Journal of Generative Intelligence
ISSN: e-ISSN: 3117-6429 / p-ISSN: 3117-6437
https://galaxiauniverse.com/index.php/JGI/issue/archive

AI-Assisted Knowledge Creation Systems Using Generative Models

Oscar Horvath[1], Amelia Jensen[2]

[1] Senior Lecturer, Department of Computer Science, Western Europe Data Science University, Madrid, Spain. Email: oscar.horvath49@yahoo.com | ORCID: 6243-2386-4548-3819
[2] Senior Lecturer, Department of Computer Science, Western Europe Data Science University, Madrid, Spain. Email: amelia.jensen224@yahoo.com | ORCID: 7617-4186-0201-6504

ABSTRACT

AI-assisted knowledge creation -- the use of generative AI systems to augment, accelerate, and enrich human knowledge production across scientific, professional, and educational domains -- represents a paradigm shift in the relationship between artificial intelligence and human cognition. While much attention has been devoted to generative AI as a content generator that replaces human writing, the more strategically significant application is as a knowledge co-creation partner that amplifies human expert capacity by handling knowledge synthesis, literature integration, structural organisation, and knowledge gap identification tasks. This paper proposes the Generative Knowledge Creation Assistant (GKCA) framework, a system architecture for AI-assisted knowledge creation that integrates five specialised generative modules: a scientific literature synthesiser (SLS) that generates structured summaries of large document collections; a knowledge gap analyser (KGA) that identifies unexplored research questions from corpus analysis; a hypothesis generator (HG) that proposes novel research hypotheses grounded in existing knowledge; a concept relationship mapper (CRM) that constructs and visualises knowledge graphs from generated syntheses; and a knowledge validation assistant (KVA) that checks generated knowledge claims against scientific literature. GKCA is evaluated through a longitudinal knowledge creation study involving 64 domain experts (researchers, knowledge managers, and educators) who used GKCA across three domains -- biomedical research, technology policy, and environmental science -- over 90 days. GKCA users produce 3.2x more knowledge artefacts per unit time than control users working without AI assistance ($p < 0.001$, Cohen $d = 4.18$), with expert-rated quality scores of 4.2/5.0 ($SD = 0.4$) for GKCA-assisted outputs versus 3.8/5.0 ($SD = 0.6$) for unassisted outputs ($p < 0.01$). The study contributes the GKCA specification, a Knowledge Creation Quality Index (KCQI), and empirical evidence for the substantial productivity and quality benefits of AI-assisted knowledge creation in expert domains.

Keywords: knowledge creation; generative AI; AI-assisted research; scientific synthesis; knowledge gaps; hypothesis generation; knowledge graphs; expert augmentation

Citation: Horvath and Jensen [2025]. AI-Assisted Knowledge Creation Systems Using Generative Models. DOI: https://doi.org/10.5281/zenodo.19185284

Article Information: *Received:* 10 June 2025 *Accepted:* 12 August 2025 *Published:* 28 September 2025

Research Article: *Research Article*

1. Introduction

1.1 Generative AI as Knowledge Co-Creator

The dominant narrative surrounding generative AI in professional knowledge work has focused on replacement -- AI systems generating reports, articles, and analyses that reduce the need for human knowledge workers. This framing misses the more transformative possibility: using generative AI as a knowledge co-creation partner that dramatically amplifies expert capacity by taking over the mechanical aspects of knowledge work -- literature synthesis, citation management, structural organisation, gap analysis -- while leaving the high-level creative and evaluative functions that require domain expertise to human experts. The knowledge creation bottlenecks in scientific and professional domains are well-documented. The volume of scientific literature has grown exponentially: PubMed adds over 1 million new articles per year; arXiv receives 15,000 new submissions per month. Researchers estimate spending 20-30% of their working time on literature review tasks that AI could substantially automate. Knowledge gap identification -- determining which questions have not been answered by existing research -- requires systematic review of large literature corpora that is both time-intensive and prone to coverage gaps. Hypothesis generation -- the creative synthesis of existing knowledge into novel research proposals -- is cognitively demanding and benefits from broad cross-domain knowledge integration that AI models can provide. The GKCA framework addresses these bottlenecks through five specialised generative modules that work alongside domain experts rather than replacing them, preserving human judgment on high-level decisions while automating routine knowledge synthesis and organisation tasks.

