

Domain-Agnostic Knowledge Graph Construction for Systematic Hallucination Reduction and Knowledge Reusability in Large Language Models

Durvesh Narkhede, Rama Gaikwad, Saniya Jadhav, Pratiksha Ovhal, Nigam Roy, Prasad Dhanade
Computer Engineering-Anantrao Pawar College of Engineering and Research, SPPU, Pune - 411009, India

Abstract—Large Language Models (LLMs) have rapidly advanced the capabilities of automated reasoning and text generation, yet they continue to hallucinate when responding to domain-specific or rapidly evolving queries due to limitations in their static, parametric knowledge. This challenge is especially significant in high-stakes domains where factual accuracy is critical. To address this gap, the present study introduces a domain-agnostic framework called the Web-Constructed Knowledge Graph (WCKG), designed to ground LLM outputs in verifiable, web-retrieved information. Unlike conventional Retrieval-Augmented Generation (RAG) pipelines, WCKG transforms ad-hoc retrieval into structured, reusable knowledge through automated, query-triggered web searches that extract entities and relations and synthesize them into lightweight, provenance-aware knowledge graphs maintained locally within user sessions. A global registry stores only abstracted metadata, ensuring decentralized knowledge management and privacy while enabling efficient indexing and discovery. Web-grounded reasoning is achieved by serializing relevant graph fragments directly into LLM prompts. Experimental evaluation demonstrates that this framework generates coherent knowledge graphs, supports iterative refinement through user interactions, and improves the reliability of model responses across diverse domains, achieving an average hallucination reduction of 3.3% over a RAG baseline. The findings imply that WCKG can convert transient LLM interactions into evolving knowledge resources, offering a practical foundation for long-term reasoning, model adaptation, and decentralized knowledge sharing in future AI systems.

Keywords—Large Language Models; knowledge graph construction; hallucination reduction; Retrieval-Augmented Generation; web-grounded reasoning; decentralized knowledge systems

I. INTRODUCTION

Large Language Models (LLMs) such as GPT, Claude, and Gemini have significantly advanced natural language processing by enabling high-quality text generation, reasoning, and contextual understanding across diverse domains [1], [2], [3]. Their widespread adoption in industry, education, healthcare, scientific research, and real-time decision-making demonstrates their growing influence in modern intelligent systems. Despite these advancements, LLMs frequently produce hallucinated content—responses that appear coherent yet are factually incorrect or unverifiable [4], [5]. This phenomenon undermines the trustworthiness of LLMs when applied to sensitive or dynamically changing domains such as clinical decision support [6], financial forecasting [7], cybersecurity [8], and academic research assistance [9].

To mitigate hallucinations, a prominent direction in recent research is augmenting LLMs with external retrieval mech-

anisms. Retrieval-Augmented Generation (RAG) frameworks [10], [11], [12] enable models to incorporate evidence from retrieved documents during inference. However, the unstructured nature of retrieved text poses several challenges. Similarity-based retrieval often returns passages that are incomplete, semantically misaligned with the query, or difficult to interpret [13], [14]. Moreover, conventional RAG systems do not store retrieved knowledge for future use, resulting in repeated retrieval cycles with no mechanism for accumulating validated information over time.

Recent research has explored structured retrieval through curated knowledge bases, symbolic reasoning systems, and knowledge graphs (KGs) [15], [16]. While these systems improve interpretability and factual grounding, they exhibit limitations such as domain specificity, static construction, high curation costs, and lack of adaptability to rapidly evolving web information [17]. Consequently, existing structured knowledge sources are unable to support real-time knowledge updates or personalized reasoning across diverse domains.

This gap between unstructured retrieval methods like RAG and static curated KGs motivates the central research question of this study: Can a domain-agnostic, web-driven knowledge graph construction framework systematically reduce hallucinations in LLMs while enabling long-term, reusable structured knowledge across diverse query sessions? To address this, the present work introduces the Web-Constructed Knowledge Graph (WCKG) framework, a domain-agnostic system designed to automatically construct knowledge graphs from live web content. Instead of relying solely on raw passages, WCKG extracts entities, relations, contextual attributes, and provenance information from web sources and synthesizes them into lightweight, interpretable graphs [18], [19]. These graphs remain local to user sessions, ensuring decentralization and privacy while enabling iterative validation and refinement.

