ENHANCING FACT-CHECKING WITH RAG AND KNOWLEDGE GRAPHS

Ashneet Khandpur Singh :: Pau Perea Paños :: Mario Reyes de los Mozos :: Gary Munnelly

ORIGINAL RESEARCH ARTICLE / DOI: 10.20901/ms.17.33.8 / SUBMITTED: 30.09.2025.

ABSTRACT *Traditional fact-checking methods, while effective, are often too slow to keep up with the speed at which false information circulates online. In recent years, artificial intelligence (AI) has gained popularity as a means of automating the fact-checking process, particularly large language models (LLMs). Although LLMs have demonstrated efficacy in assisting with verification tasks, they are constrained by factors such as the quality of training data and their ability to retrieve pertinent information for verification. They are also prone to 'hallucinations', generating plausible but false or misleading information, raising critical concerns about their reliability as independent fact-checkers. This paper explores a hybrid approach that combines retrieval-augmented generation (RAG) with knowledge graphs (KGs) and OpenCTI, a platform for cyber threat intelligence. This approach ensures that the final output is not only informative but also transparent, linking claims to authoritative sources. This makes it a valuable tool for journalists, fact-checkers, or the general public, who increasingly require timely and reliable answers in an information environment shaped by disinformation.*

KEYWORDS

DISINFORMATION, KNOWLEDGE GRAPHS, RETRIEVAL-AUGMENTATED GENERATION, AI, LLMS, THREAT INTELLIGENCE, OPENCTI

Author's note

Ashneet Khandpur Singh :: Eurecat, Technology Centre of Catalonia, Barcelona **::** ashneet.khandpur@eurecat.org
Pau Perea Paños :: Eurecat, Technology Centre of Catalonia, Barcelona **::** pau.perea@eurecat.org
Mario Reyes de los Mozos :: Eurecat, Technology Centre of Catalonia, Barcelona **::** mario.reyes@eurecat.org
Gary Munnelly :: Adapt Centre, Trinity College Dublin, Dublin **::** gary.munnelly@adaptcentre.ie

Note: This work has received funding from the European Union's Horizon Europe Research and Innovation Programme under Grant Agreement No. 101132686 (ATHENA project).

ORIGINAL RESEARCH ARTICLE / DOI: 10.20901/ms.17.33.8 / SUBMITTED: 30.09.2025.

INTRODUCTION

Disinformation and fake news spread rapidly in today's digital world, influencing public opinion, shaping political discourse, and even affecting health and financial decisions. The accessibility of information has never been greater, but neither has its potential for harm. Disinformation campaigns – ranging from climate denial to manipulated electoral narratives – exploit cognitive biases like confirmation bias and rise on algorithmic amplification across social platforms. These campaigns are no longer isolated incidents; rather, they are part of a broader phenomenon called information disorder (Gaeta et al., 2023; Gao, 2025), where false or misleading content is produced, shared, and believed at a scale never seen before.

While false information is not a new occurrence, the past events such as Brexit and the U.S. presidential election (Barbera, 2018; Lee, 2019; Rose, 2017) have increased its impact on society and its dissemination, leading some to label the current era as the post-truth era (Broda & Strömbäck, 2024; Rose, 2017). Within this landscape, it is crucial to distinguish between misinformation, which is false information disseminated with no malicious intent, and disinformation, which is purposefully produced to deceive and cause harm (Wardle, 2018). Both lead to persistent misconceptions that can skew public opinion, sabotage democratic discourse, and even encourage knowledge resistance, where individuals reject countervailing evidence (Broda & Strömbäck, 2024).

Social media platforms have transformed these dynamics by enabling misinformation and disinformation to spread faster and on a larger scale. Algorithmic recommendation systems often prioritize emotionally charged, sensational, or controversial content, which is something that false information frequently offers (Aïmeur et al., 2023). Cross-platform sharing also lets manipulated images, videos, and conspiracy theories spread rapidly from fringe communities to mainstream conversations (Olan et al., 2024).

Computational detection methods have progressed substantially, especially in recognizing linguistic and stylistic patterns of false information. However, the review of Broda & Strömbäck (2024) highlights a key gap: the majority of models are tested in controlled environments on benchmark datasets, with limited validation in real-world, multilingual, and multimodal settings. Furthermore, the strategic use of disinformation by political elites, state entities, and organized networks is still under-researched, despite evidence that these actors play a pivotal role in constructing and propagating false narratives.

Addressing these gaps requires tools that not only detect but also contextualize and explain false claims within a broader threat landscape. While prior research has explored the use of RAG and KGs for fact-checking, these approaches are often limited to standalone verification tasks and lack operational integration with cyber threat intelligence workflows.

ORIGINAL RESEARCH ARTICLE / DOI: 10.20901/ms.17.33.8 / SUBMITTED: 30.09.2025.

This study introduces a novel integration of RAG and KGs within a CTI-oriented framework, enabling fact-checking outputs to be explicitly linked to verified entities, relationships, and provenance data. By embedding this approach into platforms such as OpenCTI, our method extends beyond claim verification to support cross-domain analysis, connecting disinformation narratives to actors, campaigns, and broader influence operations. This bridges the gap between automated detection and actionable intelligence, providing transparent and explainable outputs for analysts and fact-checkers.

The main contributions of this work are:
(i) An architecture that integrates RAG-based fact-checking with knowledge graph-driven CTI systems;
(ii) A mechanism for preserving explainability and provenance across retrieval and generation stages,
(iii) A practical deployment within OpenCTI for analyzing FIMI-related disinformation.

The rest of the paper is organized as follows. Section 2 introduces fundamental background concepts of interest throughout the whole work, including large language models, knowledge graphs, retrieval-augmented generation, and cyber threat intelligence. Section 3 provides a comprehensive review of the related work on knowledge graphs, retrieval-augmented generation and threat intelligence, related to fact-checking. It situates our work within the broader context of these research areas and highlights the gaps that our system aims to address. Section 4 describes the proposed methodology and system architecture. Section 5 presents results and discussion. Finally, conclusions and future research are discussed in Section 6.

Theoretical background and key concepts

The term FIMI was coined by the European External Action Service (EEAS)[1] to describe a strategic and coordinated threat that goes beyond traditional disinformation. FIMI represents a multidimensional threat that intersects national security, information integrity, democratic governance, and hybrid warfare[2].

Unlike classical propaganda, FIMI emphasizes the manipulative intent and coordinated behavior of actors, often state or state-linked entities, which can employ tactics like content fabrication, artificial amplification, impersonation, and cross-platform campaigns to erode public trust and destabilize democratic processes.

