AI4Debunk

D6.3 FIRST REPORT ON THE BUILDING PROCESS OF THE KNOWLEDGE GRAPHS

September 2025

Funded by the European Union

Grant Agreement No.: 101135757
Call: HORIZON-CL4-2023-HUMAN-01-CNECT
Topic: HORIZON-CL4-2023-HUMAN-01-05
Type of action: HORIZON Innovation Actions

D6.3 FIRST REPORT ON THE BUILDING PROCESS OF THE KNOWLEDGE GRAPHS

Project Acronym	AI4Debunk
Project Number	101135757
Project Full Title	Participative Assistive AI-powered Tools for Supporting Trustworthy Online Activity of Citizens and Debunking Disinformation
Work package	WP 6
Task	Task 3
Due date	01/09/2025
Submission date	30/09/2025
Deliverable lead	Partner CNR
Version	0.1
Authors	Alessia D'Andrea (Partner CNR), Arianna D'Ulizia (Partner CNR), Riccardo Denni (Partner CNR), Nikos Pavlidis (DotSoft), Vachtanidis Giorgos (DotSoft), Karanasios Giorgos (Dotsoft), Elizabeth Filippidis (DotSoft)
Contributors	
Reviewers	Pascaline Gaborit (Partner Pilot4dev)
Abstract	The deliverable D6.3 - First report on the building process of the knowledge graphs – describes the building process of the two knowledge graphs, named the "unimodal" knowledge graph and the "multimodal" knowledge graph. The first "unimodal" knowledge graph consists in extracting the textual description from multimedia contents and adding this textual knowledge in the knowledge graph. The "multimodal" knowledge graph consists in embedding the multimedia contents within the knowledge

Funded by
the European Union

	graph based on multimodal feature description.
Keywords	Knowledge graph, multimodal knowledge graph, Wikidata, taxonomy, ontology

DOCUMENT DISSEMINATION LEVEL

Dissemination level	
X	PU - Public
	SEN - Sensitive

DOCUMENT REVISION HISTORY

Version	Date	Description of change	List of contributor(s)
0.1	01/08/2025	First draft	CNR, DotSoft
0.2	12/08/2025	Draft revised by the Task 6.3 participants	CNR, DotSoft
0.3	29/08/2025	Final draft version revised by the internal reviewer	CNR, DotSoft, Pilot4Dev
0.4	29/08/2025	Final version	CNR, DotSoft

STATEMENT ON MAINSTREAMING GENDER

The AI4Debunk consortium is committed to including gender and intersectionality as a transversal aspect in the project's activities. In line with EU guidelines and objectives, all partners – including the authors of this deliverable – recognise the importance of advancing gender analysis and sex-disaggregated data collection in the development of scientific research. Therefore, we commit to paying particular attention to including, monitoring, and periodically evaluating the participation of different genders in all activities developed within the project, including workshops, webinars and events but also surveys, interviews and research, in general. While applying a non-binary approach to data collection and promoting the participation of all genders in the activities, the partners will periodically reflect and inform about the limitations of their approach. Through an iterative learning process, they commit to plan and implement strategies that maximise the inclusion of more and more intersectional perspectives in their activities.

Funded by
the European Union

DISCLAIMER

The AI4Debunk project has received funding from the European Union's Horizon Europe Programme under the Grant Agreement No. 101135757.

COPYRIGHT NOTICE

How to cite this report: Alessia D'Andrea, Arianna D'Ulizia, Riccardo Denni, Nikos Pavlidis, Vachtanidis Giorgos, Karanasios Giorgos, Elizabeth Filippidis (2025). AI4Debunk D6.3: First Report on the Building Process of the Knowledge Graphs.

The AI4Debunk consortium is the following:

Participant number	Participant organisation name	Short name	Country
1	LATVIJAS UNIVERSITATE	UL	LV
2	FREE MEDIA BULGARIA	EURACTIV	BE
3	PILOT4DEV	P4D	BE
4	INTERNEWS UKRAINE	IUA	UA
5	CONSIGLIO NAZIONALE DELLE RICERCHE	CNR-IRPPS	IT
6	UNIVERSITA DEGLI STUDI DI FIRENZE	MICC/UNIFI	IT
6.1	CONSORZIO NAZIONALE INTERUNIVERSITARIO PER LE TELECOMUNICAZIONI	CNIT	IT
7	BARCELONA SUPERCOMPUTING CENTER CENTRO NACIONAL DE SUPERCOMPUTACION	BSC	ES
8	DOTSOFT OLOKLIROMENES EFARMOGES DIADIKTIOY KAI VASEON DEDOMENON AE	DOTSOFT	EL
9	UNIVERSITE DE MONS	UMONS	BE
10	NATIONAL UNIVERSITY OF IRELAND GALWAY	NUIG	IE
11	STICHTING HOGESCHOOL UTRECHT	HU	NL
12	STICHTING INNOVATIVE POWER	IP	NL
13	F6S NETWORK IRELAND LIMITED	F6S	IE

Funded by the European Union

Table of content

LIST OF FIGURES

ABBREVIATIONS

WP	Work Package
EC	European Commission
RDF	Resource Description Framework
NLP	Natural Language Processing
ML	Machine Learning
KG	Knowledge Graph
GNN	Graph Neural Network
MLP	Multi-layer Perceptrons
CNN	Convolutional Neural Network

Funded by
the European Union

EXECUTIVE SUMMARY

The deliverable D6.3 - First report on the building process of the knowledge graphs – describes the building process of the two knowledge graphs, named the “unimodal” knowledge graph and the “multimodal” knowledge graph. The first “unimodal” knowledge graph consists in extracting the textual description from multimedia contents and adding this textual knowledge in the knowledge graph. The “multimodal” knowledge graph consists in embedding the multimedia contents within the knowledge graph based on multimodal feature description.

A brief introduction on the objectives and expected outcome of Task 6.3 - Creation of the knowledge graphs – is provided. Moreover, the building process of the two knowledge graphs is discussed, focusing on (i) the creation of the taxonomy subgraph starting from a Wikidata taxonomy subgraph that contains the relevant concepts related to the fields of the two case studies (ii) the import of the false statements and the related multimedia contents, extracted in Task 6.1, into the Wikidata taxonomy subgraph; (iii) the connection of the false statements and the related contents to the taxonomy subgraph.

Funded by the European Union

1 INTRODUCTION

This deliverable introduces the objectives and expected outcomes of Task 6.3 - Creation of the knowledge graphs - within Work Package 6 - Design, creation, and adaptation of knowledge graphs. Two different knowledge graphs have been prepared, named the "unimodal" knowledge graph and the "multimodal" knowledge graph, based on the fake statements and related multimedia contents collected in Task 6.1.

