

Recognizing Similar Crises through the Application of Ontology-based Knowledge Mining

Ngoc Luyen Le^{1,2}
ngoc-luyen.le@hds.utc.fr

Marie-Hélène Abel¹
marie-helene.abel@hds.utc.fr

Elsa Negre³
elsa.negre@dauphine.fr

1: Université de technologie de Compiègne, CNRS, Heudiasyc, CS 60319 - 60203 Compiègne Cedex, France.

2: Vivocaz, 8 B Rue de la Gare, 02200, Mercin-et-Vaux, France.

3: Paris-Dauphine University, PSL Research Universities, CNRS UMR 7243, LAMSADE, Paris, France.

Abstract

Recognizing and learning from similar crisis situations is crucial for the development of effective response strategies. This study addresses the challenge of identifying similarities within a wide range of crisis-related information. To overcome this challenge, we employed an ontology-based crisis situation knowledge base enriched with crisis-related information. Additionally, we implemented a semantic similarity measure to assess the degree of similarity between crisis situations. Our investigation specifically focuses on recognizing similar crises through the application of ontology-based knowledge mining. Through our experiments, we demonstrate the accuracy and efficiency of our approach to recognizing similar crises. These findings highlight the potential of ontology-based knowledge mining for enhancing crisis recognition processes and improving overall crisis management strategies.

Keywords: Crisis Management, Knowledge Representation, Ontology, Semantic Similarity

1. Introduction

In the field of crisis management, the ability to recognize and learn similar crisis situations is of paramount importance for devising effective response strategies. By identifying similarities between different crises, decision-makers can leverage existing knowledge and experiences to inform their decision-making processes, enhance response capabilities, and mitigate the potential impact of crises (Bundy et al., 2017; 't Hart, 2023). Furthermore, the recognition of similar crises enables the identification of relevant actions and effective communication strategies specific to each crisis situation (Negre, 2013). Understanding the similarities between crises is crucial for decision-makers as it allows them to gain insights into the limitations of traditional crisis management practices and identify the most

appropriate interventions and measures based on the experiences and outcomes of past crisis situations (Boin and McConnell, 2007). This enables more targeted and tailored responses, optimizing the allocation of resources and minimizing potential disruptions. However, the task of identifying such similarities across a wide range of crisis-related information and scenarios can often be daunting and time-consuming (Boin et al., 2020).

To address this challenge, ontology-based knowledge mining techniques offer a promising solution for structuring and organizing crisis-related information and scenarios. Ontology, as a structured framework for modeling knowledge within specific domains, formalizes concepts and relationships, thereby enhancing data understanding, retrieval, and integration (Gruber, 1993). By capturing the relationships, attributes, and properties of various crisis elements, an ontology-based knowledge base becomes a valuable resource for identifying patterns and similarities within the domain of crises. The use of ontology-based knowledge mining in crisis management offers several advantages. It allows for the integration and consolidation of diverse data sources, including historical records, expert knowledge, and data feeds, to build a comprehensive knowledge base. Through modeling, reasoning, and inference techniques, ontology-based knowledge mining can uncover meaningful connections and associations, enabling the identification of similar crisis situations and providing decision-makers with valuable insights into crisis situations. (Abiteboul et al., 2011).

The recognition of similar crises necessitates the development of a systematic approach, which forms the core of our investigation in this study. In this paper, we focus on recognizing similar crises through the application of ontology-based knowledge mining. Our methodology involves the utilization of an ontology-based crisis situation knowledge base, enriched with relevant crisis-related information.

Additionally, we employ a semantic similarity measure to assess the degree of similarity between different crisis situations. By combining these components, we aim to enhance the accuracy and efficiency of crisis recognition processes, thereby enabling more effective crisis management and response strategies.

The remainder of this paper is organized as follows: Section 2 introduces works from the literature on which our approach is based. Section 3 presents our main contributions, outlining the methodology for recognizing similar crises through ontology-based knowledge mining and the use of semantic similarity measure. In section 4, we provide experiments about identifying and calculating the crises similarities based on our developed ontology. Finally, we conclude and discuss the perspectives.

2. Related Work

In this section, our focus is on exploring the utilization of ontologies for knowledge representation and investigating different approaches for modeling crisis management and identifying similar crises. We delve into the use of ontologies as a foundation for structuring and organizing crisis-related knowledge.

2.1. Crisis-Related Information Management

In this paper, we focus on crisis-related information management, specifically addressing the various aspects and challenges related to the effective structuring, organization, and utilization of information during three critical phases: pre-crisis, crisis response, and post-crisis. The utilization of ontology-based knowledge representation serves as our primary approach to advance the organization and reusability of crisis-related information.