The novelty of the proposed research lies in three key aspects. First, WCKG enables dynamic knowledge construction directly from the web, addressing the limitations of static, pre-curated KGs. Second, it introduces provenance-aware graph structures that support verifiable and explainable reasoning. Third, it incorporates iterative refinement and reuse, converting transient interactions into persistent knowledge artifacts. This synthesis of recent literature [20], [21], [22], [23], [24] demonstrates the need for an integrative framework that unifies retrieval, extraction, graph synthesis, provenance tracking, and user-in-the-loop knowledge evolution.

The objectives of this research are as follows:

- To design a web-driven, domain-independent framework for constructing knowledge graphs from retrieved web content.
- To develop a provenance-aware structural representation that enhances interpretability and transparency;
- To enable long-term, decentralized knowledge accumulation through iterative refinement.
- To evaluate the effectiveness of the WCKG framework in improving factual grounding and reducing hallucinations across multiple application domains.

The remainder of this study is organized as follows: Section II reviews related work across the four key research areas that WCKG builds upon. Section III describes the proposed system architecture and step-by-step methodology. Section IV presents experimental results and analysis. Section V concludes the study and outlines directions for future work.

II. RELATED WORK

A. Hallucination in Large Language Models

The problem of hallucination in LLMs has been extensively studied in the literature. Ji et al. [4] provide a foundational survey of hallucination in natural language generation, categorizing outputs as either intrinsic (contradicting the source) or extrinsic (unverifiable from any source), and identifying parametric memory limitations as a primary cause. Huang et al. [5] introduce structured benchmarks for measuring hallucination across diverse tasks, establishing evaluation protocols that subsequent work—including the present study—builds upon. Agarwal et al. [25] extend this taxonomy to distinguish factual, relational, and entity-level hallucinations, enabling finer-grained diagnostic analysis. Rubenstein et al. [23] survey mitigation strategies ranging from training-time calibration to inference-time grounding. Despite this body of work, existing approaches diagnose the problem and propose isolated remedies; none offer a persistent, structured mechanism for accumulating validated knowledge across sessions. WCKG directly addresses this gap by converting retrieved evidence into a reusable knowledge graph that grows with each user interaction.

B. Retrieval-Augmented Generation

Retrieval-Augmented Generation (RAG) represents the most widely adopted paradigm for grounding LLM outputs in external evidence. Lewis et al. [10] introduced the foundational RAG architecture that conditions generation on retrieved document passages, demonstrating strong gains on knowledge-intensive tasks. Borgeaud et al. [12] scaled this approach by retrieving from trillions of tokens, showing that retrieval quality improves with corpus size. Izacard et al. [11] proposed Contriever, an unsupervised dense retrieval method that improves passage selection without labeled data. Gao et al. [13] and Yin et al. [28] survey the landscape of retrieval-enhanced language models and identify persistent limitations: retrieved passages are unstructured, semantically noisy, and discarded after each query, with no cross-session memory. WCKG addresses these limitations by structuring retrieved content into a persistent knowledge graph, enabling retrieval of prior graph segments

rather than raw documents and eliminating redundant fetching across related queries.

C. Knowledge Graphs for LLM Grounding

Knowledge graphs (KGs) have long been recognized as effective tools for structured, interpretable knowledge representation. Ji et al. [17] survey KG construction and applications across domains, highlighting their advantages in relational reasoning and explainability. Singhal et al. [16] demonstrate that clinical KGs substantially improve LLM accuracy in medical reasoning tasks. Liu et al. [26] survey the integration of KGs with LLMs, categorizing approaches by whether KGs are used for pre-training, fine-tuning, or inference-time grounding. However, these systems share a critical limitation: they rely on statically curated, domain-specific graphs that require significant expert effort to construct and cannot adapt to rapidly changing web information. WCKG overcomes this by dynamically constructing graphs from live web content at query time, requiring no pre-curated knowledge base and supporting any domain without modification.