To operationalize its response to information manipulation, the EEAS has developed the EU FIMI Toolbox[1], which structures countermeasures along four pillars: situational awareness, resilience building, disruption and regulation, and external action. These include mechanisms such as the Rapid Alert System (RAS) for cross-border coordination, support for independent media and civil society to strengthen resilience, regulatory

[1] For further information on the European Union's approach to Information Integrity and Countering Foreign Information Manipulation and Interference (FIMI), see the European External Action Service (EEAS): https://www.eeas.europa.eu/eeas/information-integrity-and-countering-foreign-information-manipulation-interference-fimi_en
[2] For an overview of the concept of Foreign Information Manipulation and Interference (FIMI), see: https://www.disinformation.ch/EU_Foreign_Information_Manipulation_and_Interference_(FIMI).html

ORIGINAL RESEARCH ARTICLE / DOI: 10.20901/ms.17.33.8 / SUBMITTED: 30.09.2025.

measures like the Digital Services Act, and international partnerships with NATO and G7 countries.

Unless otherwise specified, this paper uses the term disinformation to refer to intentional influence activities.

The process of detecting disinformation campaigns, analyzing them and identifying the tactics, techniques and procedures (TTPs) that are applied is an increasingly complex process, but one that cannot be ignored, as these TTP models are of great value for making appropriate and necessary strategic decisions in a coordinated and collaborative manner through the use of CTI (Cyber Threat Intelligence) platforms. The identification of a TTP model for DISARM[3] (Disinformation Analysis and Risk Management), in the context of automating the collection of CTI campaign information, is an emerging area of research and development. The trend in scientific and technical research on TTP modelling for CTI collection automation focuses on structured CTI exchange formats, automation of TTP extraction, and integration of AI/ML for pattern recognition and contextual understanding of threats. The emphasis of existing research is on structuring CTI using common data formats, such as STIX (Structured Threat Information eXpression) and others, for interoperability, which forms the basis for automating the collection of TTPs from threat campaigns. However, uniform standards for the direct use of DISARM are less mature and still under development, focusing on interoperability and automated processing of threat intelligence. The scientific community is exploring the use of artificial intelligence and machine learning to generate, classify, and cluster TTPs from large CTI datasets, with the goal of automating the detection and enrichment of threat campaigns. Approaches include latent diffusion models and classifiers to generate model elements like TTPs, but applied mainly in associated domains (e.g., cyber threat hunting, anomaly detection). There is also considerable interest in research to integrate human input with automated systems for CTI analysis, improving the relevance and accuracy of TTP models.

Disinformation detection covers not just articles and posts but also multimedia (deepfakes, synthetic audio/video), using computer vision and voice analysis. AI-based tools for detecting fake news and disinformation campaigns have evolved considerably, with the latest solutions emphasizing advanced natural language processing (NLP) (Puczyńska & Djenouri, 2024), deep learning architectures and joint machine learning techniques. Real-time detection now leverages transformer-based models (such as BERT and GPT), sentiment and intent analysis, cross-platform data aggregation, and sophisticated uncertainty quantification (UQ) to deliver more robust and transparent results.

A knowledge graph is a structured representation of real-world entities and the relationships between them. These graphs are typically stored in a graph database, which is designed to natively capture and manage interconnected data. Within a knowledge graph, entities can represent objects, events, situations, or abstract concepts, while the relationships between these entities capture the context and meaning of how they are

[3] DISARM, the Disinformation Analysis and Risk Management framework: https://www.disarm.foundation/

ORIGINAL RESEARCH ARTICLE / DOI: 10.20901/ms.17.33.8 / SUBMITTED: 30.09.2025.

connected. A knowledge graph can be viewed as a map, where the points on the map represent things, such as people, places, or events, and the lines between them show how they are related.

Beyond this basic structure, knowledge graphs build on decades of research in the semantic web community, where standards such as RDF (Resource Description Framework) and SPARQL (SPARQL Protocol and RDF Query Language) were designed to make information machine-readable and easily queryable.

In addition to storing data and relationships, knowledge graphs incorporate organizing principles, which serve as frameworks or conceptual rules that classify and structure information. By combining facts, relationships, and these guiding principles, a knowledge graph can reveal patterns and connections that might not be obvious at first glance. This makes it a valuable tool for uncovering new insights and helping users understand complex information in a clearer, more connected way.

Large Language Models (LLMs) are a class of artificial intelligence systems designed to process and generate natural language. Trained on massive amounts of data, these models learn statistical patterns in language that allow them to perform a wide range of tasks, including text generation, summarization, question answering, and information extraction. Their ability to model linguistic context enables them to produce fluent and contextually relevant outputs across diverse domains.

LLMs are typically built using transformer-based architectures, which are well suited for capturing relationships between words and phrases across long sequences of text. Rather than relying on predefined rules or keyword matching, LLMs generate language by predicting the most likely continuation of given input based on patterns learned during training. This allows them to capture nuance, semantic relationships, and implicit reasoning in ways that traditional retrieval or rule-based systems cannot.

Once trained, LLMs can be applied to a variety of applications with minimal task-specific customization. In addition to supporting interactive tasks such as drafting text or answering questions, LLMs can be embedded within larger systems that perform multistep reasoning or automated analysis. When combined with external data sources or orchestration mechanisms, such systems may exhibit agent-like behavior, enabling them to support more complex workflows.

Despite their flexibility and expressive power, LLMs operate primarily as probabilistic models and do not inherently guarantee factual accuracy or traceable reasoning. This limitation has motivated the integration of LLMs with external knowledge sources and retrieval mechanisms, such as RAG and KGs, to improve grounding, transparency, and reliability in knowledge-intensive application.

RAG is an approach that combines the fluency of modern AI language models with the reliability of evidence-based search. Instead of generating answers from what the AI

ORIGINAL RESEARCH ARTICLE / DOI: 10.20901/ms.17.33.8 / SUBMITTED: 30.09.2025.

remembers, RAG retrieves the most relevant information from a real, trusted database. It then uses that information to craft a clear, accurate, and contextually aware response. This helps address a major shortcoming of traditional AI systems: the tendency to generate plausible but incorrect information, a problem known as hallucination.

The RAG system consists of three main parts: retriever, retriever fusion/augmentation, and generator. First, the user asks a question. The question is converted to a vector representation that is used to search for a similar context in the database. The most similar information is retrieved from the database and fed to an LLM, which then answers the question based on the context found. This response is grounded in real evidence, not guessing.

RAG emerged as a response to the limitations of traditional 'closed-book' models, which rely solely on their training data (for instance, Lewis et al. (2020)). By introducing retrieval, it becomes an 'open-book' system that can consult external, up-to-date knowledge at inference time, resulting in improved accuracy (Mansurova et al., 2024).

Throughout the 21st century, FIMI and disinformation have emerged as major strategic risks, particularly for Western democracies (Bleyer-Simon et al., 2025). These influence operations aim to mislead, disrupt, or destabilize societies by spreading false or misleading information. Addressing this challenge requires international cooperation and shared standards. In response, the European Union and the United States have promoted the development of a common language and a shared technology stack that would enable the description, exchange, and correlation of influence operations in the same way as techniques used against information security[4].

Cyber Threat Intelligence (CTI) plays a crucial role in this process. CTI involves the structured collection, analysis, and dissemination of threats in the cyber domain. Instead of just reacting when an attack happens, CTI helps organizations prepare in advance by understanding the tactics and tools that attackers are likely to use.

