A Gradio-Based Toolkit for Remote Sensing Data Fusion Literature

Ontology-Backed Curation with Basic Uncertainty Tagging

Caleb Cheruiyot[1], Jonathan P. Leidig[1] and Jiaxin Du[1,*]

[1]College of Computing, Grand Valley State University, Allendale, MI, USA

Abstract

We present a proof-of-concept toolkit focused on remote sensing data fusion literature that turns research articles into searchable, ontology-backed records and a lightweight knowledge graph. The system offers an interactive, Gradio-based interface featuring searchable cards and a simple subgraph preview. Uncertainty is incorporated as basic, human-editable tags. A small, binary text-classification component (BERT/RoBERTa) supports triage of abstracts into "fusion-related" vs. "other remote sensing" to aid curation. We demonstrate the end-to-end pipeline and provide preliminary classifier metrics on a small labeled set, emphasizing the system's scope as a practical starting point. Code, ontology, and configuration are released for reproducibility. This work supports the AI-for-science goal of developing transparent, interpretable, and uncertainty-aware AI tools for scientific knowledge integration.

Keywords

Remote Sensing, Data Fusion, Knowledge Graph, Uncertainty Tagging, Gradio, BERT

1. Introduction

Data fusion—the integration of heterogeneous data sources into a unified representation—is critical for enhancing coverage, resolution, and interpretability across diverse domains. A widely adopted structure for organizing fused data is the knowledge graph, which represents entities and their semantic relationships as nodes and edges, enabling queryable, interconnected representations [1]. However, constructing and maintaining such structured knowledge remains labor-intensive, especially when accounting for uncertainty inherent in source data. Recent research increasingly highlights the dual importance of data fusion and uncertainty modeling [2, 3]. For example, [4] demonstrated the value of integrating multi-scale measurement data to reveal latent relationships. Similarly, BUGPan [5] effectively addressed spatial uncertainty in image fusion, underscoring the necessity of managing uncertainty for robust integration. Given the resource-intensive nature of building knowledge graphs, emerging approaches aim to improve efficiency by combining ML-based toolkits [7] with strategically minimized human input [6].

Building on these developments, we present a lightweight, end-to-end toolkit that transforms remote sensing data fusion literature into searchable records and an ontology-backed knowledge graph with basic uncertainty annotations. Implemented via a Gradio based interface [1], the system supports search, graph previews, and triage using a binary text classifier. Specifically, we offer:

- An ontology-driven schema linking papers, datasets, and fusion methods in the remote sensing domain.
- A lightweight binary classifier (BERT/RoBERTa) for abstract-level triage.
- Basic uncertainty tagging to support interpretability.
- A simple, reproducible interface for search and visualization, with released artifacts for extension.

The 2nd International Workshop on Artificial Intelligence for the Science of Science (AI4SciSci 2025), co-located with ACM/IEEE-CS JCDL 2025

*Corresponding author.
cheruiyc@mail.gvsu.edu (C. Cheruiyot); leidijon@gvsu.edu (J. P. Leidig); dujia@gvsu.edu (J. Du)

[1]https://www.gradio.app

In the context of digital libraries, such lightweight AI tools can accelerate domain curation while preserving human oversight. By emphasizing interpretability and incremental uncertainty annotation, our approach aligns with AI4SciSci's vision of transparent and reproducible AI systems for scientific knowledge integration.

2. System Architecture

A modular data pipeline was designed to integrate structured data handling, semantic enrichment, machine learning, and lightweight visualization. Figure 1 illustrates the high-level architecture of the proposed toolkit. The architecture is organized into five layers: *data ingestion*, *storage*, *semantic enrichment*, *AI-based classification*, and *user interaction*. This layered design supports flexible querying, ontology-based linkage, and *editable uncertainty tagging* across datasets and fusion methods. At this stage, uncertainty metadata are qualitative and manually assigned rather than formally modeled or propagated computationally.

