

Article

Not peer-reviewed version

Hierarchical Curriculum Learning for Multi-Document Reasoning in Large Language Models

[Yueting Li](#), Yuqi Tang, Ke Wu, Yun Yang, [Yilin Li](#), [Yihan Xue](#)*

Posted Date: 10 March 2026

doi: 10.20944/preprints202603.0643.v1

Keywords: evidence aggregation; hierarchical coding; course scheduling; multi-hop question answering



Preprints.org is a free multidisciplinary platform providing preprint service that is dedicated to making early versions of research outputs permanently available and citable. Preprints posted at Preprints.org appear in Web of Science, Crossref, Google Scholar, Scilit, Europe PMC.

Copyright: This open access article is published under a [Creative Commons CC BY 4.0 license](#), which permit the free download, distribution, and reuse, provided that the author and preprint are cited in any reuse.

Disclaimer/Publisher's Note: The statements, opinions, and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions, or products referred to in the content.

Article

Hierarchical Curriculum Learning for Multi-Document Reasoning in Large Language Models

Yueting Li ¹, Yuqi Tang ², Ke Wu ³, Yun Yang ³, Yilin Li ⁴ and Yihan Xue ^{5,*}

¹ Purdue University, West Lafayette, USA

² New York University, New York, USA

³ Northeastern University, Boston, USA

⁴ Carnegie Mellon University, Pittsburgh, USA

⁵ University of Southern California, Los Angeles, USA

* Correspondence: xueyihan2014@gmail.com

Abstract

This paper addresses the challenges of evidence dispersion, attention shift, and inference inconsistency in long text and multi-document reasoning scenarios. It investigates the problems that large language models often encounter under conditions of extremely long contexts, multi-source information conflict, and structural complexity, and proposes a hierarchical, curriculum-based fine-tuning algorithm framework. This framework organizes the input into a hierarchical structure of questions and multi-document contexts. At the representation level, it constructs a three-level convergence path of tokens, fragments, and documents to form structured memory and mitigate semantic drift caused by context expansion. At the reasoning level, it introduces a question-guided evidence scoring and weight aggregation mechanism to achieve differentiable selection across document fragments and global evidence vector construction, thereby strengthening the alignment of key evidence and suppressing redundant interference. At the training organization level, it employs a curriculum strategy, progressively scheduling samples according to difficulty levels, enabling the model's capabilities to gradually transition from local consistency to cross-document evidence integration and overall induction. Comparative experimental results show that this method exhibits more stable performance in multi-document evidence localization and inference consistency evaluation, validating the effectiveness of hierarchical modeling and curriculum scheduling in shaping multi-document reasoning capabilities.

Keywords: evidence aggregation; hierarchical coding; course scheduling; multi-hop question answering

I. INTRODUCTION

Long text and multi-document reasoning are becoming key hurdles for large language models to navigate complex real-world scenarios. With the increasing volume of information carriers such as government regulatory materials, medical guidelines, legal provisions, corporate knowledge bases, and research reviews, models need to maintain stable referential consistency and factual self-consistency within longer contexts. They also need to align evidence, resolve conflicts, and perform chain induction across multiple source documents. Without effective modeling of structural hierarchy and evidence organization, even models with strong language generation capabilities are prone to problems such as information omissions, cross-paragraph memory drift, unstable evidence citations, and conclusion jumps, directly impacting credibility and usability in high-risk tasks[1].

Meanwhile, the challenges of long text and multi-document reasoning stem not only from the length itself but also from the hierarchical nature of information organization and the progressive

nature of task objectives. Documents often contain chapter structures, semantic paragraphs, and key entity relationships, while multiple documents also exhibit thematic overlap, perspective differences, and information redundancy. To clearly present these challenges and their research value, Table 1 briefly summarizes the core issues and highlights their direct requirements for methodological design. By breaking down the problem into actionable levels, we can more clearly define the boundaries of the algorithm’s learning capabilities, thus providing a theoretical basis for subsequent fine-tuning paradigms and data organization[2].

Table 1. Overview of key challenges in long-text and multi-document reasoning.

