

Empowering Students with Knowledge Graphs: A New Approach to Fact-Checking and Research

Javno kvalifikacijsko izlaganje

Ana Pasarić

RESEARCH AIM

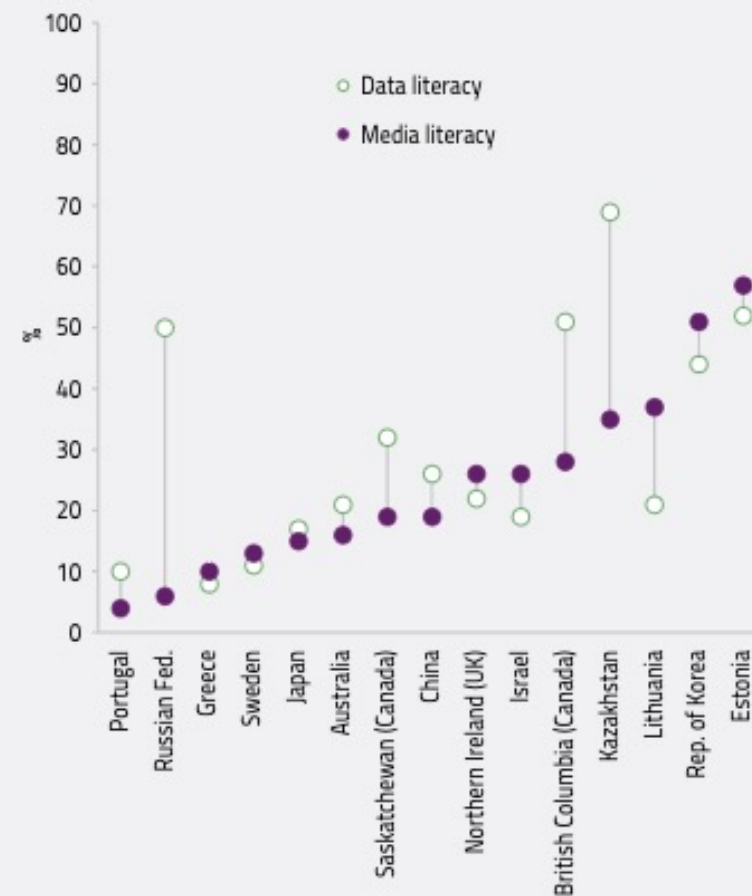


Search engines like Google are becoming commonplace platforms for information retrieval.

FIGURE 5.10:

Media and data literacy are embedded in rich countries' curricula

a. Percentage of curriculum that embeds media and data literacy competencies, selected education systems, 2019



GEM StatLink: https://bit.ly/GEM2023_fig5_10a_

Source:

<https://unesdoc.unesco.org/ark:/48223/pf0000385723.locale=en>

“Fact checking is the task of determining if a given claim holds.”

Huynh, V.P. and Papotti, P., 2019, November. A benchmark for fact checking algorithms built on knowledge bases. In Proceedings of the 28th ACM International Conference on Information and Knowledge Management (pp. 689-698).

“Fact checking is the practice of systematically publishing assessments of the validity of claims made by public officials and institutions with an explicit attempt to identify whether a claim is factual.”

Walter, N., Cohen, J., Holbert, R.L. and Morag, Y., 2020. Fact-checking: A meta-analysis of what works and for whom. *Political communication*, 37(3), pp.350-375.

“The current research landscape in automated fact checking is comprised of systems that estimate the veracity of claims based on available metadata and evidence pages.”

Atanasova, P., 2024. Generating fact checking explanations. In *Accountable and Explainable Methods for Complex Reasoning over Text* (pp. 83-103). Cham: Springer Nature Switzerland.

„...Knowledge graphs (KGs) have also been employed to provide a mechanism in which reliable third-party sources such as scientific articles and Wikipedia are used for fact-checking.“

Nikopentis, G., Mayank, M., Phukan, O.C. and Sharma, R., 2023, November. Reinforcement learning-based knowledge graph reasoning for explainable fact-checking. In Proceedings of the international conference on advances in social networks analysis and mining (pp. 164-170).

RESEARCH METHOD

Research Objective

Systematic literature review [SLR], which aims to investigate the applications of knowledge graphs [KGs] in fact-checking.

Inclusion Criteria

- (1) articles published between 2019 and 2024;
- (2) articles published in English;
- (3) with main focus on fact checking with knowledge graphs;
- (4) articles that are published in IT journals or any technology related journal articles as well as conference proceedings;

Exclusion Criteria

Studies that adhered to the following criteria:

- (1) articles that have not been peer-reviewed
- (2) articles published in other languages
- (3) poster, PowerPoint presentation or books