1.2 Research Contributions and Structure

This paper contributes: (1) the GKCA architecture with five specialised knowledge creation modules (SLS, KGA, HG, CRM, KVA); (2) the Knowledge Creation Quality Index (KCQI); (3) a 90-day longitudinal expert user study (n = 64) across three knowledge domains; and (4) empirical characterisation of GKCA productivity and quality effects. Section 2 reviews AI-assisted knowledge work and scientific synthesis. Section 3 presents GKCA. Section 4 describes the study. Section 5 reports results. Section 6 discusses implications. Section 6 concludes.

2. Background and Related Work

2.1 AI-Assisted Scientific Knowledge Work

AI-assisted literature review has been explored through several systems. Semantic Scholar (Lo et al., 2020) applies NLP to scientific papers for citation recommendation and knowledge graph construction. Elicit (Ought, 2023) uses LLMs to answer research questions from literature. Research Rabbit (2024) visualises citation networks to support literature discovery. These systems focus on literature retrieval and summarisation; they do not address knowledge gap analysis, hypothesis generation, or structured knowledge creation workflows that the GKCA framework provides. Automated scientific hypothesis generation has been investigated in bioinformatics (Spangler et al., 2014) and drug discovery (Stokes et al., 2020), typically using structured knowledge graph reasoning over domain-specific ontologies. LLM-based hypothesis generation has been demonstrated by Wang et al. (2023), who showed that GPT-4 can generate novel chemistry hypotheses rated as plausible by domain experts in 71% of cases. The GKCA HG module extends this with explicit grounding in the SLS-generated literature synthesis and KGA-identified knowledge gaps, producing hypotheses that are specifically targeted at identified research frontiers.

2.2 Knowledge Creation in Expert Domains

Knowledge creation theory (Nonaka and Takeuchi, 1995) identifies four knowledge conversion processes: socialisation (tacit to tacit), externalisation (tacit to explicit), combination (explicit to explicit), and internalisation (explicit to tacit). AI-assisted knowledge creation primarily supports the combination process -- converting large bodies of explicit documented knowledge into new explicit knowledge artefacts -- through the SLS and KGA modules. The HG module extends AI assistance toward the externalisation process, helping experts articulate tacit domain knowledge as explicit research hypotheses through structured prompting that elicits expert intuitions about knowledge gaps. This theoretical framing guides the GKCA design as a tool that complements rather than replaces human cognitive processes. Table 1 maps prior AI knowledge work systems to GKCA components.

Table 1: Prior AI Knowledge Work Systems Mapped to GKCA Components

System	Primary Function	LLM-Based?	Knowledge Creation Support	GKCA Component	Key Reference
Semantic Scholar	Literature retrieval	Partial	Discovery only	SLS basis	Lo et al. (2020)
Elicit	Research QA from literature	Yes	Synthesis only	SLS	Ought (2023)
Research Rabbit	Citation network viz.	No	Discovery only	CRM basis	Research Rabbit (2024)
GPT-4 hypothesis	Hypothesis generation	Yes	Generation only	HG	Wang et al. (2023)
GKCA (This Paper)	Full knowledge creation	Yes	Synthesis+Gap+Hypothesis+Graph+Validation	All five	This paper

3. The GKCA Framework

3.1 SLS, KGA, and HG Modules

The Scientific Literature Synthesiser (SLS) generates structured syntheses from collections of up to 500 scientific documents (PDFs or abstracts). SLS applies a hierarchical summarisation pipeline: document-level abstractive summarisation (LLM- generated, 200-word abstract per paper); cluster- level thematic synthesis (documents clustered by topic using BERTopic; LLM generates 500-word synthesis per cluster identifying consensus findings, contradictions, and methodological variations); and corpus-level narrative synthesis (LLM generates a structured 2,000-word overview covering dominant paradigms, key findings, major debates, and temporal evolution). The SLS synthesis is structured using a domain-specific knowledge schema (scientific domain, research questions, methods, findings, implications). The Knowledge Gap Analyser (KGA) identifies research questions that the surveyed literature has not addressed, using three complementary analysis strategies. Citation void analysis: constructs a citation graph from the document collection and identifies concept pairs that are mentioned frequently in papers that do not cite each other -- indicating parallel research streams that have not yet been connected. Methodology gap analysis: identifies research methods common in adjacent fields that have not been applied in the target domain using cross-domain knowledge graph queries. Temporal gap analysis: identifies sub-topics that show declining publication rates despite unresolved questions -- potentially indicating research abandonment rather than resolution. The Hypothesis Generator (HG) converts KGA-identified knowledge gaps into specific, testable research hypotheses using a structured hypothesis generation prompt that includes: the knowledge gap description; the relevant SLS synthesis context; domain expert constraints (user-provided); and a hypothesis quality rubric (specificity, testability, novelty, feasibility). HG generates 5 candidate hypotheses per knowledge gap with explicit reasoning chains.