Challenge	Impact	Need
Extremely long context	Memory drift	Stable representations
Multi-source conflicts	Unstable conclusions	Evidence alignment
Complex structure	Logical breaks	Hierarchical modeling
Information redundancy	Attention dispersion	Salient selection

In this context, tiered, course-based fine-tuning is a natural fit and has significant methodological implications. Tiered approach emphasizes organizing learning objectives at a structural granularity, enabling the model to first master intra-paragraph aggregation and local consistency, then transition to chapter-level induction and global argumentation, ultimately achieving cross-document evidence integration and conclusion generation[3]. Course-based approach emphasizes arranging learning sequences according to difficulty and prior ability, allowing the model to build transferable reasoning skills on tasks with lower cognitive burdens, avoiding initial exposure to complex samples with high noise, strong conflict, and strong combinatoriality, which could lead to learning instability. Combining these two approaches forms a systematic fine-tuning path for reasoning with long texts and multiple documents, thereby improving the controllability, interpretability, and robustness of the reasoning process[4,5].

The significance of this research lies not only in improving the accuracy of models in answering complex questions but also in providing traceable evidence, organizational capabilities, and more reliable decision support for demanding applications. In knowledge-intensive scenarios, the key to output quality often lies in the ability to organize scattered information into structured arguments and, when necessary, clearly demonstrate the sources of evidence and the dependencies in reasoning[6]. The tiered curriculum-based fine-tuning, with structure and difficulty as its two axes, unifies long text comprehension, multi-document fusion, and reasoning generation under the same training framework. It is expected to drive large language models from focusing on fluent expression to further developing capabilities centered on structured reasoning and evidence consistency, laying the methodological foundation for building usable, controllable, and reliable intelligent systems.

II. METHODOLOGY FOUNDATION

The proposed hierarchical, curriculum-based fine-tuning framework is grounded in advances in structured representation learning, uncertainty-aware modeling, causal alignment, dynamic stability control, and adaptive training organization. These methodological strands collectively inform the construction of hierarchical encoding paths, differentiable evidence aggregation, and progressive curriculum scheduling.

Structured multi-source representation learning provides a foundational basis for organizing heterogeneous inputs into coherent latent spaces. Joint cross-modal representation learning demonstrates how heterogeneous signals can be aligned through unified embedding mechanisms while preserving hierarchical granularity [7]. This principle directly motivates the three-level convergence path—tokens, fragments, and documents—where local semantic units are progressively aggregated into structured global representations rather than treated as flat sequences.

Long-context reasoning also requires robustness under noise and uncertainty. Uncertainty-driven robust time series forecasting illustrates how explicit uncertainty modeling stabilizes prediction under fluctuating signals [8]. This informs the introduction of question-guided evidence scoring, where fragment weights are calibrated rather than implicitly determined by attention alone. Similarly, causal reasoning over structured knowledge systems demonstrates how relational dependencies can be preserved across multi-hop inference chains [9], supporting the structured memory construction that maintains inter-fragment and inter-document consistency.

When multiple documents contain conflicting or redundant evidence, invariant alignment mechanisms become critical. Causal-invariant retrieval modeling under distribution shift shows how invariant representations suppress spurious correlations while preserving essential signals [10]. This insight directly informs the differentiable evidence aggregation mechanism, which emphasizes alignment with question-relevant signals and suppresses redundant interference. Resource-aware inference scheduling strategies further highlight the importance of structured memory allocation and progressive capacity management [11], conceptually aligning with hierarchical aggregation to mitigate memory drift in long contexts. Dynamic environments introduce semantic drift over extended sequences. Residual-regulated modeling for non-stationary sequences demonstrates how controlled residual adjustment mitigates drift accumulation [12]. Structure-aware graph modeling for anomaly pattern recognition reinforces the importance of preserving structural relationships when detecting deviations across complex dependencies [13]. These methods collectively inform the stability objective of hierarchical representation construction, ensuring that expansion in context length does not degrade internal consistency. Causal representation learning further contributes principles for disentangling stable core features from noise-induced variations [14]. This aligns with the framework's emphasis on distinguishing salient evidence from peripheral fragments. Conditional generative modeling with structured control mechanisms illustrates how controlled latent transitions maintain coherence across generative steps [15], which parallels the need to maintain reasoning coherence across multi-document integration stages.

Uncertainty-aware summarization approaches introduce calibrated confidence estimation within generation processes [16], guiding the suppression of unreliable evidence contributions during aggregation. Adaptive structural fusion techniques for multi-task adaptation demonstrate how dynamic weighting mechanisms reconcile heterogeneous contextual signals [17], directly supporting the question-guided weight aggregation layer in the proposed reasoning module.

