ is formally defined as the accuracy of these proxy answers against the expected Boolean truths a_j:

$$\mathcal{S}(r_i) = \frac{1}{l} \sum_{j=1}^{l} \mathcal{F}(a'ij, aj), \tag{4}$$

where $\mathcal{F}$ is an indicator function that verifies the correctness of the extracted information:

$$\mathcal{F}(a'_{ij}, \hat{a}_{ij}) = \begin{cases} 1, & \text{if } a'_{ij} = \hat{a}_{ij}, \\ 0, & \text{if } a'_{ij} \neq \hat{a}_{ij}. \end{cases} \tag{5}$$

Here, $a'_{ij} = \mathcal{R}(r_i, q_j)$ represents the answer inferred by the reward model. This scoring function $\mathcal{S}(r_i) \in [0, 1]$ effectively quantifies

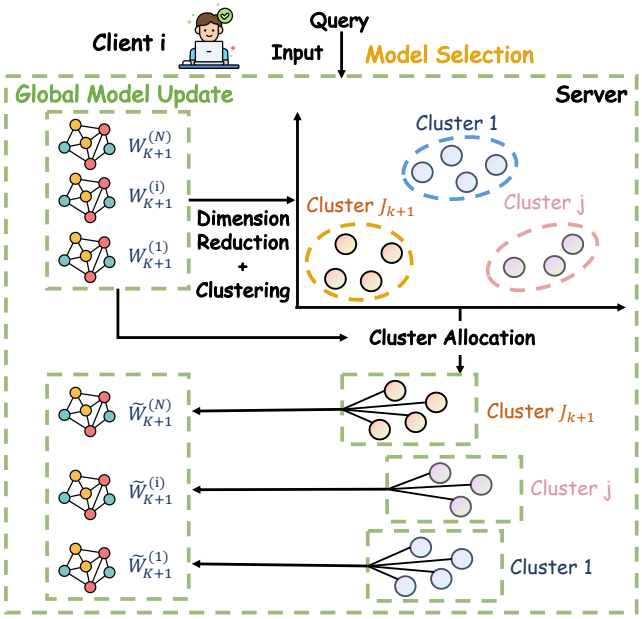

Figure 5: Overview of the Publication Clustered Federated Learning (PCFL) framework. During training, the server receives local model updates from diverse publishers and groups them into latent disciplinary clusters. During inference, the Model Selection module routes the client's input query to the most appropriate cluster model for optimal generation.

the informativeness and coverage of the report r_i in a granular, objective manner.

Preference Pair Construction via Dual-Threshold Filtering. To construct a robust training dataset for Direct Preference Optimization (DPO), simply pairing random responses is insufficient. We introduce a Dual-Threshold Filtering strategy to ensure that the training pairs are both high-quality and sufficiently distinct to provide a meaningful learning signal. For the set of k scored responses, we calculate the maximum score $S_{max} = \max_i \mathcal{S}(r_i)$ and the score variance $\sigma_s^2 = \mathrm{Var}(\{\mathcal{S}(r_1), ..., \mathcal{S}(r_k)\})$. We select data based on two criteria:

(1) **Quality Threshold (τ_{qual}):** We require $S_{max} > \tau_{qual}$. This ensures that at least one generated report is of sufficient quality to serve as a positive example.
(2) **Distinction Threshold (τ_{var}):** We require $\sigma_s^2 > \tau_{var}$. This ensures sufficient contrast between the best and worst responses, preventing the model from learning from ambiguous pairs where the quality gap is negligible.

From the filtered sets, we construct the preference pair (y_w, y_l) by selecting the response with the highest score as the winner y_w and the response with the lowest score as the loser y_l. These pairs are then used to optimize the model via DPO, aligning the generation probability with the defined reward signal.