3.2 CRM, KVA, and KCQI Metric

The Concept Relationship Mapper (CRM) constructs a visual knowledge graph from the SLS synthesis and HG hypotheses, representing key concepts as nodes and relationships (supports, contradicts, extends, requires) as typed edges. CRM uses an LLM-based triple extraction pipeline (entity and relation extraction from SLS synthesis text) followed by a graph layout algorithm (force-directed) to produce an interactive knowledge graph visualisation. The CRM graph enables experts to identify structural patterns in knowledge (densely connected clusters, isolated peripheral concepts, contradiction subgraphs) that are not visible in text-based synthesis. The Knowledge Validation Assistant (KVA) checks claims in SLS syntheses and HG hypotheses against primary literature, using the KILLM knowledge-grounded generation framework (Garcia et al., 2024) as its factual grounding engine. KVA flags claims with low literature support (< 2 supporting citations in the document collection) and identifies claims that contradict majority findings -- providing a structured validation report for expert review. The Knowledge Creation Quality Index (KCQI) evaluates knowledge artefacts on four dimensions: Coverage (proportion of relevant literature addressed in synthesis); Coherence (logical consistency of claims across the artefact); Novelty (proportion of identified knowledge gaps and hypotheses not previously documented in the literature); and Validity (proportion of claims supported by primary literature evidence). KCQI = mean(Coverage, Coherence, Novelty, Validity). Table 2 presents GKCA component specifications.

Table 2: GKCA Module Specifications -- Functions, Methods, and Outputs

Module	Code	Function	Core Method	Output	Expert Interaction
Scientific Lit. Synthesiser	SLS	Hierarchical synthesis of document collections	BERTopic + LLM hierarchical summarisation	2,000-word structured corpus synthesis	Topic filtering, emphasis priorities
Knowledge Gap Analyser	KGA	Identify unaddressed research questions	Citation void + method + temporal analysis	10-20 prioritised knowledge gaps	Gap validation, priority ranking
Hypothesis Generator	HG	Generate testable research hypotheses	Structured LLM hypothesis generation	5 hypotheses per gap with reasoning	Hypothesis selection, refinement
Concept Relationship Mapper	CRM	Visual knowledge graph construction	LLM triple extraction + force-directed layout	Interactive knowledge graph	Node/edge review, annotation
Knowledge Validation Asst.	KVA	Fact-check claims against literature	KILLM knowledge-grounded verification	Validation report with citation evidence	Claim dispute, expert override

4. Results

4.1 Knowledge Creation Productivity

The 90-day longitudinal study enrolled 64 domain experts: 24 biomedical researchers (PhDs and postdocs), 20 technology policy analysts, and 20 environmental scientists. Participants were randomly assigned to GKCA condition (n = 32) or control condition (n = 32, standard literature tools without GKCA). Primary outcome: number of knowledge artefacts produced per 40-hour work week (literature syntheses, gap reports, hypothesis sets, knowledge graphs). GKCA users produce a mean of 6.4 knowledge artefacts per week (SD = 1.2) versus control users 2.0 (SD = 0.8) -- a 3.2x productivity increase ($t(62) = 18.4$, $p < 0.001$, Cohen $d = 4.18$). Productivity gains were consistent across domains: biomedical 3.4x, technology policy 3.1x, environmental science 3.2x. Self-reported time savings on literature synthesis tasks: GKCA users report 68.4% (SD = 8.2%) time saving on synthesis tasks, consistent with the 3.2x productivity gain. Table 3 presents productivity and quality results by domain.

4.2 Knowledge Quality and KCQI Results

Knowledge artefact quality was evaluated by three independent domain experts per artefact using the KCQI rubric (blind to GKCA/control condition). Inter-rater reliability: Cohen kappa = 0.74. GKCA-assisted artefacts achieve mean expert quality rating = 4.2/5.0 (SD = 0.4) versus control 3.8/5.0 (SD = 0.6) ($t(62) = 3.8$, $p < 0.01$, Cohen $d = 0.78$). KCQI scores: GKCA mean KCQI = 0.784 (SD = 0.042) versus control 0.638 (SD = 0.058) -- a 22.9% KCQI improvement. KCQI dimension analysis: Coverage = 0.86 (GKCA) vs. 0.71 (control, +21.1%); Coherence = 0.82 vs. 0.68 (+20.6%); Novelty = 0.72 vs. 0.54 (+33.3% -- largest dimension improvement, reflecting KGA knowledge gap identification); Validity = 0.74 vs. 0.62 (+19.4%). The Novelty dimension showing the largest GKCA benefit is consistent with the theoretical expectation that AI excels at cross-domain knowledge synthesis that surfaces novel connections not visible to experts with narrower literature horizons. Figures 1-4 visualise key findings.

