KATHMANDU UNIVERSITY
SCHOOL OF SCIENCE
DEPARTMENT OF MATHEMATICS

A FINAL YEAR INTERNSHIP REPORT

[COMP - 408]

By

Gaurav Subedi (028319-20)

At

Center for Artificial Intelligence Research Nepal

SUPERVISOR:

Dr. Tek Raj Chhetri, PhD

June, 2025

ACKNOWLEDGEMENT

I express my sincere gratitude to everyone who supported me throughout my internship at the Center for Artificial Intelligence Research Nepal (CAIR-Nepal). I am especially thankful to my supervisor, Dr. Tek Raj Chhetri, PhD, for his invaluable guidance, continuous encouragement, and insightful feedback, which were instrumental in shaping the direction of my work.

This internship provided me with practical experience in the development of ontologies, construction of knowledge graphs using GraphDB, and validating the knowledge graph of COVID-19-related misinformation in Nepal. I would like to thank the entire team at CAIR-Nepal for creating a collaborative and intellectually stimulating environment that enhanced my learning.

I am also grateful to the Department of Mathematics at Kathmandu University for including this internship in the curriculum, providing a platform to apply theoretical knowledge in a real-world research setting. The opportunity to contribute to a meaningful project on misinformation has been a significant step forward in my academic and professional journey in computational mathematics and artificial intelligence.

APPROVAL

This is to certify that the internship report entitled

"From Rumors to Relations: Constructing Knowledge Graphs for Misinformation and Disinformation"

submitted by **Gaurav Subedi (028319-20)** in partial fulfillment of the requirements for the course **COMP-408 (Internship)** has been examined and approved.

Supervisor:

Dr. Tek Raj Chhetri, PhD
Center for Artificial Intelligence and Research Nepal (CAIR-Nepal)

____

(Signature)

Date: ____________________

SUMMARY

My internship at CAIR-Nepal as a Research Intern was a transformative experience, allowing me to apply computational mathematics to real-world challenges in analyzing COVID-19 misinformation in Nepal. I developed expertise in constructing knowledge graphs using GraphDB, leveraging LLMs for data collection, designing ontology to model misinformation patterns, and validating the graph using SHACL constraints. Working in a dynamic research environment, I honed technical skills in ontology development, data visualization, and validation, while collaborating with a team to address disinformation. This internship bridged theoretical learning with practical applications, enhancing my problem-solving, communication, and research skills, and preparing me for a career in AI and data science.

OVERVIEW OF THE COMPANY

The Center for Artificial Intelligence and Robotics (CAIR-Nepal) is a research-oriented organization established in early 2023 with a mission to advance multidisciplinary scientific research, particularly in the areas of artificial intelligence (AI) and emerging computing technologies. Headquartered in Nepal, CAIR-Nepal aims to foster technological innovation while promoting ethical, safe, and socially responsible applications of AI.

The organization was born out of an international collaboration following the team's success at Klimathon 2022 in Innsbruck, Austria. This formative experience, along with mentorship from the University of Innsbruck's startup incubator, served as the foundation for CAIR-Nepal's inception.

CAIR-Nepal is dedicated not only to cutting-edge research and development but also to technological literacy and public engagement. It provides educational opportunities for students and researchers, raises awareness among local communities and policymakers, and encourages the responsible use of AI technologies. Through collaboration with global experts and institutions, CAIR-Nepal seeks to create AI solutions that contribute to the betterment of society and uphold shared human values.

CONTENTS

1 Introduction

Misinformation and disinformation have emerged as significant threats to public discourse, social cohesion, and public health in the digital era. While misinformation refers to the unintentional spread of false information, disinformation involves the deliberate dissemination of falsehoods with the intent to deceive [18]. The proliferation of such content has intensified during crises, particularly during the COVID-19 pandemic and political elections, where misleading claims have led to widespread confusion, vaccine hesitancy, and mistrust in institutions [19].

The relevance of combating disinformation lies in its ability to influence behaviors, polarize societies, and undermine democratic processes. Conventional detection techniques, including fact-checking websites and social media moderation, often fall short due to the volume, velocity, and variability of the content. Hence, there is a pressing need for structured and semantically rich representations to capture the relationships and evolution of misinformation narratives [22].

This project explores the use of *knowledge graphs* as a powerful approach to model and analyze misinformation. A knowledge graph (KG) is a structured representation of facts and their interrelations, enabling machines to reason over entities and relationships. By encoding claims, their sources, topics, fact-checks, and associated metadata, KGs facilitate deep semantic understanding and querying of complex misinformation narratives [12].

This report presents the construction of a domain-specific knowledge graph focused on COVID-19 misinformation in Nepal. It covers the entire pipeline—from data collection and ontology design to semantic validation and analysis—demonstrating the value of knowledge graphs in structuring and analyzing disinformation in a low-resource and multilingual context.

1.1 Background

Misinformation and disinformation have gained renewed attention during global crises such as the COVID-19 pandemic. From false cures like "hot lemon water can cure COVID-19" to conspiracy theories linking 5G networks with virus transmission, the spread of such content has had tangible consequences [4]. In countries like Nepal, the challenge is exacerbated by language diversity, limited media literacy, and the rapid adoption of social media platforms [1].

The distinction between misinformation (unintentional) and disinformation (intentional) is critical, as it affects strategies for mitigation. While both disrupt public understanding, disinformation is often tied to organized campaigns or ideological motives. The consequences range from poor health outcomes to weakened trust in governance.

1.2 Motivation

Misinformation and disinformation have become increasingly sophisticated, exploiting the speed and scale of social media to mislead the public on critical issues such as public health, elections, and climate change. Traditional detection methods—ranging from manual fact-checking to machine learning-based classifiers—often operate at the surface level. They primarily focus on individual content items, treating each post or article as an isolated instance. While effective for classification, these approaches fall short in modeling the deeper semantic, temporal, and relational structures that underlie coordinated disinformation campaigns.

There is a growing need for structured, explainable, and interoperable representations that can capture the complex relationships between misinformation claims, their category, affected topics, and verification outcomes. Knowledge Graphs (KGs), with their ability to represent entities and their relationships in a semantically rich manner, offer a promising foundation to fill this gap. They allow for multi-hop reasoning, entity linking, temporal tracking, and integration of heterogeneous data sources, which are essential for modeling the lifecycle and propagation of false information.

Moreover, most existing tools and datasets for misinformation detection are built around English-language content, limiting their applicability in multilingual or low-resource settings such as Nepal. KGs can help mitigate this challenge by providing a language-agnostic structure for information integration and analysis, which is particularly beneficial when dealing with mixed-language content, code-switching, or region-specific narratives.

Efforts such as the Coronavirus Infectious Disease Ontology (CIDO) [11] have demonstrated the utility of ontologies and KGs in integrating domain knowledge and enabling structured reasoning in pandemic-related biomedical contexts. Inspired by such applications, this project aims to leverage a KG-based approach to systematically capture, analyze, and reason over misinformation narratives, focusing on the COVID-19 infodemic in Nepal.

1.3 Problem Statement

The key challenge in misinformation analysis lies in identifying and modeling subtle and evolving relationships among claims, category, and narratives. Current systems struggle with:

- Handling linguistic ambiguity and lack of standard terminology.
- Integrating data from diverse and noisy sources.
- Providing explainable insights and patterns from large misinformation corpora.

This project frames the problem as a semantic modeling and data integration task, where the goal is to design an ontology and populate a knowledge graph that semantically links different facets of disinformation.

1.4 Objectives

The objectives of this project are as follows:

- Design a domain-specific ontology to model COVID-19 misinformation relevant to the Nepali context.
- Construct a knowledge graph capturing claims, their variants, category, and fact-checks.
- Enable semantic querying, reasoning, and visualization to explore misinformation narratives.

2 Literature Review

This section provides a comprehensive survey of existing research and tools relevant to the construction of a knowledge graph for modeling COVID-19 misinformation in Nepal. It covers definitions and typologies of misinformation and disinformation, detection and analysis techniques, the role of knowledge graphs in information ecosystems, named entity and relation extraction methods, gaps in current approaches, and the justification for this internship project.

The review establishes the theoretical and technical foundation for the project, highlighting its relevance in addressing disinformation in a low-resource, multilingual context.

2.1 Misinformation and Disinformation: Definitions and Typologies

Misinformation and disinformation are critical concepts in the study of false information. Misinformation refers to false or inaccurate information shared without malicious intent, often due to misunderstanding or incomplete knowledge. Disinformation, in contrast, involves the deliberate creation and dissemination of false information to deceive or manipulate [7, 18]. These definitions, rooted in academic literature, are echoed in media discourse, where misinformation is often associated with errors or rumors, while disinformation is linked to propaganda or strategic deception [19].

Wardle and Derakhshan [18] propose a typology of information disorder, identifying seven categories: satire/parody, misleading content, imposter content, fabricated content, false connection, false context, and manipulated content. Satire/parody, while not always harmful, can be misinterpreted as factual. Misleading content distorts facts without outright fabrication, such as exaggerated health claims like "hot lemon water cures COVID-19." Imposter content involves impersonating credible sources, while fabricated content is entirely false, such as conspiracy theories linking 5G to COVID-19. False connection and false context misalign headlines or visuals with content, and manipulated content alters media to mislead [18]. This typology is particularly relevant to the COVID-19 infodemic, where hoaxes (e.g., fake cures) and propaganda (e.g., anti-vaccine narratives) proliferated, especially in regions like Nepal with high social media penetration [1].