Crisis-related Information Management focuses on the efficient collection, organization, and utilization of information during crisis situations. It encompasses a range of activities and processes aimed at acquiring, processing, analyzing, and disseminating information to support decision-making and response efforts (Meesters, 2021; Bowen and Lovari, 2020). Various studies have explored the application of ontologies in different aspects and situations of crisis-related information management. For instance, in the study by De Nicola et al. (2019), an emergency management ontology was created with a focus on knowledge related to smart cities and crises arising from natural or anthropogenic events. Furthermore, Elmhadhbi et al. (2020) introduced the *POLARISCO* ontology, which captures knowledge from various stakeholders at both the commandment and operational levels of disaster response. Benaben et al. (2016, 2020)

introduced the *COSIMMA* metamodel which offers a systematic abstract representation of knowledge in the field of collaborative situations in crisis management. The metamodel encompasses four interconnected dimensions: context, partners, objectives, and behavior. In the study by Le et al. (2023b), the authors proposed the development of a lightweight crisis management domain ontology by leveraging the inheritance of the ISyCri meta-model (Benaben et al., 2008). The ontology incorporates concepts related to crisis description, affected individuals, and available resources. In summary, the existing ontologies in the field of crisis management are highly reliant on specific crisis scenarios, each containing unique knowledge and information pertaining to the particular type of crisis and/or geographical area. However, these ontologies lack a systematic analysis of the model and organization for the effective management of general crisis-related information.

The structure and organization of crisis-related information enable the storage of specific characteristics related to each individual crisis. Developing a comprehensive ontology for crisis-related information management facilitates the storage, retrieval, and reuse of information, allowing for valuable insights and experiences to be extracted regarding actions, utilized resources, and communication methods. In the upcoming section, we will delve into existing research on leveraging ontology for the recognition of similar crises.

2.2. Similar Crisis Identification

In order to recognize similar crises, the application of the Case-Based Reasoning (CBR) approach can be instrumental in finding similar cases. Indeed, CBR refers to a problem-solving methodology that leverages past experiences or cases to solve new problems or make decisions. In the context of crisis management, CBR involves comparing the characteristics and attributes of different crisis situations and identifying similarities based on their shared features. By analyzing and comparing relevant factors such as the nature of the crisis, the affected population, response strategies, and outcomes, a CBR system can identify similar crisis scenarios. In recent years, several studies have employed the CBR approach to tackle the problems of identifying similar crises and planning crisis responses. For instance, in Yu et al. (2018), the authors specifically examined risk response in the field of critical infrastructure protection, with a particular emphasis on natural disasters. The research underscores the significance of historical cases as valuable resources that enable timely decision-making in the development

of response strategy plans.

The identification of similar crises can be effectively achieved by utilizing knowledge-based recommender systems, which have demonstrated their value in numerous domains, including crisis management. By analyzing historical crisis data and patterns, knowledge-based recommender systems can identify similarities and patterns among different crises, enabling the identification of similar crisis scenarios. For example, in the study conducted by Negre (2013), the author proposed a methodology for identifying similar actions and leveraging knowledge gained from past experiences to support decision-makers in crisis management. The objective is to enhance crisis management by providing decision-makers with valuable insights and lessons learned from similar crisis situations. In Le et al. (2023b), a constraint-based recommender system was proposed to optimize the deployment of citizen vehicle resources in the context of crisis management. The objective was to develop a system that not only recommends the most suitable deployment strategies but also enhances the overall efficiency of utilizing citizen vehicles during crisis situations.

CBR demonstrates a higher level of adaptability when it comes to novel crises, as it leverages case reuse and adaptation. In contrast, knowledge-based recommender systems may face challenges when dealing with new or rare crisis scenarios due to limited historical data. Additionally, CBR is typically domain-specific, specifically tailored to crisis management, while knowledge-based recommender systems have broader applicability across various domains. However, both CBR and knowledge-based recommender systems share similarities in their approach to identifying similar crisis situations. They both rely on existing knowledge and patterns, utilizing past cases or experiences to provide recommendations or solutions. Moreover, both approaches employ similarity measures that take into account crisis characteristics such as type, context, and impact to determine similarity. In the following section, we delve into the exploration of our approach for identifying similar crisis situations.

3. Our approach

In this section, we present our approach, which involves processing cycles for recognizing similar crisis situations, constructing an ontology-based crisis situation knowledge base, and developing a method for semantic similarity measurement using the ontology-based crisis situation knowledge base.

3.1. Process Cycle Formulation

We present our process cycle for recognizing similar crisis situations, as illustrated in Figure 1. At the onset of a new crisis situation, the relevant information pertaining to the event is meticulously extracted and transformed into an ontology-based representation, ensuring a standardized and structured format for further analysis. The construction of the new crisis situation adopts a metamodel derived from the ontology-based crisis situation knowledge base, which serves as a foundation for accurately reconstructing the key elements and attributes of the crisis. Moving forward, the process entails the critical step of matching and measuring the semantic similarity between the newly identified crisis situation and the existing crisis knowledge base built upon the ontology. This comparison is aimed at determining the degree of resemblance and relevance between the new crisis situation and previously encountered cases. By employing semantic similarity measurement techniques, a comprehensive list of the most pertinent and similar crisis situations is generated, offering decision-makers valuable insights and reference points for their decision-making process. Subsequently, the full information of the identified similar crisis situation cases is thoroughly evaluated and adapted to address the specific needs and requirements of decision-makers in their endeavor to mitigate the crisis. This evaluation involves a careful examination of the lessons learned, successful strategies, and best practices employed in the similar crisis situations, with the aim of providing practical guidance and solutions for problem-solving in the current crisis scenario. To ensure the currency and effectiveness of the crisis response system, the information obtained from the similar crisis situations can be integrated and updated in the ontology-based crisis situation knowledge base, enriching the repository with valuable experiences and insights. This ongoing refinement process ensures that the knowledge base remains up-to-date and capable of effectively supporting future crisis management efforts. In the final stage of the process cycle, a curated selection of the top n crisis situations, based on their relevance and similarity to the current crisis situation, is recommended to decision-makers. This curated list serves as a valuable resource, aiding decision-makers in their strategic planning, resource allocation, and overall crisis response efforts. By leveraging the knowledge and experiences encapsulated in the similar crisis situations, decision-makers can make well-informed and effective decisions to address the challenges posed by the crisis at hand.