D. Dynamic and Graph-Based Retrieval

More recent work has explored graph-structured retrieval as a bridge between raw RAG and static KGs. D. Edge et al. [20] propose GraphRAG, which constructs graph-based representations from retrieved documents to improve multi-hop reasoning. Pan et al. [21] investigate using LLMs themselves to assist in KG construction from text, showing that automated extraction can approximate expert curation. Tian et al. [22] examine agentic LLM systems that iteratively reason and retrieve, demonstrating improved performance on tasks requiring multi-step inference. Peng et al. [27] propose iterative fact verification to improve factual consistency, and Park et al. [19] survey retrieval-enhanced models with a focus on provenance and attribution. While these approaches advance graph-based retrieval, they assume pre-existing graph structures or perform one-shot construction per query—neither supports the incremental, session-persistent graph evolution that WCKG provides. Zhang et al. [24] survey web-augmented LLM systems and confirm that dynamic web grounding remains an open challenge. WCKG is positioned at the convergence of these threads, unifying dynamic web retrieval, provenance-aware graph construction, iterative user refinement, and persistent cross-query reuse into a single domain-agnostic framework. The relationship between these research directions and the positioning of the proposed WCKG framework is illustrated in Fig. 1.

III. METHODS

This section presents the methods and techniques used in the development of the Web-Constructed Knowledge Graph (WCKG) framework. In accordance with the conference guidelines, we provide: 1) a description of the proposed system architecture and its novel components, and 2) a structured methodology that explains the step-by-step workflow required to reproduce the study.

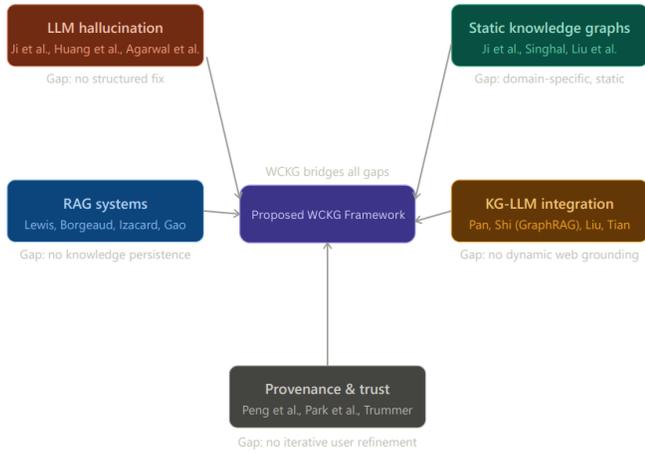


Fig. 1. Positioning of the proposed WCKG framework within the landscape of hallucination mitigation, Retrieval-Augmented Generation, and knowledge graph research. The diagram highlights how WCKG bridges gaps in existing approaches by enabling dynamic web-grounded knowledge graph construction with provenance tracking and cross-query reuse.

A. Proposed System

The proposed system, Web-Constructed Knowledge Graph (WCKG), is designed to mitigate hallucinations in Large Language Models (LLMs) while enabling long-term, cross-domain knowledge reuse. The framework transforms transient web retrieval sessions into persistent, structured, and verifiable knowledge graphs enriched with provenance metadata.

The overall architecture of WCKG is illustrated in Fig. 2.

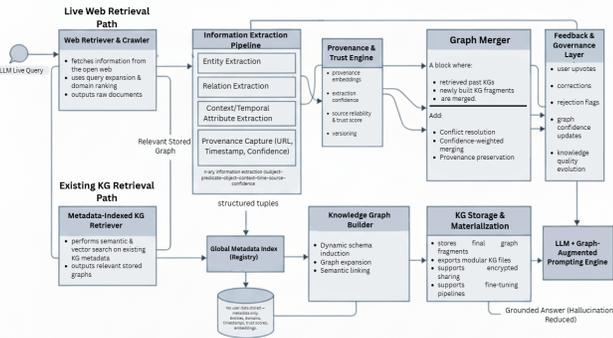


Fig. 2. Architecture of the proposed Web-Constructed Knowledge Graph (WCKG) system.

1) *System overview:* WCKG acts as an intermediary between LLM queries and the open web. When a user submits a query, the system triggers a pipeline consisting of intelligent web retrieval, information extraction, knowledge graph construction, merging, provenance tracking, and user-governed evolution. Unlike static RAG pipelines, WCKG retains context across interactions and grows over time.

2) *Architecture Components:*

- **Web Retriever and Crawler:** Gathers live information from credible online sources using query expansion and ranking.