In this deliverable, the building process of the two knowledge graphs is described. In particular, the creation of the taxonomy subgraph starting from a Wikidata taxonomy subgraph, the import of the false statements and the related multimedia contents, extracted in Task 6.1, into the Wikidata taxonomy subgraph, as well as the connection of the false statements and the related contents to the taxonomy subgraph are illustrated.

1.1 OBJECTIVES

This deliverable aims to describe the construction process of the knowledge graphs that illustrate the structure of the deceptive data. The building of the knowledge graphs starts from the construction of the Wikidata taxonomy subgraph that contains the relevant concepts related to the fields of the two case studies of the project (i.e. the war in Ukraine and the climate changes). Afterwards, the set of fake statements and the related multimedia contents (videos, images, audios, etc.), extracted in Task 6.1, are imported as a series of nodes into the Wikidata taxonomy subgraph. Finally, individual nodes of false statements and the related contents are connected to the taxonomy using the features extracted using NLP and ML techniques (e.g. score, topic, keywords, sentiment, multimodal information) in Task 6.2.
Despite the complexity of the task, the team approached each challenge with a methodical plan. Asynchronous data fetching and scalable graph construction tool, such as graph handling API, were employed to ensure efficiency, while robust error handling mechanisms addressed potential disruptions in data retrieval.
The work conducted in Task 6.3 marks a significant milestone in the development of the Debunk API. By creating a knowledge graph that combines structured taxonomies with real-world disinformation data, the project has established a strong foundation that can be utilized to identify similarity with other news sources that are under investigation.

Funded by
the European Union

1.2 EXPECTED OUTCOME

Two knowledge graphs are expected in the project at M21 (September 30, 2025), illustrating the semantic structure of the 2000 fake claims collected in the datasets released in Task 6.1. The knowledge graphs meant to be suitable for the development of a disinfoscore for AI4Debunk's fact-checking tools (Web plug-in, Disinfopedia, App and AR/VR interface) during the next steps of the AI4Debunk project.

The code of the two knowledge graphs developed is available in the AI4Debunk Github folder at this link:
https://github.com/AI4Debunk/Knowledge-Graph/tree/main/Task%206.3.2%20%26%206.3.3

2 PRELIMINARY CONCEPTS

2.1 KNOWLEDGE GRAPHS

Knowledge Graphs (KGs) (Singhal et al., 2012) serve as structured representations of knowledge, designed to capture complex relationships among entities in a way that both humans and machines can interpret. They are based on a data model that defines the formal rules and structure through which data is organized and understood, acting as a blueprint for representing various types of information. This data model is crucial as it dictates how entities (nodes) and relationships (edges) are created, maintained, and connected. KGs employ ontological frameworks, which further guide how entities are classified, interlinked, and related across domains, ensuring a unified language for interpreting diverse sources of information.

At the core of KGs, it lies the ontology—a formal, hierarchical structure that categorizes knowledge by defining entities, attributes, and relationships within a domain. This structure is vital because it provides the semantic context that KGs need to connect information in a proper way. For instance, an ontology in a KG designed for disinformation detection would include classes like “Person”, “Organization”, and “Event”, each with specific attributes (e.g., name, date, and location) and relationships (e.g., affiliation, part of, and occurred at). This level of detail allows a KG to represent new information while maintaining accuracy and relevance, essential when distinguishing reliable information from disinformation.

Taxonomy, another foundational component of KGs, organizes concepts into a layered, hierarchical structure, defining parent-child relationships between categories. In the context of

Funded by
the European Union

KGs, taxonomy serves to sort entities into predefined categories that reflect real-world classifications. For example, a KG designed to verify news stories might include a taxonomy where “Event” is a broader category that contains subcategories like “Political Event”, “Cultural Event”, and “Natural Disaster”. This layered structure supports data discovery and retrieval by allowing algorithms to traverse the hierarchy and locate relevant information efficiently.

Several studies have explored diverse methodologies for incorporating KGs into fake news detection frameworks. A significant methodology was developed by Gao et al. (2023), who proposed a commonsense knowledge graph model that compared news content against commonsense knowledge extracted from external sources. This model achieved over 90% accuracy in identifying fake news by leveraging entity disambiguation and relationship analysis. Tchechmedjiev et al. (2019) presented ClaimsKG that focused on compiling and organizing fact-checked assertions validated by trusted fact-checking organizations. This KG structure includes meta data such as the fact-checking verdict, claim source, and publication date, enabling the detection system to identify recurring false claims, map the origins of disinformation, and correlate validated claims across different domains. By providing a repository of verified information, ClaimsKG allows for effective querying and tracking of disinformation over time, enhancing the ability of models to cross-reference claims with established knowledge in real-time. Mayank et al. (2022) introduced DEAP-FAKED, another notable KG-based methodology for structuring claims with contextual metadata, including origins, sources, and supporting data. This model enables machine learning algorithms to identify patterns differentiating false claims from verified ones by evaluating source credibility and cross-referencing entities within the graph.

2.2 MACHINE LEARNING ON GRAPHS

Machine learning (ML) on graphs represents a sophisticated approach to leverage the relational and contextual nature of KGs in identifying patterns, insights, and anomalies, especially relevant for disinformation detection. Graph Neural Networks (GNNs) (Scarselli et al., 2008) are central to machine learning on graphs. Unlike traditional neural networks, which work on fixed inputs, GNNs are designed to interpret the dynamic, interconnected nature of graphs by learning from both the nodes and their edges.

In a disinformation detection context, GNNs facilitate analyzing information at multiple levels: entity-level (the node) and relationship-level (the edge). This enables the model to consider how entities are connected and influenced by each other, enhancing the ability to detect anomalies in the structure or content that may indicate potential disinformation. For example, GNNs can track patterns of consistent disinformation spread across interconnected sources, identifying nodes that frequently deviate from verified information pathways.

Funded by
the European Union

Graph Embeddings represent another crucial ML technique for KGs. Embeddings convert graph data—both nodes and edges—into continuous vector spaces, effectively translating complex, relational information into a format that simple ML models, such as Multi-layer Perceptrons (MLPs) and Convolutional Neural Networks (CNNs), can efficiently process. For instance, embedding techniques like node2vec and DeepWalk create low-dimensional representations that encapsulate the relative positions and relationships among nodes. Such embeddings allow ML models to compare nodes (entities) within the graph, facilitating quick identification of outliers or unusual relational patterns typical in disinformation. Through these methods, ML on graphs offers a layered, contextual approach to fight disinformation by enabling algorithms to recognize the relational and structural nuances within a KG. The synthesis of ML techniques and KGs not only improves model accuracy but also enhances interpretability, providing insights into how and why certain content is flagged (Zhou et al., 2020).