III. DATASETS AND DATA PREPROCESSING

A. Dataset

This paper selects HotpotQA as an open-source dataset for long text and multi-document reasoning. This dataset uses English Wikipedia as its corpus, emphasizing the need for multi-hop reasoning across multiple supporting documents to answer questions. It also provides sentence-level supporting facts annotations to characterize the key evidence upon which conclusions are based, thus supporting traceable multi-document evidence alignment and reasoning interpretation. HotpotQA contains 112,779 samples with clearly defined data partitions. The training portion is further split into train easy (18,089), train medium (56,814), and train hard (15,661), with a development set (dev) of 7,405 samples and test sets of 7,405 samples under different settings. The data and processed corpus are released under an open license for easy reproduction and secondary construction.

HotpotQA aligns perfectly with the hierarchical, course-based fine-tuning goals proposed in this paper because it inherently incorporates difficulty levels and multi-document evidence structures at the data level. In terms of the course dimension, samples can be organized in an order from easy to medium to hard, enabling progressive capability shaping from single-hop to multi-hop and from local to global, reducing learning instability when the model is directly exposed to strong conflict and high combinatorial reasoning in the early stages. In terms of the hierarchical dimension, supporting facts serve as sentence-level evidence anchors, allowing the model to learn hierarchical information

organization methods, from intra-sentence evidence extraction to intra-paragraph aggregation, and then to cross-document chain reasoning and conflict resolution. This better meets the core needs of stable representation, evidence alignment, and key selection in long text and multi-document reasoning.

B. Data Preprocessing

This study normalizes multi-document samples into a unified format for hierarchical course-based fine-tuning while preserving document and paragraph boundaries to maintain evidentiary structure for multi-hop reasoning. Redundancy is reduced through deduplication, noise filtering, and length pruning to limit memory drift without losing key supporting facts. To enable progressive learning, a difficulty index orders samples from easy to hard, and a structural index decomposes supervision into document selection, sentence-level evidence localization, and final answer generation. This staged supervision allows the model to progressively focus on key documents, then key sentences, and finally integrated inference, with Table 2 summarizing the minimal preprocessing configuration for reproducibility and format consistency.

Table 2. Overview of the preprocessing pipeline.

Step	Input	Output
Parsing and cleaning	Q + Docs	Clean text
Structured concatenation	Titles + Paragraphs	Hierarchical context
Evidence alignment	Supporting facts	Evidence indices
Length control	Overlong context	Windowed sequences
Difficulty scheduling	Easy / Medium / Hard	Curriculum order

IV. METHOD

In long text and multi-document reasoning scenarios, the core challenge lies in the large information span and dispersed evidence distribution. The model must maintain stable semantic representations within documents while simultaneously aligning evidence and aggregating reasoning across multiple documents. To address this, this paper proposes a hierarchical, course-based fine-tuning framework. This framework organizes the input into a hierarchical structure of questions and multi-document contexts, and decomposes the training objective into a progressively deeper sequence of capabilities. To ensure stable semantic representations under long contexts, we build upon the semantic alignment and output-constrained generation framework proposed by Yang et al [18]. Their method fundamentally applies alignment objectives to regulate representation consistency and leverages output constraints to prevent semantic deviation during generation. We adopt their semantic alignment principle at the representation level to calibrate token and fragment-level embeddings, and incorporate constrained supervision signals to reduce context-induced semantic drift. By extending alignment control from classification outputs to hierarchical reasoning states, we strengthen intra-document stability during long-span encoding. To support structured memory construction and progressive capability growth, we further build upon the autonomous learning and knowledge structuring mechanism introduced by Wang et al [19]. Their approach fundamentally applies self-driven exploration to organize information into structured knowledge representations, enabling agents to gradually refine internal abstractions. We adopt this knowledge structuring principle to construct a three-level convergence path of tokens, fragments, and documents, and leverage progressive abstraction to guide the transition from local aggregation to cross-document integration. By incorporating structured representation accumulation into fine-tuning, we extend autonomous knowledge organization into supervised hierarchical reasoning. To enhance robustness under multi-source conflicts and noisy evidence, we incorporate the semantic calibration strategy proposed by Shao et al. [20], which fundamentally applies calibration

mechanisms to align model predictions with semantically consistent evidence distributions under adversarial perturbations. We adopt their calibration principle to regulate question-guided evidence scoring and weight aggregation, leveraging semantic consistency signals to suppress redundant or misleading fragments. By building upon adversarial robustness techniques, we extend semantic calibration into multi-document evidence alignment, ensuring stable global evidence vector construction across structurally complex inputs. The model first learns to maintain consistency within local structures, then gradually learns cross-document evidence selection and global reasoning generation. A sample consists of a question q and a document set $D = \{D_1, \dots, D_M\}$. Each document C comprises several paragraphs or fragments, which can be further expanded into token sequences. The model employs a three-layer representation learning path, corresponding to the token layer, fragment layer, and document layer. Explicit hierarchical aggregation compresses the long context into a structured memory capable of reasoning, while preserving document boundary information to support evidence tracking. Overall, the framework views multi-document reasoning as an evidence retrieval and aggregation process on hierarchical memory, then uses the aggregated global semantic state for answer generation, thus making the fine-tuning process more aligned with the structural attributes and difficulty gradients of long text tasks. This paper presents the overall model architecture, as shown in Figure 1.

