A Survey on Efficient Large Language Model Training: From Data-centric Perspectives

Junyu Luo[1], **Bohan Wu**[1], **Xiao Luo**[2†], **Zhiping Xiao**[3†], **Yiqiao Jin**[4], **Rong-Cheng Tu**[5], **Nan Yin**[6], **Yifan Wang**[7], **Jingyang Yuan**[1], **Wei Ju**[1], **Ming Zhang**[1†]
[1] State Key Laboratory for Multimedia Information Processing,
School of Computer Science, PKU-Anker LLM Lab, Peking University
[2] University of California, Los Angeles [3] University of Washington
[4] Georgia Institute of Technology [5] Nanyang Technological University
[6] HKUST [7] University of International Business and Economics
https://github.com/luo-junyu/Awesome-Data-Efficient-LLM

Abstract

Post-training of Large Language Models (LLMs) is crucial for unlocking their task generalization potential and domain-specific capabilities. However, the current LLM post-training paradigm faces significant data challenges, including the high costs of manual annotation and diminishing marginal returns on data scales. Therefore, achieving data-efficient post-training has become a key research question. In this paper, we present the first systematic survey of data-efficient LLM post-training from a data-centric perspective. We propose a taxonomy of data-efficient LLM post-training methods, covering data selection, data quality enhancement, synthetic data generation, data distillation and compression, and self-evolving data ecosystems. We summarize representative approaches in each category and outline future research directions. By examining the challenges in data-efficient LLM post-training, we highlight open problems and propose potential research avenues. We hope our work inspires further exploration into maximizing the potential of data utilization in large-scale model training.

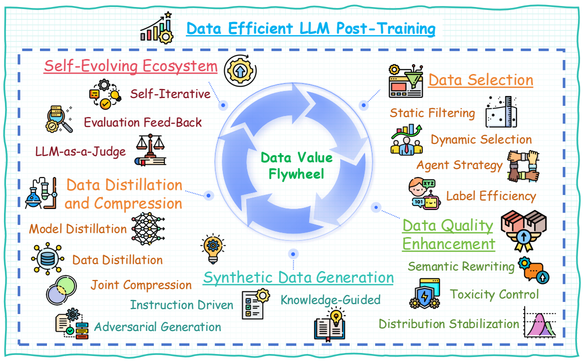

Figure 1: Illustration of the data flywheel in **Data-Efficient LLM Post-Training**, depicting the iterative cycle of data selection, data quality enhancement, synthetic data generation, knowledge distillation, and self-evolving data ecosystems to maximize model performance with minimal data requirements.

1 Introduction

Large Language Models (LLMs) post-training has emerged as a crucial stage for unlocking their domain adaptation capabilities and task generalization potential (Luo et al., 2025b). This phase has effectively enhanced models' abilities in long-context reasoning (Zelikman et al., 2022; Yuan et al., 2024c), human alignment (Rafailov et al., 2024), instruction tuning (Zhang et al., 2023b), and domain-specific adaptation (Cheng et al., 2024).

During the LLM post-training phase, data is the essential driver of model evolution. However, the current paradigm faces a severe *data dilemma*: the cost of manually annotating high-quality data is rapidly growing, while simply scaling data volume yields diminishing returns. Moreover, static datasets inherently limit models from adapting to evolving real-world knowledge. The linear dependency between data volume and model performance fundamentally stems from the inefficient data usage in traditional post-training paradigms. The recent success of DeepSeek-R1 (Guo et al., 2025), which leverages reinforcement learning for data-efficient post-training, further demonstrates the effectiveness and necessity of data-efficient approaches in achieving superior LLM performance. Our work establishes the **first systematic survey on data-efficient training** of LLMs, providing a unified, taxonomized framework to address the fragmented research landscape. Our survey reveals that breaking through efficiency bottlenecks requires establishing value extraction across the lifecycle, rather than merely expanding data scale.

Researchers have explored various approaches to fully exploit the data potential in LLM post-training (Jeong et al., 2024; Wang et al., 2024a; Pan et al., 2023b). While these methods have made notable progress in improving data efficiency, the

†Corresponding authors.

Proceedings of the 63rd Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers), pages 30904–30920
July 27 - August 1, 2025

field still lacks a comprehensive review. In this paper, we provide a comprehensive survey of data-efficient LLM post-training from a data-centric perspective. Specifically, we introduce the concept of a *data value flywheel* (as illustrated in Figure 1), which consists of five key components: data selection, data quality enhancement, synthetic data generation, data distillation and compression, and self-evolving data ecosystems. Using this framework, we present a taxonomy of existing work, summarize key components, and identify promising research directions. We hope our work serves as both a useful roadmap for newcomers and a guide for future advancements in the field.

Differences from previous surveys. While several surveys have explored a few aspects of LLMs post-training, including data selection (Wang et al., 2024b), synthetic data generation (Long et al., 2024; Tan et al., 2024), model self-feedback (Liang et al., 2024a; Pan et al., 2023a), self-evolution (Tao et al., 2024), trustworthiness (Liu et al., 2023), and time-efficiency (Wan et al., 2023), these studies primarily focus on individual aspects rather than a holistic perspective. Our survey fills this gap by systematically examining these methods through the lens of data efficiency, offering critical insights into maximizing data value extraction.

2 Taxonomy

This section categorizes data-efficient post-training methods for LLMs into five core methodologies:

- **Data Selection**: *Filtering high-value subsets from raw data.* ❶ Static Filtering: Offline selection based on data properties; ❷ Dynamic Selection: Adjusting weights based on model uncertainty; ❸ Agent Strategy: Multi-model voting for reliable selection; ❹ Labeling Efficiency: Combining active learning and semi-supervised strategies for cost-effective sample coverage.
- **Data Quality Enhancement**: *Improving the utility of existing data.* ❶ Semantic Rewriting: Enhancing expression diversity through semantic-preserving transformations and generating variants while maintaining original meaning; ❷ Toxicity Control: Correcting harmful content; ❸ Distribution Stabilization: Adjusting data characteristics for robustness
- **Synthetic Data Generation**: *Creating new training data.* ❶ Instruction-Driven: Model-generated instruction-response pairs; ❷ Knowledge-Guided: Generation with structured knowledge; ❸ Adversarial Generation: Producing challenging samples.
- **Data Distillation and Compression**: *Extracting core knowledge for efficient training.* ❶ Model Distillation: Transferring large model output distributions to smaller models while preserving key knowledge; ❷ Data Distillation: Extracting high information density samples to construct compact datasets equivalent to full-scale data; ❸ Joint Compression: Combining model architecture compression with data selection strategies for end-to-end efficiency optimization
- **Self-Evolving Data Ecosystem**: *Building self-evolution mechanisms.* ❶ Self-Iterative Optimization: Using current model to generate data; ❷ Dynamic Evaluation Feedback: Real-time monitoring and adjustment; ❸ LLM-as-a-Judge: Feedback-Driven Data Optimization;

Category	**Data Dependency**	**Compute Cost**	**Model Dependency**	**Data Value Mining**
Data Selection	++	+	+	+++
Quality Enhance.	++	++	++	++
Synthetic Generation	+	+++	+++	+
Distill. & Compress.	+	+	+++	+++
Self-Evolving	+	+++	+++	+++

Table 1: Comparison of different data-efficient post-training methods across key dimensions.

Table 1 compares the five methodologies across key dimensions, where more '+' indicates higher requirements or better performance. Data selection shows high data efficiency but requires quality source data. Quality enhancement maintains balanced requirements across dimensions. Synthetic generation and self-evolving approaches demand more compute and model resources but reduce dependency on original manually annotated datasets, as they primarily rely on teacher model outputs or self-generated data. Distillation methods excel in data efficiency while depending on model capabilities.

These five dimensions complement each other: selection filters quality data, enhancement improves utility, generation expands coverage, distillation concentrates knowledge, and self-evolution enables continuous improvement. Together, they pursue the goal of *less data, higher returns*.

3 Data Selection

Data selection is crucial for enhancing LLM post-training efficiency by identifying high-value data subsets. As shown in Figure 3, we divide existing approaches into four dimensions: (1) static filter-