3 BUILDING PROCESS OF THE UNIMODAL KNOWLEDGE GRAPH

The task of constructing the unimodal knowledge graph under Task 6.3 represents a pivotal component of the AI4Debunk project, aiming to create an advanced framework capable of effectively modeling and analyzing existing information. The overarching goal was to develop a knowledge graph (KG) that could serve as a robust analytical tool for understanding the structure of disinformation (Manos et al., 2024). By leveraging an ontology-based approach, the task sought to provide a scalable and interoperable solution for the detection and contextualization of false information across diverse domains.

Figure 1 represents the logical flow of the framework applied in the AI4Debunk project combining Knowledge Graphs and ML to detect disinformation and fake news.

Funded by
the European Union

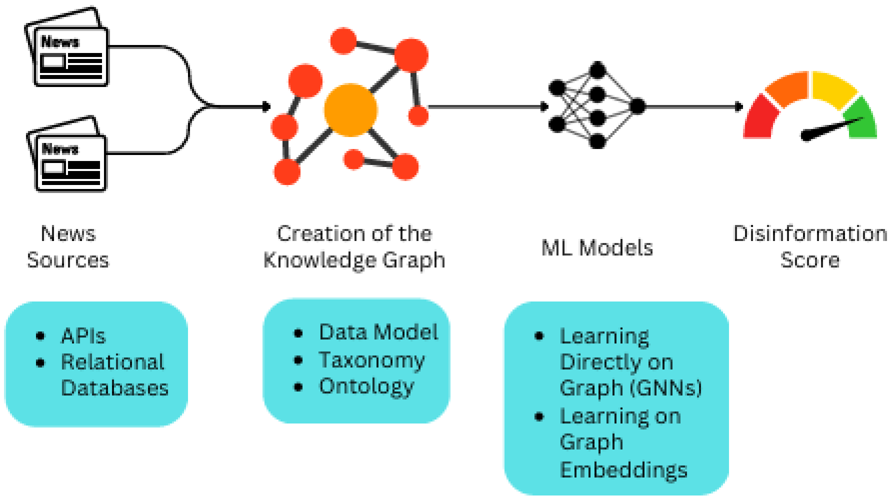

FIGURE 1. LOGICAL FLOW OF THE FRAMEWORK APPLIED IN THE AI4DEBUNK PROJECT COMBINING KNOWLEDGE GRAPHS AND ML

At the core of this framework there is the creation of the knowledge graph designed to offer a dynamic framework capable of encoding complex relationships and enhancing interpretability in detecting disinformation. This methodology differs from conventional architectures by integrating a robust, ontology-based taxonomy with an adaptive process for handling deceptive content in multimedia formats. This workflow encompasses structured steps that align graph construction with multimodal data processing, ultimately making the graph versatile for real world applications in disinformation detection.
Specifically, two different knowledge graphs have been prepared in the project, named the “unimodal” knowledge graph and the “multimodal” knowledge graph. The former is based on the fake statements and related multimedia contents collected in Task 6.1 and consists in extracting the textual description from multimedia contents and adding this textual knowledge in the knowledge graph. The latter integrate the multimodal features extracted from the fake statements and related multimedia contents in Task 6.2 and consists in embedding the multimedia contents and extracted multimodal features within the knowledge graph. This section describes the process followed for building the “unimodal” knowledge graph.

The first step of the building process consists in the development of a taxonomy subgraph, designed to encapsulate relevant concepts from two critical case studies: climate change and the Russian Invasion in Ukraine. This taxonomy serves as the backbone of the knowledge graph, enabling the structured representation of relationships and hierarchical categorizations central to these topics. To achieve this, the open Wikidata API was utilized, leveraging its rich repository of structured information. The plan also emphasized semantic alignment with Schema.org standards to ensure interoperability and facilitate clear interpretations of the graph’s structure.

Funded by the European Union

Following the creation of the taxonomy, the next step involved populating the graph with real-world examples of false statements and related multimedia content. These elements were sourced from datasets developed in Task 6.1, providing verified instances of disinformation. The integration process was tailored to preserve semantic consistency while ensuring that the graph could represent the novel relationships between deceptive content and its contextual entities.

The final stage of the task focused on establishing connections between the imported deceptive content and the taxonomy subgraph. Utilizing advanced natural language processing (NLP) and ML techniques, these connections were forged based on semantic and contextual relevance.

In the following sub-sections, a detailed description of each step is provided.

3.1 CREATION OF THE TAXONOMY SUBGRAPH

The taxonomy subgraph is the foundational layer of the unimodal knowledge graph, designed to capture and organize the hierarchical relationships and essential entities related to the two selected case studies: climate change and the Russian Invasion in Ukraine. This subgraph provides the structured framework necessary for integrating false statements and related multimedia contents in subsequent stages, enabling advanced analysis and disinformation detection. The creation process was driven by a systematic methodology that prioritized semantic consistency, scalability, and alignment with internationally recognized standards.

Scope and Objectives

The primary objective of the taxonomy subgraph was to establish a semantically rich and context-aware framework capable of representing the complex relationships inherent in the selected domains. To achieve this, it was critical to identify and extract relevant entities and their associated properties from a reliable and comprehensive knowledge source. Wikidata (www.wikidata.org), a collaborative and continuously updated knowledge base, was chosen as the primary data source due to its extensive repository of structured information and its alignment with the Resource Description Framework (RDF) standard.

The scope of the taxonomy subgraph was defined by the two case studies. For climate change, the focus was on capturing concepts such as environmental actions, scientific organizations, and policy frameworks. For the Russian Invasion in Ukraine, entities related to military events, geopolitical actors, and causal relationships were prioritized. This focused approach ensured that the taxonomy would remain relevant and practical for the specific objectives of the project.

Data Retrieval and Extraction

The data retrieval process began by identifying root entities in Wikidata that encapsulate the overarching themes of the case studies. For example, “Climate Change” (Q7942) and “Russian Invasion in Ukraine” (Q110999040) were selected as the root nodes. From these starting points, relevant properties such as subclass of (P279), instance of (P31), and has part (P527) were used to traverse the hierarchical relationships and extract additional entities. This approach ensured that the taxonomy captured both broad categorizations and detailed relationships.