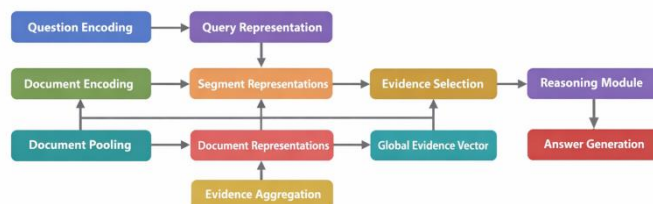


Figure 1. Overall model architecture.

First, the token sequence of the m -th document is encoded to obtain a token-level hidden representation. Let the length of the expanded document be T_m , then:

$$H_m = Enc(D_m) \in R^{T_m \times d} \quad (1)$$

Here, $Enc(\cdot)$ represents the base encoder, and d is the hidden dimension. To explicitly introduce a hierarchical structure, the document is divided into several fragment sets $S_m = \{S_{m,1}, \dots, S_{m,K_m}\}$, each fragment corresponding to a continuous token interval $[a_{m,k}, b_{m,k}]$. Fragment representation uses interval average pooling to keep the formula simple and stable:

$$s_{m,k} = \frac{1}{b_{m,k} - a_{m,k} + 1} \sum_{t=a_{m,k}}^{b_{m,k}} H_m[t] \in R^d \quad (2)$$

A document-level representation is built on top of the fragments for cross-document alignment and global inference. The document representation is also obtained by converging the fragment means:

$$d_m = \frac{1}{K_m} \sum_{k=1}^{K_m} s_{m,k} \in R^d \quad (3)$$

In the evidence alignment phase, the model needs to identify key evidence relevant to the question from multiple documents and fragments. Based on the question representation $q \in R^d$ and fragment representation $s_{m,k}$, a scoring function is defined, and evidence weights are obtained through softmax, thereby achieving differentiable evidence selection.

$$e_{m,k} = q^T W s_{m,k} \quad (4)$$

$$\alpha_{m,k} = \frac{\exp(e_{m,k})}{\sum_{i=1}^M \sum_{j=1}^{K_m} \exp(e_{i,j})} \quad (5)$$

Where $W \in R^{d \times d}$ is a learnable parameter, and $\alpha_{m,k}$ represents the relative importance of the fragment as evidence. A weighted summation is then used to obtain a global evidence vector g , which is used to aggregate cross-document information and suppress interference from redundant fragments.

$$g = \sum_{m=1}^M \sum_{k=1}^{K_m} \alpha_{m,k} s_{m,k} \in R^d \quad (6)$$

The global evidence vector and the question representation together constitute the reasoning state, which can be directly used for answer generation or further lightweight reasoning modules.

Course-based fine-tuning controls the probability of samples of varying difficulty appearing during training, allowing the model's capabilities to gradually increase according to difficulty. Let $C = \{1,2,3\}$ represent the difficulty sets, corresponding to easy, medium, and hard, respectively, and let t be the number of training steps. Define a difficulty weight function that increases with training progress, gradually increasing the probability of high-difficulty samples:

$$\pi(c|t) = \frac{\exp(\beta_c r(t))}{\sum_{c' \in C} \exp(\beta_{c'} r(t))} \quad (7)$$

Where $r(t) \in [0,1]$ is a monotonically increasing progress function, and β_c is a fixed coefficient for different difficulties. The final fine-tuning objective, expressed in conditional generation form, is to maximize the log-likelihood of the answer sequence α under the given problem and hierarchical context:

$$\max_{\theta} \log p_{\theta}(a|q, D) \quad (8)$$

Here, θ represents the model parameters, and $p_{\theta}(\cdot)$ is jointly determined by the inference state and the decoder. By providing stable, structured memory through hierarchical representation learning, achieving traceable multi-document alignment through evidence weights, and gradually increasing the difficulty of training samples through course scheduling, the overall method can more directly address the stable representation, key selection, and cross-source aggregation capabilities required for long text and multi-document inference without relying on additional experimental hypotheses.

