Ontology Learning System (OLS)

Sundos Al Subhi, Dr. Chetan Tiwari, and Dr. Armin R. Mikler
Department of Computer Science, Georgia State University

Introduction & Motivation

In the United States, floods are the second deadliest of all-weather-related hazards with approximately 98 deaths per year [2]. Planning for flooding events is critical to mitigate their impacts on society. When it comes to handling information about disasters, planning and responding agencies needs to collaborate together. However, it is hard for different groups to share information because they use different systems for managing large amounts of data and information. This makes it complicated to organize and share information in a way that everyone can understand it. Also, it is harder to coordinate effectively, which can slow down response efforts.

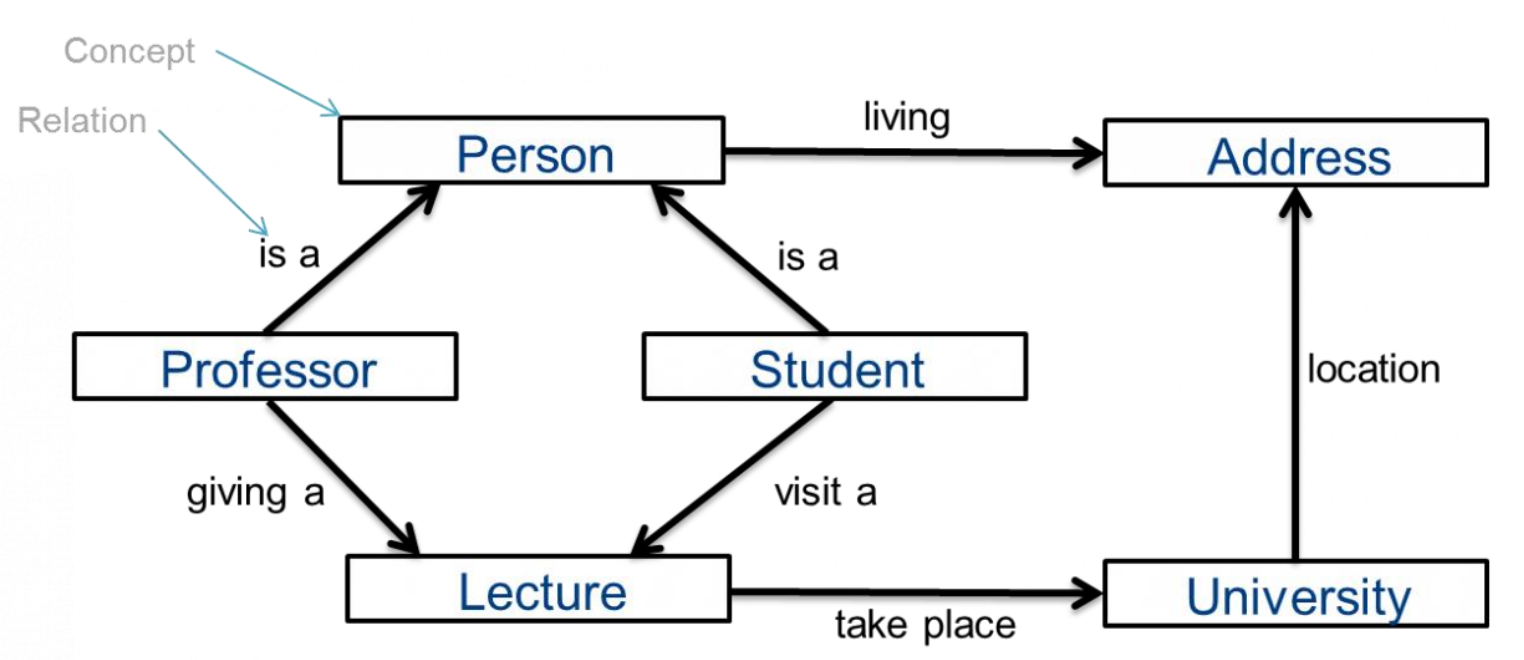


Figure 1. Example of an Ontology [3]

Ontologies present a domain of knowledge, which contains concepts and relationships between the concepts [5]. Ontologies can be used to organize unstructured data, such as text data into a formal conceptualization of a particular domain [1]. However, the process of creating hazard ontologies requires domain expert and it is a time and effort-intensive process.