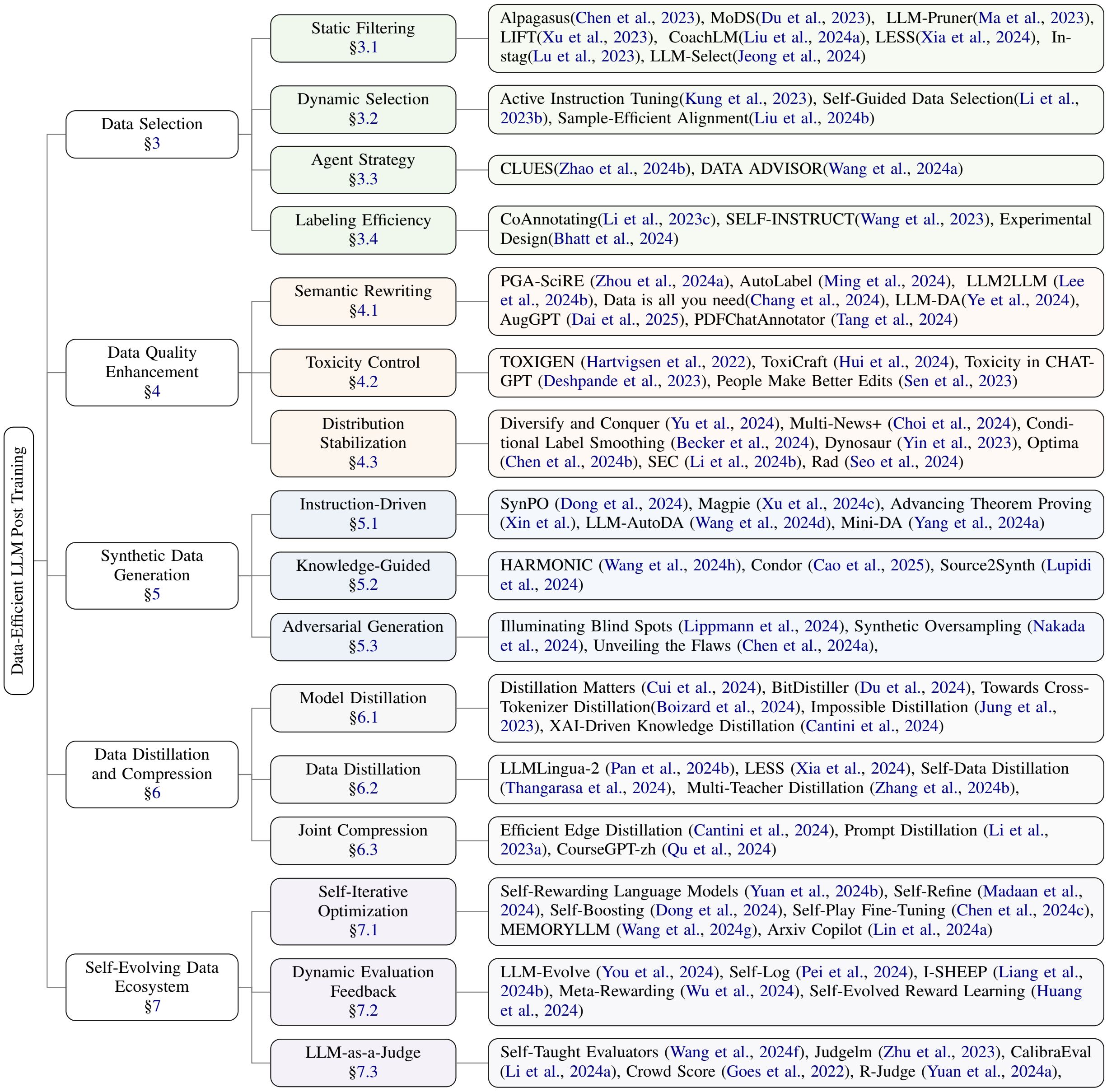

Figure 2: A taxonomy of Data-Efficient LLM Post Training.

ing based on inherent data properties, (2) dynamic selection that adapts during training, (3) agent strategy using collaborative mechanisms, and (4) labeling efficiency through human-AI collaboration.

3.1 Static Filtering

Static filtering evaluates inherent data properties offline to identify samples with high information density and representativeness.

Quality-based Filtering. Alpagasus (Chen et al., 2023) achieves comparable performance using only 17% of the original data through complexity-based filtering (instruction length, diversity, and perplexity). MoDS (Du et al., 2023) employs multi-dimensional indicators and density peak clustering, while (Kang et al., 2024) uses KL-divergence-driven selection to align domain distributions. Information-theoretic approaches (Kim and Baek, 2024) leverage entropy metrics using negative log-likelihood and inverse word frequency to identify redundant samples. ActivePrune (Azeemi et al., 2024) implements two-stage pruning through n-gram perplexity scoring followed by quantized LLM evaluation. CoT-Influx (Huang et al., 2023) employs coarse-to-fine pruning for reasoning enhancement in mathematical tasks. LLM-Select (Jeong et al., 2024) demonstrates that large language models can perform feature selection using only feature names and task descriptions, rivaling traditional data science tools.

Semantic Enhancement. LIFT (Xu et al., 2023) enhances instruction quality through automatic revi-

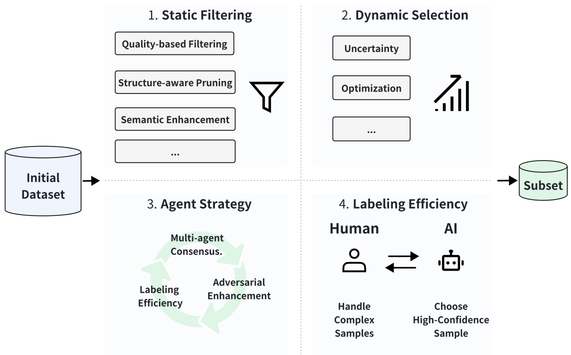

Figure 3: Overview of four major data selection approach categories: static filtering, dynamic selection, agent strategy, and labeling efficiency.

sion. InsTag (Lu et al., 2023) proposes fine-grained instruction tagging to analyze diversity and complexity in supervised fine-tuning datasets, demonstrating that model capability grows with more diverse and complex data.

3.2 Dynamic Selection

Dynamic methods adapt data weights by evaluating sample importance based on model feedback.

Uncertainty-driven Selection. Active Instruction Tuning (Kung et al., 2023) prioritizes high-uncertainty tasks through prediction entropy. Self-Guided Data Selection uses Instruction Following Difficulty (IFD) to measure loss variance and eliminate easily learned examples (Li et al., 2023b).

Optimization-based Selection. Sample-efficient alignment (Liu et al., 2024b) uses Thompson sampling to maximize contribution in preference alignment tasks. Compute-constrained data selection (Yin and Rush, 2024) optimizes between data utility and computational cost. P3 (Yang et al., 2024b) integrates policy-driven difficulty assessment with pace-adaptive selection and diversity promotion through Determinantal Point Process. LESS (Xia et al., 2024) employs optimizer-aware gradient similarity search with low-rank gradient features for targeted instruction tuning. In domain-specific applications, data pruning methods (Lin et al., 2024b) use influence and effort scores to identify representative samples.

3.3 Agent Strategy

Agent-based approaches leverage collaborative mechanisms for reliable selection.

Multi-agent Consensus. Multi-agent methods like CLUES (Zhao et al., 2024b) implement multi-model voting mechanisms based on training dynamics and gradient similarity metrics.

Adversarial Enhancement. Recent works like DATA ADVISOR (Wang et al., 2024a) uses red-team agents for safety filtering, while Automated Data Curation (Chen and Mueller, 2024) optimizes data through generator-discriminator frameworks.

3.4 Labeling Efficiency

These methods efficiently optimize annotation processes through iterative human-AI collaboration.

Human-AI Collaboration. Methods like LLMaAA (Zhang et al., 2023a) employ LLMs as annotators with uncertainty sampling. CoAnnotating (Li et al., 2023c) implements uncertainty-guided labor division between humans and AI.

Automated Generation. SELF-INSTRUCT (Wang et al., 2023) enables autonomous self-generated instruction data, while (Li et al., 2023d) uses one-shot learning for rapid sample identification.

Workflow Optimization. Recent works establish scalabel efficient annotation workflows through adaptive experimental design (Bhatt et al., 2024) and systematic curation systems (Pang et al., 2024).

3.5 Discussion

Current data selection approaches face challenges in aligning static metrics with dynamic model requirements, managing computational complexity in optimization, and achieving cross-domain generalization. Future research points toward meta-learning-based selection frameworks, causal inference for sample analysis, and efficiency-aware optimization with hardware constraints, advancing data selection toward theoretical grounding.

4 Data Quality Enhancement

As illustrated in Figure 4, enhancing data quality is critical for maximizing the effectiveness of LLM post-training (Zhou et al., 2024b). Through semantic refinement, toxicity control, and distribution stabilization, researchers aim to improve the informativeness, safety, and robustness of training data. We categorize existing methods into three directions.

4.1 Semantic Rewriting

Semantic rewriting focuses on augmenting data diversity while preserving original meaning through controlled transformations. This can be achieved through several key approaches:

Instruction Refinement. CoachLM (Liu et al., 2024a) automatically revises complex instructions to reduce ambiguity, while (Li et al., 2024c) uses

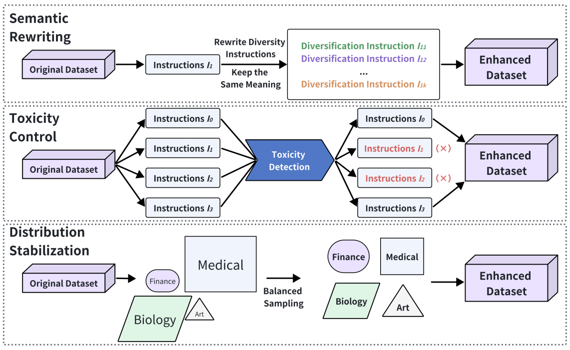

Figure 4: Three key approaches for data quality enhancement in LLM post-training: semantic rewriting for diversity, toxicity control for safety, and distribution stabilization for balanced representation.

structured prompt chains for paraphrase generation, enhancing model generalization across tasks.

Domain-Specific Augmentation. Methods like (Jia et al., 2024) use curriculum learning for metaphor detection, while PGA-SciRE (Zhou et al., 2024a) injects structured knowledge for scientific relation extraction, adapting models to specialized tasks.

Automated Enhancement. AutoLabel (Ming et al., 2024) seamlessly integrates human feedback for quality rewriting, while LLM2LLM (Lee et al., 2024b) iteratively improves low-confidence samples. LANCE (Wang et al., 2024c) enables LLMs to autonomously generate, clean, review, and annotate data, serving as continuous self-evolving data engineers. Recent studies extensively explore human-AI collaboration (Chung et al., 2023) and various data types: text (Dai et al., 2025), tabular (Banday et al., 2024), and multimodal (Tang et al., 2024). LLM-DA (Ye et al., 2024) employs contextual rewriting strategies with entity-level replacements for few-shot NER tasks, while (Zhang et al., 2025) leverages lightweight LLM generation and tree hybridization for cross-domain parsing augmentation.

4.2 Toxicity Control

Mitigating harmful content is crucial for data quality enhancement. Recent approaches focus on detection, benchmarking, and human collaboration:

Detection Frameworks. Methods like (Zhang et al., 2024a) effectively distill toxicity knowledge into compact detectors, while (Wang and Chang, 2022) strategically leverages generative prompts for zero-shot toxicity classification across diverse tasks.