Building upon the general process cycle, this study

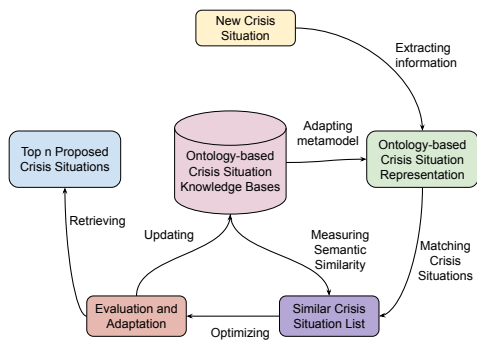


Figure 1. A process cycle for recognizing similar crisis situations.

primarily focuses on two key tasks that play a crucial role in recognizing similar crisis situations. The first task pertains to the construction of an ontology-based crisis situation knowledge base, which serves as the foundation for organizing and structuring crisis-related information. By developing this knowledge base, we aim to create a comprehensive repository of crisis situations, their characteristics, and associated knowledge, enabling efficient retrieval and analysis of relevant cases. The second task involves the implementation of a method for measuring semantic similarity between different crisis situations. This method serves as a crucial component in the process of identifying similar crisis situations from the knowledge base. By employing semantic similarity measures, we aim to capture the shared features, attributes, and context between crisis situations, enabling effective comparison and identification of similarities. In the next section, we will present the construction of the ontology-based crisis situation knowledge base.

3.2. Ontology-based Crisis Situation Knowledge Base

In this section, we delve deeper into the construction of the crisis-related information ontology, which forms an integral part of our ontology-based crisis situation knowledge base. This entails defining the ontology structure, identifying relevant crisis domains and sub-domains, and establishing the relationships and dependencies between different elements within the ontology. In our work, using an ontology-based approach offers advantages such as semantic richness, automated reasoning, and flexible extensibility. Ontologies provide a structured and meaningful representation of knowledge, enhancing data analysis and decision-making through explicit relationships and inferences. They enable interoperability, contextual understanding, and advanced querying, accommodating domain-specific constraints. While simple database

formats and RDF stores provide storage and querying capabilities, ontologies excel in capturing complex domain semantics, promoting shared understanding, and facilitating deeper insights from the data. The ontological choice allows to effectively manage the complexity of the domain, fulfill the requirement for semantic reasoning, and achieve the goal of long-term scalability and interoperability.

The process development of an ontology requires a systematic approach that involves gathering information, defining the scope, creating a conceptual model, defining the ontology schema, implementing the ontology, and evaluating and refining it as necessary. Therefore, we adopted the Agile Methodology for Ontology Development (AMOD) to develop our ontological model (Abdelghany et al., 2019). In general, the AMOD methodology comprises three key phases: (i) the pre-game phase, where the goals, requirements, and scope of the ontology are clearly defined to provide a general overview; (ii) the development phase, where the project is divided into iterative sprints¹, each dedicated to essential tasks such as conceptualization, formalization, and integration, enabling a progressive advancement in the development of the ontology; and (iii) the post-game phase, which encompasses critical activities such as evaluating and maintaining the ontology to ensure its enduring reliability and effectiveness as time progresses.

In line with the pre-game phase, our proposed ontology primarily aims to establish a standardized representation of crisis situation information and contextual details of crisis situations. This ontological model enhances information retrieval and adaptation within the crisis management system. The scope of the ontology encompasses a wide range of crisis situations and factors that contribute to the organization and structure of crucial information during crisis situations. It is designed to be flexible and adaptable, allowing for the incorporation of new crisis scenarios. Furthermore, our focus lies in creating a formal and structured representation of the crisis situation domain, with the objective of promoting a shared understanding and facilitating the use of standardized terminology among all stakeholders. By leveraging these valuable resources, our general aim is to improve the ability to locate relevant information and draw upon past experiences from similar crisis situations.

The development of the crisis-related information ontology plays a critical role in improving the organization and management of complex crisis

¹A sprint is a short period of concentrated work within a project where a team works together to accomplish specific objectives or deliverables.