Algorithm 1: Training process of FedAvg

Input: Publisher set C; Communication round T; Local epoch number E; The initial shared global model parameters w^0 on server; The local learning rate η_l; The local dataset $\mathcal{D}_i$ of the i-th publisher; The loss function $\mathcal{L}$

Output: The shared global model parameters w^T on server;

ServerGlobalUpdating(C, T, w^0):
for *each communication round* $t = 1$ *to* T **do**
  for *each publisher* $i \in C$ *in parallel* **do**
    PublisherLocalTraining(i, w^{t-1});
    Receive publisher-uploaded parameters Δw_i^t;
  end
  Perform global aggregation by:
    $w^t \leftarrow w^{t-1} + \sum_{i=1}^{|C|} \frac{|\mathcal{D}_i|}{\sum_{j=1}^{|C|} |\mathcal{D}_j|} \Delta w_i^t$
end
PublisherLocalTraining(i, w^t):
$w_i^t \leftarrow w^t$;
for *epoch* $e = 1$ *to* E **do**
  $w_i^t \leftarrow w_i^t - \eta_l \frac{\partial \mathcal{L}_i}{\partial w_i^t}$
end
$\Delta w_i^{t+1} = w_i^{t+1} - w^t$;
Send Δw_i^{t+1} to the sever;

4.3 Publisher clustered federated learning

Handling Data Heterogeneity via Clustered Objectives. In the academic publishing landscape, data is naturally distributed across different publishers (clients) that exhibit significant statistical heterogeneity (Non-IID). For instance, a publisher specializing in Medicine holds data with vastly different vocabulary and reasoning patterns compared to a publisher focused on Computer Science. In standard Federated Learning (e.g., FedAvg), a single global model is trained to minimize the average loss across all clients. However, due to the divergent domain distributions, the local updates from these publishers often point in conflicting directions. Aggregating these conflicting gradients into a single model leads to "client drift", where the global model fails to converge to an optimal solution for any specific domain.

To resolve this conflict while respecting data privacy, we adopt a Publisher Clustered Federated Learning (CFL) framework. Instead of forcing a single global model w, we aim to learn a set of K specialized models $\{w_1^*, \ldots, w_K^*\}$ corresponding to latent academic disciplines. The optimization objective is formulated to minimize the loss for each publisher i relative to its most suitable cluster model:

$$\min_{W=w_1,\ldots,w_K} \sum_{i \in C} \sum_{(x,y) \in \mathcal{D}i} \mathcal{L}(f(x; w\phi(i)), y), \tag{6}$$

where C represents the set of publishers, $\mathcal{D}_i$ is the local proprietary dataset of publisher i, and $\phi(i) \in \{1, \ldots, K\}$ is the assignment function mapping publisher i to its optimal cluster k. This formulation ensures that publishers with similar domains contribute to the

Algorithm 2: Training process of FedScholar (Ours)

Input: Publisher set C; Communication round T; Local epoch number E; The initial shared global model parameters w^0 on server; The dimension d after dimensionality reduction; The threshold c for clustering; Cluster set $C^{'} = \emptyset$; Cluster size $B^{'} = 0$;

Output: The shared global model parameters $w^{T,k}$ for each cluster k

/* Publisher local training and clustering in communication round 1 */

for *each publisher* $i \in C$ *in parallel* **do**
  `PublisherLocalTraining`(i, w^0); `// in Algorithm 1`
  Receive publisher-uploaded parameters Δw_i^1;
  Compute $\hat{\Delta} w_i^1$ by reducing Δw_i^1's dimension to d;
end
while $C \neq \cup_{b=0}^{B^{'}} C_b^{'}$ **do**
  Let $i \in C \setminus \cup_{b=0}^{B} C_b^{'}$;
  $B^{'} \leftarrow B^{'} + 1$;
  Let $C_{B^{'}}^{'} = \{i\}$ and $\mathcal{A} = C \setminus \cup_{b=0}^{B^{'}} C_b^{'}$;
  while $\mathcal{A} \neq \emptyset$ **do**
    Select $j \in \mathcal{A}$;
    for k *in* $C_{B^{'}}^{'}$ **do**
      Compute $sim(j,k)$ by Equation 7
    end
    if $sim(j,k) \geq c$ *for all* $k \in C_{B^{'}}^{'}$ **then**
      $C_{B^{'}}^{'} \leftarrow C_{B^{'}}^{'} \cup \{j\}$
    end
    $\mathcal{A} \leftarrow \mathcal{A} \setminus \{j\}$
  end
end
Perform global aggregation by:

$$w^1 = w^0 + \sum_{i=1}^{|C|} \frac{|\mathcal{D}_i|}{\sum_{j=1}^{|C|} |\mathcal{D}_j|} \Delta w_i^1$$

Initialize shared model parameters for each cluster with w^1;
/* Perform FedAvg procedure in each cluster */
for *each communication round* $t = 2$ *to* T **do**
  for *each cluster* $k \in C^{'}$ **do**
    `ServerGlobalUpdating`$(k, 1, w^{t-1,k})$; `// in Algorithm 1`
  end
end

same sub-model, mitigating gradient interference and preserving the integrity of domain-specific knowledge.