Other typologies emphasize intent and impact. For instance, Tandoc et al. [16] classify false information based on levels of facticity and deception, ranging from satire (low deception, high facticity) to fabrication (high deception, low facticity). These frameworks guide the categorization of misinformation in the project's ontology, enabling structured modeling of claims by type and intent.

2.2 Detection and Analysis Techniques

The detection and analysis of misinformation have advanced significantly with the application of machine learning (ML) and natural language processing (NLP). Supervised ML models, trained on labeled datasets like LIAR [17], FakeNewsNet [14], and PHEME [23], use features such as linguistic cues, source credibility, and network propagation patterns to classify claims as true or false. For example, the LIAR dataset, with 12,800 political statements, supports models that predict veracity based on speaker metadata and textual content [17]. Unsupervised approaches, such as clustering or anomaly detection, are used when labeled data is scarce, identifying outliers in social media posts based on sentiment or engagement metrics [22].

Transformer-based models, such as BERT and RoBERTa, have achieved state-of-the-art performance in claim detection and verification [5]. These models leverage contextual embeddings to capture nuanced linguistic patterns, outperforming traditional feature-based classifiers. For instance, Hanselowski et al. [9] used BERT for the Fact-Checking Network (FNC-1) challenge, achieving high accuracy in stance detection (e.g., whether a headline agrees or disagrees with a claim). Fact-checking APIs, such as ClaimBuster [10], PolitiFact, and Snopes, provide structured access to verified claims, enabling automated integration with detection systems. Stance detection models, which assess whether a text supports, refutes, or is neutral toward a claim, further enhance verification by analyzing user-generated content [23].

Despite these advances, most detection systems focus on English-language content and operate at the surface level, classifying individual posts without modeling their relational or temporal dynamics. This limitation is particularly pronounced in multilingual settings like Nepal, where diverse languages and cultural contexts complicate detection [1].

2.3 Knowledge Graphs in Information Ecosystems

Knowledge graphs (KGs) are structured representations of entities and their relationships, typically encoded as RDF triples or property graphs, enabling semantic querying and reasoning [12]. In information ecosystems, KGs have been applied across domains like health, news, and policy to integrate heterogeneous data and support complex analysis. In public health, the Coronavirus Infectious Disease Ontology (CIDO) [11] standardizes concepts like therapeutic agents and transmission processes, facilitating integration of COVID-19-related data. Similarly, the COVID-19 Vocabulary (COVoc) [6] provides standardized terms for geographic and medical entities, enhancing interoperability.

In media and disinformation research, KGs have modeled fake news diffusion and claim propagation. For example, Mayank et al. [13] used a graph-based approach to track deepfake dissemination, linking sources, platforms, and fact-checks. These efforts demonstrate KGs' ability to represent complex narratives, such as the spread of anti-vaccine conspiracies during the COVID-19 pandemic.

However, few KGs focus on localized disinformation, particularly in low-resource settings. Existing systems often prioritize global or English-centric datasets, limiting their applicability in multilingual contexts like Nepal. This gap motivates the development of a region-specific KG for COVID-19 misinformation, as pursued in this project.

2.4 Named Entity and Relation Extraction Approaches

Named Entity Recognition (NER) and relation extraction are critical for populating knowledge graphs. NER identifies entities like people, organizations, or locations in text, while relation extraction determines their connections (e.g., "mentions," "targets"). Transformer-based models, such as BERT-NER and Flair, achieve high accuracy in NER by leveraging contextual embeddings [2]. For example, Flair's stacked embeddings combine contextual and static word representations, improving performance on noisy social media text. In the COVID-19 context, NER can identify entities like "Patan Hospital" or "Hot Lemon Water" from misinformation claims.

Relation extraction methods range from rule-based to neural approaches. OpenIE (Open Information Extraction) uses linguistic patterns to extract relations without predefined schemas, suitable for unstructured text [15]. However, it struggles with domain-specific relations like "factCheckedBy" or "circulatesIn." Neural models, trained on datasets like TACRED [20], predict fine-grained relations using dependency parsing and contextual embeddings. Recent advances, such as those by Zhao et al. [21], integrate graph neural networks to capture multi-hop dependencies, improving relation extraction in complex narratives.

These methods, while powerful, are often trained on English-centric corpora, limiting their effectiveness in multilingual settings. Adapting them to Nepali or mixed-language content requires custom preprocessing and translation, as implemented in this project.

2.5 Gaps in Existing Approaches

Despite significant progress, several gaps persist in misinformation detection and modeling:

- **Lack of Explainability**: Transformer-based models, while accurate, often function as black boxes, offering limited insight into why a claim is flagged as false [3]. This opacity hinders trust and adoption in critical applications like public health.
- **Limited Semantic Linkage**: Most detection systems treat claims as isolated instances, ignoring their relational context (e.g., how a claim connects to sources, locations, or fact-checks). This limits their ability to model coordinated disinformation campaigns [22].
- **Lack of Multi-hop Reasoning**: Analyzing misinformation often requires reasoning across multiple entities (e.g., a claim's source, its target demographic, and its fact-check status). Current models rarely support such complex queries, which KGs can address [12].
- **Language Coverage**: Datasets and tools predominantly focus on English, with limited support for low-resource languages like Nepali, Maithili, or Bhojpuri. This restricts their applicability in diverse linguistic contexts [1].
- **Cultural Context**: Existing systems often overlook cultural nuances, such as Nepal's reliance on traditional remedies or religious beliefs, which shape misinformation narratives.

These gaps highlight the need for a semantically rich, explainable, and multilingual approach to misinformation modeling, particularly in underrepresented regions like Nepal.

2.6 Justification for the Internship Project

This internship project addresses the identified gaps by constructing a domain-specific knowledge graph for COVID-19 misinformation in Nepal, focusing on semantic linkage, explainability, and cultural relevance. Unlike traditional ML models that classify claims in isolation, the project's KG captures relationships between claims, sources, categories, and fact-checks, enabling multi-hop reasoning via SPARQL queries (Section 6.3). By integrating with ontologies like CIDO and COVoc [6, 11], the graph ensures interoperability with global knowledge bases, facilitating cross-domain analysis.

The project's focus on Nepal fills a critical gap in localized disinformation research. Most existing datasets (e.g., LIAR, FakeNewsNet) are English-centric, neglecting multilingual and low-resource settings [14, 17]. By curating 91 claims, primarily in Nepali and English, and addressing cultural narratives (e.g., Pashupatinath water, cow urine), the project provides a context-aware framework for misinformation analysis. The use of SHACL for validation enhances explainability, ensuring data integrity and transparency, while the graph's structure supports applications like policy advisory and media monitoring (Section 9).

This work not only meets the academic requirements of the B.Sc. in Computational Mathematics course (COMP 408) but also contributes a reusable, scalable tool for combating disinformation, with potential to extend to other domains and regions. By addressing semantic, linguistic, and cultural gaps, the project lays the foundation for impactful research and practical interventions in Nepal's misinformation ecosystem.

3 Preliminaries

This section provides an overview of the core concepts underpinning this project: ontologies and knowledge graphs. These concepts form the foundation for modeling and analyzing COVID-19 misinformation in a structured and semantically rich manner.

3.1 Ontologies

An ontology is a formal representation of knowledge within a specific domain, defining concepts, entities, and their relationships in a structured and machine-readable format [8]. It consists of classes (e.g., `MisinformationClaim`, `Institution`), properties (e.g., `belongs_to_category`, `located_in`), and instances that instantiate these classes. Ontologies enable semantic interoperability by standardizing terminology and relationships, allowing systems to reason over data consistently. In this project, the `misinfo:` ontology was designed to model COVID-19 misinformation in Nepal, integrating with external ontologies like the Coronavirus Infectious Disease Ontology (CIDO) [11] and COVoc [6] to ensure compatibility with global standards.

Ontologies are typically encoded using standards like the Resource Description Framework (RDF) and Web Ontology Language (OWL). RDF represents knowledge as triples (subject, predicate, object), such as <`Claim_2, references_cure, HotLemonWater`>, while OWL extends RDF with advanced reasoning capabilities, such as class hierarchies and constraints. The use of ontologies in this project enables precise modeling of misinformation claims, their sources, and their cultural context, facilitating complex queries and analysis.

3.2 Knowledge Graphs

A knowledge graph (KG) is a structured representation of real-world facts, where entities (nodes) are connected by relationships (edges) to form a graph [12]. In the context of this project, the KG encodes misinformation claims, their categories, sources, locations, and fact-checks as RDF triples, stored in a semantic triple store like GraphDB. For example, the claim "Hot Lemon Water Cures COVID-19" is represented as a node connected to nodes like `Cure_HotLemonWater`, `Location_Nepal`, and `Category_FakeCures` via relations like `references_cure` and `belongs_to_category`.

KGs enable semantic querying and multi-hop reasoning, allowing users to explore complex relationships, such as tracing a claim's origin or its impact on specific demographics. They are particularly valuable for misinformation analysis, as they capture the interconnected nature of narratives, unlike traditional databases that treat data as isolated records. In this project, the KG supports SPARQL queries (e.g., Listing 1) to uncover patterns, such as the prevalence of social media as a misinformation source or the clustering of claims around institutions like Patan Hospital.

By combining ontologies and knowledge graphs, this project creates a robust framework for modeling and analyzing misinformation, bridging computational techniques with domain-specific knowledge to address the challenges of misinformation/disinformation in Nepal.

4 Methodology

This section details the systematic approach undertaken to construct the COVID-19 misinformation knowledge graph for Nepal. The methodology ensures reproducibility by clearly describing data sourcing, preprocessing, entity and relation extraction, ontology design, and the overall graph construction pipeline.