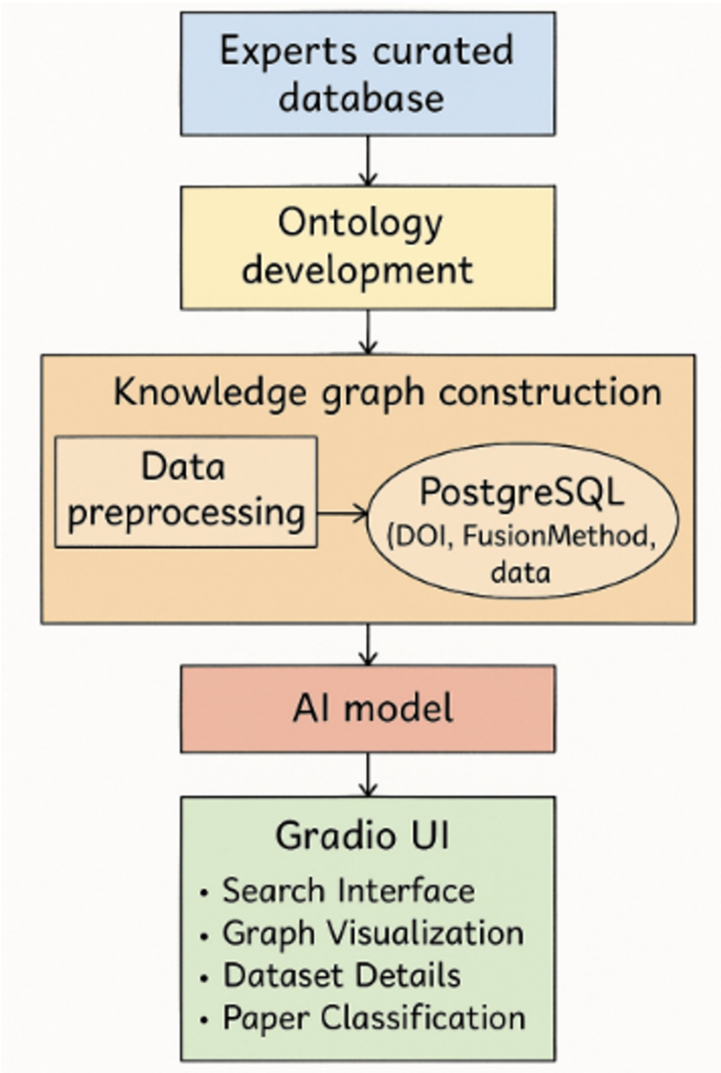

Figure 1: System architecture. The prototype centers on literature ingestion (remote sensing domain), ontology-backed storage, a small binary classifier for abstract triage, and a Gradio-based user interface supporting basic search and subgraph preview.

The pipeline begins with *data ingestion*, where tabular datasets (CSV and Excel) describing research papers, data fusion techniques, and datasets are pre-processed using pandas and stored in a `PostgreSQL` database. These structured data are exported to an RDF-compatible schema through an OWL ontology layer implemented with `OWLReady2`[2], enabling semantic relationships between papers, datasets, fusion methods, and associated uncertainty tags [7]. A semantic knowledge graph is then generated from the database using `NetworkX`[3] to visualize entity relationships and support graph-based exploration.

To assist document triage, a small `BERT`-based classifier [8] identifies whether an abstract is fusion-related or general remote sensing. This classification facilitates organization within the ontology but does not directly determine graph structure. Finally, a `Gradio`[4] interface enables users to search, filter, and interpret curated records through card views, keyword filters, subgraph previews, and classification outputs, completing the end-to-end process from ingestion to human-interpretable exploration [9].

3. Methodology

3.1. Ontology Development

An OWL ontology was designed to provide semantic structure for literature-level representation of remote sensing data fusion. Rather than aiming for comprehensive domain coverage, the ontology was intentionally kept lightweight and application-driven, with the goal of supporting search, linking, and human-editable uncertainty tagging within a small curated corpus. This design was informed by prior work on scholarly knowledge graphs [7, 6], which highlighted the trade-offs between comprehensive coverage and usability for expert-driven curation. Existing ontologies were considered; however, they proved too complex, incomplete, or insufficiently tailored to the remote sensing data fusion domain, motivating the minimal schema adopted here.