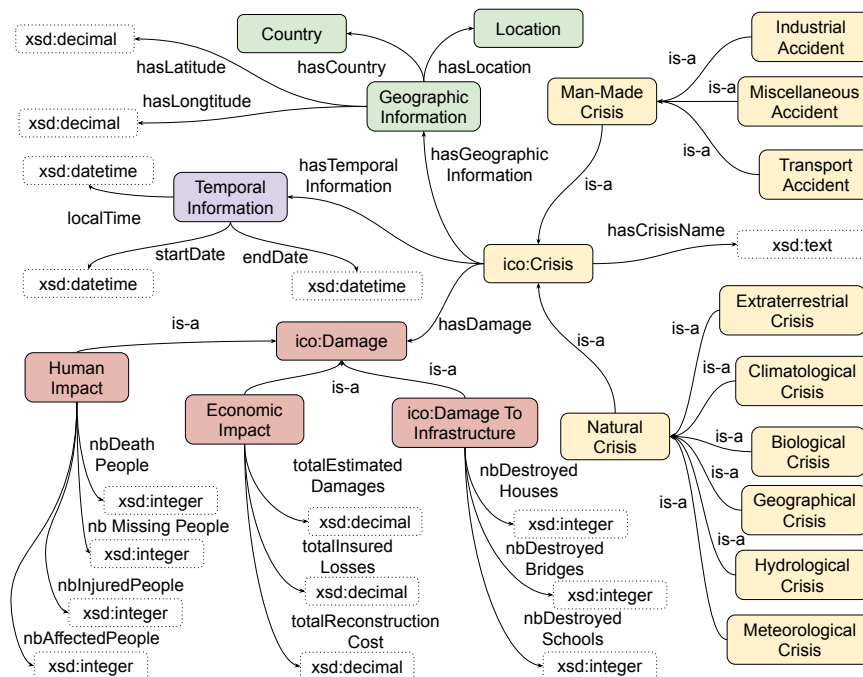


Figure 2. A dedicated ontology section for crisis-related information ontology included a hierarchy of crisis types (yellow boxes), damage (orange boxes), physical characteristics (blue boxes), geographical information (green boxes), and temporal entity (violet boxes), [“ico”: IyCri meta-model, “to”: Time Ontology].

situations. In line with the development phase, this study focuses on constructing a comprehensive domain ontology that encompasses key concepts, relationships, and knowledge associated with crisis situations. The development of our domain ontology for crisis situations builds upon the foundations of the crisis management ontology proposed by Le et al. (2023b) and the inheritance of the ISyCri meta-model (Benaben et al., 2008). Indeed, the ISyCri meta-model bears many similarities to a meta-ontology as it encapsulates essential concepts for domain ontologies, establishing a close relationship between the two (Arru et al., 2019). Therefore, we have developed our ontology independently, without relying on a general top-level ontology like SUMO (Niles and Pease, 2001), DOCLE (Gangemi et al., 2002), or BFO (Arp et al., 2015). This deliberate choice allows us to create lightweight, customized, and context-specific ontologies that are specifically tailored for the crisis management domain. The distinctiveness of the proposed ontology comes from how it's designed to match the needs of crisis-related data. Even though there might be similar classes in other ontologies as indicated in related work, in this ontology, their meanings and associations are customized to fit the specifics of crises. This ontology recognizes the varied and complex nature of crisis situations, aligning its classes and characteristics very closely with the details of crisis situations. Exploring

our developed ontology², as depicted in Figure 2, we extensively refine and enhance our proposed ontology by focusing on key entities: Crisis, Geographical Information, Temporal Entity, Physical characteristic, and Damage. In the paper, we did not explicitly differentiate between the terms ‘disaster’ and ‘crisis’ within the scope of ontological analysis. Indeed, these terms are often used interchangeably, and their distinction might not always be well-defined. Their usage can vary depending on factors such as the event’s severity, the extent of impact, and the level of preparedness and response. Firstly, crises can be classified into various typologies based on factors such as the nature of risks, associated events, and public perceptions. For example, the information service of the French government³ categorizes major crises into terrorist, health, technological, natural, and cyber crises. Additionally, the Institute for Radioprotection and Nuclear Safety (IRSN)⁴ classifies crises into individual, diffuse pollution, non-industrial collective, and technological crises related to specific sites and actions. In our work, we adopt the crisis classification proposed by the Centre for Research on the Epidemiology of Disasters (CRED) at the School of Public Health of the Université catholique

²Full version of crisis-related information ontology is published at <https://github.com/lengocluyen/OBCS-KB>.

³www.gouvernement.fr/en

⁴www.irsn.fr/EN

de Louvain in Brussels, Belgium (Guha-Sapir et al., 2016). This classification organizes crises into a hierarchical structure based on crisis types, which further include natural and man-made crises. Natural crises can be classified into several categories, including extraterrestrial, climatological, biological, geographical, hydrological, and meteorological crises. Each category is further divided into sub-categories. For example, meteorological crises can include extreme temperatures, fog, storms, and more. Geographical crises encompass earthquakes, mass movements, volcanic activity, and others. On the other hand, man-made crises are divided into industrial, miscellaneous, and transport accident crises. Similarly, these subcategories can be further categorized by other groups. For example, industrial accident crises may include explosions, gas leaks, collapses, oil spills, and more. Secondly, geographical information encompasses various aspects such as location, region, country, continent, as well as topography, specific longitude and latitude coordinates. Thirdly, temporal entity includes details regarding the starting and ending dates of the crisis, as well as the duration of the crisis. With temporal entity, the temporal entity, we utilize concepts and relationships from the time ontology⁵, ensuring consistency and reusability. Fourth, physical characteristic encompasses crucial information pertaining to the triggering origin of the crisis, as well as the magnitude scale and corresponding values associated with the event. Finally, the damages caused by the crisis event were categorized into human impact, economic impact, and damage to infrastructure. The human impact category focuses on statistics related to the number of affected, injured, missing, and deceased individuals. The economic impact category involves the total estimated damages, insured losses, and reconstruction costs. The damage to infrastructure category encompasses a wide range of damages, including destroyed or damaged houses, schools, commercial buildings, roads, bridges, and more.