4.1 Data Collection

The dataset of 91 misinformation claims primarily from Nepal (75 claims), and Kathmandu (14 claims), was collected mainly through manual methods, with APIs, web crawling, and archives as supplementary approaches.

- **Manual Collection with LLM Assistance**: Most data was manually gathered from the internet, with LLMs aiding in identifying and categorizing claims. LLMs assisted in translating Nepali content and verifying cultural accuracy.
- **Web Crawling**: Scrapy scraped news and fact-checking sites (e.g., Nepal Fact Check) for claims.

Manual collection, enhanced by LLMs, ensured cultural and linguistic accuracy for Nepali claims.

4.2 Preprocessing

The raw data, consisting of multilingual text from social media, news, and fact-checking sources, underwent a preprocessing pipeline to ensure consistency and suitability for entity and relation extraction. The pipeline addressed noise, linguistic diversity, and formatting issues:

- **Text Cleaning**: HTML tags, emojis, URLs, non-standard Unicode characters, and excessive whitespace were removed using regular expressions in Python. For example, tweets with embedded links or emojis were stripped to focus on textual content.
- **Normalization**: Text was standardized through lowercasing, punctuation unification (e.g., converting multiple punctuation marks to single periods), and expansion of contractions (e.g., "can't" to "cannot"). Nepali text was transliterated to Roman script when necessary to facilitate processing.
- **Translation**: Nepali-language claims were translated into English using LLMs (e.g., a fine-tuned transformer model), with human oversight to preserve semantic fidelity and cultural nuances, such as idiomatic expressions like "lemon pani" (hot lemon water).
- **Deduplication and Noise Removal**: Duplicate or near-duplicate claims (e.g., variations of "Hot Lemon Water cures COVID-19") were identified. Incomplete or irrelevant entries (e.g., non-COVID-related claims) were filtered out, reducing noise and ensuring a clean corpus of 91 claims.

This preprocessing pipeline produced a unified, English-based dataset suitable for downstream semantic extraction while retaining the cultural context of Nepali claims.

4.3 Entity Extraction

Entity extraction identified key named entities and concepts within the preprocessed claims, forming the nodes of the knowledge graph. This process was performed manually with assistance from a Large Language Model (LLM) to ensure accuracy and cultural relevance, particularly for Nepal-specific entities.

- **Manual Extraction with LLM Assistance**: Each of the 91 preprocessed claims was manually reviewed to identify entities, such as persons, organizations, locations, cures, and general concepts. An LLM was used to suggest potential entities by analyzing claim text for named entities (e.g., "Patan Hospital," "Kathmandu") and domain-specific terms (e.g., "Hot Lemon Water" as a cure). The LLM was prompted with instructions like: "Extract named entities (persons, organizations, locations) and COVID-19-specific concepts (cures, vaccines) from the text." Human annotators validated and refined LLM suggestions to correct errors (e.g., distinguishing "WHO" as the World Health Organization vs. a pronoun) and ensure cultural accuracy (e.g., recognizing "Pashupatinath" as a location).
- **Entity Types**: The extracted entities were categorized based on the ontology:
  - **Misinformation Claims**: 91 unique claims (e.g., "Vaccines Make You Magnetic").
  - **Institutions**: 4 entities, including Patan Hospital, WHO, Nepal Army, United Nations.
  - **Locations**: 2 entities, including Nepal and Kathmandu.
  - **Cures**: 8 entities, such as Hot Lemon Water, Cow Urine, Chloroquine.
  - **General Entities**: 8 concepts, including Vaccine, 5G, PCR Test.
  - **Demographics**: 5 groups, such as Children, Elderly, Dalit Community.
  - **Religious Groups**: 2 groups, Hindu Community, Muslim Community.
  - **Sources**: 3 types, Social Media, Traditional Beliefs, Word of Mouth.
- **Process**: For each claim, the LLM generated a list of candidate entities, which were manually reviewed and categorized. The process ensured 131 unique entities, verified for accuracy by cross-referencing fact-check metadata (e.g., Nepal Fact Check reports).

This manual, LLM-assisted approach prioritized precision and context-awareness, resulting in a comprehensive set of entities for the knowledge graph.

4.4 Relation Extraction

Relation extraction identified semantic connections between entities, forming the edges of the knowledge graph. This process was conducted manually with assistance from a large language model (LLM) to capture complex and context-dependent relationships within misinformation narratives.

- **Manual Extraction with LLM Assistance**: Relations were manually identified by analyzing the context of each claim and its associated entities. The LLM was prompted to suggest relationships based on the claim text, using instructions such as: "Identify relationships between entities in the text, such as 'mentions,' 'located_in,' or 'references_cure."' For example, given the claim "Vaccines are rejected by WHO in Kathmandu," the LLM suggested relations like `mentions_institution(Vaccine, WHO)` and located_in(Claim, Kathmandu). Human annotators reviewed, validated, and refined these suggestions to ensure accuracy, especially for domain-specific relations.

- **Relation Types**: The ontology defined eight relation types to model key aspects of misinformation claims:
  - belongs_to_category: Links claims to misinformation categories (e.g., Conspiracy, Fake Cures).
  - originates_from: Connects claims to their sources (e.g., Social Media).
  - located_in: Associates claims with geographic locations (e.g., Kathmandu).
  - references_entity: Links claims to general entities (e.g., Vaccine).
  - references_cure: Connects claims to alleged cures (e.g., Hot Lemon Water).
  - mentions_institution: Associates claims with institutions (e.g., WHO).
  - targets_demographic: Links claims to demographic groups (e.g., Elderly).
  - targets_religious_group: Connects claims to religious communities (e.g., Hindu Community).
- **Process**: Each claim was processed using the LLM to propose candidate relations. Human annotators ensured that each proposed relation aligned with the ontology and reflected evidence from the fact-checking dataset. This process resulted in 251 curated relations, including 91 instances each of belongs_to_category, originates_from, and located_in.

This manual, LLM-assisted approach balanced linguistic nuance and domain expertise, ensuring the resulting graph was both culturally grounded and semantically accurate.

4.5 Schema/Ontology Design

The ontology was developed specifically to model COVID-19 misinformation claims in the context of Nepal. However, rather than designing entirely in isolation, the model strategically incorporates elements from existing domain ontologies to ensure semantic interoperability and alignment with established biomedical standards.

The ontology is defined under the custom namespace misinfo:, with core classes like MisinformationClaim, Institution, Cure, Entity, and Location. Despite its custom design, the following external vocabularies and ontologies were integrated:

- **CIDO (Coronavirus Infectious Disease Ontology)**: This ontology was reused for standardizing concepts such as therapeutic agents (cido:TherapeuticAgent), transmission processes, and diagnostic procedures. For example, misinformation categories like FakeCures, TransmissionMisinfo, and TestingMisinfo were annotated with corresponding CIDO classes [11].
- **COVoc**: The ontology also references covoc:GeographicLocation from the COVID-19 Vocabulary (COVoc) to standardize the representation of geographic locations mentioned in claims [6].

This hybrid design—custom but semantically enriched—allowed for domain-specific flexibility while retaining compatibility with broader semantic web standards, facilitating better integration, reuse, and validation in knowledge graph construction workflows.

4.6 Ontology Requirements

Defining clear requirements was essential to ensure the ontology effectively supported the modeling and analysis of COVID-19 misinformation in Nepal. The requirements were established to address the project's objectives, the specific characteristics of the misinformation domain, and the need for interoperability with existing standards. The following key requirements guided the ontology design:

- **Domain Coverage**: The ontology must comprehensively represent entities and relationships specific to COVID-19 misinformation in Nepal, including claims, sources (e.g., social media, traditional beliefs), locations (e.g., Nepal, Kathmandu), institutions (e.g., WHO, Patan Hospital), cures (e.g., Hot Lemon Water), and demographics (e.g., Elderly, Hindu Community). It should capture cultural and linguistic nuances, such as traditional remedies or region-specific narratives.
- **Semantic Interoperability**: To enable integration with global knowledge bases, the ontology must align with existing standards like CIDO [11] for biomedical concepts and COVoc [6] for geographic and COVID-19-specific terms. This ensures compatibility with external datasets and facilitates cross-domain reasoning.
- **Flexibility and Extensibility**: The ontology should be modular to accommodate future expansions, such as additional misinformation categories, languages, or domains (e.g., political disinformation). This requires a structure that supports easy addition of new classes, properties, or instances without disrupting existing relationships.
- **Queryability**: The ontology must support complex SPARQL queries to enable multi-hop reasoning, such as tracing a claim's source, its geographic context, or its fact-check status. This requires well-defined classes and properties with clear semantics.
- **Validation Support**: The ontology must be compatible with SHACL constraints to enforce data integrity, ensuring that all triples adhere to defined schemas.
- **Cultural and Contextual Relevance**: Given Nepal's multilingual and multicultural context, the ontology must accommodate entities and relations that reflect local practices, such as traditional remedies or religious narratives (e.g., Pashupatinath water), while maintaining generalizability for broader applications.

These requirements ensured that the ontology was both fit-for-purpose for the Nepali misinformation context and aligned with semantic web standards, enabling robust knowledge graph construction and analysis.

4.7 Ontology Modeling

Ontology modeling involved creating a structured representation of the misinformation domain, defining classes, properties, and their relationships to capture the semantics of COVID-19 misinformation in Nepal. The modeling process followed a top-down approach, starting with high-level concepts and refining them into specific classes and properties.