The core classes include `FusionMethod`, `Dataset`, and `Uncertainty`, with supporting classes such as `Paper` and `Publisher`. These are connected through properties including `integrates`, `usesData`, `isPublishedIn`, and `hasUncertainty`. A numeric attribute, `hasConfidenceLevel`, allows storage of confidence-related indicators extracted from text or assigned by curators. Uncertainty modeling was motivated by existing research on scientific uncertainty detection and integrated modeling [2, 3]. Instead of formal reasoning or probabilistic propagation, uncertainty is represented as qualitative and numeric metadata to support filtering and interpretation during search.

The ontology was implemented using the `owlready2` library and serialized in RDF format for integration with the knowledge graph pipeline. Structural consistency was verified using the Owlready2 reasoner to ensure valid class and property definitions. In addition, a subset of entities and relations was manually reviewed by domain experts to confirm semantic correctness. No external ontology alignment or quantitative quality metrics were applied in this prototype. Figure 2 illustrates the ontology structure and its key relationships.

3.2. Knowledge Graph Construction

Two domain-specific knowledge graphs were constructed to support exploration of *data fusion* and *remote sensing* literature [1]. Source data consisted of expert-curated tables describing publications, datasets, and fusion methods. Preprocessing included normalization, removal of non-standard characters, case harmonization, and basic consistency checks. The cleaned data were stored in a PostgreSQL database to enable structured querying and controlled schema evolution.

A graph representation was then generated using the `NetworkX` library. Nodes correspond to entities defined in the ontology (e.g., `Paper`, `Dataset`, `FusionMethod`), while directed edges encode semantic relations such as `integrates` and `usesData`. Only valid entity links were instantiated to preserve referential integrity.

Each graph instance includes optional uncertainty annotations inherited from ontology metadata.

[2]https://owlready2.readthedocs.io/
[3]https://networkx.org
[4]https://www.gradio.app

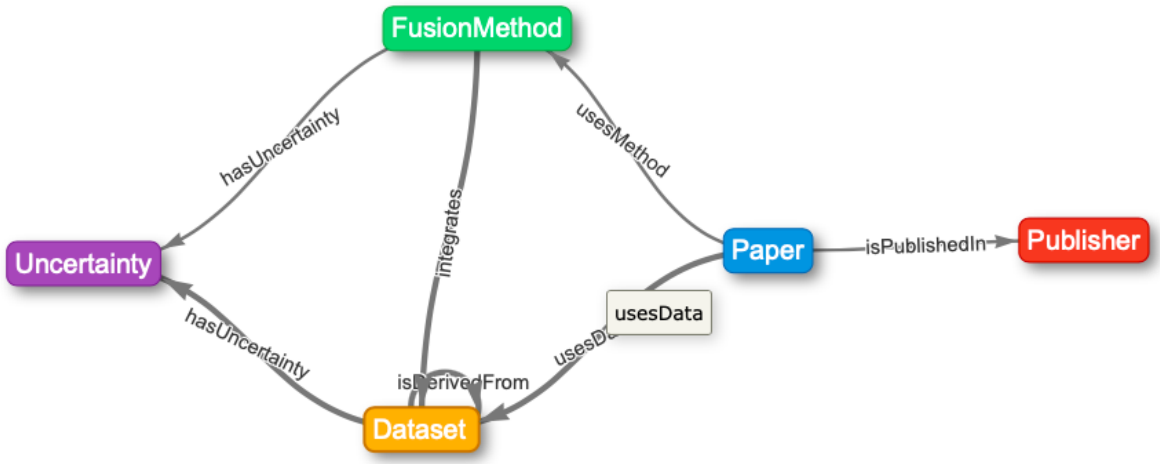

Figure 2: Data fusion ontology linking papers, datasets, and fusion methods with editable uncertainty tags; no formal uncertainty reasoning.

Graphs were exported in GraphML format and rendered for interactive inspection using `pyvis` [5]. In parallel, the relational backend supports keyword- and attribute-based queries. Separate graphs were maintained for data fusion and general remote sensing to align with the binary classifier's scope and reduce semantic drift between domains [9]. As the toolkit is a prototype, evaluation focused on structural validity, correct linkage of entities, and interface consistency rather than formal graph metrics.

3.3. Text Classification

Although most literature tagging in the system was performed manually, machine learning was incorporated to assist abstract-level triage by identifying whether a paper is related to data fusion or general remote sensing. Following human-supervised knowledge curation practices [9], a transformer-based classifier was used as a support tool rather than as a fully automated decision mechanism.