To support these efforts, several key frameworks and tools have been developed. Together, they help structure CTI activities and ensure that information can be exchanged and acted upon efficiently:

The **DISARM Framework**, inspired by the MITRE ATT&CK framework [5] (widely used for Threat Intelligence in cybersecurity), is a tool that supports organizations in identifying, analyzing, and protecting themselves against influence operations by maintaining and improving their publicly available frameworks and tools through the cataloguing of tactics, techniques and procedures specific to FIMI campaigns[6].

[4] European External Action Service, "Annex 3 – FIMI_29 May," in *Trade and Technology Council Fourth Ministerial – Annex on Foreign Information Manipulation and Interference in Third Countries*, Brussels: EEAS, 31 May 2023. [Online]. Available: https://www.eeas.europa.eu/sites/default/files/documents/2023/Annex%203%20-%20FIMI_29%20May.docx.pdf. [Accessed: Aug. 6, 2025].
[5] The MITRE Corporation, "MITRE ATT&CK®," web resource, The MITRE Corporation, McLean, VA and Bedford, MA, USA. [Online]. Available: https://attack.mitre.org/. [Accessed: Aug. 6, 2025].
[6] DISARM Foundation C.I.C., "DISARM Framework," web resource, DISARM Foundation C.I.C., London, UK. [Online]. Available: https://www.disarm.foundation/framework. [Accessed: Aug. 6, 2025].

ORIGINAL RESEARCH ARTICLE / DOI: 10.20901/ms.17.33.8 / SUBMITTED: 30.09.2025.

Structured Threat Information eXpression (STIX)[7] is the most used and standardized format for sharing cyber threat. It ensures that different systems and organizations can exchange threat information in a consistent and machine-readable way. Its popularity in intelligence sharing facilitates the integration of FIMI into existing CTI flows (Barnum, 2014).

OpenCTI platform is an open-source threat intelligence platform that helps organizations to structure and operationalize their holistic cyber threat intelligence; it allows users to import bundles in STIX format, view them and orchestrate automatic processing[8].

OVERVIEW OF THREAT KNOWLEDGE REPRESENTATION

In this section, we provide an overview of the methods and techniques for representing knowledge about threats and incidents in the field of disinformation, as well as the techniques used to leverage that knowledge, such as fact-checking, which is the focus of this paper.

One of the earliest techniques used to represent knowledge is knowledge graphs (KG), which are widely used to structure, store, and query factual information, enabling automated reasoning over entity relationships. This makes them highly relevant for detecting disinformation and supporting fact-checking workflows. Ciampaglia et al. (2015) pioneered computational fact-checking using knowledge networks by demonstrating that veracity can be estimated by measuring the distance between concepts in a graph. Building on this, Shi & Weninger (2016) proposed fact-checking via KGs augmented with semantic relatedness measures. Thorne et al. (2018) introduced the FEVER dataset, which served as the model for numerous KG-based verification pipelines despite being text-based.

In the context of misinformation detection, Tchechmedjiev et al. (2019) developed ClaimsKG, a multilingual claim-evidence knowledge graph for linking related claims and evidence across languages. The purpose of ClaimsKG is to enable users to search for verified facts about a given entity. Similarly, it can be used to infer false facts about people, organisations, etc. For stance detection and claim verification, K. Zeng et al. (2021) applied graph neural networks over entity-relation triples, showing that graph-structured features perform better than text-only baselines.

From a survey perspective, Manos et al. (2024) evaluated KG-supported fact-checking across multiple domains, demonstrating that such approaches can be integrated into both manual and automated verification workflows.

[7] OASIS Cyber Threat Intelligence Technical Committee, "Introduction to STIX," web resource, OASIS, Burlington, MA, USA, last updated Jul. 24, 2024. [Online]. Available: https://oasis-open.github.io/cti-documentation/stix/intro.html. [Accessed: Aug. 6, 2025].
[8] Filigran SAS, "Operationalizing Threat Intelligence for National Security: Executive Summary," executive summary white paper, Filigran SAS, Asnières-sur-Seine, France, Mar. 2025. [Online]. Available: https://filigran.io/app/uploads/2025/03/filigran-operationalizing-threat-intelligence-for-national-security_executive-summary.pdf. [Accessed: Aug. 6, 2025].

ORIGINAL RESEARCH ARTICLE / DOI: 10.20901/ms.17.33.8 / SUBMITTED: 30.09.2025.

Despite these advances, most KG-based techniques are limited in their ability to adapt to quickly changing narratives as they work with static or pre-built graphs. Integrating KGs with real-time retrieval pipelines, such as RAG, offers a way to ground generative outputs in structured, verified data while dynamically updating with new facts.

In the stage of reasoning and knowledge exploitation, recent advances in large-scale language models have added significant value, enabling a more sophisticated understanding of natural language. However, fact-checking tasks are severely challenged by LLMs' propensity to "hallucinate" facts. Retrieval-augmented generation, a technique introduced by Lewis et al. (2020) to ground generative models with retrieved external knowledge, demonstrated significant performance gains in knowledge-intensive NLP tasks. Building on this, Singhal et al. (2024) showed improved factual accuracy in automated fact-checking by integrating explicit evidence retrieval into RAG pipelines. More recently, in order to improve interpretability and trustworthiness, Ge et al. (2025) tackled the problem of conflicting evidence by adding source credibility scores into the RAG process. To further enhance reliability, L. Zeng et al. (2025) studied RAG behavior under misleading retrievals using the RAGuard dataset, showing that performance can degrade below zero-shot baselines if the retrieved context is deceptive.

Beyond these approaches, Khaliq et al. (2024) utilized multimodal LLMs with a reasoning method called Chain-of-RAG for fact-checking across text and images, achieving strong performance in the political domain.

These works highlight the potential of retrieval-based architectures to produce more reliable outputs, but few have explored their integration into operational fact-checking environments like OpenCTI.

Cyber Threat Intelligence platforms, such as OpenCTI, primarily focus on the aggregation and analysis of technical indicators of compromise, guided by standards like the Structured Threat Information Expression (Barnum, 2014). Recent research has extended CTI to address disinformation campaigns. The authors in Cheng et al. (2025) propose CTINexus, a framework leveraging optimized in-context learning of LLMs for data-efficient CTI knowledge extraction and high-quality cybersecurity knowledge graph construction. As the authors note, the framework requires minimal data and parameter tuning and can adapt to various ontologies with minimal data annotation. However, its evaluation is limited to controlled datasets, raising questions about scalability to noisy, multilingual, or adversarial real-world CTI data. Cotroneo et al. (2025) introduced FakeCTI, a dataset that systematically links fake news to disinformation campaigns and threat actors. González et al. (2025) developed DISINFOX, a modular and interoperable client-server platform for sharing disinformation incidents. Suomalainen et al. (2025) introduce a framework of generic metrics for identifying typical features of hybrid operations from CTI. However, integration with existing CTI standards and the computational demands of LLM-based approaches remain practical challenges for real-world deployment.