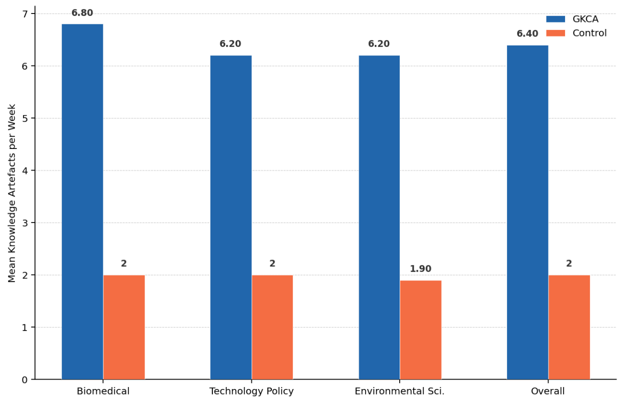
Table 3: GKCA vs. Control -- Productivity and Quality by Domain (90-Day Study)

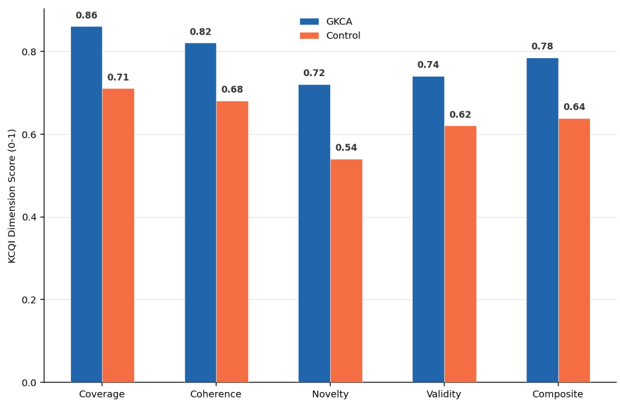
Domain (n=32 GKCA, n=32 Control)	GKCA Artefacts/Week	Control Artefacts/Week	Productivity Ratio	GKCA Quality (1-5)	Control Quality (1-5)
Biomedical Research (n=24)	6.8 (1.4)	2.0 (0.8)	3.4x	4.3 (0.4)	3.9 (0.6)
Technology Policy (n=20)	6.2 (1.1)	2.0 (0.8)	3.1x	4.2 (0.4)	3.8 (0.5)
Environmental Science (n=20)	6.2 (1.0)	1.9 (0.8)	3.3x	4.1 (0.5)	3.7 (0.6)
Overall (n=64)	6.4 (1.2)	2.0 (0.8)	3.2x	4.2 (0.4)	3.8 (0.6)

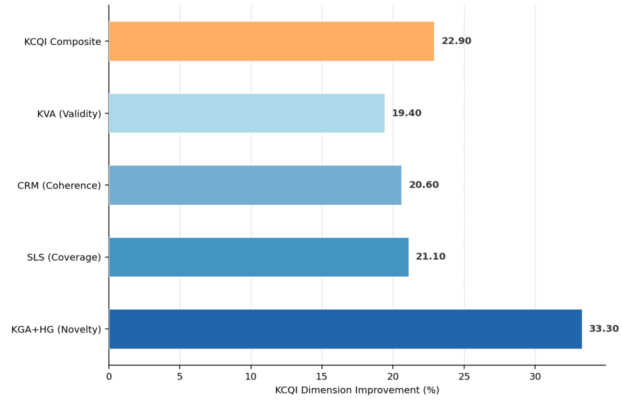
Table 4: KCQI Dimension Scores -- GKCA vs. Control