We fine-tuned a pre-trained BERT model [8] for binary sequence classification using abstract text. Tokenization was performed via the HuggingFace tokenizer, and training used mini-batch optimization with class weighting and random oversampling to address label imbalance. Experiments were also conducted using a RoBERTa variant for comparison, though BERT achieved the best overall performance.

A total of 120 abstracts were independently annotated by two domain experts, achieving strong agreement (Cohen's $\kappa = 0.86$) [10]. The dataset was divided into 80% for training and 20% for validation using stratified sampling. Because of the limited data size, a separate test set was not created, and the results are presented as a feasibility assessment rather than a benchmark evaluation.

Predictions with confidence below 0.6 were flagged and reviewed by human annotators through the Gradio interface, enabling iterative correction and refinement. Although no classical baselines (e.g., TF–IDF with SVM) were included in this prototype, future work will incorporate comparative evaluations and larger-scale validation.

3.4. Visualization Interface

The user interface was implemented using Gradio, chosen for its simplicity and flexibility in building interactive web applications. The interface provides four main views: *Search*, *Graph*, *Classify*, and

[5] https://pypi.org/project/pyvis/

Ontology. Users can query the system by research topic, fusion method, or dataset name, and results are presented as styled cards with DOI links, metadata, abstracts, and embedded uncertainty tags.

Uncertainty is visualized using color-coded highlights and tooltips to improve interpretability. The ontology view presents a structured outline of the knowledge model, while the graph view renders filtered subgraphs for interactive exploration. The classification tab allows users to submit abstracts and receive automated predictions regarding their relevance to data fusion. Figure 3 shows an overview of the interface.

Overall, the interface serves as a lightweight digital library front end [1], enabling users to explore ontology-backed literature relationships while observing uncertainty cues and classification outputs interactively.