Parameter-Based Clustering Strategy. The core challenge in CFL is determining the assignment function $\phi(i)$ without accessing the private raw data $\mathcal{D}_i$. While previous works have attempted to cluster based on local data samples, these methods often require sharing data representations or statistics, which poses privacy risks and ignores the structural heterogeneity between clients. We propose a strictly privacy-preserving approach based on the hypothesis that publishers with similar academic content will produce aligned model updates. Specifically, if two publishers belong to the same discipline (e.g., both publish Physics papers), the gradients generated during local training will exhibit high directional similarity in the parameter space. Therefore, we utilize the model updates (weights or gradients) as the features for clustering. Due to the high dimensionality of the raw parameters, direct similarity calculations failed to effectively distinguish between different publishers. Therefore, we first applied Principal Component Analysis (PCA) for dimensionality reduction. At each communication round, the server computes the pairwise cosine similarity between the projected updates Δw_i and Δw_j received from publishers i and j:

$$\text{sim}(i,j) = \frac{\Delta w_i \cdot \Delta w_j}{||\Delta w_i|| \cdot ||\Delta w_j||}. \tag{7}$$

Based on this similarity matrix, we employ a greedy clustering approach to dynamically group publishers into K clusters. Updates are then aggregated exclusively within these identified clusters to update the specific cluster models $\{w_1, \ldots, w_K\}$. This method effectively captures the latent disciplinary structure of the scientific community without ever inspecting the underlying manuscripts. The whole procedure are shown in Algorithm 2.

Table 1: Main results across models on easy, medium, and hard tasks. All results are presented as percentages. Each model runs 3 times inference and evaluation. The top two results are highlighted in bold.

Model	LOCAL	GLOBAL			
		Easy	Medium	Hard	Average
Centralized					
Base	61.03 ± 0.23	69.77 ± 0.97	59.93 ± 0.38	53.17 ± 0.46	60.91 ± 0.41
Central	**65.40** ± 0.17	**72.57** ± 0.52	**65.83** ± 1.49	**59.69** ± 0.81	**65.97** ± 0.66
Decentralized					
Local	61.63 ± 0.70	69.85 ± 1.83	61.74 ± 1.07	53.94 ± 1.28	61.77 ± 0.47
FedAvg	62.89 ± 0.29	68.91 ± 1.19	61.01 ± 0.74	55.24 ± 1.00	61.68 ± 0.60
FedProx	61.23 ± 0.37	69.42 ± 0.95	61.10 ± 1.30	54.61 ± 0.38	61.66 ± 0.47
Scaffold	61.95 ± 0.69	70.17 ± 1.54	60.40 ± 0.68	52.60 ± 1.81	60.99 ± 0.46
FedAdagrad	61.97 ± 0.92	70.24 ± 0.50	59.54 ± 0.78	**55.87** ± 1.22	61.89 ± 0.54
FedAdam	61.02 ± 1.30	68.55 ± 2.39	59.48 ± 0.48	54.92 ± 1.84	60.97 ± 0.22
FedYogi	61.64 ± 0.28	69.11 ± 0.22	59.90 ± 1.49	55.26 ± 0.48	61.41 ± 0.41
FedScholar	**63.79** ± 0.38	**70.28** ± 1.60	**61.97** ± 0.64	54.25 ± 0.95	**62.09** ± 0.46

5 Experiment

5.1 Experiment setup

Baseline Algorithms. To evaluate the performance of our proposed method, we compare it against seven representative Federated Learning (FL) algorithms. We first consider several foundational global FL methods: *FedAvg* [9], the de facto baseline in the field, which performs simple model averaging; *FedProx* [8], which introduces an L_2 proximal term to handle statistical heterogeneity; and *Scaffold* [6], designed to mitigate client drift via control variates and server-side learning rate adjustments. Additionally, we incorporate *FedAdam*, *FedYogi*, and *FedAdagrad* [13], which extend standard centralized adaptive optimizers to the federated setting. Beyond the FL paradigm, we include two non-collaborative baselines: *Local*, where models are trained independently on local data,