ORIGINAL RESEARCH ARTICLE / DOI: 10.20901/ms.17.33.8 / SUBMITTED: 30.09.2025.

Despite these advances, existing CTI solutions rarely integrate explainable fact-checking or combine retrieval and knowledge graph reasoning for disinformation verification.

METHODOLOGY

Architecture Overview

An overview of the proposed architecture is presented in Figure 1. The following subsections describe each of its main components in more detail: the knowledge graph, the RAG system, and the OpenCTI integration.

Data Sources (Knowledge Graph)

Content used to inform the language model is provided via APIs developed by the ATHENA project,[9] a Horizon Europe-funded research project aimed at identifying and analyzing online foreign information manipulation and interference (FIMI). ATHENA maintains a flexible content acquisition and analysis platform that collects online material such as news articles and social media posts. For research, this data is exposed through several APIs, collectively referred to as the knowledge graph.

The knowledge graph is, in actuality, comprised of two components. A relational database, which is used for fast read/write operations during content acquisition and analysis; and a triplestore which provides a formal representation of information in a semantic graph, enabling more advanced forms of analysis such as reasoning and network exploration. Content is initially stored in the database following a standardized *Harmonized Data Model*, and later transferred to the triplestore, where it can be enriched with semantic relationships.

The model developed by ATHENA distinguishes between two types of entity (inspired by STIX 2.1): *Observables*, which represent evidence of disinformation, such as news articles or social media posts; and *Identities*, which capture organizations involved in spreading or hosting this content. It is noteworthy that the ATHENA model does not store any personally identifiable information about individuals. The focus is on the organizational level.

In addition to raw content, ATHENA enriches its data with analytical scores such as news reliability or hate-speech indicators. The APIs provide advanced search functionality that allows complex queries across these attributes. This makes it possible for the RAG system to retrieve Observables that are not only relevant to the user query but also pre-analyzed for quality and reliability.

[9] ATHENA Consortium, "ATHENA – Protecting democracy against Foreign Information Manipulation and Interference (FIMI)," project website. [Online]. Available: https://project-athena.eu/

ORIGINAL RESEARCH ARTICLE / DOI: 10.20901/ms.17.33.8 / SUBMITTED: 30.09.2025.

RAG

A core component of our methodology is Retrieval-Augmented Generation, which combines LLMs with an external knowledge retrieval system. The concept is straightforward: rather than requiring an artificial intelligence model to respond solely from memory, we first help it retrieve pertinent information from a collection of documents and then allowing it to generate an answer based on the acquired data. In our case, this memory is built from a collection of news articles that are in the Knowledge Graph.

Upon receiving a query, the system initiates a search inside its database for the most relevant passages. Retrieval is performed using a dense, embedding-based approach. Both the question and the stored documents are transformed into a mathematical representation that allows their similarity to be measured. This step ensures that the system retrieves not only documents containing terms matching the query, but also those that are conceptually related, a common challenge in journalistic fact-checking where paraphrasing and indirect references are frequent.

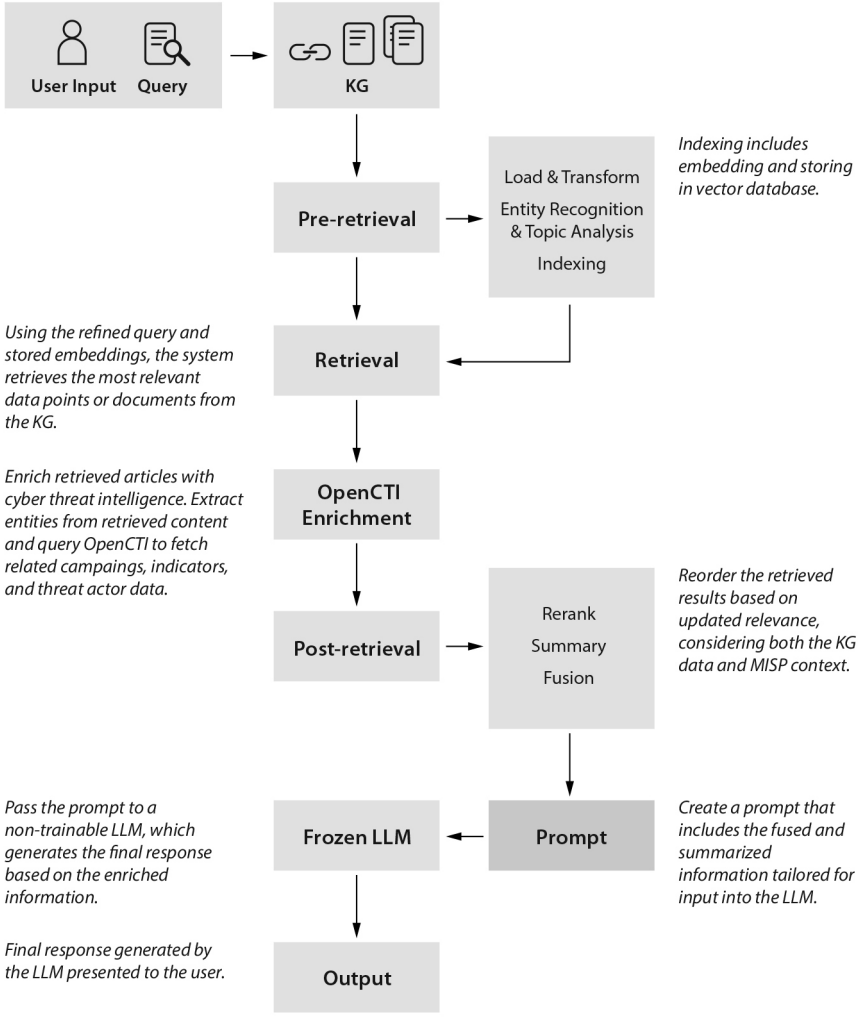
To improve retrieval precision, documents are segmented into smaller, semantically coherent units before indexing. This chunking strategy ensures that the retrieved material is focused on specific statements or events, reducing the risk that irrelevant context is introduced into the verification process. Only the most relevant segments are selected for downstream processing.

Once candidate passages are retrieved, they are filtered and ranked according to their semantic relevance to the claim. Because large language models operate within a limited context window, the retrieved passages are further processed using a smart truncation strategy. Instead of cutting text at a fixed length, truncation is applied at sentence boundaries, preserving complete factual statements. This approach improves the clarity and reliability of the model's justifications by ensuring that supporting evidence is not fragmented or taken out of context.

The selected evidence is then combined with the original claim and passed to an instruction-tuned large language model. In our experiments, we use Mixtral-8x7B and Qwen 2.5, two contemporary transformer-based models designed for reasoning and instruction following. The model is prompted to classify each claim as *True*, *False*, or *Unverifiable* and to produce a concise justification grounded explicitly in the retrieved documents. Throughout this process, source identifiers are preserved, allowing the system to link its conclusions directly to the underlying evidence.