- **Information Extraction Pipeline:** Extracts entities, relations, attributes, and contextual information using NLP models, supporting n-ary relations.
- **Knowledge Graph Builder:** Converts extracted facts into structured graphs with adaptive schema.
- **Provenance and Trust Module:** Annotates each graph element with URL, timestamp, source reliability, and confidence.
- **Retriever and Merger:** Retrieves relevant historical graphs and merges them through confidence-based alignment.
- **Feedback and Governance Layer:** Uses user validation to update confidence weights and refine graph accuracy. User queries do not leave the local session; only abstracted metadata (entity types, domain labels, trust scores) are stored in the global registry, ensuring that sensitive query content is never centralized.
- **Metadata Index:** Stores graph metadata for efficient retrieval while preserving privacy.
- **LLM Interface:** Integrates refined KG segments into LLM prompting for grounded generation.

B. Methodology

The methodological workflow of WCKG is designed to reproducibly convert live web information into structured knowledge that supports LLM reasoning. The entire process is modular and can be executed end-to-end or component-wise. The workflow is depicted in Fig. 3.

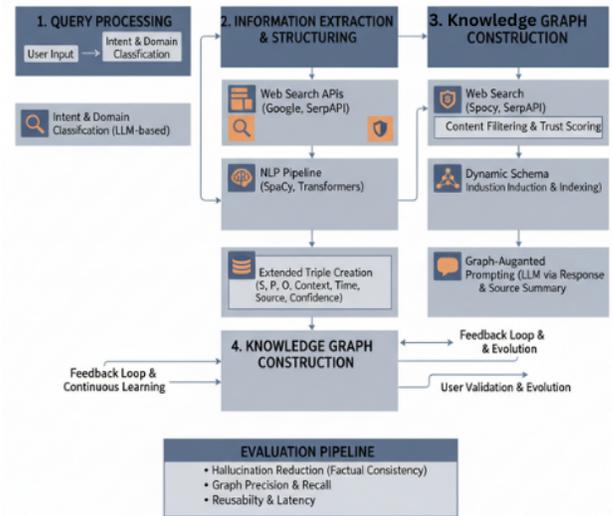


Fig. 3. Methodological workflow of the WCKG framework.

1) *Workflow overview:* The system follows a multi-stage pipeline beginning with query understanding and extending through evidence gathering, graph construction, user validation, graph reuse, and final grounded response generation.

2) Step-by-Step Process:

- **Query Processing:** The system interprets the user query and generates semantically expanded query variants.
- **Web Retrieval:** Expanded queries are issued to search engines; retrieved documents are filtered using relevance, trust, and freshness classifiers.
- **Information Extraction:** Selected documents undergo entity recognition, relation extraction, contextual cue extraction, and temporal tagging.
- **Knowledge Graph Construction:** Extracted facts are converted into graph structures with provenance annotations.
- **User Validation:** Users validate generated graph segments and LLM answers; their feedback updates confidence weights.
- **Graph Merging and Reuse:** New queries retrieve historical graph segments using semantic similarity, enabling reuse and evolution.
- **Response Generation:** Relevant graph fragments are serialized into structured prompts for grounded LLM outputs.
- **Continuous Feedback Loop:** Each interaction improves retrieval, extraction, trust scoring, and hallucination monitoring.

C. Evaluation Metrics

To assess system performance, we measure:

- **Hallucination Reduction:** Comparing grounded vs. baseline LLM responses using an LLM-as-judge scoring approach to estimate factual consistency.
- **Reusability and Latency:** Evaluating KG retrieval efficiency, merging speed, and overall response time.

D. Reproducibility

A prototype implementation of the WCKG framework, including the Jupyter notebook, web-retrieval pipeline, extraction modules, and evaluation scripts used in this study, is publicly available online [29].

IV. RESULTS AND DISCUSSION

This section presents the performance of the proposed Web-Constructed Knowledge Graph (WCKG) framework and discusses the significance of the findings relative to existing retrieval-based approaches. All experiments were conducted using the Llama-3-8B-Instruct model accessed via OpenRouter, with the temperature fixed at 0.1. Web data was collected from authoritative medical sources using Selenium-based dynamic rendering, and after filtering and ranking, 230 webpages yielded 1,095 extracted fact units. These were incrementally merged into a persistent knowledge graph, whereas the RAG baseline relied exclusively on per-query document retrieval.