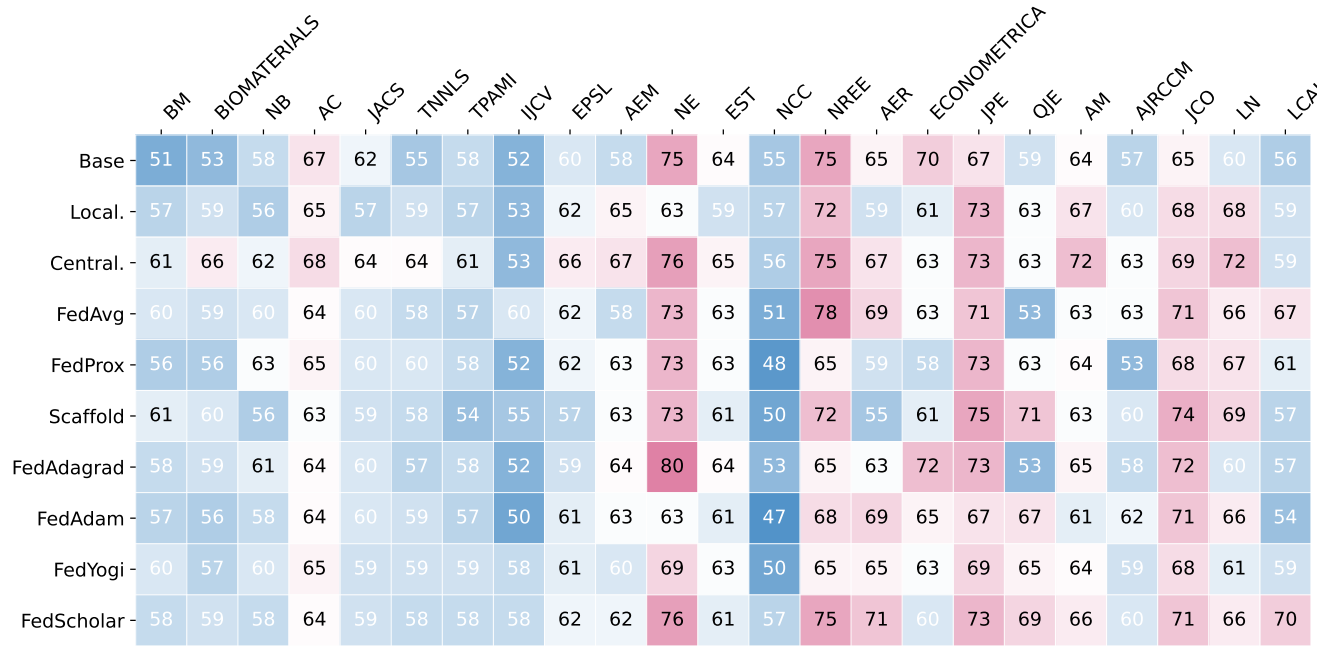

Figure 6: Performance heatmap comparison across 23 diverse academic publishers. The correspondence between publication abbreviations and full titles is provided in Appendix B. All results are presented as percentages.

and *Central*, which serves as an upper-bound performance oracle by assuming centralized access to all participants' data.

Experimental Setup. We conduct our experiments using the Qwen-2.5-7B-Instruct [17] model, implemented within the OpenFedLLM framework [22]. To facilitate reward-based training, we employ GPT-4o-mini [11] to compute training rewards, while GPT-4o-2024-11-20 [12] is utilized as the evaluator for ProxyQA scores. Our study focuses on a cross-silo FL scenario, where all participating publishers contribute to the training in every communication round.All experiments are performed with a learning rate of 0.0003. For the baseline FL algorithms, we adhere to the default hyperparameters specified in OpenFedLLM: for FedProx, the proximal term μ is set to 0.1; for FedAdagrad, FedAdam, and FedYogi, the parameters are configured as $\beta_1 = 0.9$, $\beta_2 = 0.99$, $\eta_g = 0.001$, and $\tau = 0.001$. For our proposed FedScholar algorithm, we set the reduced dimensionality to $d = 2$ and the clustering threshold to $c = 0.9$. To maintain a balance between computational efficiency and communication overhead, we fix the local training to 1 epoch and the total communication rounds to 3 across all experimental settings.