A breadth-first search algorithm was implemented to traverse Wikidata, ensuring systematic exploration of entities while avoiding redundant visits. This traversal was optimized using asynchronous HTTP requests with the qwikidata library, which enabled efficient data fetching and reduced latency. Additionally, a retry mechanism with exponential backoff was employed to address potential network disruptions and API rate limits.

To control the size and complexity of the subgraph, a configurable traversal depth was set. This allowed the team to balance comprehensiveness with manageability, ensuring that the taxonomy remained both detailed and scalable. For example, while higher levels of depth were used to capture detailed sub-entities for scientific organizations in the climate change domain, a more focused depth was applied to military operations in the Russo-Ukrainian war to maintain relevance.

<u>Semantic Standardization and Mapping</u>

A key feature of the taxonomy subgraph is its semantic consistency, achieved through mapping extracted entities to Schema.org classes. This alignment ensures that the graph adheres to a widely recognized standard, enhancing its interoperability with other datasets and systems. For instance, entities such as “Climate Change” were mapped to SCHEMA.Thing, while “Environmental Actions” were aligned with SCHEMA.Action. Similarly, events within the Russo-Ukrainian war were categorized under SCHEMA.Event, providing a clear and interpretable structure.

This semantic standardization was critical for maintaining the integrity of the taxonomy, especially when integrating external data sources or expanding the graph in the future. The use of Schema.org also facilitated better alignment with tools and platforms that rely on this standard, ensuring compatibility across diverse applications.

<u>Graph Construction and Visualization</u>

The taxonomy subgraph was constructed using RDFLib, a Python library designed for creating and managing RDF-based graphs. This tool provided the structural backbone for storing and querying the interconnected data efficiently. The extracted entities and relationships were

Funded by the European Union

serialized in multiple formats, including Turtle and JSON-LD, to support both human-readable and machine-readable applications.

For visualization, Plotly was employed to generate intuitive graphical representations of the taxonomy. Nodes representing entities and edges denoting relationships were color-coded to distinguish between different types of entities and properties. For example, nodes corresponding to root concepts such as “Climate Change” and “Russian Invasion in Ukraine” were highlighted, while subordinate entities were represented with different colours to reflect their hierarchical position.

These visualizations played a critical role in validating the accuracy of the taxonomy and communicating its structure to stakeholders. They also serve as a useful tool for exploring the relationships captured within the graph, making the taxonomy accessible to both technical and non-technical audiences.

<u>Challenges and Mitigation Strategies</u>

Several challenges were encountered during the creation of the taxonomy subgraph, primarily related to data retrieval and semantic alignment. The vastness of Wikidata presented a challenge in ensuring the relevance and accuracy of extracted entities. To address this, a rigorous filtering process was implemented, focusing on high-confidence properties and relationships. Additionally, asynchronous data fetching and efficient error handling mechanisms ensured that the process remained resilient to network disruptions and API constraints.

Semantic standardization posed another challenge, particularly when mapping diverse entities to Schema.org classes. To mitigate this, the team developed a comprehensive mapping framework that considered contextual nuances and domain-specific requirements. Regular validation against sample datasets helped ensure the accuracy and consistency of the mappings.

<u>Outcome and Future Applications</u>

Funded by the European Union

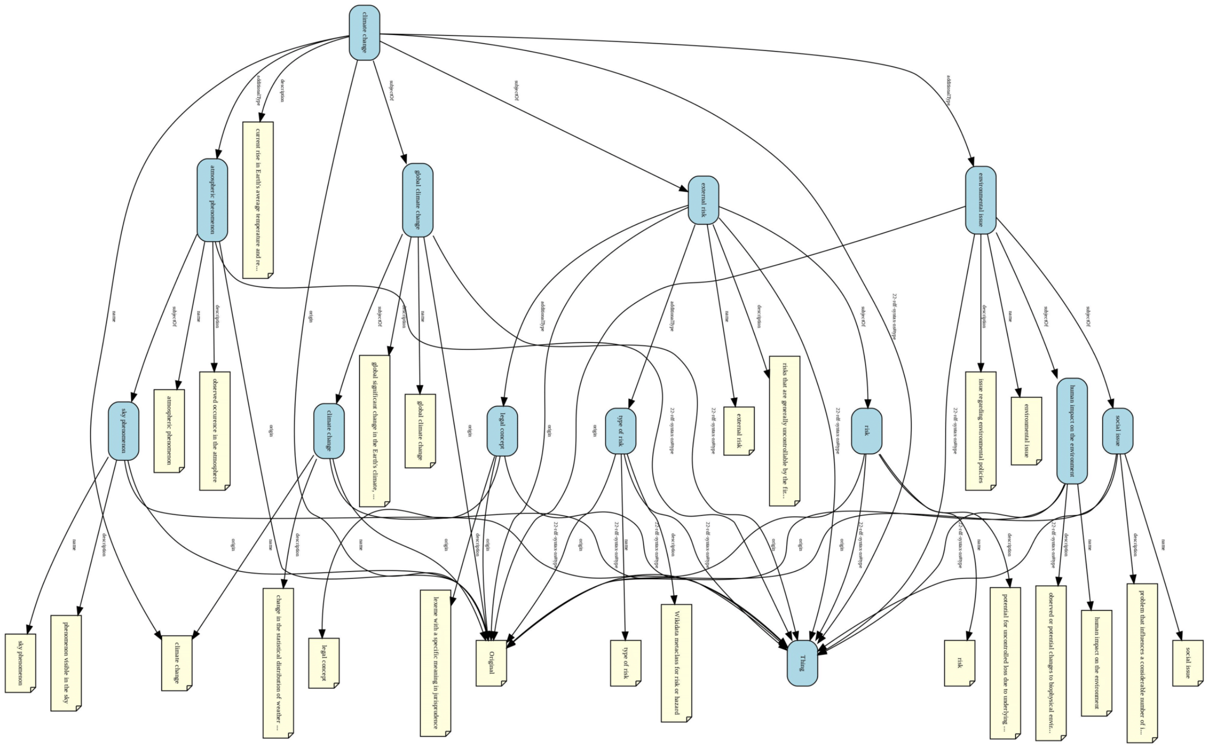

FIGURE 2. A SNAPSHOT OF THE TAXONOMY SUBGRAPH FOR "CLIMATE CHANGE" CASE STUDY

The completed taxonomy subgraph (Figure 2) provides a robust, semantically rich framework that forms the foundation of the unimodal knowledge graph. It encapsulates the hierarchical and relational complexities of the case studies, offering a clear and interpretable structure for integrating and analyzing disinformation data. By adhering to standards such as Schema.org and leveraging scalable technologies, the taxonomy is well-positioned for future expansions, including the integration of multimodal content.