By integrating retrieval with generation, the system combines the factual reliability of information retrieval with the expressive and reasoning capabilities of large language models. This design substantially reduces the risk of hallucinated content and increases transparency in automated fact-checking. For journalism and other related fields where accuracy and accountability are critical, RAG offers a practical and trustworthy framework for supporting verification tasks.

ORIGINAL RESEARCH ARTICLE / DOI: 10.20901/ms.17.33.8 / SUBMITTED: 30.09.2025.



▲ *Figure 1.*
Overview of the proposed architecture integrating Retrieval-Augmented Generation with a knowledge graph and the OpenCTI platform. The figure illustrates the end-to-end workflow, from claim ingestion and evidence retrieval to explainable fact-checking outputs, highlighting how retrieved evidence and structured CTI entities are combined to support analysis of disinformation and influence operations.

Integration with OpenCTI

While the RAG system based on news articles provides valuable evidence for fact-checking, it has an important limitation: news reports alone rarely capture the broader

ORIGINAL RESEARCH ARTICLE / DOI: 10.20901/ms.17.33.8 / SUBMITTED: 30.09.2025.

context of disinformation campaigns. To address this, we integrate OpenCTI, a platform dedicated to storing, organizing, and sharing cyber threat intelligence. By linking OpenCTI with the RAG workflow, the system gains access not only to raw reports but also to structured intelligence that has already been validated by trusted sources. This integration ensures that fact-checking benefits from both timely reporting and deeper contextual knowledge.

Proposed Architecture for Threat Intelligence

The proposed workflow for this project seeks to follow the architecture agreed upon by the European Union and the United States to preserve intelligence and data at every point in the analysis chain by combining these three technologies through RAG, which will verify each user request.

In this case, OpenCTI connectors will be used to transfer information from the OpenCTI architecture to the database used by the RAG. Connectors are the cornerstone of the OpenCTI platform and allow organizations to easily incorporate, enrich, or export data. In OpenCTI, there are connectors for enriching stored data, export/import information, and the one proposed in this architecture, the stream connector, which allows data to be linked in real time with another platform or service (in this case, the database used in the RAG).

In this way, the RAG will be able to obtain real-time updates on all cases of disinformation stored in OpenCTI. Unlike the articles and news items used by the RAG, the information in OpenCTI corresponds to cases of information that have already been analyzed (by reliable sources) and can be enriched through correlations between the two parts. For example, it is difficult to deduce the tactics and techniques used from news or articles that have been used to verify the narrative of the case; by using OpenCTI cases, this information can be obtained by correlating cases similar to those analyzed.

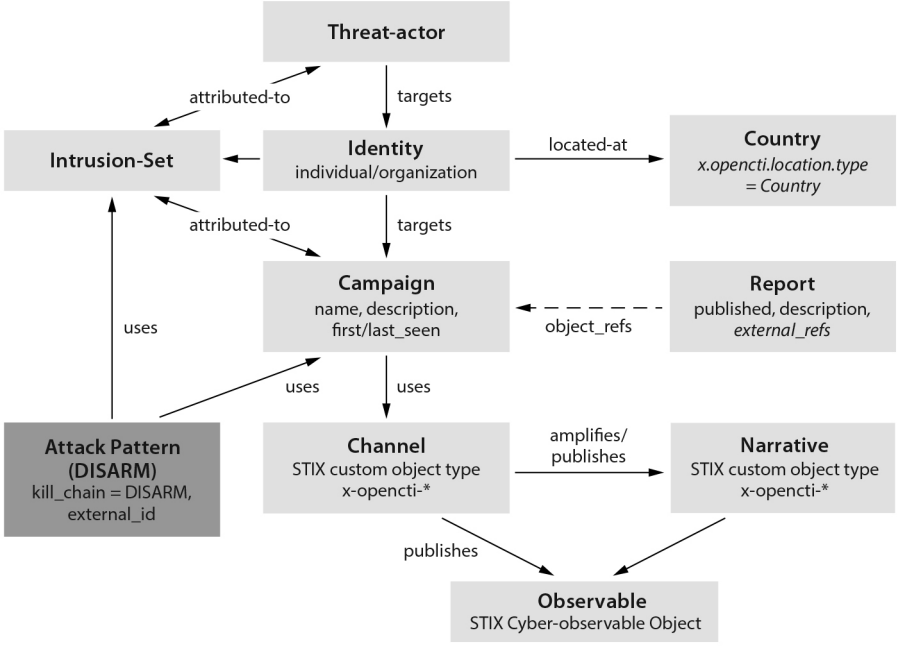
Proposed Data Structure for Threat Intelligence

Within OpenCTI, different types of entities can be created to store and link information. The correct definition of entities is vital in order to have information available in an orderly manner, allowing for subsequent analysis and correlations to create or modify new information.

Based on other research on disinformation and FIMI with OpenCTI and the characteristics of the data resulting from the RAG, it has been decided to store the information within a STIX object called 'campaign'. In STIX, campaigns are defined as a group of adversarial behaviors that describe a set of malicious activities or attacks that occur over a period against a specific set of targets. Campaigns usually have well-defined objectives and may be part of an Intrusion Set.

Within the 'campaign' object, the rest of the STIX objects will be referenced, which will provide all the information related to the disinformation or FIMI case.

ORIGINAL RESEARCH ARTICLE / DOI: 10.20901/ms.17.33.8 / SUBMITTED: 30.09.2025.



▲ *Figure 2.*
Diagram of relationships and attributes between STIX 2.1 entities used in OpenCTI for storing structureFIMI cases. The diagram shows white boxes representing each of the entities that could be obtained through an analysis of the related RAG using the STIX structure. The entity to be added to the case through subsequent analyses is shown in grey. The complete diagram shows a structured case stored in OpenCTI regardless of its origin (RAG or reliable sources).

As shown in Figure 2, the relationships between the STIX objects that will be uploaded to OpenCTI. In addition to the campaign that will encompass all objects, we find the following object names:

>Identity. Represents individuals, organizations, or classes of individuals/organizations (e.g., a politician). In OpenCTI, it corresponds to the Entities object. Its main function is to describe the individuals or organizations that are the targets of attacks.
>Country. This does not exist as a native object in STIX 2.1, so it is modelled as a STIX extension. It allows the nationality or geographical location of an Identity to be indicated. In OpenCTI, it corresponds to the object with the same name, Country.
>Intrusion-Set. A set of adversarial behaviors and resources that share common characteristics. It represents the general authorship of campaigns.
>Threatactor. Specific individuals, groups or organizations acting with malicious intent. Several Threatactors can be part of the same Intrusion-Set. A Threatactor can be directly related to the Identity object as the target of its operations.

ORIGINAL RESEARCH ARTICLE / DOI: 10.20901/ms.17.33.8 / SUBMITTED: 30.09.2025.