- **Class Hierarchy**: The ontology was structured around core classes under the `misinfo:` namespace, including:
  - `MisinformationClaim`: Represents individual claims (e.g., "Hot Lemon Water Cures COVID-19").

- `Misinfocategory`: Represents the category of the misinformation of the claim.
- `Institution`: Includes organizations like WHO and Patan Hospital.
- `Location`: Represents geographic entities, such as Nepal and Kathmandu, linked to `covoc:GeographicLocation`.
- `Cure`: Captures alleged treatments, such as Hot Lemon Water or Cow Urine, linked to `cido:TherapeuticAgent`.
- `Entity`: General concepts like Vaccine or 5G.
- `Source`, `Demographic`, `ReligiousGroup`: Represent sources (e.g., Social Media), target groups (e.g., Elderly), and religious communities (e.g., Hindu Community).

Classes were organized hierarchically, with `MisinformationClaim` as the central class, linked to others via object properties.

- **Object Properties**: Eight object properties were defined to model relationships, as detailed in Section 4.4 (e.g., `belongs_to_category`, `references_cure`, `located_in`). These properties were designed to support multi-hop reasoning, such as linking a claim to its source and location.
- **Data Properties**: Properties like `misinfo:claim_id` (unique identifier), `misinfo:title` (claim text), and `misinfo:fact_checked_truth` (fact-check result) were included to annotate instances with metadata.
- **Integration with External Ontologies**: To ensure interoperability, classes and properties were mapped to external ontologies. For example, `misinfo:Cure` was linked to `cido:TherapeuticAgent` via `skos:related`, and `misinfo:Location` was aligned with `covoc:GeographicLocation`. This mapping enhanced the ontology's compatibility with global standards.
- **Modeling Tools**: The ontology was modeled using Protégé, a widely-used ontology editor, to define classes, properties, and constraints visually. RDFLib was used to serialize the ontology in Turtle format for integration into the knowledge graph.

The resulting model, visualized in Figure 1, provided a semantically rich framework for representing misinformation claims and their relationships, tailored to the Nepali context while maintaining global interoperability.

4.8 Ontology Development

Ontology development involved the practical implementation of the modeled ontology, translating the conceptual design into a machine-readable format suitable for knowledge graph construction. This process was iterative, involving prototyping, testing, and refinement to ensure the ontology met the defined requirements.

- **Implementation in OWL**: The ontology was formalized using the Web Ontology Language (OWL) under the `misinfo:` namespace. OWL was chosen for its support for reasoning and interoperability with semantic web standards. The ontology was encoded using RDFLib in Python, generating Turtle files for classes, properties, and constraints.
- **Constraint Definition**: SHACL shapes were developed to enforce data integrity. For instance, a SHACL shape for `misinfo:MisinformationClaim` required exactly one `belongs_to_category` property and a unique `claim_id` matching a specific pattern.

These shapes were implemented using pySHACL and tested against sample triples.

- **Testing and Validation**: The ontology was tested by mapping the 91 claims and their 131 entities and 251 relations to the defined classes and properties. Validation with pySHACL ensured all triples conformed to the ontology, with no violations reported in the final `validation_report.ttl`.
- **Integration with GraphDB**: The ontology was imported into GraphDB alongside the RDF triples, enabling RDFS and OWL reasoning. This allowed inference of implicit relationships, such as subclass hierarchies, enhancing query expressiveness.

The development process resulted in a robust, validated ontology that effectively supported the construction of the knowledge graph, enabling semantic querying and analysis of misinformation in Nepal.

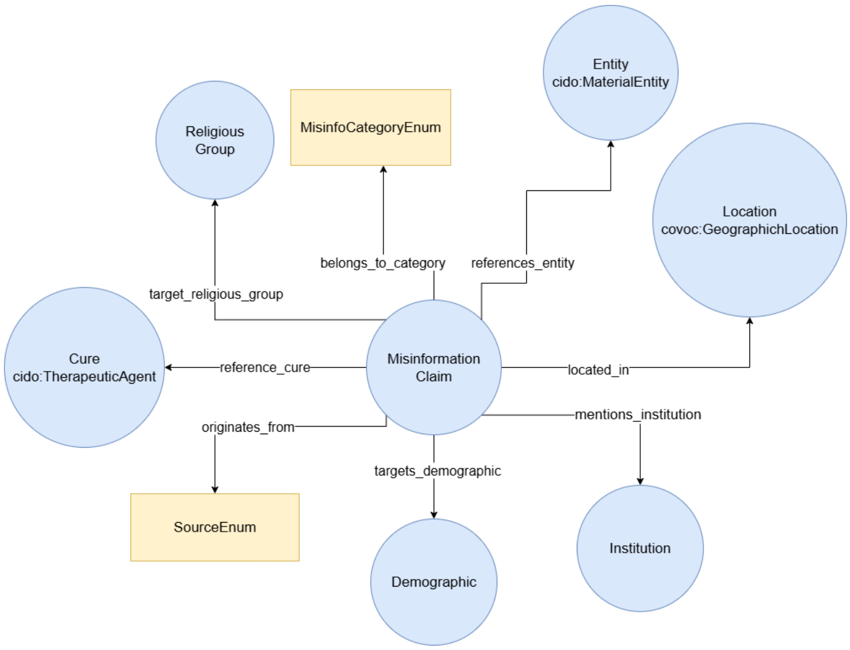

Figure 1: Ontology diagram showing core classes and relationships.

4.9 Knowledge Graph Construction Workflow

The knowledge graph construction followed a modular pipeline, transforming raw data into a validated RDF-based graph. The workflow integrates the above steps and is summarized in Figure 2.

- **Pipeline Overview**:
  1. *Data Ingestion*: Raw claims were collected via manual curation, Scrapy-based web scraping, and Twitter API queries, yielding 91 claims.
  2. *Preprocessing*: Text was cleaned, normalized, translated, and deduplicated.
  3. *Entity and Relation Extraction*: Manual extraction with LLM assistance identified 131 entities and 251 relations.
  4. *Ontology Mapping*: Entities and relations were mapped to the `misinfo:` ontology, reusing CIDO, COVoc, using RDFLib to generate RDF triples.

5. *Graph Storage*: The resulting 3,502 triples were stored in GraphDB, a semantic triple store supporting SPARQL queries and RDFS/OWL reasoning.
6. *Validation*: pySHACL validated the triples against SHACL shapes, ensuring compliance with ontology constraints.
7. *Querying and Visualization*: SPARQL queries (e.g., Listing 1) and GraphDB's visualization interface enabled analysis and exploration.

- **Format**: The KG was constructed as RDF triples, serialized in Turtle format for compatibility with semantic web standards. For prototyping, Neo4j was used to visualize entity relationships as a labeled property graph before RDF conversion.
- **Tools**: Python orchestrated the pipeline, with RDFLib for triple generation, GraphDB for storage and querying, pySHACL for validation, and Neo4j for initial visualization.

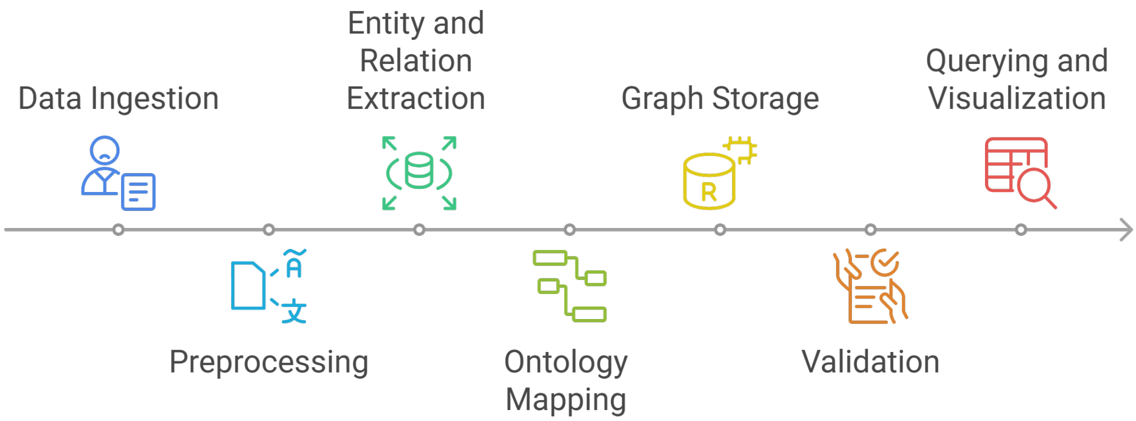

Figure 2: Workflow for constructing the COVID-19 misinformation knowledge graph, from data collection to validation.

This pipeline ensures reproducibility, with clear steps for data collection, processing, extraction, ontology mapping, and validation, producing a robust KG for analyzing misinformation in Nepal.

5 System Implementation

This section describes the practical implementation of the methodologies outlined in the previous sections. The implementation phase primarily involved designing an ontology, populating a knowledge graph from semi-structured misinformation data, and validating the structured output using SHACL constraints. The technologies adopted facilitated flexible, modular, and standards-compliant development.

5.1 Tools and Technologies

- **Python:** Served as the primary programming language for data parsing, RDF generation, SHACL validation, and integration with various APIs and libraries.
- **RDFLib:** A Python library used to represent and manipulate RDF triples. It was central

in converting data into RDF format, serializing graphs, and linking entities based on the ontology schema.

- **Scrapy**: A Python-based web scraping framework used to extract misinformation claims from news and fact-checking websites
- **Neo4j:** Used for early prototyping of the graph structure and for visualizing entity relationships before full RDF conversion.
- **GraphDB:** A semantic triple store used to store, query, and validate the knowledge graph. It supported SPARQL querying, inference using RDFS/OWL reasoning.
- **pySHACL:** A Python-based SHACL validation engine used to enforce schema-level integrity and ensure RDF triples adhered to defined SHACL shapes derived from the ontology.