3.2 IMPORT OF THE FALSE STATEMENTS AND THE RELATED MULTIMEDIA CONTENTS

The integration of false statements and multimedia content into the knowledge graph represents a critical step in the development of a comprehensive framework for detecting and analyzing disinformation. This stage focuses on populating the taxonomy subgraph with real-world examples of deceptive content, providing a dynamic and actionable resource for subsequent analysis. The process required careful planning to ensure the semantic and contextual alignment

Funded by the European Union

of the imported data with the existing taxonomy structure, enabling seamless integration and scalability.

Sourcing and Preparing the Data

The initial step in this process involved sourcing false statements and related multimedia content from reliable and open-access fact-checking platforms. These platforms were selected based on their credibility and coverage of misinformation relevant to the project's two case studies: climate change and the Russian invasion in Ukraine. Examples of false statements included fabricated narratives about climate policies and misinformation campaigns surrounding geopolitical events. The data collection effort emphasized both diversity and accuracy, ensuring that the imported content reflected a wide range of disinformation patterns.

For multimedia content, such as images, videos, and audio files, the project prioritized data that provided clear visual or auditory representations of false claims. To prepare this data for integration, each multimedia file was accompanied by metadata describing its source, context, and potential thematic connections. Textual descriptions were also extracted where necessary to facilitate semantic indexing and alignment with the taxonomy.

Integration Strategies

Two primary strategies were explored for incorporating the false statements and multimedia content into the knowledge graph:

1. Direct Embedding:
   This approach involved embedding multimedia files directly into graph nodes. While this method allows for easy retrieval of raw content, it was deemed resource-intensive and less scalable for the current phase of the project. Direct embedding was thus deprioritized for this iteration of the knowledge graph.
2. Textual Extraction and Semantic Indexing:
   This approach focused on extracting textual descriptions from multimedia elements and mapping them to relevant nodes within the taxonomy subgraph. By leveraging NLP techniques, key features such as keywords and thematic relevance were extracted from the text. These features provided the basis for linking the content to specific entities in the taxonomy. For example, a false statement about renewable energy policies was indexed under relevant concepts within the climate change taxonomy, such as "Renewable Energy" and "Policy Frameworks."

Funded by the European Union

The textual extraction method was selected for its efficiency, scalability, and alignment with the semantic structure of the taxonomy. This approach also ensures that the graph remains interoperable with external datasets and systems that rely on text-based representations.

Implementation and Tools

The integration process relied on advanced tools and methodologies to streamline data processing and maintain accuracy. Key components included:

- Natural Language Processing:
  NLP techniques were employed to preprocess and analyze the text accompanying the false statements and multimedia content. This involved tokenization, stopword removal, and stemming, ensuring that the extracted features were both relevant and precise.
- Feature Extraction:
  Semantic features such as keywords and concepts were extracted using libraries like nltk and sentence-transformers. These features were instrumental in establishing meaningful connections between the imported content and the taxonomy nodes.
- Graph Construction:
  The imported data was serialized into the knowledge graph using RDFLib. Each false statement and multimedia element was represented as a node, linked to relevant entities in the taxonomy based on its extracted features.

This structured approach allowed for the efficient integration of diverse content types, ensuring that the knowledge graph remained both comprehensive and coherent.

Challenges and Solutions

The process of importing false statements and multimedia content presented several challenges, particularly in maintaining the semantic and contextual integrity of the graph. One significant challenge was the variability in data quality and structure across different fact-checking platforms. To address this, a rigorous preprocessing pipeline was implemented, standardizing the data before integration.

Another challenge was the complexity of aligning multimedia content with taxonomy nodes. This was mitigated through the use of textual descriptions, which provided a common semantic framework for mapping multimedia elements to relevant concepts. Regular validation against the taxonomy ensured that these mappings were both accurate and contextually meaningful.

Outcome and Significance

Funded by
the European Union

The integration of false statements and multimedia content into the knowledge graph marks a substantial advancement in the project's ability to model and analyze disinformation. By systematically importing and indexing deceptive content, the graph now serves as a dynamic resource for understanding the relationships between false information and its broader contextual framework.

3.3 CONNECTION OF THE FALSE STATEMENTS AND THE RELATED CONTENTS TO THE TAXONOMY SUBGRAPH

The final step in constructing the unimodal knowledge graph involved linking the false statements and related multimedia content to the taxonomy subgraph. This process established the semantic and contextual relationships between deceptive content and the structured framework of concepts developed in the taxonomy. By creating these connections, the knowledge graph became a dynamic tool capable of encoding complex relationships and enabling advanced disinformation analysis. This section outlines the methodologies, tools, and challenges encountered in achieving these connections.

Establishing Connection Criteria

To ensure meaningful integration, connections between false statements and taxonomy entities were guided by clearly defined criteria based on semantic relevance and contextual alignment. Key features extracted from the false statements and multimedia content included:

- Keywords: Terms and phrases extracted from the text of false statements, reflecting their core themes and topics.
- Relevant Concepts: identified through NLP.
- Relationships: Implied or explicit causal, associative, or hierarchical relationships relevant to the case studies.

These features were compared against the entities and properties within the taxonomy subgraph to establish semantic and contextual matches. For example, a false statement about renewable energy policies was linked to taxonomy nodes like “Renewable Energy” (as an instance) and “Policy Frameworks” (as a broader category).

Methodology and Tools

The process of connecting nodes utilized a combination of NLP techniques, similarity metrics, and graph management tools to ensure precision and scalability.

1. Feature Embedding and Similarity Calculation:
   The first step involved converting textual data into numerical representations using pre-trained embedding models, such as sentence-transformers. This approach ensured that the semantic meaning of the text was captured in a high-dimensional vector space. To quantify the relationship between false statements and taxonomy nodes, cosine similarity was employed. Higher similarity scores indicated stronger semantic alignment, guiding the establishment of connections based on thematic relevance and keywords.
2. Node Matching and Connection:
   Each false statement and its associated multimedia content were evaluated against the taxonomy entities. Nodes were connected based on similarity scores, thematic relevance, and predefined thresholds to ensure precision. Connections were categorized as:
   a. Direct Links: Established when a strong semantic match was identified (e.g., a statement explicitly referencing climate change policies).
   b. Indirect Links: Created when the relationship was inferred based on shared features or thematic overlap (e.g., multimedia content related to climate protests linked to broader environmental actions).
3. Graph Construction and Visualization:
   The connections were implemented in the knowledge graph using RDFLib. Relationships were encoded as edges between nodes, with attributes describing the nature and strength of the connection. Visualization was achieved using Plotly, which rendered the graph with color-coded edges to reflect connection types and similarity scores. This enhanced interpretability, making it easier for stakeholders to explore the graph's structure.