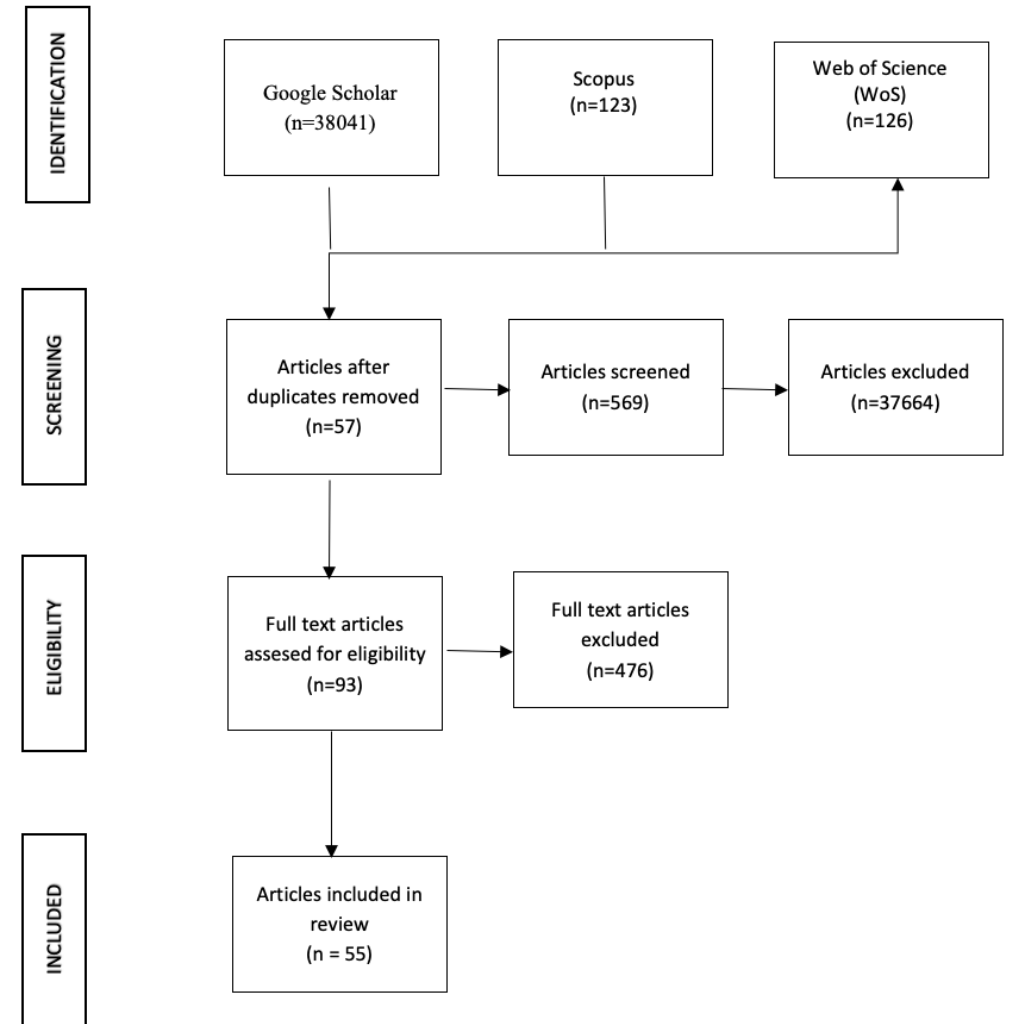


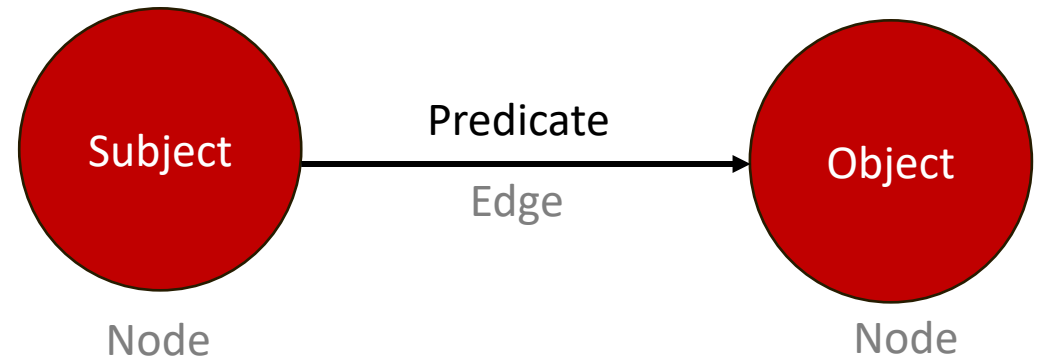
Figure 1: a flowchart of the search process

KNOWLEDGE GRAPH

- a structured representation of information comprised of triples:

<subject, predicate, object>

- consists of nodes and relationships
- nodes represent entities
- relationships describe the context and how it is connected to other nodes

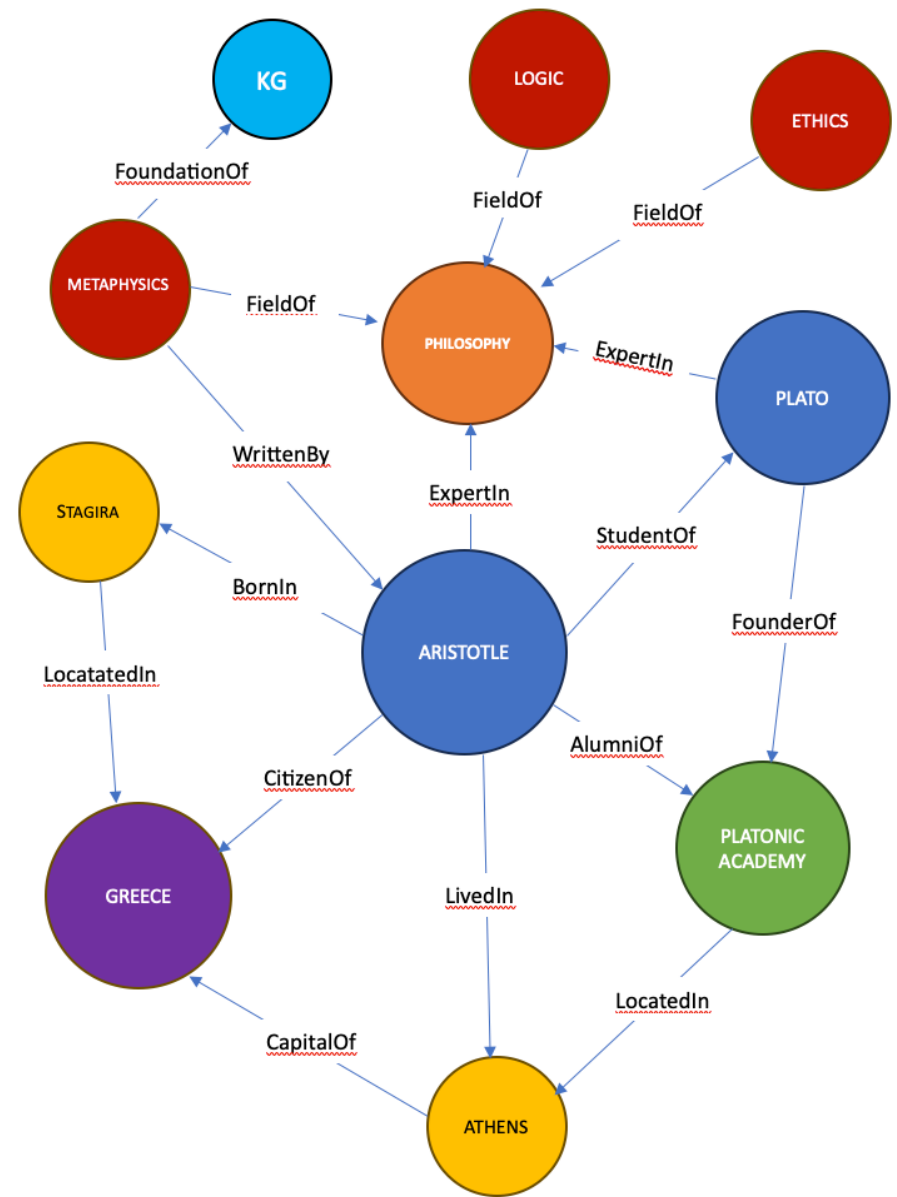


KNOWLEDGE GRAPH



- (Aristotle, AlumniOf, Platonic_Academy)
- (Aristotle, BornIn, Stagira)
- (Aristotle, CitizenOf, Greece)
- (Aristotle, LivedIn, Athens)
- (Aristotle, ExpertIn, Philosophy)
- (Metaphysics, WrittenBy, Aristotle)
- (Metaphysics, FoundationOf, Philosophy)
- (Metaphysics, FoundationOf, KG)
- (Plato, BirthPlace, Classical_Athens)
- (Plato, ExpertIn, Philosophy)
- (Plato, FounderOf, Platonic Academy)
- (Platonic Academy, LocatedIn, Athens)
- (Stagira, Location, Greece)
- (Greece, CapitalOf, Athens)
- (Logic, FieldOf, Philosophy)
- (Ethics, FieldOf, Philosophy)