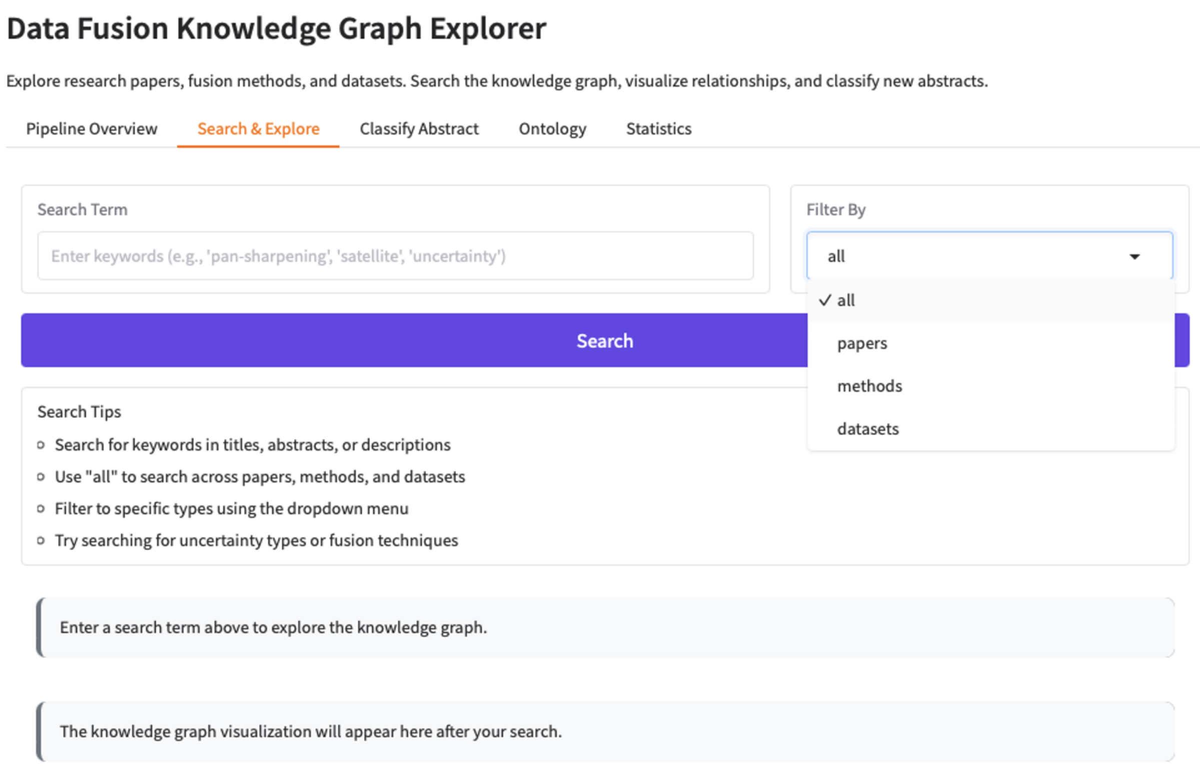

Figure 3: Overview of the user interface components.

4. Results and Evaluation

The evaluation of the developed system was conducted across multiple dimensions, including search accuracy, classification reliability, interface responsiveness, and overall system robustness. The preliminary results demonstrate that the integrated architecture performs effectively in delivering dynamic, context-aware domain insights with associated uncertainties.

4.1. Search and Query Evaluation

The search engine embedded in the user interface mitigates common user query issues through exact and partial matching via Python's `difflib.get_close_matches()` function, misspelling corrections, etc. Test cases were designed to validate core functionalities such as searching by paper title, dataset name, and fusion method using related keywords. In all tested scenarios, the system returned card-styled results with enriched metadata, abstracts, and DOI references that were deemed relevant by expert evaluators. Across 20 representative test queries spanning fusion methods, datasets, and publication titles, the system achieved full accuracy on exact matches and an average relevance of 92% for partial matches, as assessed by two domain experts. Each retrieved record displayed associated uncertainty tags [2] from the ontology, providing contextual cues for reliability and potential data ambiguity.

4.2. Text Classification Performance

To assess the effectiveness of transformer-based models for abstract triage, both `BERT-Base-Uncased` and `RoBERTa-Base` were evaluated on the labeled abstract set. Table 1 summarizes the performance results.

The fine-tuned `BERT` model achieved an accuracy of 0.966 and an F1 score of 0.963 on the validation subset, outperforming the `RoBERTa` variant. Precision was perfect for the fusion class, while recall was slightly lower, indicating conservative classification behavior. These results suggest that transformer-based classifiers are effective for abstract-level domain filtering even with limited data.

Classification outputs were linked to ontology instances such as `FusionMethod` and `Dataset`, enabling automated tagging during knowledge graph construction. Low-confidence predictions were routed to human annotators via the interface for confirmation or correction, preventing classification errors from propagating into the graph.

Because the dataset is small and manually curated, results should be interpreted as evidence of feasibility rather than generalization performance. Nonetheless, the findings confirm that lightweight fine-tuning can meaningfully support expert-driven literature organization.

Table 1
Summary of model performance on 120 labeled abstracts.

Model	Accuracy	Precision	Recall	F1 Score	Avg.
BERT-Base-Uncased	0.966	1.000	0.928	0.963	0.964
RoBERTa-Base	0.828	0.764	0.928	0.838	0.834

4.3. Interface Responsiveness and User Experience

The interface was evaluated for usability and performance. Pagination mechanisms were tested by injecting simulated datasets to ensure stable loading and filtering. Visual transitions, layout adjustments, resizing, and mobile responsiveness were validated. Functional buttons such as "View Fusion Possibilities" and "View Data Fusion Recommendations" successfully directed users to relevant contextual sections, activated embedded logic, and met anticipated use case scenarios. A small pilot evaluation with three research assistants (familiar with remote sensing data) confirmed that the color-coded uncertainty indicators improved interpretability, allowing users to distinguish between confirmed and tentative links in the graph. Participants highlighted the system's transparency and ease of use, though they also noted the need for richer visual encoding of uncertainty in future iterations. Figure 4 shows a sample knowledge graph query highlighting multiple aspects of uncertainty within search results.

5. Discussion

This work demonstrates the feasibility of a lightweight toolkit that integrates ontology-based representation, uncertainty tagging, and human-in-the-loop machine learning for organizing remote sensing data fusion literature. The system prioritizes interpretability and usability over full automation, supporting expert-guided exploration rather than autonomous knowledge extraction.