5.2 Implementation

The implementation followed a modular pipeline architecture that transformed raw misinformation data into a validated, queryable knowledge graph.

5.2.1 Data Collection and Preprocessing

The implementation began with the collection and preprocessing of misinformation claims. A total of 91 claims were gathered, with a focus on content relevant to Nepal (75 claims), particularly Kathmandu (14 claims). Collection methods combined manual techniques with automated tools:

- **Manual Collection with LLM Assistance**: The majority of claims were manually identified and curated from social media, news portals, and fact-checking platforms. Large Language Models (LLMs) were used to assist in detecting misinformation patterns, categorizing content, translating Nepali text into English, and verifying cultural and contextual accuracy.
- **Web Crawling**: The Scrapy framework was employed to crawl public news and fact-checking websites (e.g., NepalCheck).

Following collection, the raw data underwent a structured preprocessing pipeline designed to ensure consistency and readability across multilingual and informal sources:

- **Text Cleaning**: Removal of HTML tags, emojis, URLs, non-standard Unicode symbols, and excessive whitespace using regular expressions and string manipulation techniques.
- **Normalization**: Standardization processes such as lowercasing, punctuation unification, and contraction expansion were applied to bring all claims into a normalized format.
- **Translation and Semantic Fidelity**: Nepali-language content was translated to English using LLMs with human supervision. Special attention was given to idiomatic expressions and culturally embedded language to maintain semantic integrity.
- **Deduplication and Noise Filtering**: Duplicate and near-duplicate entries were removed. Entries with incomplete or low-quality information were excluded to improve downstream processing.

This phase resulted in a linguistically and semantically clean dataset, forming the foundation for subsequent entity and relation extraction in the knowledge graph construction pipeline.

5.2.2 Entity and Relation Extraction

Entities and relations were extracted manually with LLM assistance, as described in Sections 4.3 and 4.4:

- **Entity Extraction**: Each claim was manually reviewed, with an LLM used to suggest candidate entities. Prompts such as "Extract named entities (Entity, Institution, Location) and COVID-19-specific concepts (cures, vaccines) from: [claim text]" guided the LLM. For example, in the claim "Vaccines are rejected by WHO in Kathmandu" (`misinfo:Claim_86`), the LLM proposed entities such as "Vaccine" (`misinfo:Entity`), "WHO" (`misinfo:Institution`), and "Kathmandu" (`misinfo:Location`). Human annotators then validated and categorized these entities into ontology-defined classes, including `misinfo:MisinformationClaim`, `misinfo:Institution`, `misinfo:Location`, `misinfo:Cure`, `misinfo:Entity`, `misinfo:Demographic`, `misinfo:ReligiousGroup`, and `misinfo:Source`. Ambiguities were resolved during annotation (e.g., classifying "WHO" as an institution), and culturally specific terms were interpreted appropriately (e.g., "Pashupatinath" as a location). The final dataset included 131 entities: 91 claims, 4 institutions (e.g., WHO, Patan Hospital), 2 locations, 8 cures (e.g., Hot Lemon Water, Cow Urine), 8 general entities (e.g., Vaccine, 5G), 5 demographics (e.g., Elderly, Dalit Community), 2 religious groups (Hindu, Muslim), and 3 sources (Social Media, Traditional Beliefs, Word of Mouth).
- **Relation Extraction**: Relationships between entities were similarly identified with LLM assistance using prompts such as "Identify relationships between entities in: [claim text], such as belongs_to_category, located_in, references_cure." For the example above, the LLM suggested relations like `misinfo:mentions_institution(Claim_86, WHO)` and `misinfo:located_in(Claim_86, Kathmandu)`. Human annotators reviewed and validated these suggestions to ensure consistency with the ontology's eight object properties.

Relation Type	**Count**
`misinfo:belongs_to_category`	91
`misinfo:originates_from`	91
`misinfo:located_in`	91
`misinfo:references_entity`	26
`misinfo:references_cure`	7
`misinfo:mentions_institution`	15
`misinfo:targets_demographic`	4
`misinfo:targets_religious_group`	2
Total	**251**

Table 1: Summary of curated misinformation relations

5.2.3 Ontology Mapping and RDF Triple Generation

Entities and relations were mapped to the `misinfo:` ontology using RDFLib. Entities were aligned with ontology classes (e.g., "Patan Hospital" as `misinfo:Institution`, "Hot Lemon

Water" as misinfo:Cure, with skos:related linking to cido:TherapeuticAgent), and relations were mapped to object properties (e.g., misinfo:located_in linked to covoc:GeographicLocation). Data properties such as misinfo:claim_id and misinfo:title were used to annotate individual claims. External ontologies (cido:, covoc:) were reused to promote semantic interoperability.

Using RDFLib, the structured data was transformed into 3,502 RDF triples in Turtle format. For example, the claim "Hot Lemon Water Cures COVID-19" (misinfo:Claim_2) was represented with its type, title, referenced cure, geographic context, category, and source.

To support visualization and debugging, entities and relationships were initially modeled in Neo4j as a labeled property graph. This intermediate representation enabled intuitive exploration using Cypher queries prior to RDF conversion.

5.2.4 Graph Storage and Validation

The final knowledge graph, comprising 3,502 RDF triples, was deployed to Ontotext GraphDB using its Workbench interface. The repository was configured with RDFS reasoning to enable semantic inference and expose a SPARQL endpoint for querying and analysis.

To ensure data quality and conformance to the ontology design, SHACL validation was performed using the pySHACL library. Validation constraints were defined based on the misinfo: ontology, targeting key classes and properties. In particular, the shape for misinfo:MisinformationClaim enforced:

- Exactly one misinfo:belongs_to_category property (e.g., misinfo:Conspiracy);
- A unique misinfo:claim_id matching the regular expression pattern misinfo-[0-9a-f]{8};
- Mandatory presence of core properties such as misinfo:full_claim.

All triples passed validation, confirming both structural and semantic integrity of the graph. The validation report was exported in Turtle format (validation_report.ttl) for auditing and reproducibility.

5.2.5 Querying and Visualization

The knowledge graph was explored and analyzed using SPARQL queries executed within the GraphDB Workbench. Results were visualized using GraphDB's built-in Graph View interface, which provided an interactive graphical representation of RDF resources and their relationships.

For instance, a SPARQL query targeting misinformation claims associated with the location Kathmandu (Listing 1) returned 15 relevant entries, including "Patan Hospital Burns Bodies Secretly" (misinfo:Claim_6) and "Kathmandu Immune to COVID" (misinfo:Claim_16). These queries enabled the identification of geographically localized narratives.

During the early stages of development, Neo4j was employed for prototyping using labeled property graphs and Cypher queries. This allowed intuitive navigation and debugging of entity relationships. However, to maintain RDF compliance and ontology-based structure, the final visualization and analysis were conducted entirely within GraphDB.

Visual outputs, such as those shown in Figure 3, revealed patterns in misinformation spread and clusters of claims sharing similar sources, topics, or geographic focus.

6 Results and Analysis

The RDF dataset (`Covid19_misinfo_nepal_instances.ttl`), comprising 3,502 triples in Turtle format, along with the SHACL validation report (`validation_report.ttl`), is available at CAIR-Nepal Repository. This section presents the outcomes and insights derived from constructing and analyzing the COVID-19 misinformation knowledge graph for Nepal. The analysis encompasses statistical summaries of the graph's structure, visualizations of key rumor clusters, example SPARQL queries with their results, patterns in misinformation propagation, and an evaluation of the graph's quality. These findings highlight the utility of the knowledge graph in modeling complex misinformation narratives and provide a foundation for targeted interventions in Nepal's low-resource, multilingual context. ‘

6.1 Knowledge Graph Statistics

The knowledge graph was constructed using 91 misinformation claims, as detailed in the RDF dataset, and adheres to the `COVID19MisinfoNepal` ontology. The following statistics summarize the graph's structure:

- **Total Triples**: 3,502 RDF triples, as reported in the dataset analysis, capturing entities, relations, and their attributes.
- **Entities**: A total of 131 entities were identified across the ontology's classes:

Category	Description
Entities	
Misinformation Claims	91 instances representing false claims (e.g., "Vaccines Make You Magnetic")
Institutions	4 instances (e.g., WHO, Patan Hospital, Nepal Army, United Nations)
Cures	8 instances (e.g., Hot Lemon Water, Cow Urine, Chloroquine)
Generic Entities	8 instances (e.g., Vaccine, 5G, PCR Test)
Demographics	5 instances (e.g., Children, Pregnant Women, Elderly, Smokers, Dalit Community)
Religious Groups	2 instances (Hindu Community, Muslim Community)
Locations	2 instances (Nepal, Kathmandu)
Sources	3 instances (e.g., Social Media, Traditional Beliefs, Word of Mouth)
Misinformation Categories	8 distinct types (e.g., Conspiracy, Fake Cures, Medical Hoax)
Relations	
`belongs_to_category`	91 relations (one per claim)
`originates_from`	91 relations (source of each claim)
`located_in`	91 relations (location context)
`references_entity`	26 relations referencing entities like Vaccine or 5G
`references_cure`	7 relations linking to cures like Hot Lemon Water or Chloroquine
`mentions_institution`	15 relations involving institutions like WHO or Patan Hospital
`targets_demographic`	4 relations targeting groups like Children or Elderly
`targets_religious_group`	2 relations targeting Hindu or Muslim communities
Types	
Entity Classes	9 classes (e.g., `MisinformationClaim`, `Cure`)
Relation Types	9 types (e.g., `belongs_to_category`)
Misinformation Category Types	11 defined (from `MisinfoCategoryEnum`)
Source Types	3 defined (from `SourceEnum`)

Table 2: Entities, Relations, and Types in the Knowledge Graph

- **Connected Components**: The graph contains 1 main connected component, with most claims linked through shared entities (e.g., `Entity_Vaccine`, `Location_Nepal`). Isolated subgraphs exist for claims with unique entities or locations, as noted in the dataset analysis.