Adversarial Benchmarking. Frameworks such as TOXIGEN (Hartvigsen et al., 2022) and ToxiCraft (Hui et al., 2024) generate adversarial datasets to stress-test models. Studies (Luong et al., 2024; Deshpande et al., 2023; Chetnani, 2023; Oh et al., 2024) examine the relationship between model size and toxicity generation, finding that smaller models often exhibit lower toxicity rates.

Human-AI Collaboration. Research demonstrates that human intervention significantly improves toxicity detection quality (Sen et al., 2023), particularly through counterfactual data augmentation. Additional work explores covert toxicity detection (Lee et al., 2024a), data contamination (Balloccu et al., 2024), and geometric interpretability (Balestriero et al., 2024) to enhance model safety.

4.3 Distribution Stabilization

Stabilizing data distribution ensures that models generalize well across different tasks and domains. Several methods tackle issues like class imbalance, noise reduction, and domain adaptation:

Imbalance Mitigation. Approaches like Synthetic Oversampling (Nakada et al., 2024) and Diversify and Conquer (Yu et al., 2024) effectively address class imbalance through adaptive synthetic sample generation. Studies show significant improvements, with (Cai et al., 2023) demonstrating a 38% fairness boost in cross-disciplinary applications.

Noise Reduction. Multi-News+ (Choi et al., 2024) signficantly reduces annotation errors through automated label correction, while (Chen and Mueller, 2024) employs self-supervised filtering for robust fine-tuning data curation. RobustFT (Luo et al., 2024) introduces a comprehensive framework for handling noisy response data through multi-expert collaborative noise detection and context-enhanced relabeling strategies, coupled with entropy-based data selection for high-quality sample retention.

Domain Adaptation. ChatTS (Xie et al., 2024) uses Fourier transforms for time-series alignment, while (Becker et al., 2024) applies domain-specific label smoothing for clinical text. Advanced approaches like Dynosaur (Yin et al., 2023) and Optima (Chen et al., 2024b) leverage curriculum learning and multi-source coordination. RADA (Seo et al., 2024) addresses low-resource domain tasks by retrieving relevant instances from other datasets and generating contextually enhanced samples through LLM prompting.

4.4 Discussion

Semantic rewriting, toxicity control, and distribution stabilization represent key strategies for improving data quality in LLM post-training. These

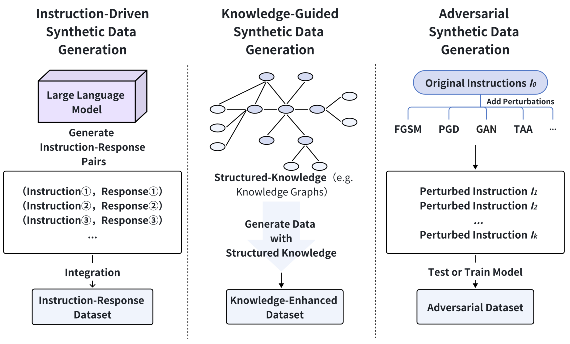

Figure 5: Three main approaches for data generation in LLM post-training: instruction-driven generation for creating instruction-response pairs, knowledge-guided generation using structured knowledge, and adversarial generation for testing model robustness.

techniques ensure the generation of diverse, high-quality data, mitigate harmful content, and stabilize data distributions to enhance model robustness. Future work should integrate these approaches into unified frameworks to maximize data diversity and model performance while reducing costs.

5 Synthetic Data Generation

Generating synthetic training data is a powerful strategy to overcome data scarcity and enhance the robustness of LLM post-training. As illustrated in Figure 5, synthetic data generation methods can be categorized into three main approaches: *Instruction-Driven*, *Knowledge-Guided*, and *Adversarial Generation*, each serving distinct purposes in enhancing model capabilities.

5.1 Instruction-Driven Synthetic Data Generation

Instruction-driven methods harness LLMs' ability to produce new examples directly from task prompts. Recent works demonstrate diverse applications: SynPO (Dong et al., 2024) generates preference pairs for alignment (12% ROUGE-L improvement), Magpie (Xu et al., 2024c) enables template-free instruction generation (98% AlpacaEval accuracy), and Advancing Theorem Proving (Xin et al.) synthesizes Lean4 proof steps, boosting GPT-4's proving capabilities by 34%.

5.2 Knowledge-Guided Synthetic Data Generation

Knowledge-guided approaches integrate external knowledge to steer data generation.

Theoretical Frameworks. Towards a Theoretical Understanding (Gan and Liu, 2024) rigorously establishes a reverse-bottleneck theory linking data diversity to enhanced model generalization.

Structured Data Synthesis. HARMONIC (Wang et al., 2024h) combines privacy-preserving tabular data generation on medical records. (Xu et al., 2024b) improves relational consistency through schema-aware fine-tuning.

Cost-Effective Strategies. (Chan et al., 2024) demonstrates hybrid generation methods reduce API costs by 70% while maintaining data utility. Source2Synth (Lupidi et al., 2024) improves factual accuracy through knowledge-graph alignment.

5.3 Adversarial Generation

Adversarial generation methods systematically probe model vulnerabilities to enhance robustness. Recent works demonstrate diverse approaches: Illuminating Blind Spots (Lippmann et al., 2024) uses agent-based simulations to generate edge cases, reducing errors by 19% on dialect variation; Unveiling Synthetic Data Flaws (Chen et al., 2024a) introduces contrastive unlearning to address data imperfections, yielding 32% quality improvements on GLUE; and ToxiCraft (Hui et al., 2024) generates subtle harmful content, revealing significant gaps in commercial safety filters.

5.4 Discussion

Each approach offers distinct trade-offs: instruction-driven methods enable rapid scaling but risk semantic drift; knowledge-guided approaches maintain fidelity through structured constraints; and adversarial generation strengthens robustness by exposing vulnerabilities. Future work should combine these strengths—for instance, merging privacy-preserving generation with adversarial testing. Key challenges persist in optimizing generation costs (Chan et al., 2024) and developing theoretical foundations (Gan and Liu, 2024).

6 Data Distillation and Compression

Data distillation and compression techniques enhance LLM post-training efficiency by reducing data complexity while preserving performance. As shown in Figure 6, this involves three complementary approaches: model distillation for knowledge transfer, data distillation for dataset compression, and joint compression for unified optimization.

6.1 Model Distillation

Model distillation transfers knowledge from large to smaller models while maintaining perfor-

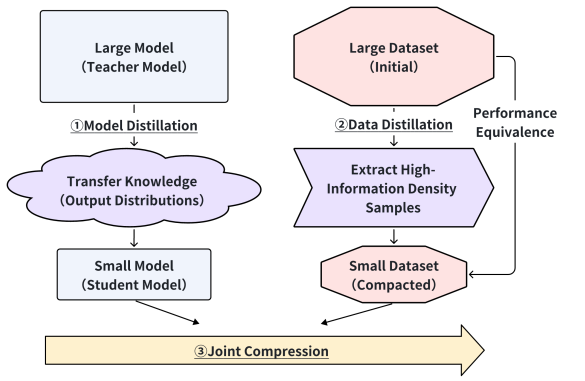

Figure 6: Data distillation and compression in LLM post-training: model distillation for knowledge transfer, data distillation for sample extraction, and joint compression for unified optimization.

mance. Recent advances include Impossible Distillation (Jung et al., 2023), which creates high-quality models from low-quality teachers, and Performance-Guided Distillation (Di Palo et al., 2024), achieving 98% accuracy with 40% reduced costs. Cross-Tokenizer Distillation (Boizard et al., 2024) enables knowledge transfer between different architectures through universal logit distillation. For edge deployment, XAI-Driven Distillation (Cantini et al., 2024) produces interpretable medical models, while BitDistiller (Du et al., 2024) enables sub-4-bit precision with minimal accuracy loss. Multistage Collaborative Distillation (Zhao et al., 2024a) improves performance through multi-teacher coordination in low-resource settings.

6.2 Data Distillation

Data distillation focuses on selecting high-information-density samples to create compact yet representative datasets. Knowledge Distillation in Automated Annotation (Pangakis and Wolken, 2024) shows LLM-generated labels can effectively train classifiers comparable to human annotations. LLMLingua-2 (Pan et al., 2024b) approaches prompt compression through token-level distillation. Domain-specific applications include Self-Data Distillation (Thangarasa et al., 2024) for model refinement, Multi-Teacher Distillation (Zhang et al., 2024b) for healthcare data integration, and techniques for reducing hallucination (McDonald et al., 2024).

6.3 Joint Compression

Joint compression combines model compression with data selection to optimize overall efficiency. Compact Language Models via Pruning and Distillation (Muralidharan et al., 2024) co-optimizes structural pruning and label smoothing, compressing LLaMA-7B to 2.8B parameters with minimal performance loss. Efficient Edge Distillation (Cantini et al., 2024) enables adaptive width scaling for edge devices through supernet training. In recommendation systems, Prompt Distillation (Li et al., 2023a) aligns ID-based and text-based representations, aiming to reduce inference time.

For multimodal applications, recent work demonstrates joint compression of graph and text encoders (Pan et al., 2024a) and curriculum-aligned prompt distillation for educational LLMs (Qu et al., 2024), achieving significant parameter reduction while maintaining performance.

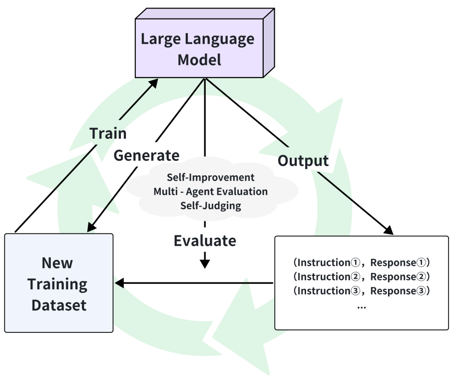

Figure 7: Self-Evolving Data Ecosystem: autonomous data generation, real-time feedback, and continuous learning.