In accordance with the post-game phase of development, our attention is directed towards testing and evaluating the ontology that has been created. This involves conducting various types of tests and assessments to ensure the quality and effectiveness of the ontology. We can perform tests and evaluations on multiple aspects, including the overall structure of the ontology, the specific instances within it, and the retrieval of information from the ontology. By conducting these rigorous evaluations, we aim to validate the ontology's functionality, reliability, and its ability to meet the intended objectives of crisis-related

⁵<https://www.w3.org/TR/owl-time/>

information management.

The implementation of the crisis-related information ontology involves the practical steps of creating and deploying the ontology within an application. This process entails converting the conceptual model of the ontology into a computer-readable format using ontology languages such as OWL (Web Ontology Language) and RDF (Resource Description Framework). In our study, we have implemented the ontology by leveraging OWL and utilizing the Protégé-OWL editor as a supportive tool (Musen, 2015). In the upcoming section, we will investigate the approach employed to measure the semantic similarity of crisis situations, based on the utilization of the constructed ontology.

3.3. Semantic Similarity Measure

We employ the term “semantic similarity” rather than simply “similarity” because it is based on the underlying meaning and relationships of the concepts represented within the ontology. The term “semantic” emphasizes the depth and context of the similarity calculation, reflecting the ontology's ability to capture the intricate semantic connections between crisis-related concepts. The measure of semantic similarity involves the comparison of two sets of triples based on their individual elements, which can be categorized into quantitative and qualitative information. This process aims to assess the degree of similarity between the sets. By examining both the numerical and descriptive aspects of the triples, we can gain insights into their similarities and differences. Consequently, we utilize the semantic similarity measure by leveraging the ontology to identify similar crises in this section.

Conventionally, an instance of crisis-related information, denoted as C , is commonly defined as a set of crisis, geographic information, temporal entity, physical characteristics and damage triples. This triple set is represented as follows:

$$C = \{a_1^C, a_2^C, \dots, a_m^C\} \quad (1)$$

where each a_t^C represents a triple, specifically $a_t^C = \langle subject_t, predicate_t, object_t \rangle$. Alternatively, the triple a_t^C in the context of Resource Description Framework is often used to represent a statement or assertion about a resource. It can be expressed as $\langle resource_t, property_t, statement_t \rangle$. By leveraging the crisis-related information ontology, we are able to extract and capture the significant information from a new crisis situation and seamlessly incorporate it into an ontology-based knowledge base. This enables us to retrieve crisis-related information by querying a subgraph RDF, thereby making valuable contributions to the development of knowledge bases specifically

tailored for crisis-related information. Our approach is inspired by the insightful research conducted by Le et al. (2022, 2023a). Through this methodology, we have the capability to measure the similarity between specific triples or sets of triples, facilitating the analysis and comparison of crisis-related information.

The measure of similarity between two specific triples is primarily based on the type of object present in each triple. In cases where the object type of the triple is qualitative, the semantic similarity of all three components of the triple, namely Qualitative Subjects, Predicates, and Objects (QSPO), will be measured as follows:

$$Sim_1(a_{s1}, a_{s2}) = \frac{\sum_{i=1}^k \bar{S}(w_{1i}, a_{s2}) + \sum_{j=1}^l \bar{S}(w_{2j}, a_{s1})}{k + l} \quad (2)$$

where $\bar{S}(w, a_s)$ represents the similarity between a word w and a QSPO a_s . With \vec{w} is a embedding vector of the word w , the calculation of the function $\bar{S}(w, a_s)$ is formally expressed as follows:

$$\bar{S}(w, a_s) = \max_{\vec{w}_q \in M} \bar{S}(\vec{w}, \vec{w}_q) \quad (3)$$

where $\vec{w}_q \in M = \{\vec{w}_1, \vec{w}_2, \dots, \vec{w}_z\}$ refers to the word embedding vector of a_s , z represents the number of words in a_s and q represents the index of the word in M within the range of 0 to z . If the objects in triples consist of numeric values, the comparison between these numbers is straightforward and involves measuring the distance between them. To compare two distinct objects, we employ the Euclidean distance as a metric because of its simplicity, interpretability, and wide applicability in various domains. Let's consider two objects denoted as a_{o1} and a_{o2} , with respective vectors $\vec{a}_{o1} = \{o_{11}, o_{12}, \dots, o_{1p}\}$ and $\vec{a}_{o2} = \{o_{21}, o_{22}, \dots, o_{2p}\}$. In this context, their semantic similarity can be defined as follows:

$$Sim_2(\vec{a}_{o1}, \vec{a}_{o2}) = \frac{1}{1 + \sqrt{\sum_{t=1}^p (o_{1t} - o_{2t})^2}} \quad (4)$$

Hence, the semantic similarity between two sets of triples, $C_1 = \{a_{11}, a_{12}, \dots, a_{1g}\}$ and $C_2 = \{a_{21}, a_{22}, \dots, a_{2g}\}$, is determined by comparing the similarity of each individual triple. The calculation is performed as follows:

$$Sim(C_1, C_2) = \frac{1}{L} \left(\sum_{u=0}^L Sim_1(a_{1u}, a_{2u}) \right) + \frac{1}{H} \left(\sum_{v=0}^H Sim_2(\vec{a}_{1v}, \vec{a}_{2v}) \right) \quad (5)$$

where L denotes the number of triples containing qualitative objects, H represents the number of triples containing quantitative objects, g represents the number of triples in C_1 and C_2 , while u and v denote the indices of the qualitative triple set and quantitative triple set, respectively.