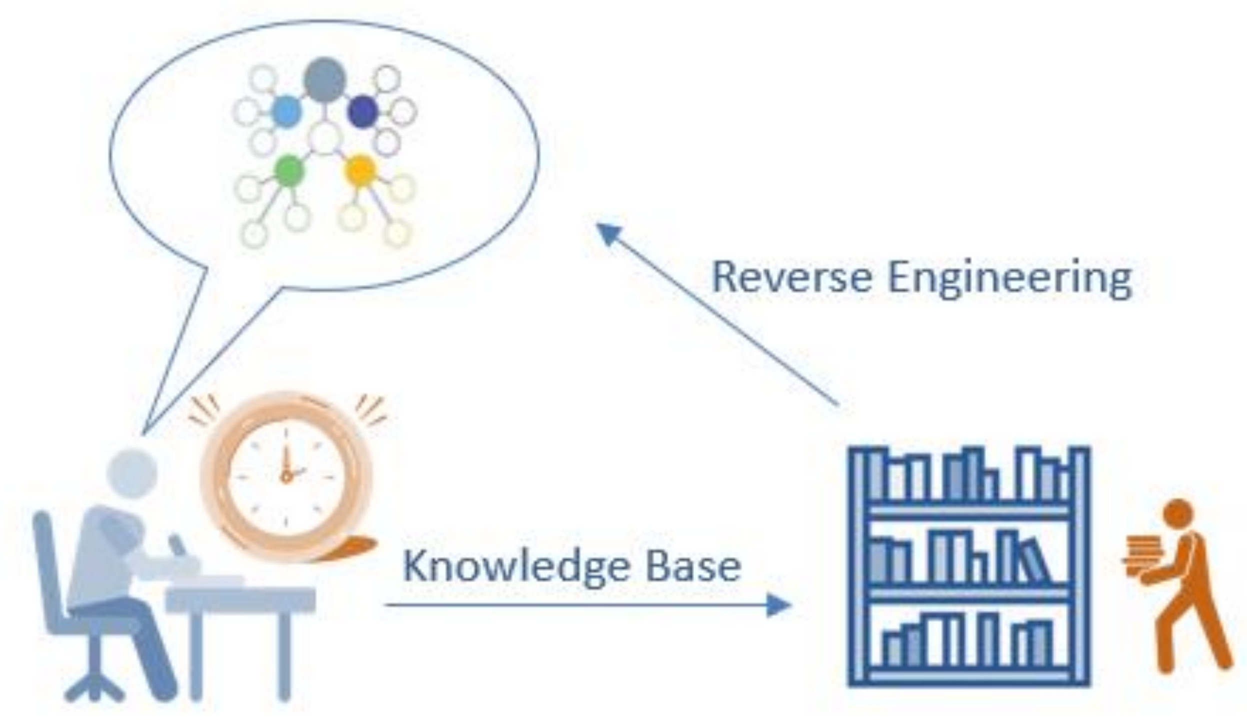


Figure 2. Research Problem

Research Objective

Research objective focuses on applying ontology learning techniques to automate the development of hazard-specific ontologies from knowledge bases of disaster-related information (e.g., scholarly articles).

Dataset & Methods

Dataset: Academic papers, technical reports, and authoritative web resources (e.g., Federal Emergency Management Agency (FEMA)).

Methods:
Ontology learning has been used to automate the construction of ontologies through the development of automated techniques to extract terms, synonyms, concepts, taxonomies, etc. from different data sources [6].

An Ontology Learning System (OLS) is a large and complex framework that encompasses various steps including data processing and information extraction. The steps within the (OLS) are the following:

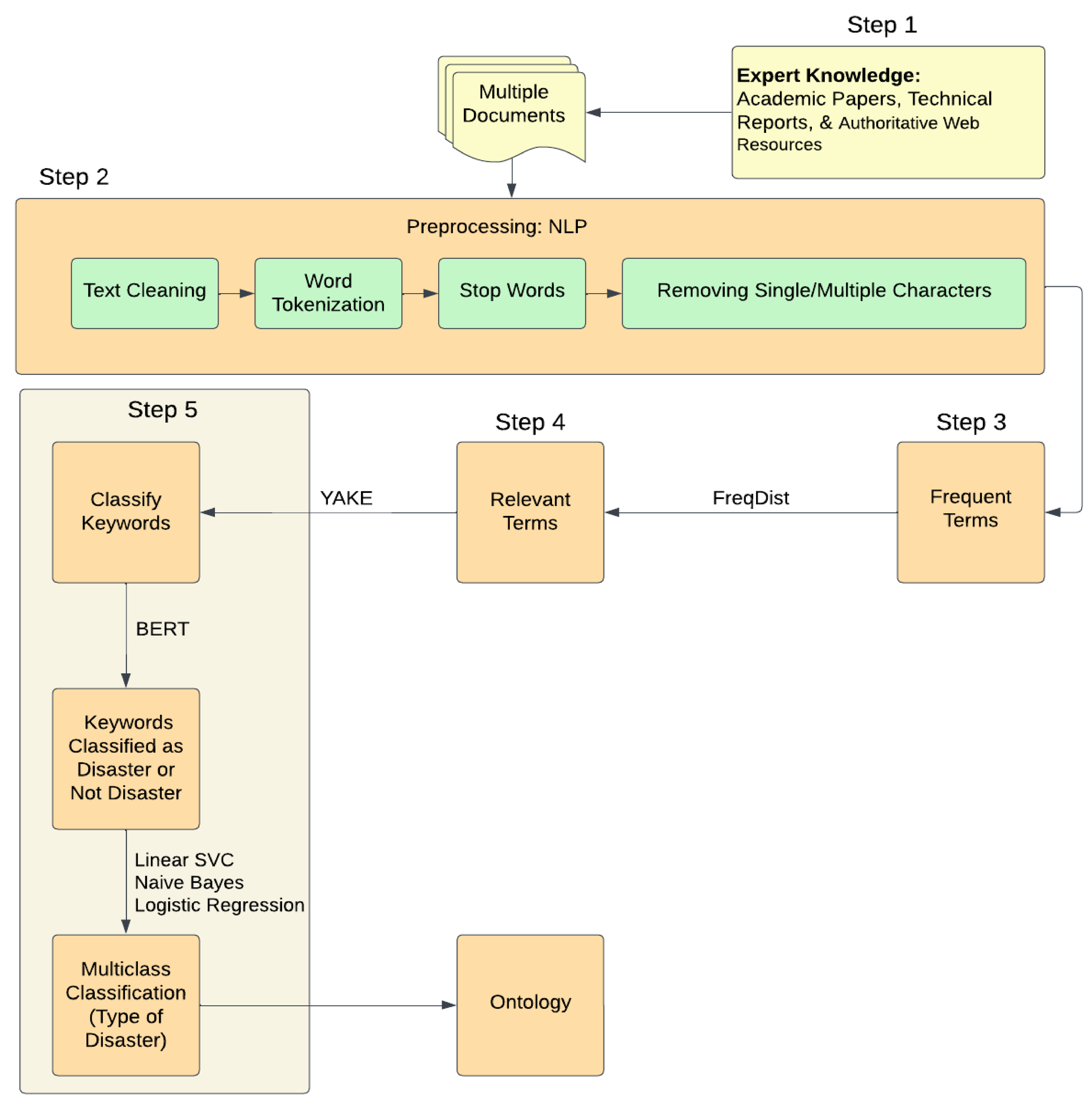


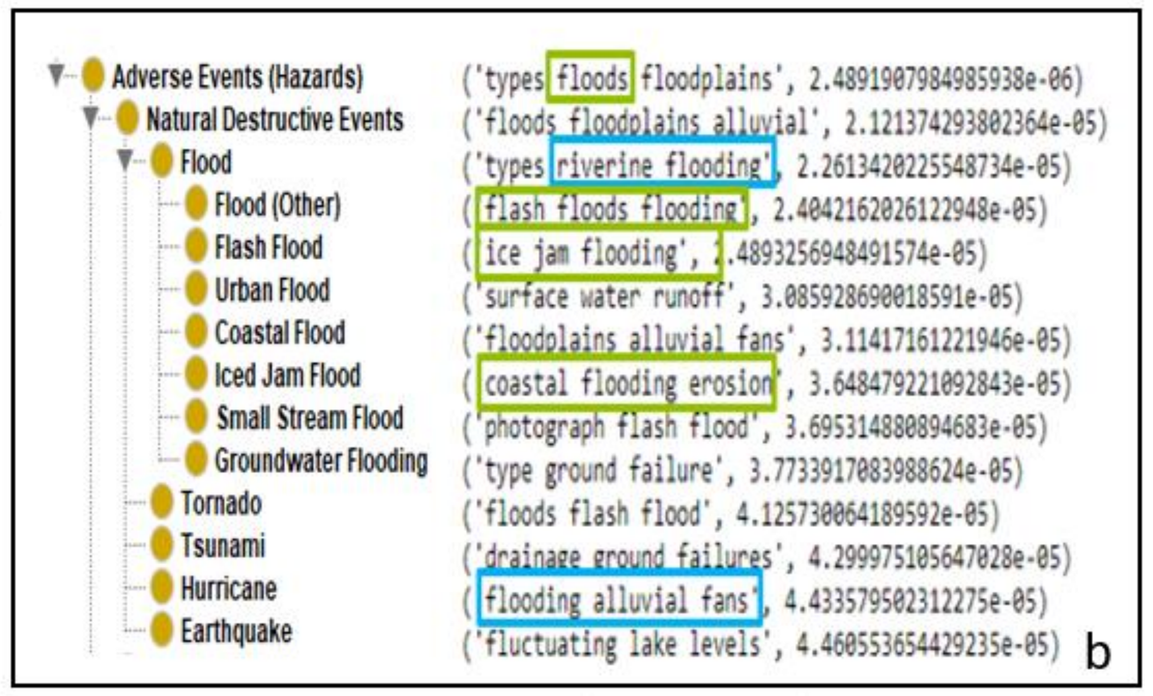
Figure 3. Overview of Ontology Learning System (OLS)

Results

The FEMA technical report results (Figure 4-a) display a frequency table with "water," "flooding," and "flood" as the top three terms. In Figure 4-b, the green squares highlight keywords aligning with ontology categories from manual flood ontology created based on the literature. Blue squares suggest keywords for potential additional categories not in the ontology. Figure 4-c outcomes classify terms like "Flash Floods Flooding" and "Ice Jam Flooding" as "Disaster," and "Type Ground Failure" as "Not a disaster," achieving 89% accuracy, precision, and recall in disaster classification.

Word	Frequency
water	98
flooding	72
flood	42
floods	42
storm	37
types	33
areas	33
ice	28
drainage	25
floodplains	24

a



Keywords	Output	Classification
'flash floods flooding'	0.95	Disaster
'ice jam flooding'	0.89	
'coastal flooding erosion'	0.87	
'type ground failure'	0.09	Not Disaster
'surface water runoff'	0.18	
'fluctuating lake levels'	0.33	

c

Figure 4. The FEMA Technical Report Results

A multiclass classification approach was tested for categorizing disaster-related keywords, with preliminary results showing Naive Bayes at 90%, Logistic Regression at 97%, and Linear SVC at 98.4% accuracy for identifying specific types of disasters (e.g., flood or hurricane).

Limitations

The results of the system have the following limitations:
(1) Limited by small datasets (one technical report and a compilation of six articles). (2) Variations in ontology structures hinder validation using a common standard [4]. (3) Findings lack validation by an expert (4) System can identify keywords and concepts but not their semantic relationships.

Contact

Sundos Al Subhi
salsubhi1@student.gsu.edu
College of Arts and Sciences
Department of Computer Science
Disaster Informatics & Computational Epidemiology

References

[1] Grigoris Antoniou and Frank van Harmelen. 2008. A Semantic Web Primer, 2nd Edition (Cooperative Information Systems). The MIT Press.
[2] CDC. Precipitation extremes: Heavy rainfall, flooding, and droughts | CDC [online]. 2020.
[3] COCOP. Knowledge management in COCOP [online]. 2024.
[4] Alon Halevy. 2005. Why your data won't mix: New tools and techniques can help ease the pain of reconciling schemas. Queue, 3(8):50–58.
[5] Ontotext. What are ontologies? [online]. 2024.
[6] Gerhard Wohlgenannt and Filip Minic. 2016. Using word2vec to build a simple ontology learning system. In International Workshop on the Semantic Web.