6.4 Discussion

These three approaches offer complementary benefits for enhancing LLM efficiency: model distillation optimizes architecture, data distillation curates high-impact samples, and joint compression unifies model-data optimization. Future research should focus on integrating these methods, particularly for edge AI and low-resource applications.

7 Self-Evolving Data Ecosystem

The Self-Evolving Data Ecosystem strategically optimizes LLM post-training through autonomous data generation, real-time feedback, and continuous learning. As shown in Figure 7, this ecosystem forms a closed loop of generation, evaluation, and adaptive training. We discuss three key components: Self-Iterative Optimization, Dynamic Evaluation Feedback, and LLM-as-a-Judge.

7.1 Self-Iterative Optimization

Self-iterative optimization enables LLMs to use their own outputs to generate new training data, refining their capabilities autonomously. Several approaches illustrate this concept:

Self-Improvement Methods. Recent works like Self-Rewarding (Yuan et al., 2024b), Self-Refine (Madaan et al., 2024), and Self-Boosting (Dong et al., 2024) enable models to autonomously improve through iterative self-optimization. Self-Play Fine-Tuning (Chen et al., 2024c) extends this by leveraging competitive self-interaction, outperforming traditional methods like DPO (Rafailov et al., 2024).

Semi-Supervised Self-Evolution. In practical deployment scenarios, models often encounter limited labeled seed data alongside abundant unlabeled domain-specific data, creating a critical challenge for effective post-training adaptation. SemiEvol (Luo et al., 2025a) addresses this challenge through a propagate-and-select framework that transfers knowledge from seed data to unlabeled samples via bi-level propagation and collaborative selection mechanisms.

Knowledge Retention. In the context of retaining knowledge while integrating new data, MemoryLLM (Wang et al., 2024g) enables continuous model updates while preserving existing knowledge. Automated Proof Generation (Chen and Mueller, 2024) and Arxiv Copilot (Lin et al., 2024a) demonstrate this capability in code verification and academic research tasks.

7.2 Dynamic Evaluation Feedback

Dynamic evaluation feedback systems allow models to make real-time adjustments based on their performance, optimizing their outputs on the fly. Key contributions include:

Multi-Agent Evaluation. The Benchmark Self-Evolving Framework (Wang et al., 2024e) and LLM-Evolve (You et al., 2024) employ multi-agent systems to evaluate and adjust LLM performance dynamically. These frameworks enable the models to self-adjust in real-time across various benchmarks, ensuring continuous evolution.

Iterative Refinement Self-Refine (Madaan et al., 2024) and Self-Log (Pei et al., 2024) employ feedback loops for iterative refinement and log parsing, optimizing the model's output without requiring external retraining.I-SHEEP (Liang et al., 2024b) offers a resource-efficient paradigm that enhances performance through self-alignment, while Interactive Evolution: A Neural-Symbolic Self-Training Framework (Xu et al., 2024a) enables LLMs to autonomously train in neural-symbolic environments.

Improved Decision Making. For improving model alignment, Meta-Rewarding (Wu et al., 2024) and Self-Evolved Reward Learning (Huang et al., 2024) leverage iterative feedback from their outputs to improve judgment skills, ensuring more accurate decision-making in complex tasks.

7.3 LLM-as-a-Judge

LLM-as-a-Judge systems represent a paradigm shift from external evaluation to self-assessment, where models evaluate their own or other models' outputs. These systems operate through three fundamental mechanisms, each addressing different evaluation challenges:

Self-Improvement through Judgment. These methods focus on improving a model's ability to assess quality. Self-Taught Evaluators (Wang et al., 2024f) and Meta-Rewarding (Wu et al., 2024) take distinct approaches: the former generates synthetic comparisons to train judgment without human data, while the latter introduces meta-judgment by having models critique their own evaluations. JudgeLM (Zhu et al., 2023) takes a different path by fine-tuning models on human preferences to create specialized evaluation models.

Debiasing Evaluation Systems. These methods address fairness concerns in automated evaluation. CalibraEval (Li et al., 2024a) recalibrates prediction distributions to mitigate position bias, while Crowd Score (Goes et al., 2022) employs multiple AI *personalities* within a single model to simulate diverse human judgments, reducing individual bias through aggregation.

Adversarial Robustness Testing. These approaches stress-test models through challenging scenarios. TOXIGEN (Hartvigsen et al., 2022) and ToxiCraft (Hui et al., 2024) create progressively more subtle toxic content to expose blind spots, while R-Judge (Yuan et al., 2024a) specifically targets situational safety risks in interactive environments rather than just content harmfulness.

7.4 Discussion

The combination of Self-Iterative Optimization, Dynamic Evaluation Feedback, and LLM-as-a-Judge creates a robust framework for autonomous LLM improvement. While these approaches show promise in reducing manual intervention, future work should focus on unifying them into scalable frameworks that generalize across diverse tasks.

8 Challenges and Future Directions

Domain-Driven Data Synthesis and Refinement. While general-purpose models like GPT are commonly used for data generation (Di Palo et al., 2024), domain-specific models can better capture professional knowledge (Lightman et al., 2023). Future work should explore domain-specific pre-trained models for generating specialized data (Luo et al., 2023; Cheng et al., 2024), along with refinement techniques to optimize data quality while reducing annotation costs.

Scalability of Large-Scale Data Synthesis. As LLM pre-training demands increasingly larger and higher-quality datasets, efficient large-scale data generation becomes crucial. Current data synthesis and augmentation methods face scalability bottlenecks. Future work should focus on developing parallel, cost-effective, and efficient data generation frameworks that meet the demands of large-scale pre-training while maintaining a balance between data diversity and relevance (Karunya et al., 2023).

Reliable Quality Assessment Metrics. Current evaluation frameworks lack standardized metrics for assessing synthetic data quality (Zendel et al., 2024). Future research should develop metrics that evaluate semantic fluency, information accuracy, and potential biases (Chundawat et al., 2022; Gerstgrasser et al., 2024) to ensure robust selection.

9 Conclusion

In this paper, we presented a systematic review of LLM post-training research from a data efficiency perspective. We established the first taxonomic framework for data-efficient post-training, encompassing five core methodologies. Through detailed analysis of representative approaches within each category, we revealed that breaking through data efficiency bottlenecks requires establishing value extraction mechanisms across the entire data lifecycle. We aimed to highlight the current state and provide valuable insights for future work in this promising field of data-efficient LLM post-training.

Limitations

While our work presents the first comprehensive framework for analyzing data-efficient LLM post-training approaches, several limitations and opportunities for future research remain. First, given the explosive growth of this field, some emerging techniques may not be fully captured in our current taxonomic system, necessitating continuous updates to maintain comprehensiveness. Second, while data efficiency is crucial, the proposed methods may face additional challenges regarding trustworthiness and scalability that warrant further investigation. Furthermore, the synergistic effects and interaction mechanisms between different data efficiency enhancement techniques remain underexplored, calling for the development of cross-method optimization theories. We anticipate these open challenges will inspire deeper theoretical innovations and practical breakthroughs.

Acknowledgments

This paper is partially supported by grants from the National Key Research and Development Program of China with Grant No. 2023YFC3341203 and the National Natural Science Foundation of China (NSFC Grant Number 62276002). The authors are grateful to the anonymous reviewers for their efforts and insightful suggestions to improve this article.

References

Abdul Hameed Azeemi, Ihsan Ayyub Qazi, and Agha Ali Raza. 2024. Language model-driven data pruning enables efficient active learning. *arXiv preprint arXiv:2410.04275*.

Randall Balestriero, Romain Cosentino, and Sarath Shekkizhar. 2024. Characterizing large language model geometry helps solve toxicity detection and generation. In *Forty-first International Conference on Machine Learning*.

Simone Balloccu, Patrícia Schmidtová, Mateusz Lango, and Ondřej Dušek. 2024. Leak, cheat, repeat: Data contamination and evaluation malpractices in closed-source llms. *arXiv preprint arXiv:2402.03927*.

Banooqa Banday, Kowshik Thopalli, Tanzima Z Islam, and Jayaraman J Thiagarajan. 2024. On the role of prompt construction in enhancing efficacy and efficiency of llm-based tabular data generation. *arXiv preprint arXiv:2409.03946*.

Luca Becker, Philip Pracht, Peter Sertdal, Jil Uboreck, Alexander Bendel, and Rainer Martin. 2024. Conditional label smoothing for llm-based data augmentation in medical text classification. In *2024 IEEE Spoken Language Technology Workshop (SLT)*, pages 833–840. IEEE.

Gantavya Bhatt, Yifang Chen, Arnav M Das, Jifan Zhang, Sang T Truong, Stephen Mussmann, Yinglun Zhu, Jeffrey Bilmes, Simon S Du, Kevin Jamieson, et al. 2024. An experimental design framework for label-efficient supervised finetuning of large language models. *arXiv preprint arXiv:2401.06692*.

Nicolas Boizard, Kevin El Haddad, Céline Hudelot, and Pierre Colombo. 2024. Towards cross-tokenizer distillation: the universal logit distillation loss for llms. *arXiv preprint arXiv:2402.12030*.

Xunxin Cai, Meng Xiao, Zhiyuan Ning, and Yuanchun Zhou. 2023. Resolving the imbalance issue in hierarchical disciplinary topic inference via llm-based data augmentation. In *2023 IEEE International Conference on Data Mining Workshops (ICDMW)*, pages 1424–1429. IEEE.

Riccardo Cantini, Alessio Orsino, and Domenico Talia. 2024. Xai-driven knowledge distillation of large language models for efficient deployment on low-resource devices. *Journal of Big Data*, 11(1):63.

Maosong Cao, Taolin Zhang, Mo Li, Chuyu Zhang, Yunxin Liu, Haodong Duan, Songyang Zhang, and Kai Chen. 2025. Condor: Enhance llm alignment with knowledge-driven data synthesis and refinement. *arXiv preprint arXiv:2501.12273*.