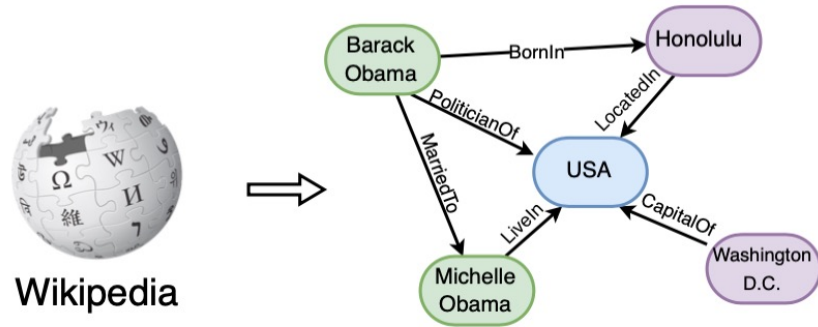
Factual triples in knowledge base



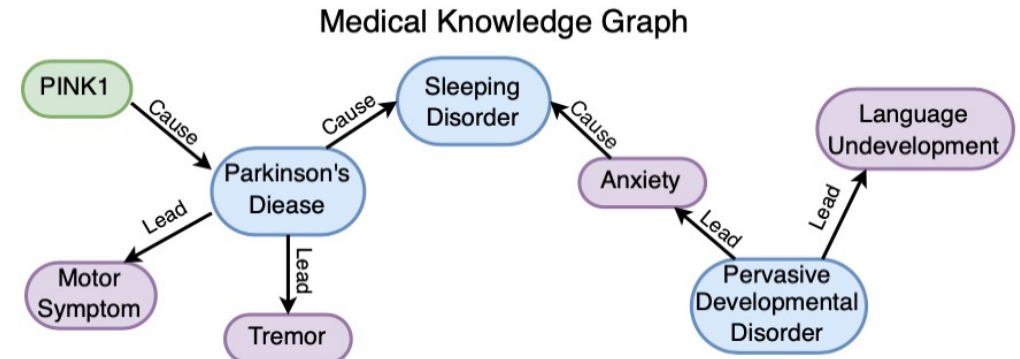
Entities and relations in knowledge graph