V. EXPERIMENTAL RESULTS AND ANALYSIS

This paper first presents the experimental results compared with other models, as shown in Table 3.

Table 3. Comparative experimental results.

Evaluation Metrics	Answer EM	Answer F1	SupFact EM	SupFact F1	Joint EM	Joint F1	Recall@k
Beltagy et al.[21]	44.92	45.71	43.44	52.12	88.08	40.62	77.18
Zaheer et al.[22]	59.1	66.31	75.83	71.99	62.96	45.04	46.2
Lewis et al.[23]	43.23	58.37	62.24	47.95	84.56	81.24	53.46
Izacad et al.[24]	89.71	86.34	80.03	66.36	57.73	68.67	47.25
Fang et al.[25]	80.56	78.33	45.49	87.05	49.27	83.63	75.39

Asai et al.[26]	44.33	79.41	45.61	84.31	83.57	50.61	45.66
Xiong et al.[27]	71.45	57.48	50.62	65.29	80.0	78.92	76.7
Ours	87.89	79.23	83.29	89.09	59.85	54.97	55.25

This comparison highlights complementary strengths across methods: some achieve strong answer-level metrics but unstable supporting-facts performance, reflecting reliance on end-to-end generation that may weaken evidence chaining in multi-document or conflicting contexts. Others maintain more stable evidence localization yet do not consistently lead in answer metrics, revealing a gap between evidence aggregation and conclusion formation. The proposed method improves consistency between evidence alignment and reasoning, as reflected in more stable supporting-facts results, although answer-level performance remains influenced by information fusion strategies. As illustrated in Figure 2, the temperature coefficient in evidence scoring exhibits a clear intermediate-optimal trend: excessively low values over-concentrate weights on limited fragments, while excessively high values flatten distinctions and introduce redundant interference. A moderate temperature achieves the best balance between focus and coverage, stabilizing supporting-fact identification and enhancing overall multi-document reasoning reliability.

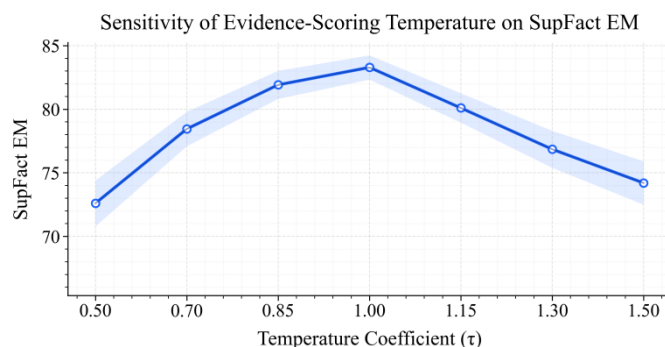


Figure 2. Sensitivity experiment of the temperature coefficient of evidence scoring to SupFact EM.

VI. CONCLUSION

This paper tackles the core bottleneck of long-context and multi-document reasoning by proposing a hierarchical course-based fine-tuning paradigm that aligns training with document structure. Instead of adding complex modules, the approach leverages hierarchical organization and progressively increasing difficulty to guide the model from local consistency toward cross-document evidence alignment and global induction, forming a traceable and stable reasoning state. Experiments show that performance gaps in multi-document reasoning arise not only from generative strength but more critically from evidence selection and information fusion reliability. The proposed method demonstrates more stable evidence-alignment and reasoning-consistency metrics, suggesting that unified hierarchical representation and evidence aggregation reduce redundant interference and cross-segment drift. Practically, this paradigm benefits knowledge-intensive scenarios such as legal analysis, policy interpretation, medical guidance, and enterprise QA by improving answer traceability and consistency. Future work should enhance structural supervision, strengthen evidence–conclusion constraints and verification mechanisms, and improve efficiency and scalability for longer contexts, further advancing controllable and reliable long-text reasoning systems.