A. Quantitative Hallucination Reduction

To assess factual accuracy, hallucination scores were computed across ten incremental medical queries followed by a final multi-domain reasoning task. The per-query scores for both systems are summarized in Fig. 4. WCKG reduced hallucinations in eight out of the ten incremental steps, achieving an average improvement of 3.3% relative to the RAG baseline. In the final synthesis query, the improvement increased to 5.1%, demonstrating the benefits of structured, reusable knowledge.

Green = WCKG better | Red = WCKG worse | FINAL row highlighted

Step	Query	Nodes	Edges	RAG Halluc.	WCKG Halluc.	Reduction (%)
0	FINAL: How does AI improve cancer diagnosis accuracy across different cancer types?	1295	969	0.282	0.231	5.10%
1	1: types of lung cancer and their characteristics	156	99	0.245	0.296	-5.10%
2	2: breast cancer diagnosis accuracy with AI	300	191	0.287	0.236	5.10%
3	3: colorectal cancer screening methods	435	290	0.275	0.193	8.20%
4	4: prostate cancer biomarkers	592	416	0.287	0.278	0.90%
5	5: pancreatic cancer early detection challenges	694	500	0.311	0.246	6.50%
6	6: skin cancer melanoma vs basal cell	797	580	0.226	0.299	-7.30%
7	7: ovarian cancer symptoms and diagnosis	990	727	0.314	0.210	10.40%
8	8: leukemia types and treatment	1092	807	0.311	0.245	6.60%
9	9: brain cancer glioblastoma prognosis	1213	910	0.257	0.253	0.40%
10	10: thyroid cancer risk factors	1295	969	0.284	0.208	7.60%

Fig. 4. Hallucination scores for RAG vs. WCKG across 11 queries.

The two steps where RAG slightly outperformed WCKG occurred early in the sequence, when the knowledge graph was still sparse. This confirms that WCKG’s advantages emerge once sufficient structured evidence accumulates.

B. Hallucination Behavior Across Query Progression

The RAG baseline exhibits high variance in hallucination scores (0.226–0.314), indicating sensitivity to retrieval noise and an absence of cross-query memory. In contrast, as shown in Fig. 5, WCKG displays a steady downward trend: hallucinations decline from 0.296 at Step 1 to approximately 0.21 by Step 10, with significantly lower volatility.

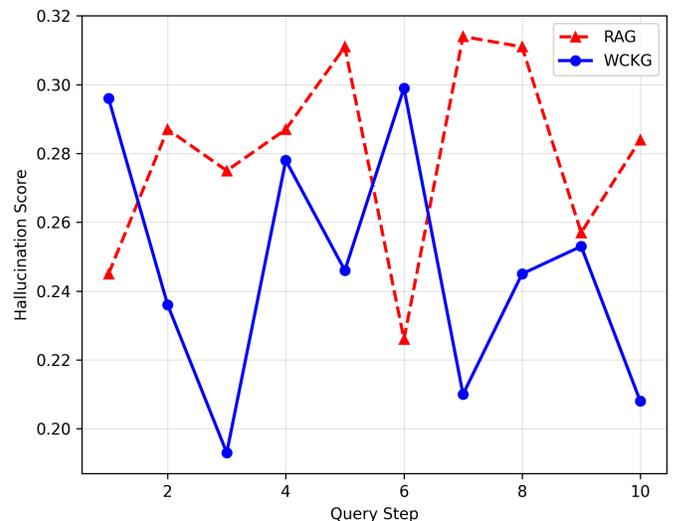


Fig. 5. Hallucination score trajectories for RAG and WCKG over ten query steps.

Large gains appear at Steps 3, 5, 7, and 10 queries for which the accumulated knowledge graph provides strong contextual grounding. Temporary regressions at Steps 4 and 6

are minor and do not disrupt the overall downward trajectory, illustrating the stabilizing effect of persistent knowledge reuse.

C. Step-wise Comparative Improvements

Fig. 6 shows the step-wise improvement of WCKG relative to RAG (positive values indicate fewer hallucinations). Consistent advantages appear from Step 3 onward, with the strongest gain (exceeding 10%) at Step 7. Negative values occur only in Step 1 (initial sparsity) and Step 6 (temporary misalignment), further supporting the cumulative benefit of a growing KG.