KNOWLEDGE GRAPHS CATEGORIES

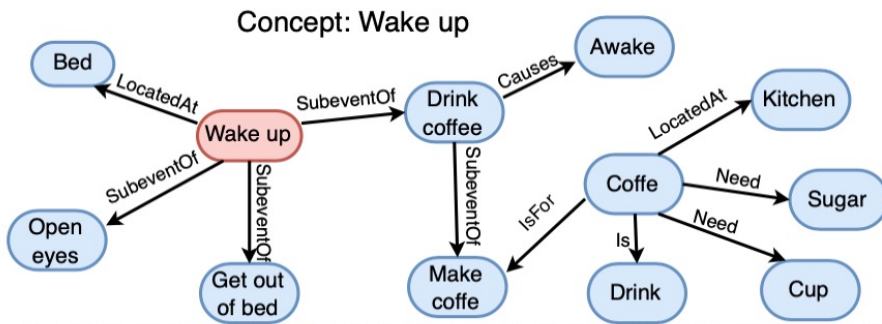
Encyclopedic Knowledge Graphs



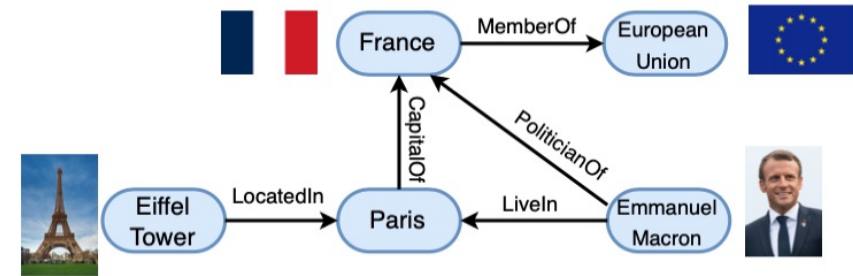
Domain-specific Knowledge Graphs



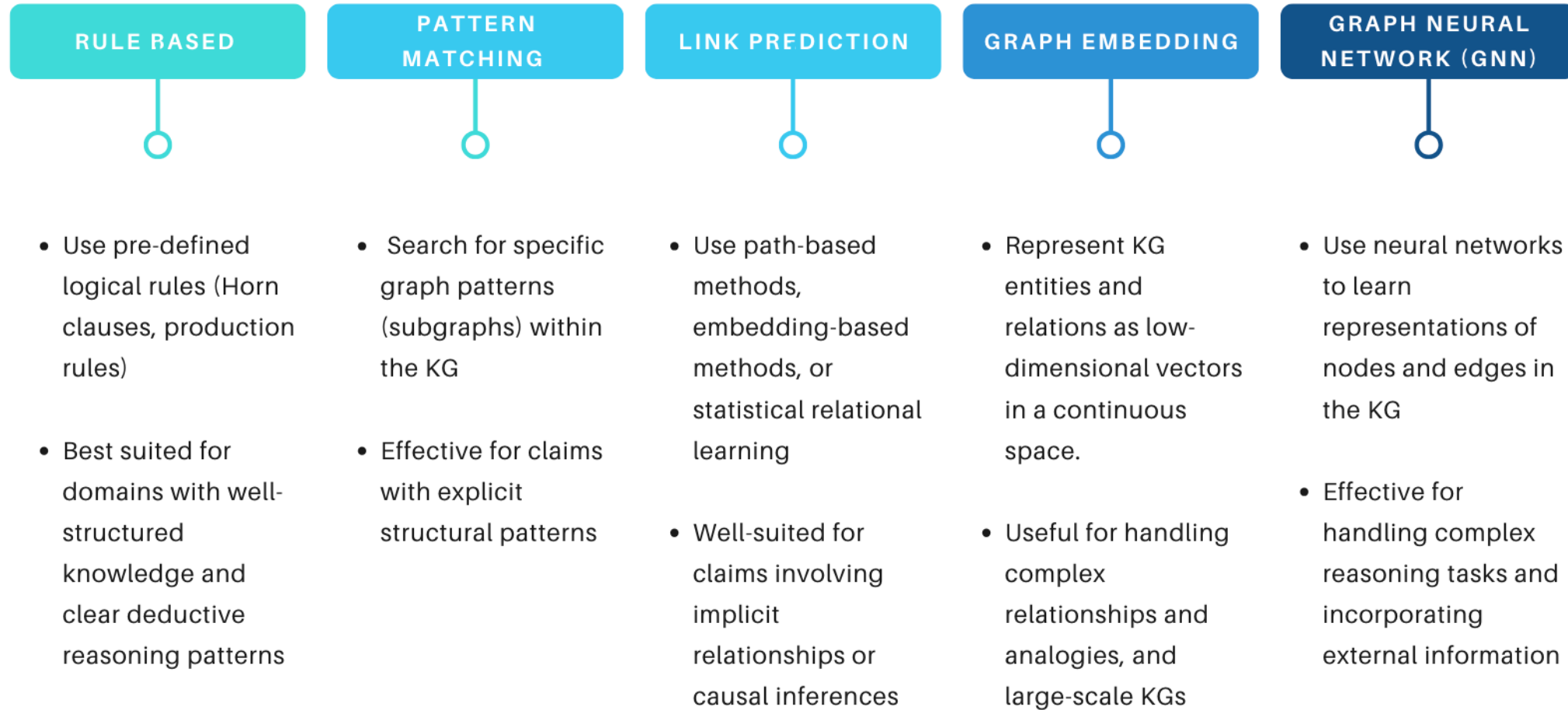
Commonsense Knowledge Graphs



Multi-modal Knowledge Graphs



FACT CHECKING TECHNIQUES WITH KGs



SYSTEMATIC LITERATURE REVIEW - FACT CHECKING WITH KGs



RULE BASED

- Tracy: Tracing Facts over Knowledge Graphs and Text, Gad-Elrab et. alt., (2019)
- ExFaKT: A Framework for Explaining Facts over Knowledge Graphs and Text, Gad-Elrab et. alt., (2019)
- Explainable Fact Checking with Probabilistic Answer Set Programming, Ahmadi et. alt., (2019)
- User Guidance for Efficient Fact Checking, Tam et. alt., (2019)
- Evidence-Aware Multilingual Fake News Detection, Hammouchi et.alt., (2022)
- Science Checker: Extractive-Boolean Question Answering For Scientific Fact Checking, Rakotoson et. alt., (2022)

PATTERN MATCHING

- ClaimsKG: A Knowledge Graph of Fact-Checked Claims, Tchechmedjiev et. alt. (2019)
- Discovering Patterns for Fact Checking in Knowledge Graphs, Lin et. alt., (2019)
- A Hybrid Linguistic and Knowledge-Based Analysis Approach for Fake News Detection on SocialMedia, Seddari et.alt., (2022)
- Text2KGBench, a benchmark for ontology-driven knowledge graph generation from text, Mihindukulasooriya et. alt., (2023)

LINK PREDICTION

- A Benchmark for Fact Checking Algorithms Built on Knowledge Bases, Huynh et. alt., (2019)
- ClaimLinker: Linking Text to a Knowledge Graph of Fact-checked Claims, Maliaroudakis et. alt., (2021)
- DialoKG: Knowledge-Structure Aware Task-Oriented Dialogue Generation, Rony et. alt., (2022)
- Using Multiple RDF Knowledge Graphs for Enriching ChatGPT Responses, Mountantonakis and Tzitzikas, (2023)
- Reinforcement Learning-based Knowledge Graph Reasoning for Explainable Fact-checking, Nikopensus et.alt., (2023)
- FakeKG: a knowledge graph of fake claims for improving automated fact-checking, Shahi (2023)

GRAPH EMBEDDING

- GEAR: Graph-based Evidence Aggregating and Reasoning for Fact Verification, Zhou et. alt., (2019)
- FACE-KEG: FACT Checking Explained using Knowledge Graphs, Vedula & Parthasarathy, (2021)

GRAPH NEURAL NETWORK (GNN)

- TISCO: Temporal Scoping of Facts, Rulaa et.alt., (2019)
- Barack's Wife Hillary: Using Knowledge Graphs for Fact-Aware Language Modeling, Logan IV et. alt., (2019)
- Can Knowledge Graph Embeddings Tell Us What Fact-checked Claims Are About?, Beretta et. alt., (2020)
- Modularized transfer learning with multiple knowledge graphs for zero-shot commonsense reasoning, Kim et. alt., (2022)
- Multi-Hop Open-Domain Question Answering over Structured and Unstructured Knowledge, Feng et. alt. (2022)
- Joint Knowledge Graph Completion and Question Answering, Liu et. alt. (2022)
- Neural methods for logical reasoning over knowledge graphs, Amayuelas et. alt. (2022)
- FactKB: Generalizable Factuality Evaluation using Language Models Enhanced with Factual Knowledge, Feng et. alt. (2023)