These statistics reflect a moderately dense graph, with each claim connected to at least three other nodes (category, source, location), enabling robust semantic querying and analysis.

6.2 Visualization Examples

The knowledge graph was visualized using GraphDB's built-in visualization interface. Each misinformation claim is represented as a node, connected to entities such as geographic locations, misinformation categories, and the platforms on which they originated.

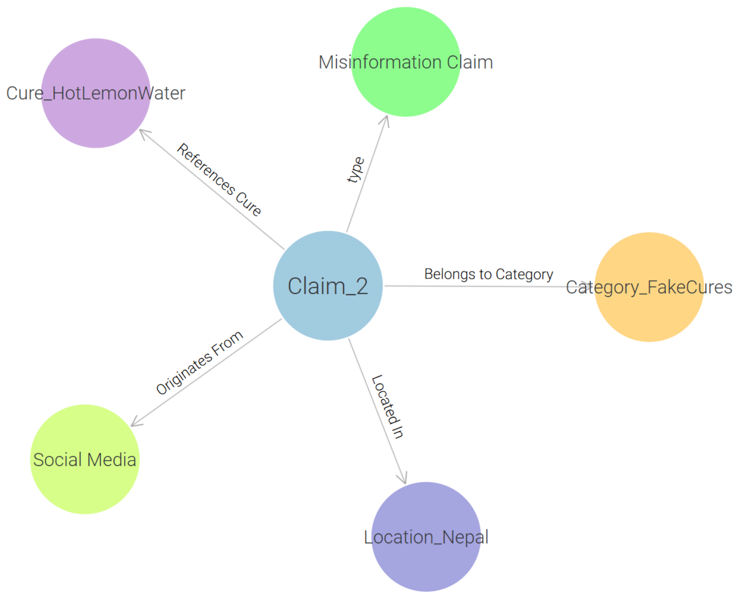

Figure 3: Visualization of the claim “Hot lemon water cures COVID-19” showing its connections to cure type, source platform, location, and fact-check.

This visual representation helps trace the semantic structure of each claim, identify common sources of misinformation, and explore patterns across claims.

6.3 Sample Queries and Outcomes

SPARQL queries were executed in GraphDB to extract actionable insights. Below are three representative queries and their outcomes:

Query 1: Claims in Kathmandu

Retrieve claims associated with Kathmandu.

```
SELECT ?claim ?title WHERE {
  ?claim a misinfo:MisinformationClaim ;
         misinfo:located_in misinfo:Location_Kathmandu ;
         misinfo:title ?title .
}
```

Listing 1: SPARQL query for claims in Kathmandu

Outcome: Returned 15 claims, including *”Patan Hospital Burns Bodies Secretly”* (`misinfo:Claim_6`), *”Kathmandu Immune to COVID”* (`misinfo:Claim_16`), *”Patan Hospital Sells Vaccines”* (`misinfo:Claim_86`), and *“Patan Hospital Uses Fake Meds”* (`misinfo:Claim_91`), among 11 others targeting Kathmandu, primarily involving Patan Hospital-related misinformation.

Query 2: Fact-Checked Claims

Identify claims with fact-check information.

```
SELECT ?claim ?title ?factCheck WHERE {
  ?claim a misinfo:MisinformationClaim ;
         misinfo:fact_checked_truth ?factCheck ;
         misinfo:title ?title .
}
```

Listing 2: SPARQL query for fact-checked claims

Outcome: Retrieved 91 claims, all associated with fact-checks, including *"Hot Lemon Water Cures COVID-19"* (`misinfo:Claim_2`, fact-checked as "No scientific evidence supports this") and *"Vaccines Make You Magnetic"* (`misinfo:Claim_1`, "No magnetic ingredients; debunked by WHO").

6.4 Insightful Patterns

Analysis of the knowledge graph revealed key patterns in misinformation sources, propagation routes, and clustering, providing insights into the spread of COVID-19 misinformation in Nepal:

- **Frequent Misinformation Sources**: `misinfo:Source_SocialMedia` was the primary origin for 83 of 91 claims (91%), highlighting its role as the dominant vector for misinformation dissemination in Nepal. Traditional Beliefs (3 claims) and Word of Mouth (5 claims) were less prevalent but significant for culturally rooted misinformation, particularly in rural or low-literacy communities.
- **Propagation Routes**: The "Hot Lemon Water" claim (`misinfo:Claim_2`) appeared in multiple variations (e.g., "Hot lemon water kills the virus", "Drink lemon water in the morning to prevent COVID"), unified in the graph via the `Cure_HotLemonWater` entity, which was linked to 3 claims. This structure enabled tracing narrative evolution across social media platforms. Similarly, vaccine-related claims (20 claims) formed a dense network around `Entity_Vaccine`, indicating widespread propagation of anti-vaccine narratives through online channels.
- **Clustering**:
  - **Vaccine Misinformation**: 20 claims referenced `Entity_Vaccine` (e.g., "Vaccines Cause Infertility", "Vaccines Contain Microchips"), forming a significant anti-vaccine narrative cluster. These claims were often linked to social media sources and fact-checked as false, reflecting a critical challenge for public health campaigns.
  - **Institutional Distrust**: Patan Hospital was targeted in 11 of the 15 Kathmandu-specific claims (e.g., "Patan Hospital Burns Bodies Secretly" (`misinfo:Claim_6`), "Patan Hospital Sells Vaccines" (`misinfo:Claim_86`), "Patan Hospital Uses Fake Meds" (`misinfo:Claim_91`)), and WHO was referenced in 4 claims, contributing to the 15 `mentions_institution` relations. This cluster indicates widespread distrust in healthcare institutions, particularly in urban areas like Kathmandu.
  - **Cultural and Religious Narratives**: Claims such as "Pashupatinath Water Immune" (`misinfo:Claim_29`) and "Vaccines Contain Pork" (`misinfo:Claim_17`) targeted religious groups (Hindu and Muslim communities), leveraging cultural beliefs to propagate misinformation. The `FakeCures` category, including `Cure_HotLemonWater`,

was prominent, reflecting reliance on traditional remedies in Nepal-focused misinformation.

– **Geographic Clustering**: 15 claims were linked to `Location_Kathmandu`, with 11 involving Patan Hospital, suggesting heightened misinformation activity in urban areas, likely driven by greater social media penetration. This cluster included diverse narratives, from institutional distrust (e.g., "Patan Hospital Fakes COVID Tests" (`misinfo:Claim_76`)) to cultural claims (e.g., "Kathmandu Immune to COVID" (`misinfo:Claim_16`)).

These patterns underscore the knowledge graph's ability to uncover interconnected misinformation themes, enabling targeted interventions such as digital literacy campaigns on social media and community-based education to address cultural and religious misinformation.

7 Discussion

This section reflects on the outcomes of constructing a knowledge graph for COVID-19 misinformation in Nepal, evaluates the methodology, and explores the broader implications of the project. The discussion is organized into four subsections: interpretation of the results, strengths of the approach, limitations, and ethical considerations.

7.1 Interpretation of Results

The construction of the COVID-19 misinformation knowledge graph revealed several critical patterns that provide insights into the nature and spread of false information in Nepal. The graph, comprising 3,502 RDF triples, 131 entities, and 251 relations, effectively captured the semantic structure of 91 misinformation claims, primarily sourced from social media (83 claims, 91%) but also from traditional beliefs and word of mouth. This dominance of social media as a misinformation vector underscores its role as a primary channel for rapid dissemination, particularly in urban areas like Kathmandu, where 15 claims were localized, 11 of which targeted Patan Hospital. This pattern suggests a concentrated narrative of institutional distrust, particularly toward healthcare providers, which aligns with global trends of skepticism toward medical institutions during the pandemic [19].

Another significant finding was the clustering of misinformation around specific themes, such as vaccines (20 claims) and fake cures (e.g., "Hot Lemon Water Cures COVID-19"). The vaccine-related cluster, centered around `Entity_Vaccine`, highlighted narratives like "Vaccines Cause Infertility" and "Vaccines Contain Microchips," which were debunked through fact-checks linked via `fact_checked_truth` relations. These clusters indicate coordinated or organic propagation of anti-vaccine sentiment, likely amplified by social media algorithms. Similarly, the prominence of culturally rooted misinformation, such as claims involving Pashupatinath water or cow urine, reflects the interplay of traditional beliefs with modern misinformation channels, a dynamic particularly relevant in Nepal's diverse socio-cultural context [1].

The SPARQL queries demonstrated the graph's utility in extracting actionable insights. For instance, Query 1 (Section 6.3) identified 15 Kathmandu-specific claims, revealing localized patterns of distrust, while Query 2 confirmed that all 91 claims were fact-checked, ensuring the graph's reliability for verification purposes. These patterns suggest that knowledge graphs can not only map misinformation but also guide targeted interventions, such as digital literacy campaigns on social media or community-based education to address culturally specific narratives.