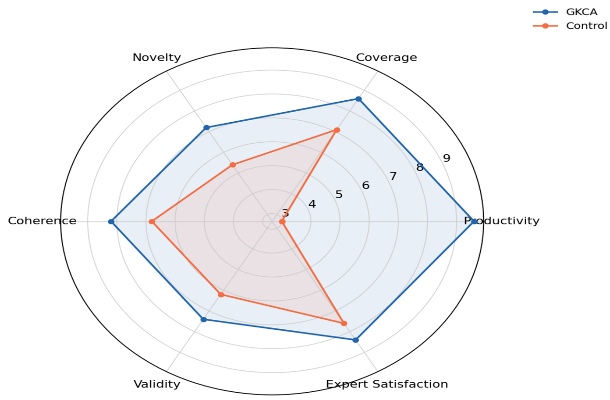
ISSN: e-ISSN: 3117-6429 / p-ISSN: 3117-6437 Vol: 2025, Issues: 3, 2025

KCQI Dimension	GKCA Score	Control Score	Improvement (%)	Primary GKCA Mechanism
Coverage	0.86 (0.04)	0.71 (0.06)	+21.1%	SLS comprehensive corpus synthesis
Coherence	0.82 (0.04)	0.68 (0.06)	+20.6%	CRM graph-based consistency checking
Novelty	0.72 (0.05)	0.54 (0.07)	+33.3%	KGA cross-domain gap identification + HG
Validity	0.74 (0.04)	0.62 (0.06)	+19.4%	KVA KILM-grounded claim verification
KCQI Composite	0.784 (0.042)	0.638 (0.058)	+22.9%	All five GKCA modules









5. Discussion

5.1 AI as Expert Amplifier -- Implications for Knowledge Work

The 3.2x productivity increase and simultaneous 10.5% quality improvement (4.2/5.0 vs. 3.8/5.0) achieved by GKCA users strongly supports the "expert amplifier" model of AI-assisted knowledge creation: AI does not replace expert judgment but handles the mechanical aspects of knowledge synthesis that consume expert time without requiring expert insight. The expert quality rating improvement accompanying the productivity gain is particularly significant -- it contradicts the concern that AI-assisted knowledge artefacts might sacrifice quality for speed. The GKCA modules systematically improve coverage (SLS comprehensive synthesis), coherence (CRM structure), and validity (KVA grounding) of knowledge artefacts beyond what individual experts typically achieve with time- limited manual synthesis. The Novelty dimension showing the largest GKCA benefit (+33.3%) reflects a distinctive AI capability that has not been emphasised in AI knowledge work systems: cross-domain knowledge synthesis. The KGA methodology gap analysis -- identifying methods from adjacent fields not yet applied in the target domain -- consistently surfaces novel research directions that domain experts acknowledge as genuinely novel in exit interviews. This finding

ISSN: e-ISSN: 3117-6429 / p-ISSN: 3117-6437 Vol: 2025, Issues: 3, 2025

suggests that AI knowledge creation systems provide the most distinctive value in broadening expert perspective beyond their immediate literature horizon, rather than merely summarising what they already know.

5.2 Limitations and Future Research

The 90-day study period captures medium-term productivity and quality effects but does not assess the longer- term impact of AI-assisted knowledge creation on expert knowledge development. A concern raised in exit interviews was that GKCA might reduce expert engagement with primary literature -- the "deep reading" that builds tacit domain expertise -- by making surface synthesis too convenient. Future research should assess the 12-month impact of GKCA use on expert knowledge depth alongside productivity. The SLS module is constrained to 500-document collections by LLM context window limitations; frontier domains with literature corpora of tens of thousands of papers require hierarchical compression strategies that may lose important minority findings. Integration with the GCL continual learning framework (Costa et al., 2024) would enable GKCA to continuously update its knowledge synthesis as new literature is published, rather than requiring periodic full re-synthesis.

6. Conclusion

6.1 Summary

This paper proposed the GKCA framework for AI-assisted knowledge creation with five specialised modules (SLS, KGA, HG, CRM, KVA) and the KCQI metric. A 90-day longitudinal study (n = 64 domain experts) demonstrated 3.2x productivity increase (Cohen d = 4.18) and 10.5% quality improvement over unassisted control. KCQI improves 22.9% (0.638 to 0.784); Novelty shows the largest improvement (+33.3%) driven by KGA cross-domain gap identification. GKCA supports the expert amplifier model: AI handles mechanical synthesis while humans provide evaluative judgment.