>Attack Pattern. A type of TTP that describes how adversaries attempt to compromise their targets. In the FIMI context, they are modelled following the DISARM framework.
>Channel. Object not defined in STIX, therefore modelled using a STIX extension. Represents the means or spaces through which actors disseminate information (e.g. social networks, forums, etc.).
>Narrative. As with the Channel object, it is implemented as a STIX extension to link to the object within OpenCTI. It allows specific narratives used in disinformation campaigns to be described, making it easier to track which narratives have been and are being exploited by actors.
>Observable. Represents an immutable object, such as IPv4 addresses, domains, emails, files, or user accounts. In STIX, it is modelled as a Cyber-observable Object, while in OpenCTI it is a separate object.

It should be noted that when we talk about OpenCTI entities and relationships, we are referring to information stored in STIX format. This format is not compatible with the databases used for RAG, so the connector will be responsible for generating new entities that can be stored in a database readable by RAG.

In this case, the relationships are intrinsic to each campaign. In other words, once the information for a campaign has been stored, the relationships between entities will be the same (for example, within the same campaign, the entity 'Identity' will always have a 'located-at' relationship with 'country'). This feature makes it easier to convert data from STIX format to a format readable by the RAG.

Therefore, during the threat intelligence data flow, there will be two parts. Firstly, the entities and relationships stored within OpenCTI in STIX format. Once new data enriches the previously analyzed data, the connector will be responsible for transferring it in real time to the database used by the RAG in an appropriate format for subsequent processing. In this way, the RAG will always have at its disposal the new data from investigations carried out in real time.

RESULTS & DISCUSSION

The evaluation presented in this section is intended to illustrate the behavior and explainability of the proposed system in realistic analyst-facing scenarios, rather than to provide a comprehensive quantitative benchmark.

The system was tested using a collection of approximately 80,000 news articles and disinformation-related cases extracted from OpenCTI. Queries were designed to reflect realistic fact-checking requests that a journalist, policy analyst, or informed citizen might ask. The responses were evaluated on three main criteria: accuracy (whether the system classified claims as *True*, *False*, or *Unverifiable* correctly), justification (whether explanations referred back to context in a meaningful way), and clarity (whether answers were easy to interpret).

ORIGINAL RESEARCH ARTICLE / DOI: 10.20901/ms.17.33.8 / SUBMITTED: 30.09.2025.

One example query asked: *"Did the government pass a new cybersecurity law last month?"* The system returned *Unverifiable*, with the justification that the retrieved context discussed US cyber operations against Russia, but not any legislative action. This result illustrates a central strength of the workflow: rather than guessing or fabricating an answer, the system chose caution. For fact-checking applications, this behavior is crucial because it preserves trust. An unverifiable response is far more reliable than a false claim of certainty.

```
question = "Did the government pass a new cybersecurity law last month?"
✓ 1m 8.9s

Answer:
 Unverifiable.  The given context discusses a one-day suspension of cyber
operations by the United States against Russia, but it does not mention anything
about a new cybersecurity law being passed. Therefore, the claim about a new
cybersecurity law is unverifiable based on the provided context.
```

▲*Figure 3.*
Example RAG response for an unverifiable claim.
The output illustrates how the system classifies a claim as unverifiable
when retrieved evidence is insufficient or inconclusive,
while still providing the reasoning behind the classification.

When tested with queries that corresponded directly to article content, the system successfully retrieved the relevant chunks and classified them accurately. For instance, when asked whether a specific political figure made a documented statement, the model identified the relevant article excerpts, confirmed the statement as *True* (as shown in the figure below), and cited the corresponding article ID. This shows that the retrieval mechanism functions well when the supporting evidence exists in the database.

```
question = "Did Lukushin say that negotiations between Russia and Ukraine will progress only after pressure on Kiev?"
✓ 1m 48.8s

Answer:
 **True.** Vladimir Lukushin stated that the progress of negotiations between
Russia and Ukraine will be possible only if the Western allies put pressure on
Kiev. This is directly quoted from the given context.   **Justification:** The
context mentions, "The progress of negotiations between Russia and Ukraine will
be possible only if the Western allies put pressure on Kiev. This was stated by
Vladimir Lukushin..." Therefore, the claim is accurate. **[Article
68a3ae3b-29b0-417d-a8ae-073606738f41]**
```

▲*Figure 4.*
Example RAG response for a claim classified as true.
The figure demonstrates how the system retrieves relevant evidence,
grounds the response in verified sources, and generates an explainable
justification supporting the truthfulness of the claim.

ORIGINAL RESEARCH ARTICLE / DOI: 10.20901/ms.17.33.8 / SUBMITTED: 30.09.2025.

In cases where OpenCTI information was relevant, the integration enriched the answers. For example, when evaluating claims about disinformation campaigns, the system not only checked the news coverage but also cross-referenced OpenCTI cases. This allowed it to highlight known tactics and patterns that aligned with or contradicted the claim. Such contextual reinforcement is valuable for users who need both verification and an understanding of broader dynamics.

```
question = "Were the elections of 2020 in the US considered fraud?"
✓ 2m 13.5s

Answer:
 False. The provided context is from a source that is presenting claims made by
Trump regarding the 2020 election being rigged and stolen. However, multiple
fact-checking articles (sources provided with IDs:
b5f7a3e3-8e3e-4d7b-98ab-e8c9cc49eac2, 25d043c8-edbd-4b0b-916b-cae1edbed44b,
etc.) confirm that these claims are baseless and have been thoroughly debunked
by election officials, courts, and independent auditors. There is no credible
evidence of widespread voter fraud or any systemic issues that would have
altered the outcome of the election. Therefore, the elections of 2020 in the US
were not considered fraudulent according to verifiable evidence.
```

▲ *Figure 5.*
Example RAG response for a claim classified as false.
The output shows how conflicting or corrective evidence is retrieved and linked to the claim, enabling the system to generate an explainable refutation supported by external sources and structured knowledge.

The comparison between using articles alone and combining them with OpenCTI revealed several advantages. Articles provide timeliness and narrative detail, but they are often incomplete when it comes to technical or tactical insights. OpenCTI fills this gap by providing structured knowledge derived from multiple validated sources. By merging them, the system delivers answers that are both current and context-aware. This dual-source approach also strengthens the robustness of fact-checking. News articles may vary in quality or bias, while OpenCTI cases follow stricter validation standards. When both sources point in the same direction, confidence in the result increases.

Another important outcome is the system's ability to justify its responses. Each answer is accompanied by a short explanation and references to specific article IDs or intelligence cases. This is not a trivial feature, as in many AI systems, the reasoning behind outputs remains hidden, which reduces user trust. By contrast, our workflow provides traceable evidence. This transparency is especially important in fact-checking, where credibility is paramount.

ORIGINAL RESEARCH ARTICLE / DOI: 10.20901/ms.17.33.8 / SUBMITTED: 30.09.2025.

CONCLUSIONS

This study has presented a workflow that combines political news articles with threat intelligence from OpenCTI, linked through a retrieval-augmented generation (RAG) architecture. The system demonstrates that fact-checking can be enhanced when different types of evidence are integrated and leveraged by a language model.