Yung-Chieh Chan, George Pu, Apaar Shanker, Parth Suresh, Penn Jenks, John Heyer, and Sam Denton. 2024. Balancing cost and effectiveness of synthetic data generation strategies for llms. *arXiv preprint arXiv:2409.19759*.

Kaiyan Chang, Kun Wang, Nan Yang, Ying Wang, Dantong Jin, Wenlong Zhu, Zhirong Chen, Cangyuan Li, Hao Yan, Yunhao Zhou, et al. 2024. Data is all you need: Finetuning llms for chip design via an automated design-data augmentation framework. In *Proceedings of the 61st ACM/IEEE Design Automation Conference*, pages 1–6.

Jie Chen, Yupeng Zhang, Bingning Wang, Wayne Xin Zhao, Ji-Rong Wen, and Weipeng Chen. 2024a. Unveiling the flaws: Exploring imperfections in synthetic data and mitigation strategies for large language models. *arXiv preprint arXiv:2406.12397*.

Jiuhai Chen and Jonas Mueller. 2024. Automated data curation for robust language model fine-tuning. *arXiv preprint arXiv:2403.12776*.

Lichang Chen, Shiyang Li, Jun Yan, Hai Wang, Kalpa Gunaratna, Vikas Yadav, Zheng Tang, Vijay Srinivasan, Tianyi Zhou, Heng Huang, et al. 2023. Alpagasus: Training a better alpaca with fewer data. *arXiv preprint arXiv:2307.08701*.

Weize Chen, Jiarui Yuan, Chen Qian, Cheng Yang, Zhiyuan Liu, and Maosong Sun. 2024b. Optima: Optimizing effectiveness and efficiency for llm-based multi-agent system. *arXiv preprint arXiv:2410.08115*.

Zixiang Chen, Yihe Deng, Huizhuo Yuan, Kaixuan Ji, and Quanquan Gu. 2024c. Self-play fine-tuning converts weak language models to strong language models. *arXiv preprint arXiv:2401.01335*.

Daixuan Cheng, Shaohan Huang, and Furu Wei. 2024. Adapting large language models via reading comprehension. In *The Twelfth International Conference on Learning Representations*.

Yash Prakash Chetnani. 2023. Evaluating the impact of model size on toxicity and stereotyping in generative llm. Master's thesis, State University of New York at Buffalo.

Juhwan Choi, Jungmin Yun, Kyohoon Jin, and Young-Bin Kim. 2024. Multi-news+: Cost-efficient dataset cleansing via llm-based data annotation. *arXiv preprint arXiv:2404.09682*.

Vikram S Chundawat, Ayush K Tarun, Murari Mandal, Mukund Lahoti, and Pratik Narang. 2022. A universal metric for robust evaluation of synthetic tabular data. *IEEE Transactions on Artificial Intelligence*, 5(1):300–309.

John Joon Young Chung, Ece Kamar, and Saleema Amershi. 2023. Increasing diversity while maintaining accuracy: Text data generation with large language models and human interventions. *arXiv preprint arXiv:2306.04140*.

Yu Cui, Feng Liu, Pengbo Wang, Bohao Wang, Heng Tang, Yi Wan, Jun Wang, and Jiawei Chen. 2024. Distillation matters: empowering sequential recommenders to match the performance of large language models. In *Proceedings of the 18th ACM Conference on Recommender Systems*, pages 507–517.

Haixing Dai, Zhengliang Liu, Wenxiong Liao, Xiaoke Huang, Yihan Cao, Zihao Wu, Lin Zhao, Shaochen Xu, Fang Zeng, Wei Liu, et al. 2025. Auggpt: Leveraging chatgpt for text data augmentation. *IEEE Transactions on Big Data*.

Ameet Deshpande, Vishvak Murahari, Tanmay Rajpurohit, Ashwin Kalyan, and Karthik Narasimhan. 2023. Toxicity in chatgpt: Analyzing persona-assigned language models. *arXiv preprint arXiv:2304.05335*.

Flavio Di Palo, Prateek Singhi, and Bilal Fadlallah. 2024. Performance-guided llm knowledge distillation for efficient text classification at scale. *arXiv preprint arXiv:2411.05045*.

Qingxiu Dong, Li Dong, Xingxing Zhang, Zhifang Sui, and Furu Wei. 2024. Self-boosting large language models with synthetic preference data. *arXiv preprint arXiv:2410.06961*.

Dayou Du, Yijia Zhang, Shijie Cao, Jiaqi Guo, Ting Cao, Xiaowen Chu, and Ningyi Xu. 2024. Bitdistiller: Unleashing the potential of sub-4-bit llms via self-distillation. *arXiv preprint arXiv:2402.10631*.

Qianlong Du, Chengqing Zong, and Jiajun Zhang. 2023. Mods: Model-oriented data selection for instruction tuning. *arXiv preprint arXiv:2311.15653*.

Zeyu Gan and Yong Liu. 2024. Towards a theoretical understanding of synthetic data in llm post-training: A reverse-bottleneck perspective. *arXiv preprint arXiv:2410.01720*.

Matthias Gerstgrasser, Rylan Schaeffer, Apratim Dey, Rafael Rafailov, Henry Sleight, John Hughes, Tomasz Korbak, Rajashree Agrawal, Dhruv Pai, Andrey Gromov, et al. 2024. Is model collapse inevitable? breaking the curse of recursion by accumulating real and synthetic data. *arXiv preprint arXiv:2404.01413*.

Fabricio Goes, Zisen Zhou, Piotr Sawicki, Marek Grzes, and Daniel G Brown. 2022. Crowd score: A method for the evaluation of jokes using large language model ai voters as judges. *arXiv preprint arXiv:2212.11214*.

Daya Guo, Dejian Yang, Haowei Zhang, et al. 2025. Deepseek-r1: Incentivizing reasoning capability in llms via reinforcement learning. *arXiv preprint arXiv:2501.12948*.

Thomas Hartvigsen, Saadia Gabriel, Hamid Palangi, Maarten Sap, Dipankar Ray, and Ece Kamar. 2022. Toxigen: A large-scale machine-generated dataset for adversarial and implicit hate speech detection. *arXiv preprint arXiv:2203.09509*.

Chenghua Huang, Zhizhen Fan, Lu Wang, Fangkai Yang, Pu Zhao, Zeqi Lin, Qingwei Lin, Dongmei Zhang, Saravan Rajmohan, and Qi Zhang. 2024. Self-evolved reward learning for llms. *arXiv preprint arXiv:2411.00418*.

Xijie Huang, Li Lyna Zhang, Kwang-Ting Cheng, M Yang, and Mao Yang. 2023. Fewer is more: Boosting llm reasoning with reinforced context pruning. *arXiv preprint arXiv:2312.08901*.

Zheng Hui, Zhaoxiao Guo, Hang Zhao, Juanyong Duan, and Congrui Huang. 2024. Toxicraft: A novel framework for synthetic generation of harmful information. *arXiv preprint arXiv:2409.14740*.

Daniel P Jeong, Zachary C Lipton, and Pradeep Ravikumar. 2024. Llm-select: Feature selection with large language models. *arXiv preprint arXiv:2407.02694*.

Kaidi Jia, Yanxia Wu, and Rongsheng Li. 2024. Curriculum-style data augmentation for llm-based metaphor detection. *arXiv preprint arXiv:2412.02956*.

Jaehun Jung, Peter West, Liwei Jiang, Faeze Brahman, Ximing Lu, Jillian Fisher, Taylor Sorensen, and Yejin Choi. 2023. Impossible distillation: from low-quality model to high-quality dataset & model for summarization and paraphrasing. *arXiv preprint arXiv:2305.16635*.

Feiyang Kang, Hoang Anh Just, Yifan Sun, Himanshu Jahagirdar, Yuanzhi Zhang, Rongxing Du, Anit Kumar Sahu, and Ruoxi Jia. 2024. Get more for less: Principled data selection for warming up fine-tuning in llms. *arXiv preprint arXiv:2405.02774*.

S Karunya, M Jalakandeshwaran, Thanuja Babu, and R Uma. 2023. Ai-powered real-time speech-to-speech translation for virtual meetings using machine learning models. In *2023 Intelligent Computing and Control for Engineering and Business Systems (ICCEBS)*, pages 1–6. IEEE.

Minsang Kim and Seungjun Baek. 2024. Measuring sample importance in data pruning for training llms from a data compression perspective. *arXiv preprint arXiv:2406.14124*.

Po-Nien Kung, Fan Yin, Di Wu, Kai-Wei Chang, and Nanyun Peng. 2023. Active instruction tuning: Improving cross-task generalization by training on prompt sensitive tasks. In *Proceedings of the 2023 Conference on Empirical Methods in Natural Language Processing*, pages 1813–1829.

Dong-Ho Lee, Hyundong Cho, Woojeong Jin, Jihyung Moon, Sungjoon Park, Paul Röttger, Jay Pujara, and Roy Ka-Wei Lee. 2024a. Improving covert toxicity detection by retrieving and generating references. In *Proceedings of the 8th Workshop on Online Abuse and Harms (WOAH 2024)*, pages 266–274.

Nicholas Lee, Thanakul Wattanawong, Sehoon Kim, Karttikeya Mangalam, Sheng Shen, Gopala Anumanchipalli, Michael W Mahoney, Kurt Keutzer, and Amir Gholami. 2024b. Llm2llm: Boosting llms with novel iterative data enhancement. *arXiv preprint arXiv:2403.15042*.

Haitao Li, Junjie Chen, Qingyao Ai, Zhumin Chu, Yujia Zhou, Qian Dong, and Yiqun Liu. 2024a. Calibraeval: Calibrating prediction distribution to mitigate selection bias in llms-as-judges. *arXiv preprint arXiv:2410.15393*.

Lei Li, Yongfeng Zhang, and Li Chen. 2023a. Prompt distillation for efficient llm-based recommendation. In *Proceedings of the 32nd ACM International Conference on Information and Knowledge Management*, pages 1348–1357.