SYSTEMATIC LITERATURE REVIEW - FACT CHECKING WITH KGs



RULE BASED				
AUTHOR'S AND YEAR	Technique	Deals with	Dataset	Result
1 Tracy: Tracing Facts over Knowledge Graphs and Text Gad-Elrab et. al. (2019)	Rule-based rewriting and semantic trace generation using knowledge graphs	generating explanations	YAGO Wikidata	This research introduces Tracy, a tool that employs background knowledge rules to generate human-interpretable explanations for candidate facts. These explanations, or semantic traces, support knowledge graph curation.
2 EfaKT: A Framework for Explaining Facts over Knowledge Graphs and Text Gad-Elrab et. al. (2019)	Rule-based reasoning	generating explanations	YAGO DBpedia Wiki (5.5 million Wikipedia articles split into sentences and indeed using Elasticsearch) Web (web pages retrieved by Bing API)	This research introduces EfaKT, a framework for generating human-interpretable explanations for the veracity of candidate facts. Background knowledge is encoded in Horn clauses to decompose complex claims into simpler, verifiable components. By drawing evidence from both knowledge graphs and textual sources, EfaKT demonstrably improves fact-checking recall while maintaining high precision. The research further highlights the effectiveness of EfaKT's explanations in supporting human fact-checkers and its potential to enhance the performance of automated fact-checking systems.
3 Explainable Fact Checking with Probabilistic Answer Set Programming Ahmadi et. al. (2019)	Explainable claim verification using knowledge graph	claims	DBpedia	The research introduces a fact-checking system that uses knowledge graphs (KGs) for explainable claim verification. By employing probabilistic logic programs to analyze claims against KG information and discovered rules, the system overcomes limitations of KGs through web text mining, ultimately achieving higher accuracy and explainability.
4 User Guidance for Efficient Fact Verification Tam et. al. (2019)	User-guided streaming fact checking framework with probabilistic prioritization	claims	Wikipedia Healthcare forum Snopes	This research introduces a probabilistic framework for user-guided fact-checking in a streaming environment. The system emphasizes user input alongside automated credibility inference to prioritize claims for validation, strategically minimizing user effort. Evaluations demonstrate significant reductions in validation workload while achieving high-precision knowledge base construction compared to baseline methods while maintaining accuracy above 90%.
5 Evidence-Aware Multilingual Fake News Detection Hammouchi et. al. (2022)	a multilingual framework for evaluating the veracity of online news based on factual evidence and source credibility	Covid-19 related news	NFact Constraint	The research introduces a multilingual framework for detecting fake news. It leverages external evidence from Google searches, assigns credibility scores to sources based on website reputation metrics, and analyzes claim-evidence relationships. The framework achieves superior performance (F1 scores of 0.85 and 0.97) compared to existing methods in the context of COVID-19 news verification.
6 Science Checker: Extractive-Boolean Question Answering For Scientific Fact Checking Rakotonson et. al. (2022)	a multi-task approach for verifying the scientific questions based on a joint reasoning from facts and evidence in research articles	question answering	PubMedQA BioRxiv	The research proposes a multi-task learning framework that mimics a limited knowledge base (KB) by extracting key evidence from research articles through summarization and extractive question answering (EQA). While not a full-fledged KB system or employing complex knowledge graph (KG) reasoning, the framework performs joint reasoning over the extracted facts to answer the question. This highlights the feasibility of utilizing text-based scientific knowledge sources to support fact checking tasks.

PATTERN MATCHING				
Author and Year	Technique	Deals with	Dataset	Result
7 Discovering Patterns for Fact Checking in Knowledge Graphs Lin et. al. (2019)	Graph Fact Checking rule approach	text analysis and crowd sourcing	YAGO (version 2.5) DBpedia (version 3.8) Wikidata MAG Offshore GDELT	The research proposes a new framework for fact checking in knowledge graphs that utilizes GFCs (Graph Fact Checking rules). These rules incorporate graph patterns that can effectively distinguish between true and false facts based on generalized fact statements.
8 ClaimsKG: A Knowledge Graph of Fact-Checked Claims Tshechmediev et. al. (2019)	ClaimsKG approach utilizes a knowledge graph to store and query fact-checked claims (veracity, source, and date)	fact-checked claims	aficheck.org factscan.ca politiFact.com snopes.com checkyourfact.com truthification.com DBpedia	This research introduces ClaimsKG, a knowledge graph, for fact-checking, stance detection and societal debate analysis. The model facilitates structured queries of fact-checked claims, including truth labels, source attribution, and publication dates. ClaimsKG is constructed through a semi-automated pipeline that harvests data from reputable fact-checking websites, annotates claims with relevant entities, and integrates the information using standardized formats.
9 A Hybrid Linguistic and Knowledge-Based Analysis Approach for Fake News Detection on SocialMedia Seddari et. al. (2022)	Hybrid fake news detection system that combines linguistic and knowledge-based approaches	linguistic and fact-verification features	Buzzfeed Political News	This research introduces a hybrid fake news detection system that integrates linguistic features (title length, readability, lexical richness, and sentiment analysis) with a novel knowledge-based approach named fact-verification features. These features leverage external knowledge sources for veracity assessment, including website reputation metrics, coverage analysis (number of covering sources), and fact-check verdicts (true/false) obtained from established platforms.
10 Text2KBench, a benchmark for ontology-driven knowledge graph generation from text Mihindukulasooriya et. al. (2023)	Text2KBench, a benchmark framework for knowledge graphs generation from text guided by ontologies	knowledge graph generation from text	Wikidata-TeKGen DBpedia-WebNLG	This research introduces Text2KBench, a benchmark for evaluating Large Language Models (LLMs) ability to construct knowledge graphs from text guided by ontologies. It extends relation extraction by incorporating ontology constraints and instructions.