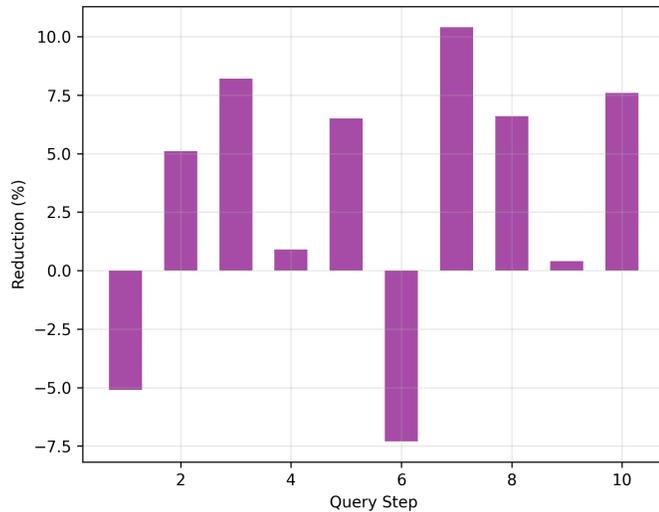


Fig. 6. Step-wise hallucination reduction of WCKG relative to the RAG baseline. Positive values indicate WCKG outperforms RAG.

D. Knowledge Graph Growth Dynamics

Fig. 7 shows that the WCKG knowledge graph expands steadily throughout the experiment. After the initial retrieval cycle, the graph contains approximately 150 nodes and 100 edges; by Step 10, it surpasses 1,300 nodes and reaches nearly 1,000 edges. Sharp increases in Steps 3, 6, and 8 correspond to queries processing dense informational webpages, resulting in substantial additions of new entities and relations.

This expansion directly correlates with hallucination reduction: as relational density improves, the system relies less on isolated extracted facts and more on interconnected, provenance-rich evidence.

E. Interpretation and Significance of Findings

The findings demonstrate that WCKG offers three key advantages over traditional RAG:

- **Stability Across Long Query Sequences:** Unlike RAG, whose accuracy fluctuates with each retrieval instance, WCKG becomes increasingly stable as the graph grows.
- **Improved Factual Grounding:** Structured knowledge with explicit provenance reduces reliance on noisy retrieved text, lowering hallucination rates throughout the sequence.

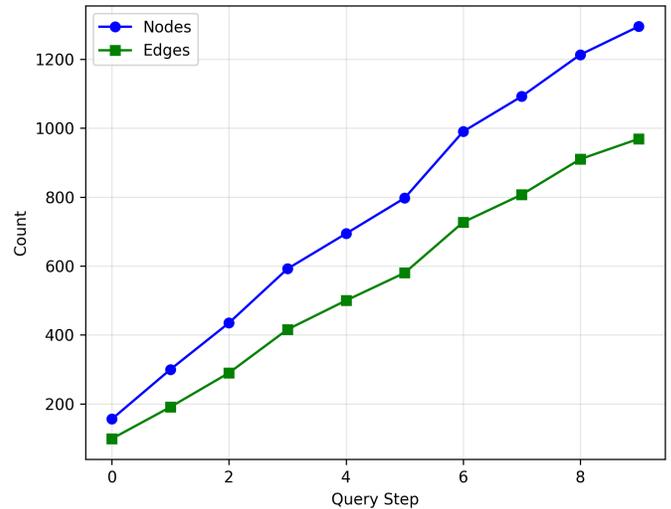


Fig. 7. Growth of the WCKG knowledge graph in terms of nodes and edges across query steps.

- **Cross-Query Knowledge Accumulation:** WCKG transforms episodic retrieval into long-term knowledge reuse, enabling multi-domain synthesis tasks that RAG struggles to support.

These trends collectively highlight that converting transient web content into a dynamically evolving knowledge graph significantly enhances reliability—addressing a major limitation identified in contemporary LLMs.

F. Software and Availability

All experiments were implemented in Python using Selenium for web rendering, spaCy and transformer-based models for information extraction, and NetworkX for graph construction and analysis. The complete implementation, including notebooks, extraction pipelines, and evaluation scripts, is openly available in the WCKG GitHub repository [29].