Ming Li, Yong Zhang, Zhitao Li, Jiuhai Chen, Lichang Chen, Ning Cheng, Jianzong Wang, Tianyi Zhou, and Jing Xiao. 2023b. From quantity to quality: Boosting llm performance with self-guided data selection for instruction tuning. *arXiv preprint arXiv:2308.12032*.

Minzhi Li, Taiwei Shi, Caleb Ziems, Min-Yen Kan, Nancy F Chen, Zhengyuan Liu, and Diyi Yang. 2023c. Coannotating: Uncertainty-guided work allocation between human and large language models for data annotation. *arXiv preprint arXiv:2310.15638*.

Xinjin Li, Yu Ma, Yangchen Huang, Xingqi Wang, Yuzhen Lin, and Chenxi Zhang. 2024b. Synergized data efficiency and compression (sec) optimization for large language models. In *2024 4th International Conference on Electronic Information Engineering and Computer Science (EIECS)*, pages 586–591. IEEE.

Yichuan Li, Kaize Ding, Jianling Wang, and Kyumin Lee. 2024c. Empowering large language models for textual data augmentation. *arXiv preprint arXiv:2404.17642*.

Yunshui Li, Binyuan Hui, Xiaobo Xia, Jiaxi Yang, Min Yang, Lei Zhang, Shuzheng Si, Junhao Liu, Tongliang Liu, Fei Huang, et al. 2023d. One shot learning as instruction data prospector for large language models. *arXiv preprint arXiv:2312.10302*.

Xun Liang, Shichao Song, Zifan Zheng, Hanyu Wang, Qingchen Yu, Xunkai Li, Rong-Hua Li, Yi Wang, Zhonghao Wang, Feiyu Xiong, et al. 2024a. Internal consistency and self-feedback in large language models: A survey. *arXiv preprint arXiv:2407.14507*.

Yiming Liang, Ge Zhang, Xingwei Qu, Tianyu Zheng, Jiawei Guo, Xinrun Du, Zhenzhu Yang, Jiaheng Liu, Chenghua Lin, Lei Ma, et al. 2024b. I-sheep: Self-alignment of llm from scratch through an iterative self-enhancement paradigm. *arXiv preprint arXiv:2408.08072*.

Hunter Lightman, Vineet Kosaraju, Yura Burda, Harri Edwards, Bowen Baker, Teddy Lee, Jan Leike, John Schulman, Ilya Sutskever, and Karl Cobbe. 2023. Let's verify step by step. *arXiv preprint arXiv:2305.20050*.

Guanyu Lin, Tao Feng, Pengrui Han, Ge Liu, and Jiaxuan You. 2024a. Arxiv copilot: A self-evolving and efficient llm system for personalized academic assistance. In *Proceedings of the 2024 Conference on Empirical Methods in Natural Language Processing: System Demonstrations*, pages 122–130.

Xinyu Lin, Wenjie Wang, Yongqi Li, Shuo Yang, Fuli Feng, Yinwei Wei, and Tat-Seng Chua. 2024b. Data-efficient fine-tuning for llm-based recommendation. In *Proceedings of the 47th International ACM SIGIR Conference on Research and Development in Information Retrieval*, pages 365–374.

Philip Lippmann, Matthijs TJ Spaan, and Jie Yang. 2024. Illuminating blind spots of language models with targeted agent-in-the-loop synthetic data. *arXiv preprint arXiv:2403.17860*.

Yang Liu, Yuanshun Yao, Jean-Francois Ton, Xiaoying Zhang, Ruocheng Guo Hao Cheng, Yegor Klochkov, Muhammad Faaiz Taufiq, and Hang Li. 2023. Trustworthy llms: A survey and guideline for evaluating large language models' alignment. *arXiv preprint arXiv:2308.05374*.

Yilun Liu, Shimin Tao, Xiaofeng Zhao, Ming Zhu, Wenbing Ma, Junhao Zhu, Chang Su, Yutai Hou, Miao Zhang, Min Zhang, et al. 2024a. Coachlm: Automatic instruction revisions improve the data quality in llm instruction tuning. In *2024 IEEE 40th International Conference on Data Engineering (ICDE)*, pages 5184–5197. IEEE.

Zichen Liu, Changyu Chen, Chao Du, Wee Sun Lee, and Min Lin. 2024b. Sample-efficient alignment for llms. *arXiv preprint arXiv:2411.01493*.

Lin Long, Rui Wang, Ruixuan Xiao, Junbo Zhao, Xiao Ding, Gang Chen, and Haobo Wang. 2024. On llms-driven synthetic data generation, curation, and evaluation: A survey. *arXiv preprint arXiv:2406.15126*.

Keming Lu, Hongyi Yuan, Zheng Yuan, Runji Lin, Junyang Lin, Chuanqi Tan, Chang Zhou, and Jingren Zhou. 2023. instag: Instruction tagging for analyzing supervised fine-tuning of large language models. In *The Twelfth International Conference on Learning Representations*.

Junyu Luo, Xiao Luo, Xiusi Chen, Zhiping Xiao, Wei Ju, and Ming Zhang. 2025a. Semi-supervised fine-tuning for large language models. In *Findings of the Association for Computational Linguistics: NAACL 2025*, pages 2795–2808.

Junyu Luo, Xiao Luo, Kaize Ding, Jingyang Yuan, Zhiping Xiao, and Ming Zhang. 2024. Robustft: Robust supervised fine-tuning for large language models under noisy response. *Preprint*, arXiv:2412.14922.

Junyu Luo, Weizhi Zhang, Ye Yuan, Yusheng Zhao, Junwei Yang, Yiyang Gu, Bohan Wu, Binqi Chen, Ziyue Qiao, Qingqing Long, et al. 2025b. Large language model agent: A survey on methodology, applications and challenges. *arXiv preprint arXiv:2503.21460*.

Ziyang Luo, Can Xu, Pu Zhao, Qingfeng Sun, Xiubo Geng, Wenxiang Hu, Chongyang Tao, Jing Ma, Qingwei Lin, and Daxin Jiang. 2023. Wizardcoder: Empowering code large language models with evol-instruct. *arXiv preprint arXiv:2306.08568*.

Tinh Son Luong, Thanh-Thien Le, Linh Ngo Van, and Thien Huu Nguyen. 2024. Realistic evaluation of toxicity in large language models. *arXiv preprint arXiv:2405.10659*.

Alisia Lupidi, Carlos Gemmell, Nicola Cancedda, Jane Dwivedi-Yu, Jason Weston, Jakob Foerster, Roberta Raileanu, and Maria Lomeli. 2024. Source2synth: Synthetic data generation and curation grounded in real data sources. *arXiv preprint arXiv:2409.08239*.

Xinyin Ma, Gongfan Fang, and Xinchao Wang. 2023. Llm-pruner: On the structural pruning of large language models. *Advances in neural information processing systems*, 36:21702–21720.

Aman Madaan, Niket Tandon, Prakhar Gupta, Skyler Hallinan, Luyu Gao, Sarah Wiegreffe, Uri Alon, Nouha Dziri, Shrimai Prabhumoye, Yiming Yang, et al. 2024. Self-refine: Iterative refinement with self-feedback. *Advances in Neural Information Processing Systems*, 36.

Daniel McDonald, Rachael Papadopoulos, and Leslie Benningfield. 2024. Reducing llm hallucination using knowledge distillation: A case study with mistral large and mmlu benchmark. *Authorea Preprints*.

Xuran Ming, Shoubin Li, Mingyang Li, Lvlong He, and Qing Wang. 2024. Autolabel: Automated textual data annotation method based on active learning and large language model. In *International Conference on Knowledge Science, Engineering and Management*, pages 400–411. Springer.

Saurav Muralidharan, Sharath Turuvekere Sreenivas, Raviraj Bhuminand Joshi, Marcin Chochowski, Mostofa Patwary, Mohammad Shoeybi, Bryan Catanzaro, Jan Kautz, and Pavlo Molchanov. 2024. Compact language models via pruning and knowledge distillation. In *The Thirty-eighth Annual Conference on Neural Information Processing Systems*.

Ryumei Nakada, Yichen Xu, Lexin Li, and Linjun Zhang. 2024. Synthetic oversampling: Theory and a practical approach using llms to address data imbalance. *arXiv preprint arXiv:2406.03628*.

Sejoon Oh, Yiqiao Jin, Megha Sharma, Donghyun Kim, Eric Ma, Gaurav Verma, and Srijan Kumar. 2024. Uniguard: Towards universal safety guardrails for jailbreak attacks on multimodal large language models. *arXiv:2411.01703*.

Bo Pan, Zheng Zhang, Yifei Zhang, Yuntong Hu, and Liang Zhao. 2024a. Distilling large language models for text-attributed graph learning. In *Proceedings of the 33rd ACM International Conference on Information and Knowledge Management*, pages 1836–1845.

Liangming Pan, Michael Saxon, Wenda Xu, Deepak Nathani, Xinyi Wang, and William Yang Wang. 2023a. Automatically correcting large language models: Surveying the landscape of diverse self-correction strategies. *arXiv preprint arXiv:2308.03188*.

Yu Pan, Ye Yuan, Yichun Yin, Zenglin Xu, Lifeng Shang, Xin Jiang, and Qun Liu. 2023b. Reusing pretrained models by multi-linear operators for efficient training. *Advances in Neural Information Processing Systems*, 36:3248–3262.

Zhuoshi Pan, Qianhui Wu, Huiqiang Jiang, Menglin Xia, Xufang Luo, Jue Zhang, Qingwei Lin, Victor Rühle, Yuqing Yang, Chin-Yew Lin, et al. 2024b. Llmlingua-2: Data distillation for efficient and faithful task-agnostic prompt compression. *arXiv preprint arXiv:2403.12968*.

Jinlong Pang, Jiaheng Wei, Ankit Parag Shah, Zhaowei Zhu, Yaxuan Wang, Chen Qian, Yang Liu, Yujia Bao, and Wei Wei. 2024. Improving data efficiency via curating llm-driven rating systems. *arXiv preprint arXiv:2410.10877*.