LINK PREDICTION					
Author and Year	Technique	Deals with	Dataset	Result	
11 Huynh et. al. (2019)	A benchmark for Fact Checking Algorithms Built on Knowledge Bases	a benchmark framework for evaluating fact-checking algorithms, allowing comparisons with knowledge base link prediction methods	textual claims against trustful Knowledge Bases (KBs)	Knowledge bases under the Open World Assumption (OWA)	The research introduces a benchmark framework by evaluating diverse algorithms on controlled data with varying properties from knowledge bases, the framework reveals that data complexity has a greater impact on performance than algorithm choice. While all methods performed well in simple scenarios, qualitative differences emerged with more complex datasets. AUROC scores across these experiments ranged from 0.5 to 0.95 for a fixed algorithm, knowledge base predicate, and data split.
12 ClaimLinker: Linking Text to a Knowledge Graph of Fact-Checked Claims Malliaroudakis et. al. (2021)	A knowledge-graph based approach to unsupervised textual factual verification	text analysis	bookmarket Java library Web service	bookmarket	The research introduces ClaimLinker - a web service and API for linking text snippets to a knowledge graph of fact-checked claims. This scalable and unsupervised approach empowers direct exploration of information veracity, fostering seamless integration of fact-checking functionalities within online domains like journalism and social media.
13 DialogKG: Knowledge-Structure Aware Task-Oriented Dialogue Generation Romy et. al. (2022)	Task-oriented dialogue system using knowledge graph	knowledge bases	Stanford Multi-Domain Dataset (SMD) CamRest	Stanford Multi-Domain Dataset (SMD) CamRest	This research introduces DialogKG, a task-oriented dialogue system that leverages a knowledge graph with structure-aware embedding and knowledge-weighted attention to select relevant information and generate human-like, factually accurate responses.
14 Using Multiple RDF Knowledge Graphs for Enriching ChatGPT Responses Mountantonakis and Tzitzikas (2023)	GPT_LoDS, a prototype for enriching and fact-checking ChatGPT responses using multiple RDF knowledge graphs	ChatGPT responses	LODsymdehis KG (contains data from over 400 individual RDF Knowledge Graphs with more than 412 million entities)	LODsymdehis KG (contains data from over 400 individual RDF Knowledge Graphs with more than 412 million entities)	This research introduces GPT_LoDS, a prototype that tackles factual inconsistencies in ChatGPT. It leverages LODsymdehis KG, a massive knowledge graph, to fact-check responses by linking entities to real-world data from hundreds of RDF Knowledge Graphs.
15 Reinforcement Learning-Based Reasoning for Explainable Fact-Checking Nikopenisus et. al. (2023)	Reinforcement learning-based knowledge graph approach using multi-hop reasoning agent for explainable fact-checking	Claim from policy-based agent	F815K-277 NELL-995	F815K-277 NELL-995	The research combines reinforcement learning with a knowledge graph (KG) to achieve explainable fact classification. They approach fact-checking as a multi-step reasoning problem. An agent trained with reinforcement learning explores the KG to find paths that support or refute a claim. These paths are then used in a voting mechanism to determine the truthfulness of the statement, with the final verdict remaining transparent and open to user evaluation.
16 FakeKG: a knowledge graph of fake claims for improving automated fact-checking Shahi (2023)	Semantic-technology-based scalable AFC solution	Claims from social media	The AMUSE framework (Shahi and Majchrzak 2022; Shahi and Nandini 2020) was used to collect and store over 113,000 fact-checked data points in a MySQL database, encompassing 40 languages from 105 countries	The AMUSE framework (Shahi and Majchrzak 2022; Shahi and Nandini 2020) was used to collect and store over 113,000 fact-checked data points in a MySQL database, encompassing 40 languages from 105 countries	The research proposes a scalable Knowledge Graph-Based approach leveraging semantic technology to address limitations in current fact-checking by incorporating social media claims and prioritizing expert workload based on existing knowledge.

GRAPH EMBEDDING					
Author and Year	Technique	Deals with	Dataset	Result	
17 GEAR: Graph-based Evidence Aggregating and Reasoning for Fact Verification Zhou et. al. (2019)	Graph-based Evidence Aggregating and Reasoning for Fact Verification	Explainable fact checking	Claims	FEVER DBpedia FEVER MultiFC COVID-19	This research proposes a GEAR framework that uses a graph-based approach and BERT to aggregate evidence for fact verification effectively. The key challenge is integrating and reasoning over multiple evidence pieces for fact-checking.
18 Vedula & Parthasarathy (2021)	FACE-KEG: Fact Checking Explained using Knowledge Graphs	Explainable fact checking	Claims	DBpedia FEVER MultiFC COVID-19	This research introduces FACE-KEG, explainable fact-checking through knowledge graphs. It constructs a claim-specific graph from a vast knowledge base, using a novel graph transforming encoder to analyze its structure. Simultaneously, FACE-KEG retrieves and encodes relevant textual context. Combined analysis of structured knowledge and surrounding text allows FACE-KEG to determine claim veracity and generate human-understandable explanations.
GRAPH NEURAL NETWORK (GNN)					
19 TISCO: Temporal Scoping of Facts Rulaa et. al. (2019)	Time-Aware framework	Knowledge bases	Web of Data	Web of Data	The research introduces Temporal Information Scoping (TISCO), framework that leverages a three-step process of matching, selection, and merging evidence to determine the valid time interval for existing facts within knowledge bases. This approach achieves an F-measure of up to 80% when applied to factual data from DBpedia.
20 Barack's Wife Hillary: Using Knowledge Graphs for Fact-Aware Language Modeling Logan IV et. al. (2019)	Neural fact checking	Generating entities and facts from a knowledge graph	Linked Wikitext-2	Linked Wikitext-2	This research introduces the Knowledge Graph Language Model (KGLM), a neural language model that uses an external knowledge graph to enhance text generation. Trained on the newly created Linked Wikitext-2 dataset, KGLM incorporates factual knowledge (entities, details, rare entities).
21 Can Knowledge Graph Embeddings Tell Us What Fact-Checked Claims Are About? Beretta et. al. (2020)	Neural-graph embeddings features for claim topic prediction	claims	DBpedia	DBpedia	This research investigated whether knowledge graph (KG) embeddings, capturing claim-topic relationships, could improve topic classification compared to text embeddings. Combining both approaches made a minor improvement, the use of KG embeddings alone resulted in significantly lower accuracy. Authors suggest that the KG model, optimized for local link prediction within the graph structure, might not capture the more complex relationships crucial for topic classification tasks.
22 Modularized transfer learning with multiple knowledge graphs for zero-shot commonsense reasoning Kim et. al. (2022)	Unsupervised zero-shot multiple-choice QA system using knowledge graphs	question answering	SocialQA (SQA) CommonsenseQA (CSQA) Abductive NLI (a-NLI) PhysicalQA (PIQA) Grande (WG)	SocialQA (SQA) CommonsenseQA (CSQA) Abductive NLI (a-NLI) PhysicalQA (PIQA) Grande (WG)	This research introduces a modularized transfer learning framework for zero-shot commonsense reasoning fusing the knowledge from multiple knowledge graphs.
23 Multi-Hop Open-Domain Question Answering over Structured and Unstructured Knowledge Feng et. al. (2022)	DEHG, Document-Entity Heterogeneous Graph Network	question answering	HybridQA	HybridQA	This study proposes DEHG, a Document-Entity Heterogeneous Graph Network framework, that tackles open-domain question answering by integrating structured and unstructured information through a graph constructor. It utilizes context encoding, multi-hop reasoning on the heterogeneous network, and answer decoding to achieve superior performance on the HybridQA dataset.
24 Joint Knowledge Graph Completion and Question Answering Liu et. al. (2022)	BiNet, a neural network, tackles knowledge graphs by combining knowledge completion and multi-hop question answering	question answering	MetaQA WebQuestionsSP SimpleQuestions	MetaQA WebQuestionsSP SimpleQuestions	This research introduces BiNet, a neural network model that achieves state-of-the-art performance in both knowledge graph completion (KG-C) and multi-hop question answering (KGQA) tasks. BiNet accomplishes this by employing a shared embedding space and answer scoring module, enabling the tasks to learn from each other through the exchange of latent features.
25 Neural methods for logical reasoning over knowledge graphs Amayuelas et. al. (2022)	Neural networks for logical reasoning over knowledge graphs	Answering First-Order Logical (FOL) queries	Query2Box BetaE FB15k-237 NELL995	Query2Box BetaE FB15k-237 NELL995	This research proposes a neural embedding framework for FOL reasoning over knowledge graphs. It utilizes flexible logical operators (conjunction, disjunction, negation) built with Multi-Layer Perceptrons (MLPs) or MLP-Mixer blocks. This enables processing complex multi-hop FOL queries, including negation, overcoming limitations of prior methods.
26 FactKB: Generalizable Factualty Evaluation using Language Models Enhanced with Factual Knowledge Feng et. al. (2023)	FACTKB: A Generalizable Factualty Evaluation System Based on Language Models with Augmented Factual Knowledge	knowledge bases	YAGO FactCollect dataset CovidFact HealthVer SciFact Wikidata ConceptNet Atomic KGAP UMLS	YAGO FactCollect dataset CovidFact HealthVer SciFact Wikidata ConceptNet Atomic KGAP UMLS	This research introduces a FACTKB, a framework for factualty evaluation, utilizes language models pre-trained on factual knowledge extracted from external knowledge bases. By employing three complementary pre-training objectives focused on entity-specific facts, auxiliary knowledge, and compositional knowledge walks, FACTKB achieves state-of-the-art performance on both in-domain news summarization benchmarks and out-of-domain scientific literature datasets.