The measure of semantic similarity for different crisis situations entails the comparison of sets of triples representing crisis-related information in the ontology. This measure takes into account both quantitative and qualitative information present in the objects of the triples. For qualitative objects, the semantic similarity is calculated by utilizing numerical word vectors extracted from pre-trained word embeddings such as GloVe (Pennington et al., 2014). This enables the computation of the semantic relatedness between words. Conversely, for quantitative objects, the similarity is evaluated by computing the Euclidean distance between their numerical vectors. The overall semantic similarity between two sets of triples is then determined by comparing the similarity of individual triples within the sets.

4. Experiments

In this section, we present the details of the data used for enriching the knowledge bases and provide an overview of the results obtained from applying our method for measuring semantic similarity on the enriched knowledge base. By showcasing the outcomes of our experiments, we aim to demonstrate the efficacy and practical applicability of our approach in identifying similar crisis situations.

4.1. Enrichment of the Knowledge Base

We present our initial experiments centered around extracting and enriching our ontology-based knowledge base in the domain of crisis-related information. To accomplish this, we harness the valuable EM-DAT database Guha-Sapir et al. (2016). The EM-DAT database provides a comprehensive insights into various aspects of crises, including crucial details such as the dates and locations of events, the extent of their impact on affected populations (including figures for the number of individuals affected, injured, or deceased), economic losses incurred, and other relevant factors essential for comprehending the repercussions of these crises. Moreover, the database offers valuable information concerning the response and relief efforts implemented in the aftermath of these crisis situations. By incorporating data sourced from the EM-DAT database, we significantly enhance the depth and richness of our ontology-based knowledge base, thereby fostering a more comprehensive understanding of crisis-related information.

The database contains a comprehensive record of 25,982 crisis situations that occurred worldwide from 1903 to 2022. These events encompass a diverse range of crisis types, including natural crisis such as earthquakes, floods, storms, droughts, and epidemics,

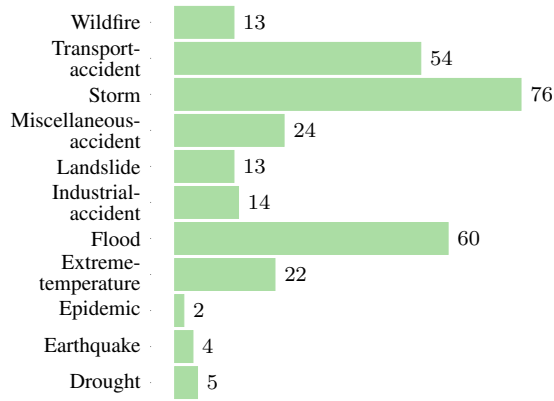


Figure 3. Number of crisis situations that occurred in France, classified by crisis type, from 1903 to 2022

as well as man-made crises such as industrial accidents, transportation incidents, and chemical spills. The coverage of the database spans across various countries, providing a global perspective on crisis occurrences. In this work, we specifically focus on the crisis situations that have occurred in France during this specified time period. Within this scope, we have identified approximately 287 distinct crisis situations that have taken place in France. To provide a visual representation of the distribution, a statistical breakdown of these diverse types of crisis situations is visually presented in Figure 3. This figure shows an informative overview of the different types of crises that have unfolded in France.

By mining crisis situations from the EM DAT database and integrating them with a domain ontology for crisis-related information, our knowledge base pertaining to various crisis situations have experienced a substantial enhancement. This process of mining and integration has facilitated the expansion and refinement of the organization of crisis-related information by incorporating structured and standardized data from a wide array of sources. Consequently, we have achieved notable improvements in semantic information retrieval within the knowledge bases.

The statistical data presented in table 1 demonstrates the substantial progress made in the crisis-related information ontology. By extracting and enriching a total of 287 crisis situations that took place in France, we have substantially broadened the ontology’s coverage, depth, and specificity. These advancements have greatly enhanced the ontology’s value in facilitating similar crisis recognition. They enable the extraction, simulation, and learning from past experiences, offering valuable insights to stakeholders and decision-makers involved in crisis situations, both within France and globally.

While RDF triples format is often associated with ontologies due to its ability to represent structured data

Table 1. Statistics on the Structure and Instances of the Crisis-Related Information Ontology

Knowledge base statistics	Quantitative
Number of Class	100
Number of Individual	2540
Number of Object Property	17
Number of Data Property	40
Number of Statements	100001

and relationships, it’s important to note that an RDF triples store can be created without incorporating an ontology. However, in our work, we have chosen to adopt an ontology-based approach to enhance the organization, semantics, and meaningful relationships within the domain of crisis-related information. This approach facilitates more robust analysis and reasoning in the subsequent stages of our work. In the next section, we present the experiments we conducted to recognize similar crises using the achieved knowledge base.