Nicholas Pangakis and Samuel Wolken. 2024. Knowledge distillation in automated annotation: Supervised text classification with llm-generated training labels. *arXiv preprint arXiv:2406.17633*.

Changhua Pei, Zihan Liu, Jianhui Li, Erhan Zhang, Le Zhang, Haiming Zhang, Wei Chen, Dan Pei, and Gaogang Xie. 2024. Self-evolutionary group-wise log parsing based on large language model. In *2024 IEEE 35th International Symposium on Software Reliability Engineering (ISSRE)*, pages 49–60. IEEE.

Zheyan Qu, Lu Yin, Zitong Yu, Wenbo Wang, et al. 2024. Coursegpt-zh: an educational large language model based on knowledge distillation incorporating prompt optimization. *arXiv preprint arXiv:2405.04781*.

Rafael Rafailov, Archit Sharma, Eric Mitchell, Christopher D Manning, Stefano Ermon, and Chelsea Finn. 2024. Direct preference optimization: Your language model is secretly a reward model. *Advances in Neural Information Processing Systems*, 36.

Indira Sen, Dennis Assenmacher, Mattia Samory, Isabelle Augenstein, Wil van der Aalst, and Claudia Wagner. 2023. People make better edits: Measuring the efficacy of llm-generated counterfactually augmented data for harmful language detection. *arXiv preprint arXiv:2311.01270*.

Minju Seo, Jinheon Baek, James Thorne, and Sung Ju Hwang. 2024. Retrieval-augmented data augmentation for low-resource domain tasks. *arXiv preprint arXiv:2402.13482*.

Zhen Tan, Dawei Li, Song Wang, Alimohammad Beigi, Bohan Jiang, Amrita Bhattacharjee, Mansooreh Karami, Jundong Li, Lu Cheng, and Huan Liu. 2024. Large language models for data annotation and synthesis: A survey. In *Proceedings of the 2024 Conference on Empirical Methods in Natural Language Processing*, pages 930–957.

Yi Tang, Chia-Ming Chang, and Xi Yang. 2024. Pdfchatannotator: A human-llm collaborative multimodal data annotation tool for pdf-format catalogs. In *Proceedings of the 29th International Conference on Intelligent User Interfaces*, pages 419–430.

Zhengwei Tao, Ting-En Lin, Xiancai Chen, Hangyu Li, Yuchuan Wu, Yongbin Li, Zhi Jin, Fei Huang, Dacheng Tao, and Jingren Zhou. 2024. A survey on self-evolution of large language models. *arXiv preprint arXiv:2404.14387*.

Vithursan Thangarasa, Ganesh Venkatesh, Mike Lasby, Nish Sinnadurai, and Sean Lie. 2024. Self-data distillation for recovering quality in pruned large language models. *arXiv preprint arXiv:2410.09982*.

Zhongwei Wan, Xin Wang, Che Liu, Samiul Alam, Yu Zheng, Jiachen Liu, Zhongnan Qu, Shen Yan, Yi Zhu, Quanlu Zhang, et al. 2023. Efficient large language models: A survey. *arXiv preprint arXiv:2312.03863*.

Fei Wang, Ninareh Mehrabi, Palash Goyal, Rahul Gupta, Kai-Wei Chang, and Aram Galstyan. 2024a. Data advisor: Dynamic data curation for safety alignment of large language models. *arXiv preprint arXiv:2410.05269*.

Jiahao Wang, Bolin Zhang, Qianlong Du, Jiajun Zhang, and Dianhui Chu. 2024b. A survey on data selection for llm instruction tuning. *arXiv preprint arXiv:2402.05123*.

Peidong Wang, Ming Wang, Zhiming Ma, Xiaocui Yang, Shi Feng, Daling Wang, and Yifei Zhang. 2024c. Language models as continuous self-evolving data engineers. *arXiv preprint arXiv:2412.15151*.

Pengkun Wang, Zhe Zhao, HaiBin Wen, Fanfu Wang, Binwu Wang, Qingfu Zhang, and Yang Wang. 2024d. Llm-autoda: Large language model-driven automatic data augmentation for long-tailed problems. In *The Thirty-eighth Annual Conference on Neural Information Processing Systems*.

Siyuan Wang, Zhuohan Long, Zhihao Fan, Zhongyu Wei, and Xuanjing Huang. 2024e. Benchmark self-evolving: A multi-agent framework for dynamic llm evaluation. *arXiv preprint arXiv:2402.11443*.

Tianlu Wang, Ilia Kulikov, Olga Golovneva, Ping Yu, Weizhe Yuan, Jane Dwivedi-Yu, Richard Yuanzhe Pang, Maryam Fazel-Zarandi, Jason Weston, and Xian Li. 2024f. Self-taught evaluators. *arXiv preprint arXiv:2408.02666*.

Yau-Shian Wang and Yingshan Chang. 2022. Toxicity detection with generative prompt-based inference. *arXiv preprint arXiv:2205.12390*.

Yizhong Wang, Yeganeh Kordi, Swaroop Mishra, Alisa Liu, Noah A Smith, Daniel Khashabi, and Hannaneh Hajishirzi. 2023. Self-instruct: Aligning language models with self-generated instructions. In *Proceedings of the 61st Annual Meeting of the Association for Computational Linguistics*, pages 13484–13508.

Yu Wang, Yifan Gao, Xiusi Chen, Haoming Jiang, Shiyang Li, Jingfeng Yang, Qingyu Yin, Zheng Li, Xian Li, Bing Yin, et al. 2024g. Memoryllm: Towards self-updatable large language models. *arXiv preprint arXiv:2402.04624*.

Yuxin Wang, Duanyu Feng, Yongfu Dai, Zhengyu Chen, Jimin Huang, Sophia Ananiadou, Qianqian Xie, and Hao Wang. 2024h. Harmonic: Harnessing llms for tabular data synthesis and privacy protection. *arXiv preprint arXiv:2408.02927*.

Tianhao Wu, Weizhe Yuan, Olga Golovneva, Jing Xu, Yuandong Tian, Jiantao Jiao, Jason Weston, and Sainbayar Sukhbaatar. 2024. Meta-rewarding language models: Self-improving alignment with llm-as-a-meta-judge. *arXiv preprint arXiv:2407.19594*.

Mengzhou Xia, Sadhika Malladi, Suchin Gururangan, Sanjeev Arora, and Danqi Chen. 2024. Less: Selecting influential data for targeted instruction tuning. *arXiv preprint arXiv:2402.04333*.

Zhe Xie, Zeyan Li, Xiao He, Longlong Xu, Xidao Wen, Tieying Zhang, Jianjun Chen, Rui Shi, and Dan Pei. 2024. Chatts: Aligning time series with llms via synthetic data for enhanced understanding and reasoning. *arXiv preprint arXiv:2412.03104*.

Huajian Xin, Daya Guo, Zhihong Shao, ZZ Ren, Qihao Zhu, Bo Liu, Chong Ruan, Wenda Li, and Xiaodan Liang. Advancing theorem proving in llms through large-scale synthetic data. In *The 4th Workshop on Mathematical Reasoning and AI at NeurIPS'24*.

Fangzhi Xu, Qiushi Sun, Kanzhi Cheng, Jun Liu, Yu Qiao, and Zhiyong Wu. 2024a. Interactive evolution: A neural-symbolic self-training framework for large language models. *arXiv preprint arXiv:2406.11736*.

Shengzhe Xu, Cho-Ting Lee, Mandar Sharma, Raquib Bin Yousuf, Nikhil Muralidhar, and Naren Ramakrishnan. 2024b. Are llms naturally good at synthetic tabular data generation? *arXiv preprint arXiv:2406.14541*.

Yang Xu, Yongqiang Yao, Yufan Huang, Mengnan Qi, Maoquan Wang, Bin Gu, and Neel Sundaresan. 2023. Rethinking the instruction quality: Lift is what you need. *CoRR*.

Zhangchen Xu, Fengqing Jiang, Luyao Niu, Yuntian Deng, Radha Poovendran, Yejin Choi, and Bill Yuchen Lin. 2024c. Magpie: Alignment data synthesis from scratch by prompting aligned llms with nothing. *arXiv preprint arXiv:2406.08464*.

Shuangtao Yang, Xiaoyi Liu, Xiaozheng Dong, and Bo Fu. 2024a. Mini-da: Improving your model performance through minimal data augmentation using llm. In *Proceedings of the Fifth Workshop on Data Science with Human-in-the-Loop (DaSH 2024)*, pages 25–30.

Yingxuan Yang, Huayi Wang, Muning Wen, Xiaoyun Mo, Qiuying Peng, Jun Wang, and Weinan Zhang. 2024b. P3: A policy-driven, pace-adaptive, and diversity-promoted framework for data pruning in llm training. *arXiv preprint arXiv:2408.05541*.

Junjie Ye, Nuo Xu, Yikun Wang, Jie Zhou, Qi Zhang, Tao Gui, and Xuanjing Huang. 2024. Llm-da: Data augmentation via large language models for few-shot named entity recognition. *arXiv preprint arXiv:2402.14568*.

Da Yin, Xiao Liu, Fan Yin, Ming Zhong, Hritik Bansal, Jiawei Han, and Kai-Wei Chang. 2023. Dynosaur: A dynamic growth paradigm for instruction-tuning data curation. *arXiv preprint arXiv:2305.14327*.

Junjie Oscar Yin and Alexander M Rush. 2024. Compute-constrained data selection. *arXiv preprint arXiv:2410.16208*.

Jiaxuan You, Mingjie Liu, Shrimai Prabhumoye, Mostofa Patwary, Mohammad Shoeybi, and Bryan Catanzaro. 2024. Llm-evolve: Evaluation for llm's evolving capability on benchmarks. In *Proceedings of the 2024 Conference on Empirical Methods in Natural Language Processing*, pages 16937–16942.