The ontology was intentionally designed as a minimal, application-driven schema to enable linking, filtering, and navigation across papers, datasets, and fusion methods. Structural consistency was verified using a reasoner, and concepts were refined through iterative curation. Uncertainty is represented as

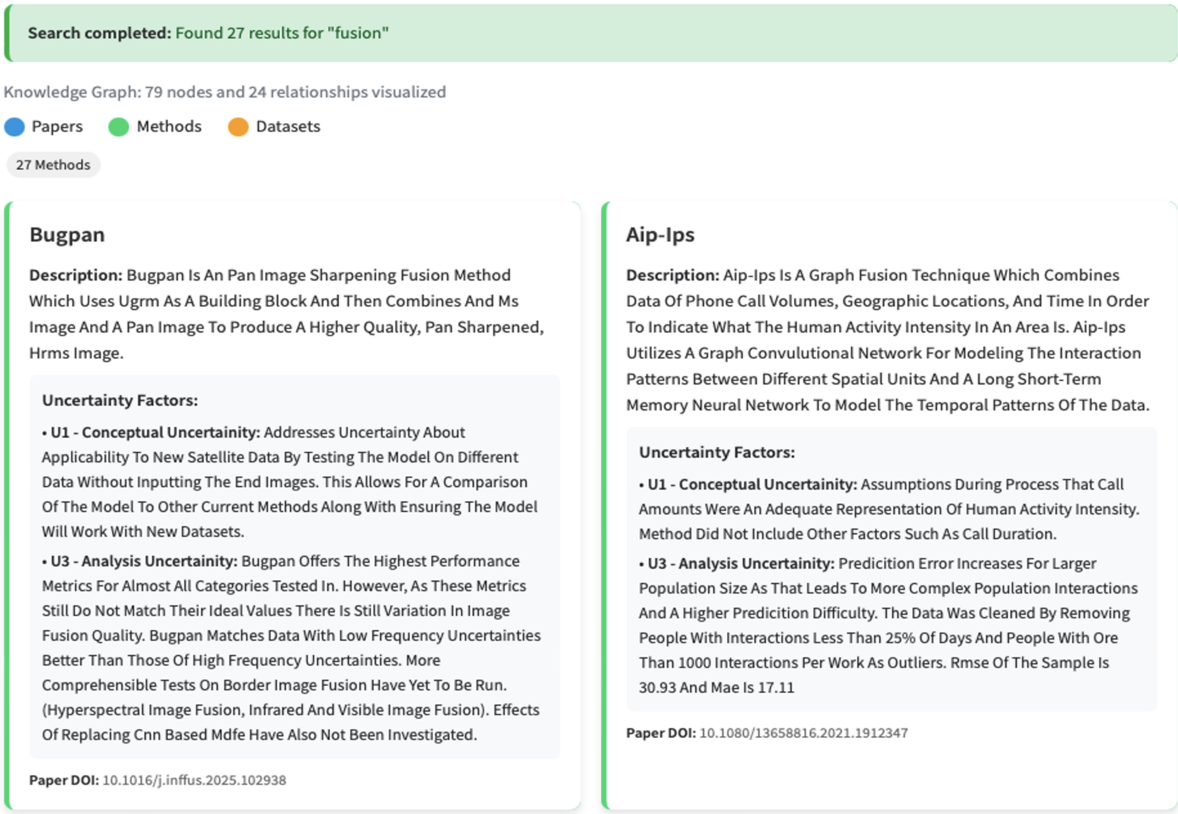

Figure 4: Sample test case searching by data fusion method.

qualitative metadata to support interpretability rather than formal inference, consistent with prior work on uncertainty modeling in scientific knowledge systems [2, 3].

The classifier serves as a triage mechanism, with expert validation of low-confidence predictions to prevent error propagation. While evaluation is limited in scale, the results indicate that lightweight fine-tuning can effectively support expert workflows.

Limitations include the small dataset, absence of large-scale benchmarking, and lack of automated uncertainty reasoning. These reflect the prototype nature of the system and motivate future work on ontology alignment, uncertainty modeling, and scaling.

6. Conclusion and Future Work

This paper presented a Gradio-based toolkit for curating remote sensing data fusion literature into searchable, ontology-backed records with editable uncertainty annotations. A lightweight transformer-based classifier supports abstract-level triage, with expert validation integrated to preserve reliability and interpretability.