4.2. Measuring Similarity between Crises

In the second experiment, we conducted SPARQL queries to retrieve the crisis-related information of each crisis from the ontology-based knowledge base in the form of RDF graphs. Utilizing our method for measuring semantic similarity between RDF graphs, we compared each crisis event with one another, resulting in a similarity score. A higher similarity score indicates a greater degree of similarity between two crisis situations. Taking a closer look at the

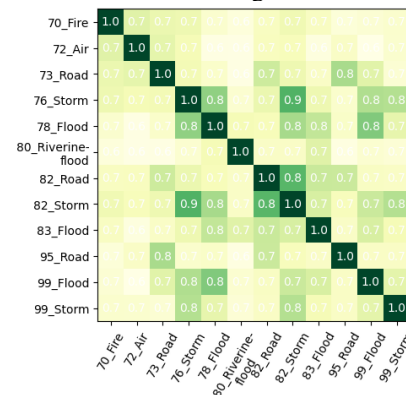


Figure 4. Similarity scores of the twelve randomly extracted crisis situations.

specific cases, we conducted a random comparison of twelve crisis situations by measuring similarity on their RDF triples. The main objective of this comparison was to evaluate the effectiveness of our similarity measure in capturing the similarities among these events. The computation results, as depicted in Figure 4, clearly demonstrate that crises of a similar type tend to exhibit higher similarity scores. This observation suggests that the semantic characteristics encapsulated within the RDF triples greatly facilitate the

identification and measure of similarities between crisis situations, particularly when they share comparable attributes or informative structures. For instance, when examining the specific events “80_RiverineFlood” and “83_Flood”, we can observe that the former has a higher similarity score, as illustrated in Figure 4. This further supports our claim that our approach effectively captures and quantifies the semantic relationships within the crisis-related information, thereby enhancing our ability to discern similarities between different crisis situations.

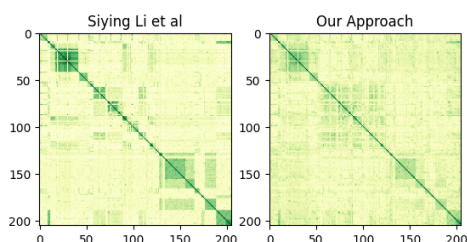


Figure 5. Comparison heat chart illustrating the similarity scores of our approach and that of Li et al. (2021)

To gain a comprehensive understanding of our similarity measure approach, we conducted a comparison with the approach proposed by Li et al. (2021). Specifically, we compared 204 crisis situations that occurred in France using both approaches and visualized the results in the form of a heat chart, as depicted in Figure 5. In general, our approach exhibits a smoother grouping of similar crises compared to Li’s approach. In terms of numerical statistics, our approach yielded significantly higher scores, with 41,376 identified crisis similarities, whereas Li’s approach only yielded 649. This indicates that our approach is more effective in capturing and quantifying the similarities between crisis situations, leading to a more comprehensive and accurate representation of crisis relationships. By conducting two comprehensive experiments focused on enriching the ontology-based knowledge base with crisis-related information and measuring semantic similarity between crises, our approach has demonstrated its effectiveness in identifying similar crises. The results obtained provide compelling evidence of the suitability of our approach in this purpose

5. Conclusion and Perspectives

In this paper, we conducted an investigation to recognize similar crises through the application of ontology-based knowledge mining. Our research encompasses two main aspects: (i) the development and enrichment of a crisis-related information ontology to facilitate the construction of the ontology-based knowledge base for crisis situations; and (ii) the

application of ontology-based knowledge mining into the implementation of semantic similarity measures for computing relatedness between crisis situations. Through a systematic process cycle for identifying similar crisis situations, our experiments demonstrated the successful adaptation of existing crisis-related data into our ontology-based knowledge bases. Furthermore, our approach provided compelling evidence of its effectiveness in computing similarity scores between crisis situations. It is essential to acknowledge the limitations of our work, particularly in practical applications where the knowledge bases should vary in terms of the information they contain regarding different crisis scenarios, actions, and corresponding communication. Therefore, it is crucial to proactively construct a comprehensive framework that leverages established standard ontologies for the knowledge base. Additionally, the similarity calculations rely on crisis instances from the ontology, which also shapes data organization. While our evaluation confirms semantic similarity’s effectiveness, it is necessary to showcase how the ontology improves outcomes. Thus, we plan to compare results with and without ontology integration, highlighting its impact on similarity assessment.

Acknowledgement

This work was funded by the French Research Agency (ANR) and by the company Vivocaz under the project France Relance - preservation of R&D employment (ANR-21-PRRD-0072-01).