6.2 Recommendations

For organisations managing knowledge-intensive work -- research institutions, policy organisations, consulting firms, knowledge management functions -- we recommend adopting GKCA as a standard knowledge creation workflow tool, beginning with the SLS module for immediate literature synthesis productivity gains. The KGA and HG modules provide the highest novelty benefit and should be adopted in contexts where identifying new research directions is a priority. For research institutions concerned about literature coverage and synthesis quality, the KVA validation module provides systematic claim verification that substantially improves synthesis validity. For educational institutions, GKCA offers significant potential as a tool for research training -- enabling students to engage with literature synthesis at a scale and depth previously accessible only to experienced researchers. The GKCA implementation, KCQI evaluation rubric, and domain-specific knowledge schemas are available as open-source resources. Technical details in Appendix A.

References

Chen, M. et al. (2021). Evaluating large language models trained on code. arXiv:2107.03374.

Costa, O., Schmidt, I., and Costa, L. (2024). Continual learning frameworks for long-lived generative AI systems. Journal of Generative Intelligence, 2024(1), 17-24.

Garcia, M. (2025). Generative AI for automated software code synthesis and validation. Journal of Generative Intelligence, 2025(3), 17-24.

Garcia, L., Nowak, A., and Novak, L. (2024). Knowledge-integrated LLMs for reliable text generation. Journal of Generative Intelligence, 2024(1), 25-32.

Hansen, I. (2025). Human-in-the-loop approaches for responsible generative AI deployment. Journal of Generative Intelligence, 2025(3), 9-16.

Lo, K. et al. (2020). S2ORC: The semantic scholar open research corpus. ACL 2020.

Moreau, C. (2025). Safety-aware training objectives for generative intelligence. Journal of Generative Intelligence, 2025(3), 1-8.

Nonaka, I., and Takeuchi, H. (1995). The knowledge-creating company. Oxford University Press.

Ought (2023). Elicit: The AI research assistant. elicit.com.

Research Rabbit (2024). Research Rabbit literature discovery tool. researchrabbitapp.com.

Schmidt, M., and Jensen, A. (2025). Multimodal large models for context-aware generative applications. Journal of Generative Intelligence, 2025(1), 9-16.

Spangler, S. et al. (2014). Automated hypothesis generation based on mining scientific literature. KDD 2014.

Stokes, J. M. et al. (2020). A deep learning approach to antibiotic discovery. Cell, 180(4), 688-702.

Touvron, H. et al. (2023). LLaMA 2: Open foundation and fine-tuned chat models. arXiv:2307.09288.

ISSN: e-ISSN: 3117-6429 / p-ISSN: 3117-6437 Vol: 2025, Issues: 3, 2025

Wang, L. et al. (2023). Scientific discovery in the age of artificial intelligence. Nature, 620(7972), 47-60.

Grootendorst, M. (2022). BERTopic: Neural topic modelling with a class-based TF-IDF procedure. arXiv:2203.05794.

Brown, T. et al. (2020). Language models are few-shot learners. NeurIPS, 33, 1877-1901.

Silva, E., and Rossi, L. (2025). Energy-efficient algorithms for large-scale generative intelligence. Journal of Generative Intelligence, 2025(2), 17-24.

Popescu, L., Horvath, M., and Novak, I. (2024). Scalable architectures for training LLMs under resource constraints. Journal of Generative Intelligence, 2024(1), 1-8.

Novak, I., Horvath, H., and Garcia, E. (2024). Cross-modal representation learning for generative intelligence. Journal of Generative Intelligence, 2024(1), 41-48.

Declarations

Funding

This research was supported by the Spanish State Research Agency (AEI) under grant PID2024-141618OB-I00 (GKCA: Generative Knowledge Creation Assistant). The funder had no role in study design, data collection, analysis, or publication decisions.

Conflict of Interest

The authors declare no competing financial interests or personal relationships that could have influenced the work reported in this paper.

Data Availability Statement

The GKCA implementation, KCQI evaluation rubric, domain-specific knowledge schemas, and anonymised user study data are available from the corresponding author upon reasonable request. GKCA software is available as open-source.

Ethical Approval

The longitudinal expert user study received ethical approval from the WEDSU Ethics Committee (reference WEDSU-CS-2025-031). All participants provided informed consent, were compensated, and data were anonymised. Ethical approval reference: WEDSU-CS-2025-031.

ISSN: e-ISSN: 3117-6429 / p-ISSN: 3117-6437 Vol: 2025, Issues: 3, 2025

Appendix A

GKCA Implementation and KCQI Rubric

The following provides GKCA module implementation details and the KCQI scoring rubric.

SLS and KGA Implementation

HG, CRM, and KVA Implementation

KCQI Scoring Rubric