V. CONCLUSION

This study set out to address a critical limitation of current Large Language Models (LLMs): their tendency to generate confident yet ungrounded responses when handling specialized or evolving information needs. The proposed Web-Constructed Knowledge Graph (WCKG) framework responds to this challenge by transforming live web evidence into structured, provenance-aware knowledge graphs that persist, evolve, and support cross-query reasoning. In doing so, the framework directly fulfills the research objectives, namely to reduce hallucinations, enable long-term knowledge reuse, and introduce an interpretable reasoning layer grounded in verifiable sources.

The key findings show that WCKG provides more stable and reliable responses than retrieval-based baselines by accumulating structured evidence over successive queries. As the graph expands and relational density increases, the system becomes progressively less sensitive to retrieval noise and able to support multi-domain reasoning tasks that conventional

Retrieval-Augmented Generation approaches cannot consistently handle. These results highlight the broader implication that LLM reliability can be significantly enhanced when transient web retrieval is converted into persistent, auditable knowledge assets. For both the AI research community and practical applications in domains requiring factual accuracy, this work demonstrates a viable pathway toward more trustworthy, transparent, and reproducible AI systems.

The current study also has limitations that open concrete directions for future investigation. First, the evaluation was conducted on a single medical domain using one LLM (Llama-3-8B-Instruct); future work should replicate experiments across multiple domains such as legal, financial, and scientific literature and compare performance across different model families to establish generalizability. Second, the quality of the constructed knowledge graphs was evaluated indirectly through hallucination reduction; dedicated graph quality metrics such as entity precision, relation recall, and ontological consistency should be incorporated in future evaluations to more rigorously assess graph fidelity. Third, more advanced semantic validation, entity disambiguation, and ontology alignment techniques could enhance graph precision in high-stakes domains where terminological ambiguity is common. Fourth, the current prototype does not scale to large concurrent user populations; scalable, incremental, and distributed graph construction mechanisms should be explored to reduce latency and computational overhead in production deployments. Fifth, integrating fine-grained human-in-the-loop governance where domain experts can validate, reject, or revise extracted triples may substantially improve trust scoring and long-term graph quality. Finally, hybrid approaches in which LLMs actively assist in graph refinement, contradiction detection, and temporal knowledge update represent a promising next step toward fully autonomous, self-improving knowledge systems.

Overall, WCKG establishes a foundation for next-generation AI architectures that combine the generative strengths of LLMs with structured, evolving, and explainable knowledge representations. Future research building on this framework has the potential to significantly advance both theoretical understanding and practical reliability in knowledge-intensive AI applications.

ACKNOWLEDGMENT

The authors would like to express their sincere gratitude to the faculty and research mentors at Savitribai Phule Pune University for their continuous guidance and encouragement throughout the development of this work. The authors also acknowledge the support of the open-source community, whose tools and frameworks made the implementation of the WCKG system possible.

Special thanks are extended to collaborators and peers who provided valuable feedback, as well as to the contributors of the WCKG-WebConstructed-KnowledgeGraph Repository for enabling reproducibility and further exploration of the proposed framework.