7.2 Strengths of the Approach

The knowledge graph approach offers several strengths, particularly in its interoperability, semantic expressiveness, and potential for integration with broader systems. By adhering to semantic web standards (RDF, OWL, SHACL), the graph ensures interoperability with existing ontologies like CIDO and COVoc [6, 11]. This compatibility allows the graph to be extended or integrated with global knowledge bases, enabling cross-domain reasoning, such as linking misinformation claims to biomedical data on COVID-19 treatments or transmission. For example, the reuse of `cido:TherapeuticAgent` for fake cures like chloroquine enhances the graph's alignment with standardized biomedical knowledge, facilitating reuse by researchers worldwide.

The semantic expressiveness of the graph is another key strength. By modeling entities (e.g., `MisinformationClaim`, `Institution`) and relations (e.g., `belongs_to_category`, `references_entity`), the graph captures complex relationships, such as the linkage between a claim, its source, and its geographic context. This structure supports multi-hop reasoning, as demonstrated by SPARQL queries that trace claims through institutions, demographics, or locations. Such expressiveness enables nuanced analysis of misinformation propagation, far surpassing traditional classification models that treat claims as isolated instances [22].

Furthermore, the graph's modular design and use of GraphDB as a triple store allow for seamless integration with other data sources, such as real-time social media feeds or additional fact-checking databases. This potential for dynamic updates makes the approach scalable and adaptable to evolving misinformation landscapes, a critical feature given the rapid spread of false narratives during crises.

7.3 Limitations

Despite its strengths, the approach has notable limitations. First, the knowledge graph represents a static snapshot of 91 claims collected during the internship period. Misinformation is inherently dynamic, with narratives evolving rapidly as new events unfold or public sentiment shifts. The static nature of the dataset limits its ability to capture temporal dynamics, such as how claims like "Hot Lemon Water Cures COVID-19" evolved over time or how new variants emerged post-collection. Future iterations could incorporate streaming data pipelines to enable real-time updates.

Second, the dataset's coverage is geographically and thematically limited. While 75 claims are Nepal-specific and 15 are tied to Kathmandu, the dataset does not fully represent rural areas or other urban centers, potentially missing region-specific narratives. Additionally, the focus on COVID-19 excludes other misinformation domains (e.g., political or climate-related disinformation), which could share similar propagation patterns. Expanding the dataset to include diverse regions and topics would enhance the graph's generalizability.

Finally, the language scope is constrained primarily to English and Nepali, with translations relying on LLMs and human verification. This approach, while effective for the collected claims, may not scale well for Nepal's linguistic diversity. Limited coverage of low-resource languages restricts the graph's applicability in fully multilingual contexts, a significant challenge in Nepal's heterogeneous linguistic landscape [1].

7.4 Ethical Considerations

The construction and use of a misinformation knowledge graph raise several ethical considerations. One primary concern is the risk of mislabeling claims as misinformation, which could inadvertently harm individuals or communities. For instance, labeling a culturally significant practice (e.g., using traditional remedies like hot lemon water) as misinformation without nuanced explanation risks alienating communities or dismissing valid cultural practices. To mitigate this, the project relied on fact-checked data and transparent validation via SHACL constraints, but future work should involve community consultation to ensure cultural sensitivity.

Data privacy is another critical issue. Although the dataset was sourced from public platforms (e.g., social media, news sites), aggregating and analyzing claims could inadvertently expose personal or sensitive information, particularly if linked to specific demographics or religious groups. The project anonymized data during preprocessing, but robust privacy safeguards, such as differential privacy techniques, could further reduce risks.

Finally, the use of public opinion data to construct the graph raises ethical questions about representation and bias. Social media, the primary source for 91% of claims, may overrepresent urban, tech-savvy populations, potentially skewing the graph toward their narratives while underrepresenting rural or marginalized voices. Ensuring equitable representation requires broader data collection strategies, including offline sources like community surveys, and transparent documentation of data biases.

These ethical considerations highlight the need for responsible AI practices in misinformation research. By prioritizing transparency, cultural sensitivity, and inclusivity, the knowledge graph can serve as a tool for public good while minimizing potential harm.

8 Conclusion

This internship project at the Center for Artificial Intelligence Research Nepal (CAIR-Nepal) focused on constructing a knowledge graph to model and analyze COVID-19 misinformation in Nepal. This concluding section summarizes the objectives and achievements, highlights the technical and conceptual contributions, and reflects on the personal learning experiences gained during the process. The work underscores the potential of knowledge graphs to address complex misinformation challenges in a low-resource, multilingual context.

8.1 Recap of Objectives and Achievements

The primary objectives of this project, as outlined in Section 1, were to fulfill the requirements of the B.Sc. in Computational Mathematics course (COMP 408), design a domain-specific ontology for COVID-19 misinformation in Nepal, construct a knowledge graph capturing claims, their categories, sources, and fact-checks, and enable semantic querying and visualization to explore misinformation narratives. These objectives were successfully met through a structured pipeline involving data collection, ontology design, graph construction, and validation.

The project achieved the creation of a knowledge graph with 3,502 RDF triples, 131 entities, and 251 relations, representing 91 misinformation claims, primarily from Nepal (75 claims) and Kathmandu (15 claims). The ontology, built under the `misinfo:` namespace and integrated with CIDO and COVoc, provided a robust framework for modeling entities (e.g., `MisinformationClaim`, `Cure`) and relations (e.g., `belongs_to_category`, `located_in`).

The graph, implemented in GraphDB and validated using SHACL constraints, supported SPARQL queries that revealed key patterns, such as the dominance of social media as a misinformation source (91% of claims) and clusters of anti-vaccine narratives and institutional distrust, particularly targeting Patan Hospital. Visualizations and queries demonstrated the graph's ability to trace semantic connections, fulfilling the goal of enabling nuanced analysis of misinformation dynamics.

8.2 Summary of Contributions

The project made significant technical and conceptual contributions to the study of misinformation in Nepal. Technically, it developed a reusable pipeline for knowledge graph construction, integrating tools like Python, RDFLib, Neo4j, GraphDB, and pySHACL. The pipeline encompassed data collection, preprocessing (text cleaning, normalization, translation), ontology design, and semantic validation. The resulting graph, with its 3,502 triples and structured relations, provides a scalable framework for modeling misinformation, adaptable to other domains or regions. The use of SHACL for validation ensured data integrity, while SPARQL queries enabled actionable insights, such as identifying Kathmandu-specific claims or fact-checked narratives.

Conceptually, the project advanced the application of knowledge graphs to misinformation analysis in a low-resource, multilingual context. By focusing on Nepal, it addressed a gap in existing research, which often prioritizes English-language content [22]. The ontology's integration with global standards like CIDO and COVoc ensures interoperability, allowing future researchers to link local misinformation data with global biomedical knowledge. The identification of patterns—such as social media's role in amplifying anti-vaccine narratives and the cultural embedding of fake cure claims—offers a foundation for targeted interventions, such as digital literacy campaigns or community education programs. These contributions highlight the value of semantic technologies in combating disinformation and supporting public health efforts.

8.3 Personal Learning Experience

The internship provided a transformative learning experience, bridging theoretical knowledge from computational mathematics with practical applications in natural language processing (NLP), disinformation analysis, and knowledge graph construction. Working with NLP techniques, particularly LLM-assisted translation and entity extraction, deepened my understanding of processing multilingual and noisy text data, a critical skill given Nepal's linguistic diversity. The challenge of handling informal social media content and culturally nuanced claims underscored the importance of combining automated tools with human verification to ensure accuracy.

Engaging with disinformation research revealed the complexity of misinformation ecosystems, where cultural, social, and technological factors intersect. I learned how narratives evolve, such as the variants of the "Hot Lemon Water" claim, and how they exploit trust or skepticism to spread. This insight emphasized the need for context-aware approaches to misinformation detection, particularly in low-resource settings.

Constructing the knowledge graph was a pivotal learning opportunity, introducing me to semantic web technologies (RDF, OWL, SPARQL) and graph databases like GraphDB. Designing the ontology required balancing domain-specific needs with interoperability, while SHACL validation taught me the importance of data integrity in semantic systems. These skills, combined with hands-on experience in Python, Neo4j, and visualization tools, have equipped me to tackle real-world problems in AI and data science. The collaborative environment at CAIR-Nepal further honed my communication and teamwork abilities, preparing me for a career in research and technology development.

In conclusion, this project not only achieved its technical objectives but also demonstrated the power of knowledge graphs to structure and analyze complex misinformation narratives. The experience has solidified my interest in computational approaches to societal challenges and inspired me to pursue further work in AI-driven solutions for public good.

9 Future Work

The knowledge graph constructed for COVID-19 misinformation in Nepal represents a significant step toward understanding and combating disinformation in a low-resource, multilingual context. However, its static nature and limited scope suggest several avenues for improvement and expansion. This section outlines potential technical enhancements, application areas, research opportunities, and deployment considerations to extend the project's impact and utility.

9.1 Technical Improvements

To enhance the knowledge graph's functionality and robustness, several technical improvements are proposed. First, integrating real-time data collection would address the static nature of the current dataset, which captures 91 claims from a fixed period. By incorporating streaming APIs from social media platforms or news aggregators, the graph could dynamically update with new misinformation claims as they emerge, enabling timely analysis of evolving narratives, such as variants of the "Hot Lemon Water" claim. This would require robust preprocessing pipelines to handle high-velocity data while maintaining semantic accuracy.