Evaluation Strategies. To provide a comprehensive performance analysis, we construct both local and global evaluation datasets. Specifically, each local dataset is partitioned into training and testing subsets. The global evaluation set is then formed by aggregating these individual test sets, resulting in a total of 100 test samples. Following established benchmarks in federated learning [24, 25], we evaluate all algorithms using two distinct strategies:

(1) **Global Evaluation (GLOBAL):** Performance is assessed on the aggregate global test set to determine the model's ability to incorporate knowledge from other clients. Higher GLOBAL scores indicate that the federated model is approaching the performance of *centralized training*.
(2) **Local Evaluation (LOCAL):** Performance is measured on each local test set and then averaged across all clients. LOCAL is particularly relevant for real-world applications as it reflects the model's efficacy on client-specific tasks without necessitating the centralization of private data.

5.2 Does FedScholar outperform state-of-the-art baselines?

FedScholar significantly outperforms all privacy-preserving baselines by effectively addressing the data heterogeneity inherent in the publishing landscape. Unlike conventional federated approaches (e.g., FedAvg, Scaffold) that enforce a single global model across all clients—a strategy that ignores the distinct data distributions of different publishers and dilutes domain-specific knowledge—FedScholar adaptively aggregates publishers with similar disciplinary characteristics. This advantage is empirically evident in Table 1, where FedScholar achieves the highest Average score (62.09%) among all decentralized methods. Crucially, our method demonstrates superior capability in retaining domain expertise, securing a remarkable Local performance of 63.79%, significantly surpassing standard FedAvg (62.89%). These results confirm that by mitigating the conflict between diverse academic disciplines through clustered aggregation, FedScholar establishes a new benchmark for high-quality, privacy-compliant academic report generation.

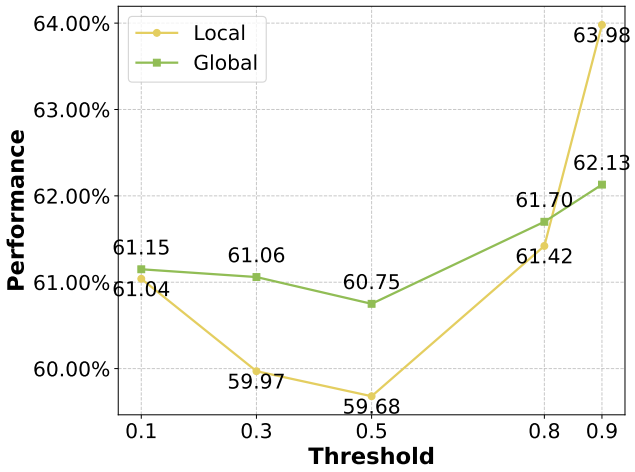

Figure 7: Ablation study on the clustering threshold.

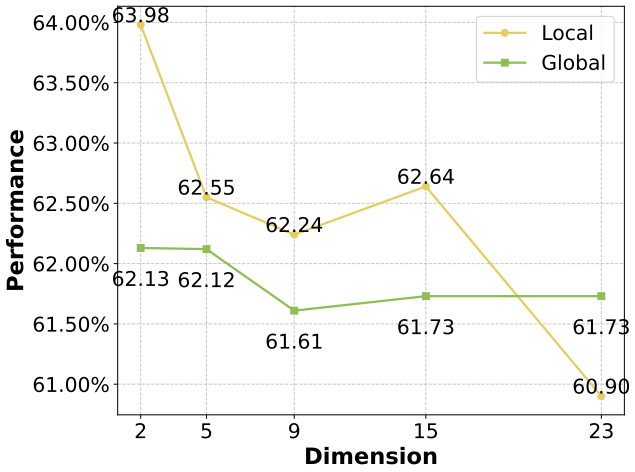

Figure 8: Ablation study on the dimension reduction.