Future work will focus on expanding ontology coverage, introducing uncertainty provenance tracking, and incorporating baseline model comparisons. Additional directions include structured user evaluations, multi-label classification, and scaling the framework to larger literature collections. Interoperability with external knowledge sources will also be explored.

Artifacts (code, ontology, and configuration) are available at https://doi.org/10.6084/m9.figshare.29914361 to support reproducibility and community engagement.

Acknowledgments

We extend our sincere gratitude to Timothy Mugambi, Nancy Odhiambo, and Braiden Betway for their invaluable contributions to this research project. This work was supported by seed funding from the College of Computing at Grand Valley State University.

Declaration on Generative AI

During the preparation of this manuscript, the author(s) used Generative AI tools for limited purposes such as grammar and language refinement. All technical content, interpretation of results, and conclusions were created by the author(s). The author(s) reviewed and edited the output as needed and take(s) full responsibility for the accuracy, originality, and integrity of the work in compliance with the CEUR Generative AI policy.

References

[1] A. Oelen, M. Y. Jaradeh, M. Stocker, and S. Auer, "Generate FAIR literature surveys with scholarly knowledge graphs," in *Proceedings of the ACM/IEEE Joint Conference on Digital Libraries (JCDL 2020)*, Wuhan, China, June 2020, pp. 97–106, doi: 10.1145/3383583.3398520.

[2] M. Kirchner, H. Mitter, U. A. Schneider, M. Sommer, K. Falkner, and E. Schmid, "Uncertainty concepts for integrated modeling – Review and application for identifying uncertainties and uncertainty propagation pathways," *Environmental Modelling & Software*, vol. 135, p. 104905, 2021, doi: 10.1016/j.envsoft.2020.104905.

[3] P. K. Ningrum, P. Mayr, and I. Atanassova, "UnScientify: Detecting scientific uncertainty in scholarly full text," in *Proceedings of the Joint Workshop on the 4th EEKE and the 3rd AII, co-located with JCDL 2023*, Santa Clara, CA, USA, June 2023, pp. 52–58.

[4] A. P. Wright, S. Davidoff, and D. H. Chau, "Nested fusion: A method for learning high-resolution latent structure of multi-scale measurement data on Mars," in *Proceedings of the 30th ACM SIGKDD Conference on Knowledge Discovery and Data Mining (KDD 2024)*, Barcelona, Spain, August 2024, pp. 5969–5978, doi: 10.1145/3637528.3671596.

[5] J. Hou et al., "Bidomain uncertainty gated recursive network for pan-sharpening," *Information Fusion*, vol. 118, p. 102938, 2025, doi: 10.1016/j.inffus.2025.102938.

[6] A. Oelen, M. Stocker, and S. Auer, "TinyGenius: Intertwining natural language processing with microtask crowdsourcing for scholarly knowledge graph creation," in *Proceedings of the ACM/IEEE Joint Conference on Digital Libraries (JCDL 2022)*, 2022, doi: 10.1145/3529372.3533285.

[7] H. Kroll, J. Pirklbauer, and W. Balke, "A toolbox for the nearly-unsupervised construction of digital library knowledge graphs," in *Proceedings of the ACM/IEEE Joint Conference on Digital Libraries (JCDL 2021)*, pp. 21–30, 2021, doi: 10.1109/JCDL52503.2021.00014.

[8] J. Devlin, M.-W. Chang, K. Lee, and K. Toutanova, "BERT: Pre-training of deep bidirectional transformers for language understanding," *arXiv preprint arXiv:1810.04805*, 2018, doi: 10.48550/ARXIV.1810.04805.

[9] M. Gohsen and B. Stein, "Assisted knowledge graph authoring: Human-supervised knowledge graph construction from natural language," in *Proceedings of the ACM SIGIR Conference on Human Information Interaction and Retrieval (CHIIR 2024)*, Sheffield, UK, March 2024, pp. 376–380, doi: 10.1145/3627508.3638340.

[10] J. Lee and E. Koh, "Teamwork dimensions classification using BERT," in *Proceedings of the International Conference on Artificial Intelligence in Education (AIED 2023): Posters and Industry Track*, Tokyo, Japan, July 2023, pp. 254–259, doi: 10.1007/978-3-031-36336-8-39.