Second, improving entity disambiguation would enhance the graph's precision. The current implementation relies on manual curation and LLM-assisted entity extraction, which may struggle with ambiguous entities (e.g., distinguishing "WHO" as the World Health Organization versus a homonym). Advanced NLP techniques, such as contextual embeddings or knowledge base linking, could improve entity resolution, ensuring accurate connections across claims, institutions, and locations [21].

Finally, expanding multilingual support is critical for Nepal's diverse linguistic landscape. The current graph primarily handles English and Nepali claims, with translations validated manually. Incorporating automated translation and entity extraction for additional languages would broaden coverage. Leveraging multilingual LLMs or cross-lingual embeddings could streamline this process, ensuring the graph captures misinformation across Nepal's linguistic diversity while maintaining cultural nuance.

9.2 Application Areas

The knowledge graph has promising applications in policy advisory, media monitoring, and academic research. For policy advisory, the graph's insights into misinformation patterns—such as the dominance of social media (91% of claims) or institutional distrust (e.g., 11 Kathmandu-specific claims targeting Patan Hospital)—can inform targeted interventions. Policymakers could use these findings to design digital literacy campaigns or collaborate with social media platforms to curb false narratives, particularly during public health crises [19].

In media monitoring, the graph's SPARQL query capabilities and visualizations (e.g., Figure 3) enable real-time tracking of misinformation trends. Media organizations or fact-checking agencies could integrate the graph to monitor emerging narratives, prioritize debunking efforts, or identify high-impact claims, such as anti-vaccine clusters. This would enhance the efficiency of fact-checking in resource-constrained environments.

For academic research, the graph serves as a reusable framework for studying misinformation dynamics. Its interoperability with ontologies like CIDO and COVoc [6, 11] allows researchers to extend the model to other domains or regions, fostering comparative studies on disinformation propagation. The graph's structured data also supports machine learning applications, such as training models for automated claim detection.

9.3 Research Opportunities

The project opens several research avenues, particularly in analyzing the temporal evolution of misinformation and developing cross-lingual knowledge graphs. Studying the temporal dynamics of rumors, such as how the "Vaccines Cause Infertility" narrative evolved over time, would require extending the graph to include temporal attributes (e.g., `claim_timestamp`). This could enable longitudinal analysis of misinformation spread, revealing how narratives adapt to events like vaccine rollouts or policy changes.

Cross-lingual knowledge graphs present another opportunity. The current graph's focus on English and Nepali limits its applicability in Nepal's multilingual context. Developing a cross-lingual framework, potentially by aligning entities across languages using ontologies or embeddings, would allow the graph to capture code-switching or region-specific narratives. This could build on existing work in multilingual misinformation detection, adapting it for low-resource settings [1].

Additionally, integrating the graph with global misinformation datasets (e.g., FakeNewsNet [14]) could enable comparative studies, examining how misinformation in Nepal aligns with or diverges from global trends. Such research could inform strategies for combating coordinated disinformation campaigns across borders.

9.4 Deployment Considerations

Deploying the knowledge graph for practical use requires addressing hosting, scalability, and accessibility. Hosting the graph on a cloud-based semantic triple store, such as GraphDB, would ensure reliable access for stakeholders like researchers, policymakers, or fact-checkers. Scalability is a key concern, as expanding the graph to include more claims or languages would increase computational demands. Optimizing query performance and indexing strategies in GraphDB could mitigate this, ensuring efficient SPARQL queries even with larger datasets.

Providing open-access APIs would enhance the graph's utility, allowing external applications to query misinformation data in real time. For example, a RESTful API could expose endpoints for retrieving claims by category, location, or source, enabling integration with fact-checking platforms or public health dashboards. However, API access must balance openness with security, implementing authentication to prevent misuse while ensuring data availability for legitimate users.

Finally, community engagement is essential for deployment. Making the graph accessible to local stakeholders, such as Nepal's fact-checking organizations or community leaders, requires user-friendly interfaces and documentation in local languages. This would maximize the graph's impact in combating misinformation while fostering trust and collaboration in Nepal's diverse socio-cultural context.

References

[1] Rabin Adhikari, Safal Thapaliya, Nirajan Basnet, Samip Poudel, Aman Shakya, and Bishesh Khanal. COVID-19-related Nepali tweets classification in a low resource setting. In *Proceedings of The Seventh Workshop on Social Media Mining for Health Applications, Workshop & Shared Task*, pages 209–215, Gyeongju, Republic of Korea, October 2022. Association for Computational Linguistics. URL https://aclanthology.org/2022.smm4h-1.52.

[2] Alan et al. Akbik. Contextual string embeddings for sequence labeling. *COLING*, 2018.

[3] Pepa Atanasova, Jakob Grue Simonsen, Christina Lioma, and Isabelle Augenstein. A diagnostic study of explainability techniques for text classification. In Bonnie Webber, Trevor Cohn, Yulan He, and Yang Liu, editors, *Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing (EMNLP)*, Online, November 2020. Association for Computational Linguistics. URL https://aclanthology.org/2020.emnlp-main.263/.

[4] Climate and Development Knowledge Network (CDKN). Nepal: Managing misinformation during covid-19. https://cdkn.org/story/feature-nepal-managing-misinformation-during-covid-19, 2020.

[5] Jacob Devlin, Ming-Wei Chang, Kenton Lee, and Kristina Toutanova. Bert: Pre-training of deep bidirectional transformers for language understanding. *arXiv preprint arXiv:1810.04805*, 2018.

[6] Ikram El Houari et al. Covoc and covtriage: novel resources to support literature triage. *Bioinformatics*, 2023. Collaborative ontology for COVID-19 literature annotation, ELIXIR, SIB Literature Services.

[7] Don Fallis. A conceptual analysis of disinformation. *iConference*, 2009.

[8] Thomas R. Gruber. A translation approach to portable ontology specifications. *Knowledge Acquisition*, 5(2):199–220, 1993. doi: 10.1006/knac.1993.1008.

[9] Andreas Hanselowski, Avinesh P. V. S., Benjamin Schiller, Felix Caspelherr, Debanjan Chaudhuri, Christian M. Meyer, and Iryna Gurevych. A retrospective analysis of the fake news challenge stance-detection task. In *Proceedings of COLING 2018*, pages 1859–1874, 2018.

[10] Naeemul et al. Hassan. Claimbuster: The first-ever end-to-end fact-checking system. *Proceedings of VLDB*, 2017.

[11] Yongqun He, Haofen Yu, and et al. Cido: an ontology for coronavirus disease knowledge and data integration. *Journal of Biomedical Semantics*, 11(1):1–14, 2020.

[12] Aidan Hogan, Eva Blomqvist, Michael Cochez, Claudia D'Amato, Gerard De Melo, Claudio Gutierrez, José Emilio Labra Gayo, Sabrina Kirrane, Sebastian Neumaier, Axel Polleres, Roberto Navigli, Axel-Cyrille Ngonga Ngomo, Sabbir M. Rashid, Anisa Rula, Lukas Schmelzeisen, Juan Sequeda, Steffen Staab, and Antoine Zimmermann. Knowledge graphs. *arXiv preprint arXiv:2003.02320*, 2021. URL https://arxiv.org/abs/2003.02320.

[13] Mohit Mayank, Shakshi Sharma, and Rajesh Sharma. Deap-faked: Knowledge

graph-based approach for fake news detection. https://arxiv.org/abs/2107.10648, 2021.

[14] Kai Shu, Deepak Mahudeswaran, Suhang Wang, and Dongwon Lee. Fakenewsnet: A data repository with news content, social context and dynamic information for studying fake news on social media. *Big Data*, 2020.

[15] Gabriel Stanovsky, Julian Michael, Luke Zettlemoyer, and Ido Dagan. Supervised open information extraction. In *Proceedings of the 2018 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*, volume 1, pages 885–895, New Orleans, Louisiana, 2018. Association for Computational Linguistics. URL https://aclanthology.org/N18-1081/.

[16] Edson C Tandoc Jr, Zheng Wei Lim, and Richard Ling. Defining "fake news": A typology of scholarly definitions. *Digital Journalism*, 6(2):137–153, 2018.

[17] William Yang Wang. Liar, liar pants on fire: A new benchmark dataset for fake news detection. In *ACL*, 2017.

[18] Claire Wardle and Hossein Derakhshan. Information disorder: Toward an interdisciplinary framework for research and policy making. *Council of Europe report*, 2017.

[19] John Zarocostas. How to fight an infodemic. *The Lancet*, 395(10225):676, 2020.

[20] Yuhao Zhang, Victor Zhong, Danqi Chen, Gabor Angeli, and Christopher D. Manning. Position-aware attention and supervised data improve slot filling. In *Proceedings of the 2017 Conference on Empirical Methods in Natural Language Processing*, pages 35–45, Copenhagen, Denmark, 2017. Association for Computational Linguistics. doi: 10.18653/v1/D17-1004. URL https://aclanthology.org/D17-1004/.

[21] Xiaoyan Zhao, Yang Deng, Min Yang, Lingzhi Wang, Rui Zhang, Hong Cheng, Wai Lam, Ying Shen, and Ruifeng Xu. A comprehensive survey on relation extraction: Recent advances and new frontiers. *ACM Computing Surveys*, 56(11):1–39, 2024. doi: 10.1145/3674501. URL https://doi.org/10.1145/3674501.

[22] Xinyi Zhou and Reza Zafarani. A survey of fake news: Fundamental theories, detection methods, and opportunities. *ACM Computing Surveys (CSUR)*, 53(5):1–40, 2020.

[23] Arkaitz et al. Zubiaga. Analysing how people orient to and spread rumours in social media by looking at conversational threads. *PloS one*, 2016.