5.3 Does FedScholar robustly adapt to diverse academic domain?

FedScholar demonstrates superior stability across heterogeneous domains, effectively preventing the performance collapse often observed in standard federated learning. As illustrated in Figure 6, the difficulty of academic report generation varies significantly by discipline. Standard methods like FedAvg struggle with this statistical heterogeneity, causing performance in"hard" domains (e.g. NCC) to plummet to as low as 50.9% due to conflicting gradient updates. In contrast, FedScholar leverages clustered aggregation to shield these vulnerable domains, restoring their performance to 56.9%—an improvement of over 6%. By raising the minimum performance floor across all 23 publishers from FedAvg's 50.9% to 56.8%, FedScholar proves it can robustly adapt to diverse data distributions without sacrificing the quality of niche or difficult academic subjects.

6 Ablation study

6.1 How Sensitive is FedScholar to the Clustering Threshold?

A stricter clustering threshold is crucial for maintaining high-quality domain specialization. We investigate the impact of the similarity threshold c in our Publisher CFL module by varying it across $\{0.1, 0.3, 0.5, 0.8, 0.9\}$. As observed in the Figure 7, lower thresholds (e.g., $c = 0.5$) imply a high tolerance for dissimilarity, which inadvertently groups unrelated publishers into the same cluster. This "loose" clustering introduces gradient conflicts and noise, causing the Local score to drop to as low as 59.68%. Conversely, increasing the threshold enforces rigorous similarity criteria, ensuring that only publishers with highly aligned disciplinary distributions are aggregated. Consequently, the model achieves peak performance at $c = 0.9$, recording a Local score of 63.98% and a Global score of 62.13%. This confirms that minimizing inter-cluster noise via a high threshold is more effective than broad, indiscriminate collaboration.

6.2 Does Higher Dimensionality Improve Clustering Precision?

Lower dimensionality actually yields better performance by filtering out noise. We further examine the effect of the feature dimension d used during the publisher clustering process, varying d across $\{2, 5, 9, 15, 23\}$. As shown in Figure reffig:dimension, contrary to the intuition that more features capture more detail, our results show a negative correlation between dimension size and model performance. Specifically, as d increases from 2 to 23, the Local score declines from a peak of 63.98% to 60.90%. This suggests that higher dimensions retain excessive noise and irrelevant fluctuations in the gradient updates, which obscures the true disciplinary similarities between publishers and degrades clustering quality. However, it is worth noting that the performance variance caused by dimension d (a range of $\sim$ 3%) is relatively smaller compared to the impact of the similarity threshold c, indicating that while dimensionality reduction is beneficial, strict thresholding remains the primary driver of clustering success.

7 Conslusion

In this paper, we introduced FedScholar, a comprehensive framework designed to tackle three critical impediments in automated academic report generation. Specifically, we addressed the barrier of distributed proprietary data via Publication Clustered Federated Learning (PCFL), which aggregates domain-specific knowledge while preserving privacy. To mitigate the severe scarcity of high-quality long-form training corpora, we developed a cost-effective prompt-driven automatic annotation pipeline. Furthermore, we overcame the difficulty of optimizing open-ended generation by designing a novel objective reward signal that replaces subjective evaluation with verifiable grounding. Extensive experiments are across 10 academic disciplines on 23 top-tier publications. Experiments confirm that these contributions collectively enable FedScholar to significantly outperform state-of-the-art baselines, paving the way for trustworthy and privacy-centric autonomous research agents.

References

[1] Durmus Alp Emre Acar, Yue Zhao, Ramon Matas Navarro, Matthew Mattina, Paul N Whatmough, and Venkatesh Saligrama. 2021. Federated learning based on dynamic regularization. *arXiv preprint arXiv:2111.04263* (2021).

[2] Pushpita Chatterjee, Debashis Das, and Danda B Rawat. 2023. Federated learning empowered recommendation model for financial consumer services. *IEEE Transactions on Consumer Electronics* 70, 1 (2023), 2508–2516.

[3] Avishek Ghosh, Jichan Chung, Dong Yin, and Kannan Ramchandran. 2020. An efficient framework for clustered federated learning. *Advances in neural information processing systems* 33 (2020), 19586–19597.

[4] Zhihan Guo, Jiele Wu, Wenqian Cui, Yifei Zhang, Minda Hu, Yufei Wang, and Irwin King. 2025. From General Reward to Targeted Reward: Improving Open-ended Long-context Generation Models. In *Proceedings of the 2025 Conference on Empirical Methods in Natural Language Processing*, Christos Christodoulopoulos, Tanmoy Chakraborty, Carolyn Rose, and Violet Peng (Eds.). Association for Computational Linguistics, Suzhou, China, 5151–5166. doi:10.18653/v1/2025.emnlp-main.260

[5] Li Huang, Andrew L. Shea, Huining Qian, Aditya Masurkar, Hao Deng, and Dianbo Liu. 2019. Patient clustering improves efficiency of federated machine learning to predict mortality and hospital stay time using distributed electronic medical records. *Journal of Biomedical Informatics* 99 (2019), 103291.