FACT CHECKING WITH KGs



SCALABILITY

KGs can efficiently handle large-scale datasets



EXPLAINABILITY

KG have the ability to provide clear and understandable justifications for its decisions or predictions



TRUSTWORTHINES

Degree to which users can rely on the information and inferences derived from the KG



INTREPRETABILITY

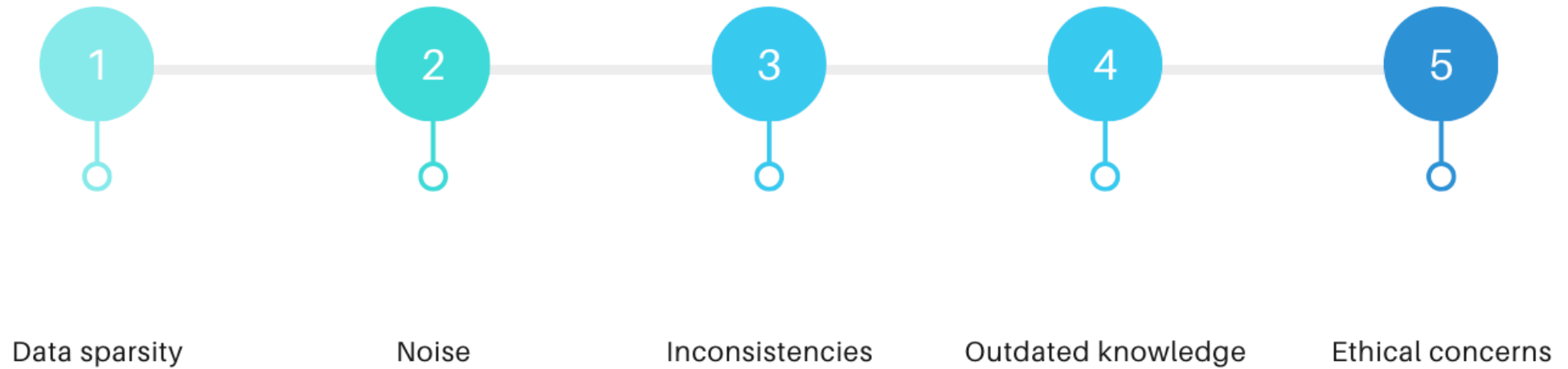
Ability of KGs to make the decision-making process transparent and understandable to humans

DATASET

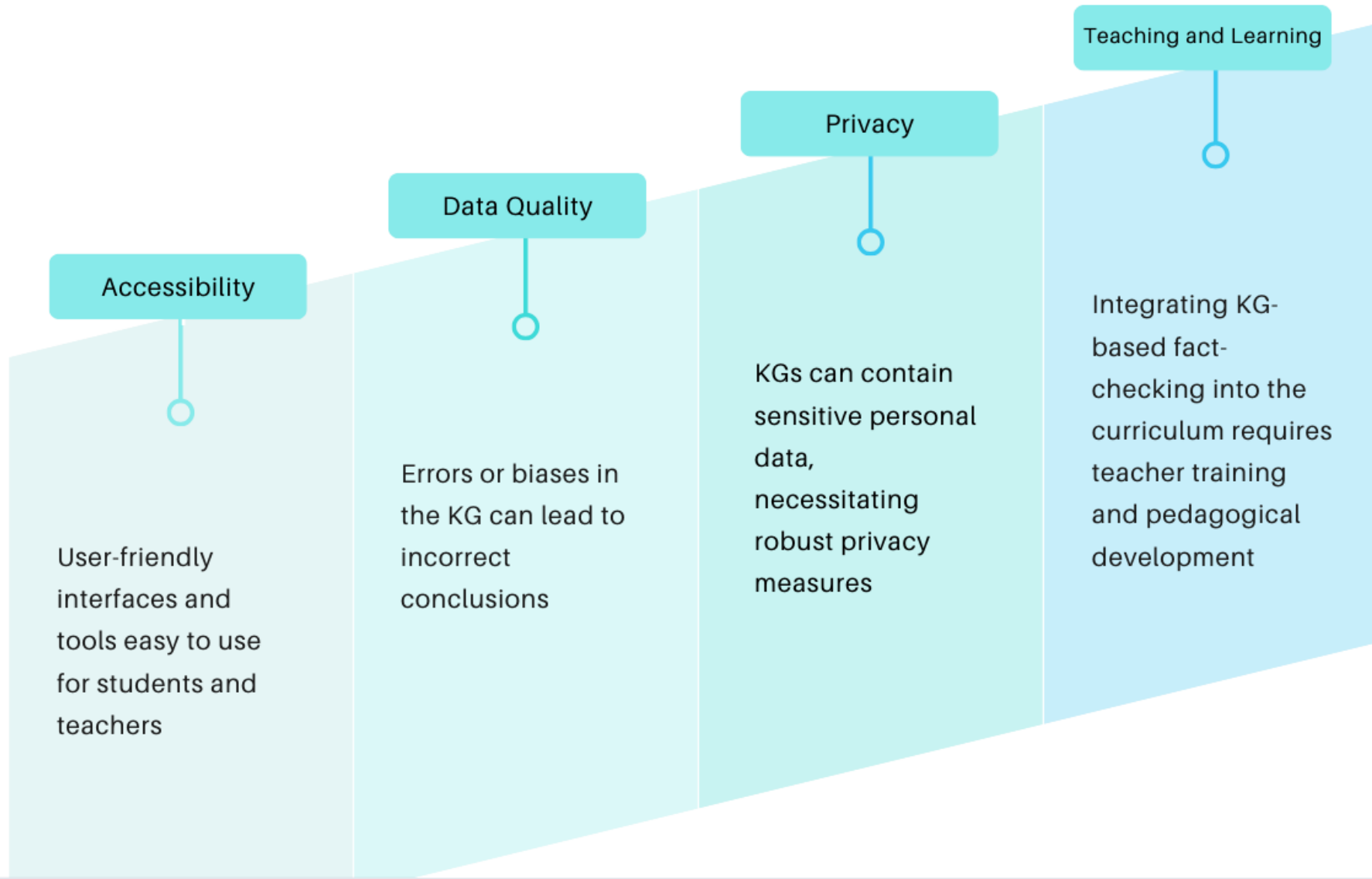
DATASET	ENTITIES	VERDICT	SOURCES	LANGUAGE
FEVER	185,445	3 Classes	Wikipedia	En
DBpedia	6.0 M	3 Classes	Wikipedia	En
YAGO	17 M	2 Classes	Wikipedia, WordNet, WikiData, GeoNames	Many
Wikidata	110 M	3 Classes	Wikipedia	Many
Snopes	6,422	3 Classes	Fact Check	En
NELL-314	13,965	3 Classes	ClueWeb09	En
SciFact	1,409	3 Classes	Science	En
CovidFact	4,086	2 Classes	Forum	En
KnowLife	656,607	3 Classes	Scientific literature	Many
SemMedDB	154,216	3 Classes	Biomedical literature	En
XFact	31,189	7 Classes	Fact Check	Many
MultiFC	36,534	2-27 Classes	Fact Check	En