Challenges and Mitigation Strategies

The process of connecting nodes presented several challenges, primarily related to the variability and ambiguity inherent in disinformation content. Key challenges included:

- Ambiguity in Content: False statements often lack clear contextual boundaries, making it difficult to identify definitive connections. To address this, multiple features (e.g., keywords and thematic relevance) were used in combination, increasing the robustness of the connections.
- Scalability: The large volume of data required efficient processing and storage. Asynchronous operations and optimized data structures ensured scalability, even with extensive datasets.

Funded by the European Union

- Alignment Across Domains: Ensuring consistent mapping between content from different case studies and the taxonomy required careful validation. Regular reviews and feedback loops with project partners helped maintain alignment.

Outcomes and Impact

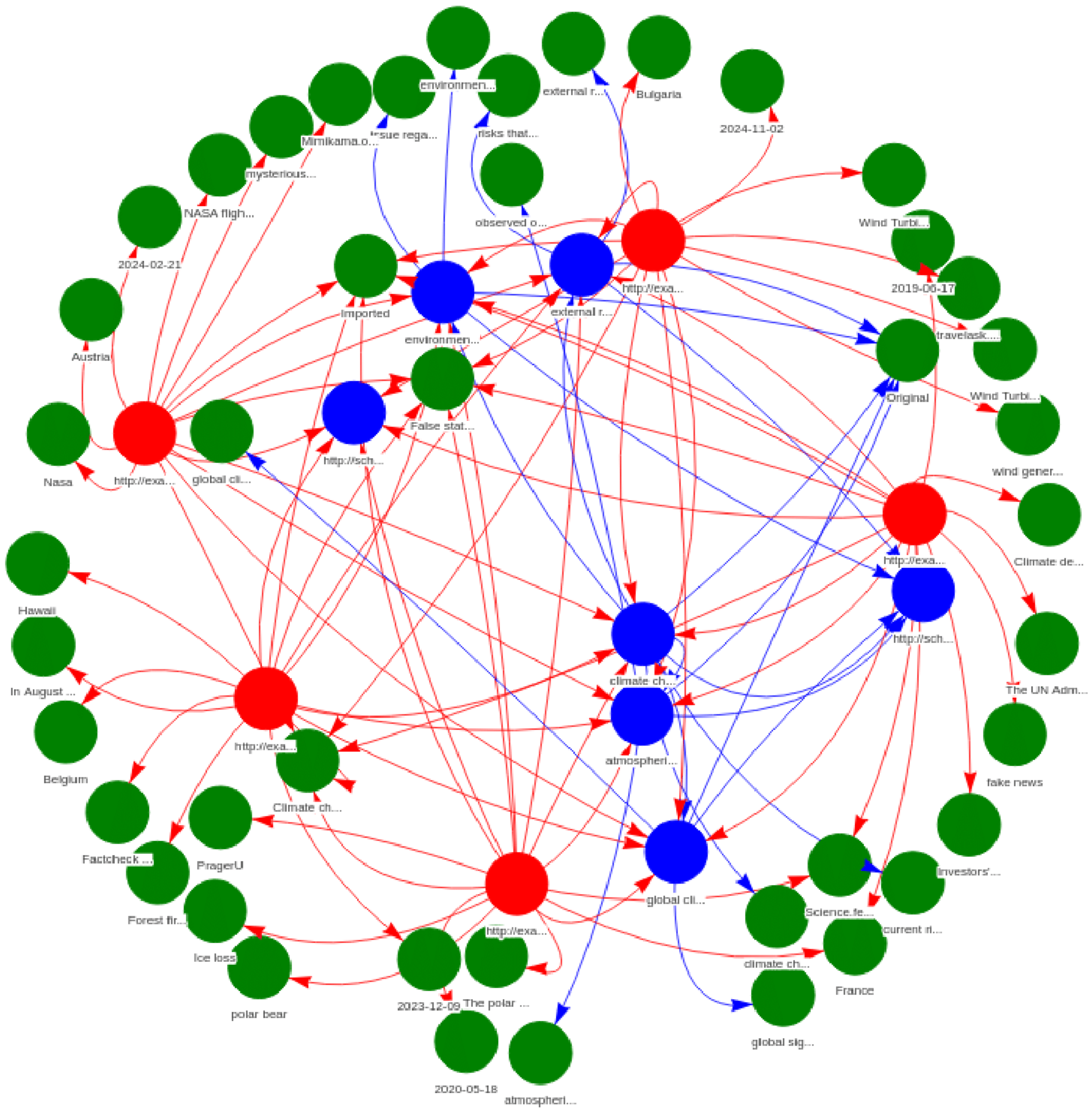

FIGURE 3. A SNAPSHOT OF THE KNOWLEDGE GRAPH FOR "CLIMATE CHANGE" CASE STUDY

The connections established between false statements, multimedia content, and the taxonomy subgraph significantly enhanced the utility of the knowledge graph. Figure 3 shows a snapshot of the knowledge graph for "Climate Change". These links provide a rich, context-aware framework for analyzing the relationships between deceptive content and its broader informational ecosystem. By encoding these connections, the graph supports advanced disinformation detection and contextualization, enabling applications such as:

Funded by the European Union

- Semantic Querying: The graph can be queried to identify patterns, relationships, and clusters of misinformation.
- Predictive Analysis: The structured connections lay the groundwork for applying machine learning models to predict the spread and impact of disinformation.

The completed unimodal knowledge graph, with its interconnected nodes and edges, is now a robust tool for addressing real-world challenges in disinformation detection. The current structure will also serve as the foundation for future enhancements, including the integration of multimodal data once feature extraction from Task 6.2 is completed.

Funded by
the European Union

4 BULDING PROCESS OF THE MULTIMODAL KNOWLEDGE GRAPH

The construction of the Multimodal Knowledge Graph is a systematic, multi-stage process designed to create a rich, interconnected data structure for disinformation analysis. The process begins by establishing a foundational, structured taxonomy of relevant concepts, which then serves as the semantic backbone for the graph. Into this taxonomy, unstructured and semi-structured data from real-world disinformation claims are ingested. This includes not only textual information but also metadata and pointers to associated multimedia content and their extracted features. Finally, semantic links are algorithmically forged between the ingested claims and the foundational taxonomy, creating a cohesive and query-able knowledge graph. This entire pipeline is automated by the provided Python script, ensuring reproducibility and scalability. The following subsections provide a detailed, step-by-step explanation of this building process.