[6] Sai Praneeth Karimireddy, Satyen Kale, Mehryar Mohri, Sashank Reddi, Sebastian Stich, and Ananda Theertha Suresh. 2020. Scaffold: Stochastic controlled averaging for federated learning. In *International conference on machine learning*. PMLR, 5132–5143.

[7] Chul Min Lee, Joaquín Delgado Fernández, Sergio Potenciano Menci, Alexander Rieger, and Gilbert Fridgen. 2023. Federated Learning for Credit Risk Assessment.. In *HICSS*. 386–395.

[8] Tian Li, Anit Kumar Sahu, Manzil Zaheer, Maziar Sanjabi, Ameet Talwalkar, and Virginia Smith. 2020. Federated optimization in heterogeneous networks. *Proceedings of Machine learning and systems* 2 (2020), 429–450.

[9] Brendan McMahan, Eider Moore, Daniel Ramage, Seth Hampson, and Blaise Aguera y Arcas. 2017. Communication-efficient learning of deep networks from decentralized data. In *Artificial intelligence and statistics*. PMLR, 1273–1282.

[10] Jaehoon Oh, Sangmook Kim, and Se-Young Yun. 2021. Fedbabu: Towards enhanced representation for federated image classification. *arXiv preprint arXiv:2106.06042* (2021).

[11] OpenAI. 2024. GPT-4o mini. https://www.openai.com/. Language model.

[12] OpenAI. 2024. Hello gpt-4o. https://openai.com/index/hello-gpt-4o/. Language model.

[13] Sashank J. Reddi, Zachary Charles, Manzil Zaheer, Zachary Garrett, Keith Rush, Jakub Konečný, Sanjiv Kumar, and Hugh Brendan McMahan. 2021. Adaptive Federated Optimization. In *9th International Conference on Learning Representations, ICLR 2021, Virtual Event, Austria, May 3-7, 2021*. OpenReview.net. https://openreview.net/forum?id=LkFG3lB13U5

[14] Karthik V Sarma, Stephanie Harmon, Thomas Sanford, Holger R Roth, Ziyue Xu, Jesse Tetreault, Daguang Xu, Mona G Flores, Alex G Raman, Rushikesh Kulkarni, et al. 2021. Federated learning improves site performance in multicenter deep learning without data sharing. *Journal of the American Medical Informatics Association* 28, 6 (2021), 1259–1264.

[15] Felix Sattler, Klaus-Robert Müller, and Wojciech Samek. 2020. Clustered federated learning: Model-agnostic distributed multitask optimization under privacy constraints. *IEEE transactions on neural networks and learning systems* 32, 8 (2020), 3710–3722.

[16] Ross Taylor, Marcin Kardas, Guillem Cucurull, Thomas Scialom, Anthony Hartshorn, Elvis Saravia, Andrew Poulton, Viktor Kerkez, and Robert Stojnic. 2022. Galactica: A Large Language Model for Science. arXiv:2211.09085 [cs.CL] https://arxiv.org/abs/2211.09085

[17] Qwen Team. 2024. Qwen2. 5: A party of foundation models, September 2024. *URL https://qwenlm. github. io/blog/qwen2* 5 (2024).

[18] Zhiyuan Wen, Jiannong Cao, Zian Wang, Beichen Guo, Ruosong Yang, and Shuaiqi Liu. 2025. Interactivesurvey: An llm-based personalized and interactive survey paper generation system. *arXiv preprint arXiv:2504.08762* (2025).

[19] WordsRated. 2025. Number of Academic Papers Published per Year. Accessed: 2025-08-27.

[20] Dong Yang, Ziyue Xu, Wenqi Li, Andriy Myronenko, Holger R Roth, Stephanie Harmon, Sheng Xu, Baris Turkbey, Evrim Turkbey, Xiaosong Wang, et al. 2021. Federated semi-supervised learning for COVID region segmentation in chest CT using multi-national data from China, Italy, Japan. *Medical image analysis* 70 (2021), 101992.