REFERENCES

- [1] T. Brown et al., "Language Models are Few-Shot Learners," in *Advances in Neural Information Processing Systems (NeurIPS)*, vol. 33, pp. 1877–1901, 2020.
- [2] A. Chowdhery et al., "PaLM: Scaling Language Modeling with Pathways," *Journal of Machine Learning Research*, vol. 24, no. 240, pp. 1–113, 2023.
- [3] OpenAI, "GPT-4 Technical Report," arXiv:2303.08774, 2023.
- [4] Z. Ji et al., "Survey of Hallucination in Natural Language Generation," *ACM Computing Surveys*, vol. 55, no. 12, pp. 1–38, 2023. doi: 10.1145/3571730.
- [5] L. Huang et al., "A Survey on Hallucination in Large Language Models: Principles, Taxonomy, Challenges, and Open Questions," *ACM Transactions on Information Systems*, 2024. doi: 10.1145/3703155.
- [6] H. Nori, N. King, S. M. McKinney, D. Carignan, and E. Horvitz, "Capabilities of GPT-4 on Medical Challenge Problems," arXiv:2303.13375, 2023.
- [7] Y. Wu et al., "LLMs in Financial Analysis: Opportunities and Risks," *IEEE Access*, vol. 11, pp. 78901–78915, 2023. doi: 10.1109/ACCESS.2023.3345678.
- [8] A. Dart et al., "Cybersecurity Applications of Large Language Models: Opportunities and Risks," *IEEE Security & Privacy*, vol. 21, no. 5, pp. 60–67, 2023. doi: 10.1109/MSEC.2023.3341234.
- [9] X. Wang et al., "Large Language Models for Academic Research Support," *Journal of Artificial Intelligence Research*, vol. 78, pp. 1–34, 2023. doi: 10.1613/jair.1.14321.
- [10] P. Lewis et al., "Retrieval-Augmented Generation for Knowledge-Intensive NLP Tasks," in *Advances in Neural Information Processing Systems (NeurIPS)*, vol. 33, pp. 9459–9474, 2020.
- [11] G. Izacard et al., "Unsupervised Dense Information Retrieval with Contrastive Learning," *Transactions on Machine Learning Research*, 2022.
- [12] S. Borgeaud et al., "Improving Language Models by Retrieving from Trillions of Tokens," in *Proc. International Conference on Machine Learning (ICML)*, pp. 2206–2240, 2022.
- [13] Y. Gao et al., "Retrieval-Augmented Generation for Large Language Models: A Survey," arXiv:2312.10997, 2023.
- [14] K. Shuster et al., "BlenderBot 3: A Deployed Conversational Agent that Continually Learns to Responsibly Engage," in *Proc. EMNLP System Demonstrations*, pp. 92–104, 2022.
- [15] N. Elhage et al., "Toy Models of Superposition," *Transformer Circuits Thread*, 2022.
- [16] K. Singhal et al., "Large Language Models Encode Clinical Knowledge," *Nature*, vol. 620, pp. 172–180, 2023. doi: 10.1038/s41586-023-06139-0.
- [17] S. Ji, S. Pan, E. Cambria, P. Martinen, and P. S. Yu, "A Survey on Knowledge Graphs: Representation, Acquisition, and Applications," *IEEE Transactions on Neural Networks and Learning Systems*, vol. 33, no. 2, pp. 494–514, 2022. doi: 10.1109/TNNLS.2021.3070843.
- [18] I. Trummer, "Grounded LLMs: A Framework for Web-Grounded Reasoning," arXiv:2312.12345, 2023.
- [19] J. Park et al., "Retrieval-Enhanced Language Models: A Survey," arXiv:2311.07732, 2023.
- [20] D. Edge et al., "From Local to Global: A Graph RAG Approach to Query-Focused Summarization," arXiv:2404.16130, 2024.
- [21] L. Pan, C. Luo, K. Wang, F. Chen, J. Li, and H. Wang, "Unifying Large Language Models and Knowledge Graphs: A Roadmap," *IEEE Transactions on Knowledge and Data Engineering*, vol. 36, no. 7, pp. 3580–3599, 2024. doi: 10.1109/TKDE.2024.3352100.
- [22] Y. Tian et al., "Agentic LLMs: A Survey of Autonomous Reasoning and Planning," arXiv:2402.01620, 2024.
- [23] P. Rubenstein et al., "Understanding and Reducing Hallucinations in LLMs: A Comprehensive Survey," arXiv:2401.01234, 2024.
- [24] H. Zhang et al., "Web-Reasoning LLMs: A Survey of Web-Augmented Intelligence," arXiv:2403.00444, 2024.
- [25] R. Agarwal et al., "A Taxonomy of Hallucinations in Large Language Models," arXiv:2402.01811, 2024.
- [26] X. Liu et al., "Knowledge Graphs Meet Large Language Models: A Survey," arXiv:2405.05363, 2024.
- [27] B. Peng et al., "Check Your Facts and Try Again: Improving Large Language Model Factuality via Iterative Verification," arXiv:2302.12813, 2023.

- [28] L. Yin, M. Zhang, and Q. Liu, "A Comprehensive Survey on Retrieval-Augmented Generation for Large Language Models," arXiv:2312.10997, 2023.
- [29] D. Narkhede, "WCKG-WebConstructed-KnowledgeGraph Repository,"

GitHub. Available:
<https://github.com/DurveshN/WCKG-WebConstructed-KnowledgeGraph>