4.1 CREATION OF THE TAXONOMY SUBGRAPH

The taxonomy subgraph is the foundational layer of the multimodal knowledge graph, designed to capture and organize the hierarchical relationships and essential entities related to the two selected case studies: climate change and the Russian Invasion in Ukraine. This subgraph provides the structured framework necessary for integrating false statements and related multimedia contents in subsequent stages, enabling advanced analysis and disinformation detection. The creation process, implemented in the build_knowledge_graph function, was driven by a systematic methodology that prioritized semantic consistency, scalability, and alignment with internationally recognized standards.

<u>Scope and Objectives</u>

The primary objective of the taxonomy subgraph was to establish a semantically rich and context-aware framework capable of representing the complex relationships inherent in the selected domains. To achieve this, it was critical to identify and extract relevant entities and their associated properties from a reliable and comprehensive knowledge source. Wikidata was chosen as the primary data source due to its extensive repository of structured information and its alignment with the Resource Description Framework (RDF) standard. The scope was defined by the two case studies, focusing on concepts like environmental actions for climate change and geopolitical actors for the Ukraine invasion.

Funded by
the European Union

Data Retrieval and Extraction

The data retrieval process began by identifying root entities in Wikidata that encapsulate the overarching themes of the case studies: “Climate Change” (**Q125928**) and “2022 Russian invasion of Ukraine” (**Q110999040**). From these starting points, a breadth-first search algorithm was implemented to traverse Wikidata, ensuring systematic exploration of entities. This traversal was optimized using asynchronous HTTP requests with the aiohttp and asyncio libraries, which enabled efficient, concurrent data fetching while a Semaphore respected API rate limits. To control the complexity of the subgraph, a configurable traversal depth was set, balancing comprehensiveness with manageability. For each entity, the qwikidata library parsed the fetched JSON data, extracting labels, descriptions, and relationships based on a predefined list of properties (e.g., P31 "instance of", P279 "subclass of").

Semantic Standardization and Mapping

A key feature of the taxonomy subgraph is its semantic consistency, achieved through mapping extracted entities to Schema.org classes via the map_wikidata_to_schema function. This alignment ensures that the graph adheres to a widely recognized standard, enhancing its interoperability. For instance, an entity identified as a "human" (Q5) in Wikidata is mapped to SCHEMA.Person. This standardization was critical for maintaining the integrity of the taxonomy. Furthermore, each entity sourced from Wikidata was marked with a custom property (MYNS.origin) set to "Original" to distinguish this foundational data from the disinformation claims added later.

Graph Construction

The taxonomy subgraph was constructed using rdflib, a Python library for managing RDF-based graphs. Extracted entities and relationships were added as nodes and edges (triples), respectively, forming a coherent graph structure. Each entity was assigned a standard URI based on its Wikidata ID. The resulting graph was then serialized in multiple formats, including Turtle and JSON-LD, to support both human-readable inspection and machine-readable applications, ensuring the taxonomy is portable and can be used by other systems.

4.2 IMPORT OF THE FALSE STATEMENTS AND THE RELATED MULTIMEDIA CONTENTS

Once the foundational taxonomy subgraph is established, the next stage involves populating the graph with real-world data. This is achieved by importing the disinformation claims collected in Task 6.1. This process, handled by adding_disinformation_to_graph function, transforms each row of the input CSV file into a rich, interconnected node within the knowledge graph, explicitly designed to accommodate multimodal information.

The import process is executed as follows:

1. **Data Ingestion:** The script begins by reading the dataset of false statements from a CSV file using a custom smart_read_csv utility function, which robustly handles different delimiters to load the data into a Pandas DataFrame.
2. **Node Creation for Each Claim:** The script iterates through each row of the DataFrame. For each row, which represents a single disinformation claim, a new entity (node) is created in the knowledge graph.
    a. A unique URI is generated for the statement (e.g., http://example.org/fake_statement_{id}).
    b. This new node is assigned the type SCHEMA.CreativeWork, identifying it as a piece of content.
    c. To differentiate these nodes from the foundational taxonomy, they are assigned the MYNS.origin property with the literal value "Imported".
3. **Mapping of All Available Metadata:** The core of the import process is a comprehensive mapping of every column in the CSV file to a specific property (predicate) in the graph. This is defined in the meta dictionary. This ensures that all collected data is preserved in a structured format. The mapping covers:
    a. **Core Information:** The text of the claim is mapped to SCHEMA.name, the URL to SCHEMA.url, and the publication date to SCHEMA.datePublished.
    b. **Attribution:** Information like the author, source, and country are mapped to SCHEMA.author, SCHEMA.publisher, and SCHEMA.countryOfOrigin, respectively.
    c. **Analysis and Classification:** Pre-existing analysis data, such as topics, keywords, and sentiment scores from the dataset, are mapped to properties like SCHEMA.about, SCHEMA.keywords, and MYNS.sentiment.
    d. **Multimodal Feature Pointers:** This is a critical step for creating a truly multimodal graph. The script explicitly maps columns that contain filenames or paths to multimedia analysis results. For example, the face_recognition_json column is mapped to the MYNS.faceRecognitionFile property, and audio_processing_json is mapped to MYNS.audioProcessingFile.
4. **Data Type Handling:** During the import, the script correctly formats the data by converting values to their appropriate RDF literal types. For example, dates are converted to the XSD.date type, and numeric IDs are converted to XSD.integer, ensuring data integrity and enabling type-specific queries.

By the end of this stage, the knowledge graph is no longer just an abstract taxonomy. It is populated with nodes representing real-world disinformation claims, with each node containing a rich set of textual, attributional, and, most importantly, multimodal metadata. The graph now serves as a central index, linking each textual claim to the file locations of its associated visual and auditory feature analyses.

4.3 CONNECTION OF THE FALSE STATEMENTS AND THE RELATED CONTENTS TO THE TAXONOMY SUBGRAPH

The final and most advanced step in the building process is to create meaningful, semantic connections between the newly imported disinformation claims (the "Imported" nodes) and the structured concepts within the foundational taxonomy (the "Original" nodes). This step transforms the graph from a collection of disconnected datasets into a truly integrated knowledge base where unstructured claims are contextually anchored to established entities. This process relies heavily on Natural Language Processing (NLP) to bridge the gap between text and structured data.