The system proves to be a practical tool for fact-checking, as it is able to deliver clear verdicts in the form of *True*, *False*, or *Unverifiable*, while also providing explanations and references that allow users to understand the basis of each response. This makes it not only useful for validating political claims but also transparent and accountable in how it reaches its conclusions. By combining unstructured reporting with structured intelligence, the workflow bridges the narrative detail of news articles with the validated contextual knowledge provided by OpenCTI. This integration makes the approach more robust than relying on a single source type, and it allows the system to situate claims within a broader landscape of disinformation tactics and patterns.

Nevertheless, the work is not without its challenges. The breadth and quality of the underlying databases directly shape the accuracy of the results, and vague or ambiguous queries can limit the system's ability to provide precise answers. The constraints of processing length also mean that not all relevant details can always be included in the analysis. These limitations suggest promising directions for future research, including the expansion of data sources, improved handling of ambiguous claims, and more advanced methods for dynamically selecting contextual information.

Despite these strengths, several limitations emerged. First, the coverage of the databases directly affects the system's performance. If no article or OpenCTI case exists about a specific claim, the system has no evidence to work with and must return *Unverifiable*. While this is the correct behavior, it may frustrate users who expect a definitive answer.

Second, ambiguity in queries poses difficulties. For instance, the question "Did the government pass a new law?" is too vague. The system may retrieve irrelevant documents or struggle to determine which government is being referred to. At present, this is handled by returning *Unverifiable*, but in future iterations, the model could prompt the user for clarification before attempting a response.

Third, it is an architecture that is highly dependent on the available information and the communities that publish it. Therefore, the system may encounter biased or subjective information that can greatly affect the analyses and objectives obtained. Recent events provide a clear example: having completely objective, real-time data on unfolding events is often an impossible task given the large amount of disinformation circulating on social media. This can hinder or create bias for recent cases of disinformation. Therefore, well-established research and thoroughly verified articles are required to enrich the

ORIGINAL RESEARCH ARTICLE / DOI: 10.20901/ms.17.33.8 / SUBMITTED: 30.09.2025.

architecture. This process takes time and can lead to the first limitation mentioned above (lack of articles or research).

Finally, length constraints sometimes limit the richness of justifications. Because language models process a finite amount of context, long articles must be truncated. While the current system uses a smart truncation method, this occasionally omits secondary details that could be useful in nuanced fact-checking.

In spite of these constraints, this study underscores the potential of combining news data, cybersecurity intelligence, and retrieval-augmented AI to create systems that are accurate, explainable, and trustworthy. Rather than replacing human judgment, the proposed approach serves as a supportive tool that strengthens the fight against disinformation by grounding responses in verifiable evidence.

References

>Aïmeur, E., Amri, S., & Brassard, G. (2023). Fake news, disinformation and misinformation in social media: A review. *Social Network Analysis and Mining*, *13*(1), 30. https://doi.org/10.1007/s13278-023-01028-5

>Barberá, P. (2018). Explaining the Spread of Misinformation on Social Media: Evidence from the 2016 U.S. Presidential Election. http://pablobarbera.com/static/barbera-CP-note.pdf.

>Barnum, S. (2014, February 20). STIX Whitepaper | *STIX Project Documentation*. https://stixproject.github.io/getting-started/whitepaper/

>Bleyer-Simon, K., Horowitz, M., Botan, M., Brautovic, M., Brogi, E., Bucholtz, I., Culloty, E., Gavurova, B., Grabovac, A., Ioakem, S., Kermer, J., Kiely, K., Leonidou, P., Mallia, M., Moreno, J., Nenadić, I., Paisana, M., Palmer, M., Reviglio, U., ... Weikmann Edited, T. (n.d.). *How is disinformation addressed in the member states of the European Union? - 27 Country cases* [Review of How is disinformation addressed in the member states of the European Union? - 27 Country cases]. Retrieved July 17, 2026, from https://edmo.eu/wp-content/uploads/2025/06/EDMO-Report_How-is-disinformation-addressed-in-the-member-states-of-the-European-Union_27-country-cases.pdf

>Broda, E., & Strömbäck, J. (2024). Misinformation, Disinformation, and Fake News: Lessons from an Interdisciplinary, Systematic Literature Review. *Annals of the International Communication Association, 48*(2), 139-166. https://doi.org/10.1080/23808985.2024.2323736

>Cheng, Y., Bajaber, O., Tsegai, S. A., Song, D., & Gao, P. (2025). CTINexus: Automatic Cyber Threat Intelligence Knowledge Graph Construction Using Large Language Models (No. arXiv:2410.21060). *arXiv*. https://doi.org/10.48550/arXiv.2410.21060

>Ciampaglia, G. L., Shiralkar, P., Rocha, L. M., Bollen, J., Menczer, F., & Flammini, A. (2015). Computational fact checking from knowledge networks. *PLOS ONE*, *10*(6), e0128193. https://doi.org/10.1371/journal.pone.0128193

>Cotroneo, D., Natella, R., & Orbinato, V. (2025). Elevating Cyber Threat Intelligence against disinformation campaigns with LLM-based Concept Extraction and the FakeCTI Dataset (No. arXiv:2505.03345). *arXiv*. https://doi.org/10.48550/arXiv.2505.03345

>Gaeta, A., Loia, V., Lomasto, L., & Orciuoli, F. (2023). A novel approach based on rough set theory for analyzing information disorder. *Applied Intelligence*, *53*(12), 15993-16014. https://doi.org/10.1007/s10489-022-04283-9

>Gao, Y. (2025). Information disorder in news language: A semantic analysis and journalistic ethics discussion, L. Chen (Ed.), *2025 6th International Conference on Economics, Education and Social Research* (ICEESR 2025). Paris, France, from 2025-03-30 to 2025-03-31. ISBN: 978-1-915228-85-7. https://doi.org/: 10.25236/iceesr.2025.012

ORIGINAL RESEARCH ARTICLE / DOI: 10.20901/ms.17.33.8 / SUBMITTED: 30.09.2025.