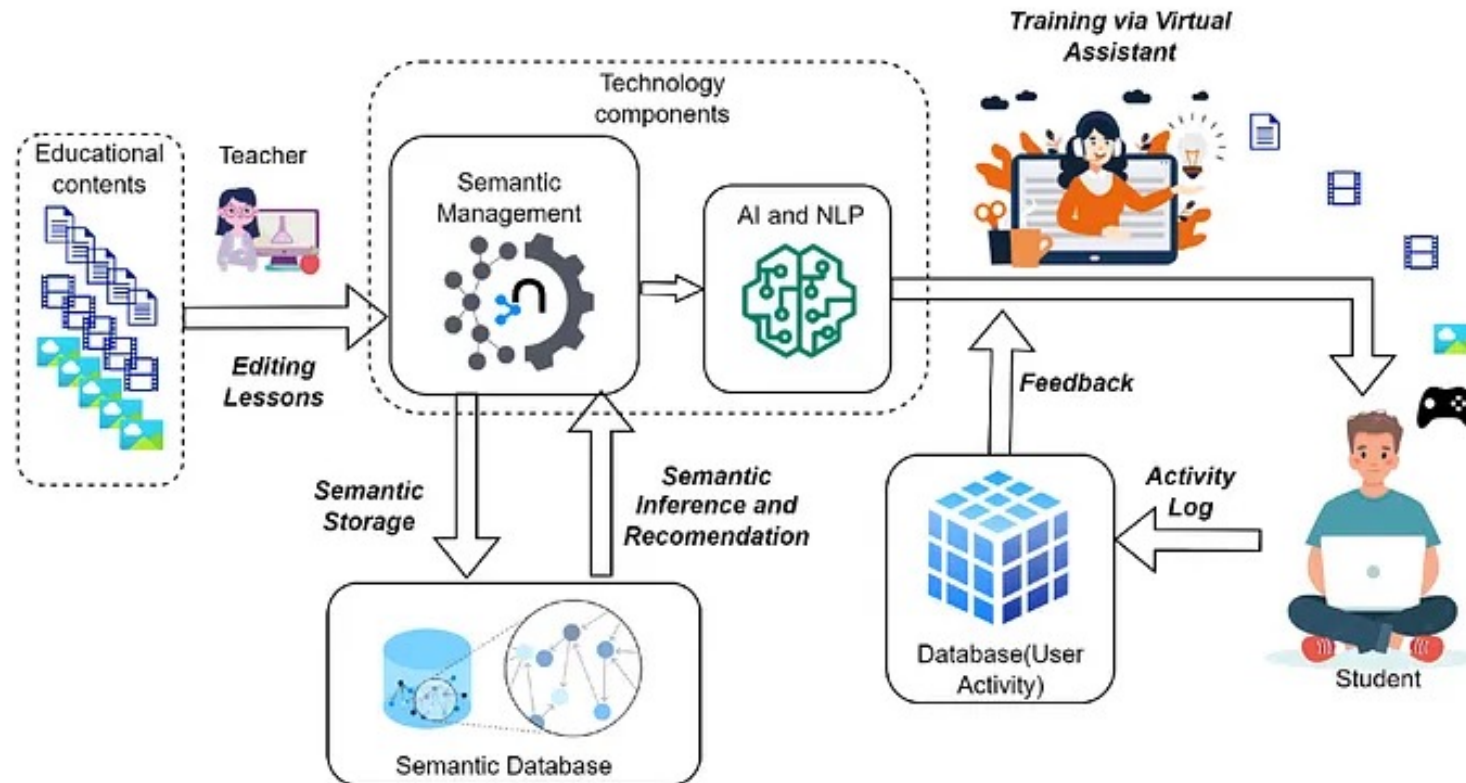
CHALLENGES AND LIMITATIONS




Challenges of integrating KGs into educational settings



POTENTIAL EDUCATIONAL USE CASES – FUTURE RESEARCH



- **Developing Educational KGs:** Creating KGs specifically designed for educational purposes
- **User-Friendly Interfaces:** Developing intuitive interfaces that allow students to explore and query KGs without technical expertise
- **Critical Thinking Skills:** Incorporating activities that teach students how to evaluate the quality of KG data and information
- **Supporting Research Projects:** Provide students with tools to explore research topics and verify information using KGs
- **Collaboration:** Collaborating with KG developers, educators, and librarians to create effective KG-based learning experiences



FIDIT

Sveučilište u Rijeci

**Fakultet informatike
i digitalnih tehnologija**

UNIRI



Pitanja?