Simon Yu, Liangyu Chen, Sara Ahmadian, and Marzieh Fadaee. 2024. Diversify and conquer: Diversity-centric data selection with iterative refinement. *arXiv preprint arXiv:2409.11378*.

Tongxin Yuan, Zhiwei He, Lingzhong Dong, Yiming Wang, Ruijie Zhao, Tian Xia, Lizhen Xu, Binglin Zhou, Fangqi Li, Zhuosheng Zhang, et al. 2024a. R-judge: Benchmarking safety risk awareness for llm agents. *arXiv preprint arXiv:2401.10019*.

Weizhe Yuan, Richard Yuanzhe Pang, Kyunghyun Cho, Sainbayar Sukhbaatar, Jing Xu, and Jason Weston. 2024b. Self-rewarding language models. *arXiv preprint arXiv:2401.10020*.

Ye Yuan, Chengwu Liu, Jingyang Yuan, Gongbo Sun, Siqi Li, and Ming Zhang. 2024c. A hybrid rag system with comprehensive enhancement on complex reasoning. *arXiv preprint arXiv:2408.05141*.

Eric Zelikman, Yuhuai Wu, Jesse Mu, and Noah Goodman. 2022. Star: Bootstrapping reasoning with reasoning. *Advances in Neural Information Processing Systems*, 35:15476–15488.

Oleg Zendel, J Shane Culpepper, Falk Scholer, and Paul Thomas. 2024. Enhancing human annotation: Leveraging large language models and efficient batch processing. In *Proceedings of the 2024 Conference on Human Information Interaction and Retrieval*, pages 340–345.

Jiang Zhang, Qiong Wu, Yiming Xu, Cheng Cao, Zheng Du, and Konstantinos Psounis. 2024a. Efficient toxic content detection by bootstrapping and distilling large language models. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 38, pages 21779–21787.

Ruoyu Zhang, Yanzeng Li, Yongliang Ma, Ming Zhou, and Lei Zou. 2023a. Llmaaa: Making large language models as active annotators. *arXiv preprint arXiv:2310.19596*.

Shengyu Zhang, Linfeng Dong, Xiaoya Li, Sen Zhang, Xiaofei Sun, Shuhe Wang, Jiwei Li, Runyi Hu, Tianwei Zhang, Fei Wu, et al. 2023b. Instruction tuning for large language models: A survey. *arXiv preprint arXiv:2308.10792*.

Yuzhe Zhang, Huan Liu, Yang Xiao, Mohammed Amoon, Dalin Zhang, Di Wang, Shusen Yang, and Chai Quek. 2024b. Llm-enhanced multi-teacher knowledge distillation for modality-incomplete emotion recognition in daily healthcare. *IEEE Journal of Biomedical and Health Informatics*.

Ziyan Zhang, Yang Hou, Chen Gong, and Zhenghua Li. 2025. Data augmentation for cross-domain parsing via lightweight llm generation and tree hybridization. In *Proceedings of the 31st International Conference on Computational Linguistics*, pages 11235–11247.

Jiachen Zhao, Wenlong Zhao, Andrew Drozdov, Benjamin Rozonoyer, Md Arafat Sultan, Jay Yoon Lee, Mohit Iyyer, and Andrew McCallum. 2024a. Multistage collaborative knowledge distillation from a large language model for semi-supervised sequence generation. In *Proceedings of the 62nd Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, pages 14201–14214.

Wanru Zhao, Hongxiang Fan, Shell Xu Hu, Wangchunshu Zhou, and Nicholas Donald Lane. 2024b. CLUES: Collaborative private-domain high-quality data selection for llms via training dynamics. In *The Thirty-eighth Annual Conference on Neural Information Processing Systems (NeurIPS)*.

Yang Zhou, Shimin Shan, Hongkui Wei, Zhehuan Zhao, and Wenshuo Feng. 2024a. Pga-scire: Harnessing llm on data augmentation for enhancing scientific relation extraction. *arXiv preprint arXiv:2405.20787*.

Yue Zhou, Chenlu Guo, Xu Wang, Yi Chang, and Yuan Wu. 2024b. A survey on data augmentation in large model era. *arXiv preprint arXiv:2401.15422*.

Lianghui Zhu, Xinggang Wang, and Xinlong Wang. 2023. Judgelm: Fine-tuned large language models are scalable judges. *arXiv preprint arXiv:2310.17631*.

A Statistics

To demonstrate the research momentum in data-efficient LLM post-training, we conducted a statistical analysis of the surveyed papers. As shown in Figure 8, there has been a remarkable growth trajectory in this field: from merely 3 publications in 2022 to 31 papers in 2023, followed by a substantial surge to 158 papers in 2024, with 23 additional publications already recorded by February 2025. This trend clearly indicates the academic community's growing interest in this research direction, with the momentum continuing to accelerate. The rapid growth also underscores the critical importance of data-efficient post-training approaches in the LLM domain.

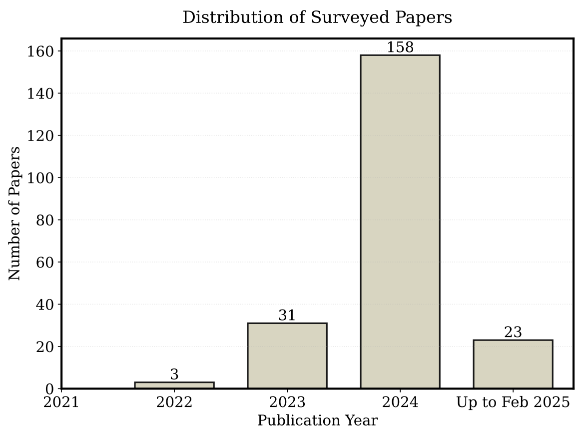

Figure 8: Distribution on publication year of surveyed papers.

Furthermore, we performed a word frequency analysis on the titles of all surveyed papers and generated a word cloud visualization (Figure 9). The word cloud reveals key methodological focal points in current research, with *augmentation*, *synthetic*, *generation*, and *alignment* emerging as prominent themes. The significant presence of terms like *fine-tuning*, *distillation*, and *efficient* underscores the field's emphasis on optimizing model training processes. These visualizations demonstrate the centrality of data-centric approaches and synthetic data methodologies in advancing LLM post-training efficiency.

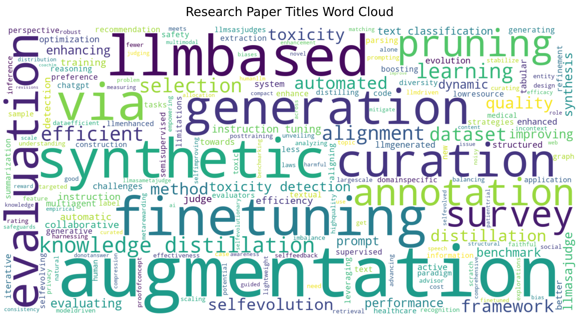

Figure 9: Word cloud of research paper titles.

The analysis demonstrates that data-efficient approaches to LLM post-training represent not only an emerging trend but also a fundamental research direction with significant implications for the advancement of language models.

B Takeaway Insights

B.1 Key Findings

Recent advancements in data-efficient LLM post-training reveal fundamental principles governing data-model interactions:

(1) The ***data flywheel*** paradigm integrates selection, augmentation, and evolution into a closed-loop lifecycle. This self-reinforcing mechanism enables continuous quality improvement through iterative refinement, transcending traditional linear data consumption

(2) **Value-centric data curation** outperforms scale-driven approaches in low-resource scenarios. Techniques like adaptive importance weighting and uncertainty-aware sampling maximize information density per training instance

(3) **Model-data co-optimization** enables joint improvements in efficiency and performance through innovations like dynamic token pruning and parameter-efficient adaptation

B.2 Paradigm Shifts

The field is witnessing fundamental changes in data utilization:

(1) Evolution from static datasets to **dynamic value-flow ecosystems** where data continuously evolves through model feedback. This necessitates new frameworks for monitoring data quality and lineage across iterations

(2) Emergence of **human-AI collaborative frameworks** combining automated generation with expert oversight. These hybrid pipelines leverage LLMs for initial labeling while preserving human judgment for critical cases

(3) Development of **cross-modal distillation** techniques that maintain semantic fidelity while reducing architectural constraints through learned alignment spaces

B.3 Critical Limitations

Current approaches face several key challenges:

(1) **Limited domain expertise** in data synthesis and refinement, where general-purpose models may fail to capture specialized knowledge and nuances required for professional domains

(2) **Scalability bottlenecks** in large-scale data generation, particularly in balancing computational costs with the need for diverse, high-quality datasets for pre-training

(3) Absence of **standardized metrics** for assessing synthetic data quality, especially in evaluating semantic fluency, information accuracy, and potential biases

B.4 Future Directions

Addressing these limitations requires advances in:

(1) **Domain-specific** pre-trained models and refinement techniques that can better capture professional knowledge while optimizing data quality and reducing annotation costs

(2) **Parallel and cost-effective frameworks** for large-scale data generation that maintain an optimal balance between data diversity and relevance

(3) **Robust evaluation metrics** and frameworks that can reliably assess synthetic data quality across different domains and use cases

C Acknowledgment of AI Assistance in Writing and Revision

We acknowledge the use of LLMs for grammar checking and language enhancement. This usage complies with the ACL Policy on AI Writing Assistance. All content and technical contributions remain original to the authors.

D Literature Review Summary

To provide a comprehensive overview of the surveyed literature, we present a detailed summary table of all referenced papers. The table includes seven key fields for each paper: **Title** (the paper's full title), **Citation** (reference key), **TLDR** (a brief summary of the paper's main contributions), **Category** (the paper's primary research direction within data-efficient LLM post-training), **Year** (publication year), **Venue** (publication venue), and **Link** (direct link to the paper). This structured compilation offers readers quick access to the original papers, enables easy tracking of research evolution across different categories, and facilitates future research by providing a comprehensive reference database of the field's development.