>Ge, Z., Wu, Y., Chin, D. W. K., Lee, R. K.-W., & Cao, R. (2025). Resolving Conflicting Evidence in Automated Fact-Checking: A Study on Retrieval-Augmented LLMs (No. arXiv:2505.17762). *arXiv*. https://doi.org/10.48550/arXiv.2505.17762
>González, F. S., Pastor-Galindo, J., & Ruipérez-Valiente, J. A. (2025). Toward interoperable representation and sharing of disinformation incidents in cyber threat intelligence (No. arXiv:2502.20997). *arXiv*. https://doi.org/10.48550/arXiv.2502.20997
>Khaliq, M. A., Chang, P., Ma, M., Pflugfelder, B., & Miletić, F. (2024). RAGAR, Your Falsehood Radar: RAG-Augmented Reasoning for Political Fact-Checking using Multimodal Large Language Models (No. arXiv:2404.12065). *arXiv*. https://doi.org/10.48550/arXiv.2404.12065
>Lee, T. (2019). The global rise of "fake news" and the threat to democratic elections in the USA. *Public Administration and Policy*, *22*(1), 15-24. https://doi.org/10.1108/PAP-04-2019-0008
>Lewis, P., Perez, E., Piktus, A., Petroni, F., Karpukhin, V., Goyal, N., Küttler, H., Lewis, M., Yih, W., Rocktäschel, T., Riedel, S., & Kiela, D. (2020). Retrieval-Augmented Generation for Knowledge-Intensive NLP Tasks. *Advances in Neural Information Processing Systems*, 33, 9459-9474. https://proceedings.neurips.cc/paper/2020/hash/6b493230205f780e1bc26945df7481e5-Abstract.html
>Manos, A. (2024). Knowledge Graphs and Machine Learning in Fake News and Disinformation Detection. *2024 International Conference on Engineering and Emerging Technologies (ICEET)*, 1–6. https://doi.org/10.1109/ICEET65156.2024.10913780
>Mansurova, A., Mansurova, A., & Nugumanova, A. (2024). QA-RAG: Exploring LLM reliance on external knowledge. *Big Data and Cognitive Computing, 8*(9), 115. https://doi.org/10.3390/bdcc8090115
>Olan, F., Jayawickrama, U., Arakpogun, E. O., Suklan, J., & Liu, S. (2024). Fake news on social media: The impact on society. *Information Systems Frontiers, 26*(2), 443-458. https://doi.org/10.1007/s10796-022-10242-z
>Puczyńska, J., & Djenouri, Y. (2024). AI in disinformation detection. *Applied Cybersecurity & Internet Governance, 3*(2), 211-232. https://doi.org/10.60097/ACIG/200200
>Rose, J. (2017). Brexit, Trump, and post-truth politics. *Public Integrity, 19*(6), 555-558. https://doi.org/10.1080/10999922.2017.1285540
>Shi, B., & Weninger, T. (2016). Discriminative predicate path mining for fact checking in knowledge graphs. *Knowledge-Based Systems, 104*, 123-133. https://doi.org/10.1016/j.knosys.2016.04.015
>Singhal, R., Patwa, P., Patwa, P., Chadha, A., & Das, A. (2024). Evidence-backed fact checking using RAG and few-shot in-context learning with LLMs. *arXiv* (No. arXiv:2408.12060). . https://doi.org/10.48550/arXiv.2408.12060
>Suomalainen, J., Ford, R.Q., Kolehmainen, K., & Kuusijärvi, J. (2025). Cyber Threat Intelligence for Hybrid Attacks: Leveraging LLMs and Data Spaces. *IEEE Access,* 13, 202467-202481. https://doi.org/10.1109/ACCESS.2025.3636275
>Tchechmedjiev, A., Fafalios, P., Boland, K., Gasquet, M., Zloch, M., Zapilko, B., Dietze, S., & Todorov, K. (2019). ClaimsKG: A knowledge graph of fact-checked claims. In C. Ghidini, O. Hartig, M. Maleshkova, V. Svátek, I. Cruz, A. Hogan, J. Song, M. Lefrançois, & F. Gandon (Eds.), *The Semantic Web – ISWC 2019* (pp. 309–324). Springer International Publishing. https://doi.org/10.1007/978-3-030-30796-7_20
>Thorne, J., Vlachos, A., Christodoulopoulos, C., & Mittal, A. (2018). FEVER: A large-scale dataset for Fact Extraction and VERification . *arXiv* (No. arXiv:1803.05355). https://doi.org/10.48550/arXiv.1803.05355
>Wardle, C. (2018a, July 6). Definitions matter. These tools will give you the words you need to talk about information disorder. *First Draft News*. https://firstdraftnews.org/articles/infodisorder-definitional-toolbox/
>Zeng, K., Li, C., Hou, L., Li, J., & Feng, L. (2021). A comprehensive survey of entity alignment for knowledge graphs. *AI Open, 2*, 1-13. https://doi.org/10.1016/j.aiopen.2021.02.002
>Zeng, L., Gupta, R., Motwani, D., Yang, D., & Zhang, Y. (2025). Worse than zero-shot? A fact-checking dataset for evaluating the robustness of RAG against misleading retrievals. *arXiv* (No. arXiv:2502.16101). https://doi.org/10.48550/arXiv.2502.16101

ORIGINAL RESEARCH ARTICLE / DOI: 10.20901/ms.17.33.8 / SUBMITTED: 30.09.2025.

UNAPRJEĐENJE UTVRĐIVANJA ČINJENICA PRIMJENOM RAG-a I GRAFOVA ZNANJA

Ashneet Khandpur Singh :: Pau Perea Paños :: Mario Reyes de los Mozos :: Gary Munnelly

Sažetak *Tradicionalne metode utvrđivanja činjenica, iako učinkovite, često su prespore da bi držale korak s brzinom kojom se lažne informacije šire internetom. Posljednjih je godina umjetna inteligencija stekla popularnost kao sredstvo automatizacije procesa utvrđivanja činjenica, osobito veliki jezični modeli (engl.* large language models*). Iako su veliki jezični modeli pokazali učinkovitost u pružanju podrške pri zadatcima utvrđivanja činjenica, njihova je primjena ograničena čimbenicima kao što su kvaliteta podataka za učenje i sposobnost pronalaženja relevantnih informacija potrebnih za utvrđivanje činjenica. Također su skloni „halucinacijama", odnosno generiranju uvjerljivih, ali netočnih ili obmanjujućih informacija, što izaziva ozbiljnu zabrinutost u pogledu njihove pouzdanosti kao samostalnih alata za utvrđivanje činjenica. U ovom se radu istražuje hibridni pristup koji objedinjuje generiranje potpomognuto dohvatom informacija (engl.* retrieval-augmented generation*, RAG), grafove znanja (engl.* knowledge graphs*) i OpenCTI, platformu za prikupljanje i analizu obavještajnih podataka o kibernetičkim prijetnjama. Takav pristup osigurava da konačni rezultat ne bude samo informativan nego i transparentan, povezujući tvrdnje s vjerodostojnim izvorima. Time predstavlja vrijedan alat za novinare, stručnjake za utvrđivanje činjenica i širu javnost, kojima su sve potrebniji pravodobni i pouzdani odgovori u informacijskom okružju oblikovanom dezinformacijama.*

KLJUČNE RIJEČI

DEZINFORMACIJE, GRAFOVI ZNANJA, GENERIRANJE POTPOMOGNUTO DOHVATOM INFORMACIJA, UMJETNA INTELIGENCIJA, VELIKI JEZIČNI MODELI, OBAVJEŠTAJNI PODATCI O PRIJETNJAMA, OPENCTI

Bilješka o autorima

Ashneet Khandpur Singh :: Eurecat, Tehnološki centar Katalonije, Barcelona **::**
ashneet.khandpur@eurecat.org
Pau Perea Paños :: Eurecat, Tehnološki centar Katalonije, Barcelona **::**
pau.perea@eurecat.org
Mario Reyes de los Mozos :: Eurecat, Tehnološki centar Katalonije, Barcelona **::**
mario.reyes@eurecat.org
Gary Munnelly :: Centar Adapt, Trinity College Dublin **::**
gary.munnelly@adaptcentre.ie