The connection process is orchestrated within the add_disinformation_to_graph function and involves the following sub-steps:

1. **Comprehensive Keyword Extraction:** Before attempting to link claims, the script first analyzes the textual content of all valid claims in the dataset. It uses the “extract_keywords”_comprehensive function, which employs the KeyBERT library. KeyBERT leverages the powerful SentenceTransformer model (paraphrase-multilingual-MiniLM-L12-v2) to identify keywords and keyphrases that are semantically representative of each claim's content, going beyond simple statistical frequency.
2. **Semantic Embedding and Similarity Calculation:** With the keywords extracted, the script then processes each disinformation claim one by one to find its place within the taxonomy. This is handled by the “match_keywords_to_wikidata” function.
    a. **Vectorization:** For a given claim, its extracted keywords are converted into high-dimensional numerical vectors (embeddings) using the same SentenceTransformer model. Concurrently, the names and descriptions of all entities in the Wikidata taxonomy cache are also converted into embeddings.
    b. **Similarity Scoring:** The script then calculates the **cosine similarity** between the vector for each keyword and the vectors for all entities in the taxonomy. This mathematical operation produces a score from -1 to 1 (or 0 to 1, depending on the embedding space) that quantifies the semantic closeness between the keyword and the taxonomic concept. A score near 1 indicates a very strong semantic match.

Funded by the European Union

3. **Creating the Link (Edge):** For each keyword, the script identifies the taxonomy entity with the highest similarity score.
    a. **Thresholding:** A connection is only established if this highest score exceeds a predefined threshold (set to 0.4 in the script). This threshold acts as a quality filter, preventing the creation of spurious links based on weak or ambiguous semantic relationships.
    b. **Adding the about Relationship:** If a match is confirmed, a new edge (triple) is added to the graph. This edge connects the disinformation claim node (the subject) to the matched taxonomy entity node (the object) using the SCHEMA.about property. For example, a false claim about melting ice caps might be linked via SCHEMA.about to the "Polar ice cap" entity in the taxonomy.

This final step completes the knowledge graph. The result is a powerful, hybrid structure where unstructured, real-world disinformation claims are not only richly described with their own metadata (including multimodal pointers) but are also intelligently linked to a verified, structured, and hierarchical ontology. This enables sophisticated queries, such as finding all false claims related to a specific organization, event, or scientific concept.

5 CONCLUSIONS

This deliverable introduces the meticulous process of creating and building two knowledge graphs—unimodal and multimodal—and capable of modelling, structuring, and analyzing disinformation interoperably and structured. The deliverable captures the methodological process used in establishing a semantic base that not only captures the richness of misleading material but also enables its analysis on text and multimedia features. The unimodal graph focuses on the text representation and extraction of false news, while the multimodal graph incorporates visuals and audio as inputs to construct a richer data model that better captures the multimodal nature of today's disinformation.

One of the central parts of this process was the creation of a taxonomy subgraph, built over a subset of Wikidata, containing concepts relating to the two project case studies: the war in Ukraine and climate change. The taxonomy was completed and augmented with false statements and related multimedia content found under Task 6.1. These entities were then semantically linked on the basis of elements such as topics, keywords, sentiment scores, and multimodal descriptors, which were collected using NLP and machine learning algorithms from Task 6.2. Employing scalable tools such as graph-handling APIs and the use of asynchronous data fetching and exception handling also contributed toward developing technical soundness for knowledge graph construction.

Funded by the European Union

The output of Task 6.3 is a milestone to the final creation of AI4Debunk's toolkit of disinformation detection tools, including the Disinfoscore metric developed in WP8 and AI-powered applications (e.g., Web plug-in, Disinfopedia, AR/VR portals) developed in WP10. The two knowledge graphs will give a semantic description of over 2000 disinformation false claims and enable sophisticated cross-referencing and examination of disinformation patterns. All this work conducted in this deliverable is finally a good mix of ontological modeling, artificial intelligence, and multimodal data analysis to produce a very meaningful resource for furthering transparency and trust in digital information systems.

As part of future work, the knowledge graphs will be validated using benchmark comparisons with current state-of-the-art models, for example, the ones discussed in Section 2.1. While theoretical assumptions and design choices make a robust starting point, empirical validation is required to demonstrate improved accuracy. For the next few months, different experiments will be carried out to validate the performance of the knowledge graphs and to fine-tune them as planned in Task 7.2.

Funded by
the European Union

REFERENCES

D’Ulizia, A., Caschera, M. C., Ferri, F., & Grifoni, P. (2021). Fake news detection: a survey of evaluation datasets. *PeerJ Computer Science*, *7*, e518.

Gao, X., Chen, W., Lu, L., Cui, Y., Dai, X., Dai, L., ... & Liu, H. (2023). An interpretable fake news detection method based on commonsense knowledge graph. *Applied Sciences*, *13*(11), 6680.

Manos, A., Filippidou, D. E., Pavlidis, N., Karanasios, G., Vachtanidis, G., D'Ulizia, A., & D'Andrea, A. (2024, December). Knowledge graphs and machine learning in fake news and disinformation detection. In *2024 International Conference on Engineering and Emerging Technologies (ICEET)* (pp. 1-6). IEEE.

Mayank, M., Sharma, S., & Sharma, R. (2022, November). DEAP-FAKED: Knowledge graph based approach for fake news detection. In *2022 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining (ASONAM)* (pp. 47-51). IEEE.

Scarselli, F., Gori, M., Tsoi, A. C., Hagenbuchner, M., & Monfardini, G. (2008). The graph neural network model. *IEEE transactions on neural networks*, *20*(1), 61-80.

Singhal A. et al., “Introducing the knowledge graph: things, not strings,” Official Google blog, vol. 5, no. 16, p. 3, 2012.

Taipalus, T. (2023). Database management system performance comparisons: A systematic literature review. *Journal of Systems and Software*, 111872.

Tchechmedjiev, A., Fafalios, P., Boland, K., Gasquet, M., Zloch, M., Zapilko, B., ... & Todorov, K. (2019, October). ClaimsKG: A knowledge graph of fact-checked claims. In *International Semantic Web Conference* (pp. 309-324). Cham: Springer International Publishing.

Zhang X, Ghorbani AA. (2019). An overview of online fake news: characterization, detection, and discussion. *Information Processing & Management* 57(2):102025

Zhou, J., Cui, G., Hu, S., Zhang, Z., Yang, C., Liu, Z., ... & Sun, M. (2020). Graph neural networks: A review of methods and applications. *AI open*, *1*, 57-81.

Funded by the European Union