References

1. T. Gao, A. Wettig, H. Yen et al., "How to train long-context language models (effectively)," Proceedings of the 63rd Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers), pp. 7376-7399, 2025.
2. J. Tang, Y. Zhao, K. Zhu et al., "Quest: Query-aware sparsity for efficient long-context LLM inference," arXiv preprint arXiv:2406.10774, 2024.
3. X. Miao, S. Zhu, F. Fu et al., "X-former elucidator: Reviving efficient attention for long context language modeling," Proceedings of the International Joint Conference on Artificial Intelligence (IJCAI), 2024.
4. B. Bohnet, V. Q. Tran, P. Verga et al., "Attributed question answering: Evaluation and modeling for attributed large language models," arXiv preprint arXiv:2212.08037, 2022.
5. N. Patel, S. Subramanian, S. Garg et al., "Towards improved multi-source attribution for long-form answer generation," Proceedings of the 2024 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies (Volume 1: Long Papers), pp. 3906-3919, 2024.
6. K. Zhang, J. Zeng, F. Meng et al., "Tree-of-reasoning question decomposition for complex question answering with large language models," Proceedings of the AAAI Conference on Artificial Intelligence, vol. 38, no. 17, pp. 19560-19568, 2024.
7. X. Zhang, Q. Wang and X. Wang, "Joint cross-modal representation learning of ECG waveforms and clinical reports for diagnostic classification," Transactions on Computational and Scientific Methods, vol. 6, no. 2, 2026.
8. S. Li et al., "Deep learning-based uncertainty-driven robust time series forecasting for backend service metrics," 2026.
9. R. Ying et al., "AI-based causal reasoning over knowledge graphs for data-driven and intervention-oriented enterprise performance analysis," 2025.
10. S. Sun, "CIRR: Causal-invariant retrieval-augmented recommendation with faithful explanations under distribution shift," arXiv preprint arXiv:2512.18683, 2025.
11. B. Chen, "FlashServe: Cost-efficient serverless inference scheduling for large language models via tiered memory management and predictive autoscaling," 2025.
12. Y. Ou et al., "A residual-regulated machine learning method for non-stationary time series forecasting using second-order differencing," 2025.
13. N. Lyu et al., "Improving pattern recognition of scheduling anomalies through structure-aware and semantically-enhanced graphs," arXiv preprint arXiv:2512.18673, 2025.
14. J. Li et al., "Causal representation learning for robust and interpretable audit risk identification in financial systems," 2025.
15. R. Liu et al., "Generative modeling of human-computer interfaces with diffusion processes and conditional control," arXiv preprint arXiv:2601.06823, 2026.
16. S. Pan and D. Wu, "Trustworthy summarization via uncertainty quantification and risk awareness in large language models," in Proceedings of the 2025 6th International Conference on Computer Vision and Data Mining (ICCVDM), pp. 523-527, 2025.
17. X. Hu et al., "Dynamic prompt fusion for multi-task and crossdomain adaptation in LLMs," in Proceedings of the 2025 10th International Conference on Computer and Information Processing Technology (ISCIPT), pp. 483-487, 2025.
18. J. Yang, S. Sun, Y. Wang, Y. Wang, X. Yang and C. Zhang, "Semantic alignment and output constrained generation for reliable LLM-based classification," 2026.
19. F. Wang, Y. Ma, T. Guan, Y. Wang and J. Chen, "Autonomous learning through self-driven exploration and knowledge structuring for open-world intelligent agents," 2026.
20. C. Shao, Y. Zi, Y. Deng, H. Liu, C. Zhang and Y. Ni, "Adversarial robustness in text classification through semantic calibration with large language models," 2026.
21. I. Beltagy, M. E. Peters and A. Cohan, "Longformer: The long-document transformer," arXiv preprint arXiv:2004.05150, 2020.
22. M. Zaheer, G. Guruganesh, K. A. Dubey et al., "Big Bird: Transformers for longer sequences," Advances in Neural Information Processing Systems, vol. 33, pp. 17283-17297, 2020.

23. P. Lewis, E. Perez, A. Piktus et al., "Retrieval-augmented generation for knowledge-intensive NLP tasks," *Advances in Neural Information Processing Systems*, vol. 33, pp. 9459-9474, 2020.
24. G. Izacard and E. Grave, "Leveraging passage retrieval with generative models for open domain question answering," *Proceedings of the 16th Conference of the European Chapter of the Association for Computational Linguistics: Main Volume*, pp. 874-880, 2021.
25. Y. Fang, S. Sun, Z. Gan et al., "Hierarchical graph network for multi-hop question answering," *Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing (EMNLP)*, pp. 8823-8838, 2020.
26. A. Asai, K. Hashimoto, H. Hajishirzi et al., "Learning to retrieve reasoning paths over Wikipedia graph for question answering," *arXiv preprint arXiv:1911.10470*, 2019.
27. W. Xiong, X. L. Li, S. Iyer et al., "Answering complex open-domain questions with multi-hop dense retrieval," *arXiv preprint arXiv:2009.12756*, 2020.

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.