[21] Qiang Yang, Yang Liu, Yong Cheng, Yan Kang, Tianjian Chen, and Han Yu. 2019. *Federated Learning*. Morgan & Claypool Publishers. doi:10.2200/S00960ED2V01Y201910AIM043

[22] Rui Ye, Wenhao Wang, Jingyi Chai, Dihan Li, Zexi Li, Yinda Xu, Yaxin Du, Yanfeng Wang, and Siheng Chen. 2024. Openfedllm: Training large language models on decentralized private data via federated learning. In *Proceedings of the 30th ACM SIGKDD conference on knowledge discovery and data mining*. 6137–6147.

[23] Qiyuan Zhang, Fuyuan Lyu, Zexu Sun, Lei Wang, Weixu Zhang, Wenyue Hua, Haolun Wu, Zhihan Guo, Yufei Wang, Niklas Muennighoff, Irwin King, Xue Liu, and Chen Ma. 2025. A Survey on Test-Time Scaling in Large Language Models: What, How, Where, and How Well? arXiv:2503.24235 [cs.CL] https://arxiv.org/abs/2503.24235

[24] Yukun Zhang, Guanzhong Chen, Zenglin Xu, Jianyong Wang, Dun Zeng, Junfan Li, Jinghua Wang, Yuan Qi, and Irwin King. 2024. Fedcvd: The first real-world federated learning benchmark on cardiovascular disease data. *arXiv preprint arXiv:2411.07050* (2024).

[25] Zhuo Zhang, Xiangjing Hu, Jingyuan Zhang, Yating Zhang, Hui Wang, Lizhen Qu, and Zenglin Xu. 2023. Fedlegal: The first real-world federated learning benchmark for legal nlp. In *Proceedings of the 61st Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*. 3492–3507.

[26] Yuxiang Zheng, Dayuan Fu, Xiangkun Hu, Xiaojie Cai, Lyumanshan Ye, Pengrui Lu, and Pengfei Liu. 2025. Deepresearcher: Scaling deep research via reinforcement learning in real-world environments. *arXiv preprint arXiv:2504.03160* (2025).

A Prompt

Annotation Prompt

You are an expert academic assistant. Given an academic paper title and content, please complete the following two tasks and output the result in strict JSON format.

Task 1: Generate a Professional Research Question

Generate a single, comprehensive research question that requires a detailed response.
- The question must be based ONLY on the provided paper content.
- Do NOT explicitly mention the paper title.
- The question must be open-ended and comprehensive.
- Crucial Requirement: The question must implicitly or explicitly cover these 5 dimensions:
1. Research Background
2. Research Problem
3. Research Methodology
4. Experimental Performance
5. Contributions to the field.

Task 2: Generate Checklist Question-Answer Pairs

Identify key aspects of the paper and create more than 15 checklist question-answer pairs.
- These questions serve as verification points for the research question above.
- Focus on key concepts and logic, avoiding trivial details like specific hyper-parameter numbers.
- Do NOT mention the limitations of the paper.
- The answer must be a Boolean string: "True" or "False".

Output Format

You must output a valid JSON object with the following structure:

```
{
"research_question": "String",
"checklist": [
{
"question": "String",
"answer": "True/False"
},
...
```

]
}

Academic Paper

Paper Title: PAPER TITLE
Paper Content: PAPER CONTENT

B Publication

We collect 23 top-tier publications: NB (Nature Biotechnology), AMR (American Economic Review), AC (Automation In Construction), JPE (Journal Of Political Economy), JCO (Journal Of Clinical Oncology), ES&T(Environmental Science & Technology), AJRCCM (American Journal Of Respiratory And Critical Care Medicine), TPAMI (IEEE Transactions On Pattern Analysis And Machine Intelligence), AEM (Advanced Energy Materials), ECONOMETRICA, JTACS (Journal Of The American Chemical Society), BIOMATERIALS, EPSL (Earth And Planetary Science Letters), NE (Nature Energy), QJE (Quarterly Journal Of Economics), NC (Nature Climate Change), TNNLS (IEEE Transactions On Neural Networks And Learning Systems), BM (Bioactive Materials), AM (Advanced Materials), LN (Lancet Neurology).

Received 18 December 2025