References

- Abdelghany, A. S., Darwish, N. R., and Hefni, H. A. (2019). An agile methodology for ontology development. *International Journal of Intelligent Engineering and Systems*, 12(2):170–181.
- Abiteboul, S., Manolescu, I., Rigaux, P., Rousset, M.-C., and Senellart, P. (2011). *Ontologies, RDF, and OWL*, page 143–170. Cambridge University Press.
- Arp, R., Smith, B., and Spear, A. D. (2015). *Building ontologies with basic formal ontology*. Mit Press.
- Arru, M., Negre, E., and Rosenthal-Sabroux, C. (2019). To Alert or Not to Alert? That Is the Question. In *52nd Hawaii International Conference on System Sciences (HICSS 52)*, Proceedings of the 52nd Hawaii International Conference on System Sciences — 2019, pages 649–658, Wailea, HI, United States.
- Benaben, F., Fertier, A., Montarnal, A., Mu, W., Jiang, Z., Truptil, S., Barthe-Delanoë, A.-M., Lauras, M., Mace-Ramete, G., Wang, T., et al. (2020). An ai framework and a metamodel

- for collaborative situations: Application to crisis management contexts. *Journal of Contingencies and Crisis Management*, 28(3):291–306.
- Benaben, F., Hanachi, C., Lauras, M., Couget, P., and Chapurlat, V. (2008). A metamodel and its ontology to guide crisis characterization and its collaborative management. In *Proceedings of the 5th International Conference on Information Systems for Crisis Response and Management (ISCRAM)*, Washington, DC, USA, May, pages 4–7.
- Benaben, F., Lauras, M., Truptil, S., and Salatgé, N. (2016). A metamodel for knowledge management in crisis management. In *2016 49th Hawaii International Conference on System Sciences (HICSS)*, pages 126–135. IEEE.
- Boin, A., Lodge, M., and Luesink, M. (2020). Learning from the covid-19 crisis: an initial analysis of national responses. *Policy Design and Practice*, 3(3):189–204.
- Boin, A. and McConnell, A. (2007). Preparing for critical infrastructure breakdowns: the limits of crisis management and the need for resilience. *Journal of contingencies and crisis management*, 15(1):50–59.
- Bowen, S. A. and Lovari, A. (2020). *Crisis Management*, pages 1–10. Springer International Publishing, Cham.
- Bundy, J., Pfarrer, M. D., Short, C. E., and Coombs, W. T. (2017). Crises and crisis management: Integration, interpretation, and research development. *Journal of management*, 43(6):1661–1692.
- De Nicola, A., Melchiori, M., and Villani, M. L. (2019). Creative design of emergency management scenarios driven by semantics: An application to smart cities. *Information Systems*, 81:21–48.
- Elmhadhbi, L., Karray, M.-H., Archimède, B., Otte, J. N., and Smith, B. (2020). Promes: An ontology-based messaging service for semantically interoperable information exchange during disaster response. *Journal of Contingencies and Crisis Management*, 28(3):324–338.
- Gangemi, A., Guarino, N., Masolo, C., Oltramari, A., and Schneider, L. (2002). Sweetening ontologies with dolce. In *Knowledge Engineering and Knowledge Management: Ontologies and the Semantic Web: 13th International Conference, EKAW 2002 Sigüenza, Spain, October 1–4, 2002 Proceedings 13*, pages 166–181. Springer.
- Gruber, T. R. (1993). A translation approach to portable ontology specifications. *Knowledge acquisition*, 5(2):199–220.
- Guha-Sapir, D., Below, R., and Hoyois, P. (2016). Em-dat: the cred/ofda international disaster database.
- Le, N. L., Abel, M.-H., and Gouspillou, P. (2022). Apport des ontologies pour le calcul de la similarité sémantique au sein d’un système de recommandation. In *Ingénierie des Connaissances (Evènement affilié à PFIA’22 Plate-Forme Intelligence Artificielle)*, Saint-Étienne, France.
- Le, N. L., Abel, M.-H., and Gouspillou, P. (2023a). Improving semantic similarity measure within a recommender system based-on rdf graphs. In *Proceedings of the 6th International Conference on Information Technology & Systems*.
- Le, N. L., Zhong, J., Negre, E., and Abel, M.-H. (2023b). Constraint-based recommender system for crisis management simulations. In *The 56th Hawaii International Conference on System Sciences*, Hawaii, United States.
- Li, S., Abel, M.-H., and Negre, E. (2021). Ontology-based semantic similarity in generating context-aware collaborator recommendations. In *2021 IEEE 24th International Conference on Computer Supported Cooperative Work in Design (CSCWD)*, pages 751–756. IEEE.
- Meesters, K. (2021). Crisis information management: From technological potential to societal impact. *The New Common: How the COVID-19 Pandemic is Transforming Society*, pages 153–159.
- Musen, M. A. (2015). The protégé project: a look back and a look forward. *AI matters*, 1(4):4–12.
- Negre, E. (2013). Towards a knowledge (experience)-based recommender system for crisis management. In *2013 Eighth International Conference on P2P, Parallel, Grid, Cloud and Internet Computing*, pages 713–718. IEEE.
- Niles, I. and Pease, A. (2001). Towards a standard upper ontology. In *Proceedings of the international conference on Formal Ontology in Information Systems-Volume 2001*, pages 2–9.
- Pennington, J., Socher, R., and Manning, C. D. (2014). Glove: Global vectors for word representation. In *Proceedings of the 2014 conference on empirical methods in natural language processing (EMNLP)*.
- Yu, F., Li, X.-Y., and Han, X.-S. (2018). Risk response for urban water supply network using case-based reasoning during a natural disaster. *Safety Science*.
- ‘t Hart, P. (2023). Teaching crisis management before and after the pandemic: Personal reflections. *Teaching Public Administration*, 41(1):72–81.