

empathi: An ontology for Emergency Managing and Planning about Hazard Crisis

Manas Gaur[†], Saeedeh Shekarpour^{*}, Amelie Gyrard[†], Amit Sheth[†]

^{*} University of Dayton, Dayton, Ohio, USA, sshekarpour1@udayton.edu

[†] Kno.e.sis Center, Dayton, Ohio, USA, {manas, amelie, amit}@knoesis.org

Abstract—In the domain of emergency management during hazard crises, having sufficient situational awareness information is critical. It requires capturing and integrating information from sources such as satellite images, local sensors and social media content generated by local people. A bold obstacle to capturing, representing and integrating such heterogeneous and diverse information is lack of a proper ontology which properly conceptualizes this domain, aggregates and unifies datasets. Thus, in this paper, we introduce *empathi* ontology which conceptualizes the core concepts describing the domain of emergency managing and planning of hazard crises. Although *empathi* has a coarse-grained view, it considers the necessary concepts and relations being essential in this domain. This ontology is available at <https://w3id.org/empathi/>.

Index Terms—Ontology, Vocabularies, Crisis Management, Hazard Domain, Emergency, Ontology Quality, Knowledge Reuse, Disaster Management.

I. INTRODUCTION

We can not change the nature. However, we can promote our planning, preparation, and response strategies about crises happening in the three phases of hazards, i.e. (i) pre-hazard, (ii) in-hazard and (iii) post-hazard. Currently, a substantial body of situational information collected from sources such as satellite images, sensors, social media content generated by people who are involved in crisis reporting and response, etc. Indeed, harnessing and exploiting this *hazard-related Big Data* is an essential means towards situational awareness which helps to manage crises, threads and risks of hazards in each phase. Despite the availability of such data, still there is a significant deficiency in **representing, annotating** and more importantly **integrating** this heterogeneous hazard-related big data. This deficiency can be relieved by providing an ontology which *conceptualizes* and *organizes* situational and environmental awareness data (events, activities) subjected to hazards of any kind. Our investigation in the state-of-the-art hazard-related conceptualization (i.e., taxonomy, vocabulary and ontology) revealed existing of a few works which mainly conceptualized either hazard domain or crisis management domain from a limited perspective or for a particular type. For example, Humanitarian eXchange Language (HXL) [1] and EDXL-RESCUER Ontology [2] are mainly concerned with help and rescue aspect of crisis management domain. However, the broader and more diverse nature of this domain requires a **comprehensive** and **abstract** modeling and representation. An abstract model of ontology is needed to adapt with evolving ontology standards and existing ontologies. To capture contextual details of a domain, comprehensiveness of the

ontology is required. Another deficiency is related to lack of incorporating relations into the conceptualization. It is needed to semantically relate concepts within the disaster domain to define the context of the problem. For example, *Affected Population, Organization, and Service* are three individual concepts of a disaster scenario, however, they need to be linked to define the context. We define *needHelp* as the relevant relation linking these concepts. Thus, they indeed should be called vocabulary or taxonomy rather than ontology (e.g., Management of Crisis vocabulary (MOAC)). These shortages motivated us to introduce an ontology that takes into consideration the prior conceptualizations (taxonomies and vocabularies) while relying on a promoted representation. Also, it brings new concepts and relations which were ignored previously although playing an essential role in capturing situational awareness (e.g., surveillance information, human sensing report, humanitarian event: prayer, concepts synonyms and instances). This paper presents our contribution in providing an ontology for Emergency Managing and Planning about Hazard Crises (**empathi**). The ontology will be used in automatic recognition of disaster concepts mentioned in social media conversation thus facilitating real-time monitoring by enforcement agencies.

The rest of the paper is organized as follows: Section II presents the relevant state-of-the-art hazard or crisis vocabularies and compare their main features. Section III lists the external vocabularies which are integrated within *empathi*. Thereafter, Section IV introduces *empathi* along with its major top concepts. Section VII reviews the related work. We close with the conclusion and the future work.

II. STATE-OF-THE-ART VOCABULARIES

In this section, we present an overview of the state-of-the-art vocabularies concerned with hazard domain as well as crisis management domain. With this respect, Table I represents a succinct comparative study of these vocabularies. The first column (i.e., **Vocabulary**) states the name of the state-of-the-art vocabulary or ontology and *empathi* is included in the last row (based on recency). The second column (i.e., **Domain Coverage**) mentions the particular areas of the hazard-related domain or crisis management domain which it covers. The third column (i.e., **URI**) checks whether or not the URI of the vocabulary is dereferenceable. The fourth column (i.e., **F for File**) specifies the available format of the vocabulary (i.e., OWL, RDF, TTL). The fifth column (i.e., **D for Documentation**) shows whether the vocabulary has online documentation

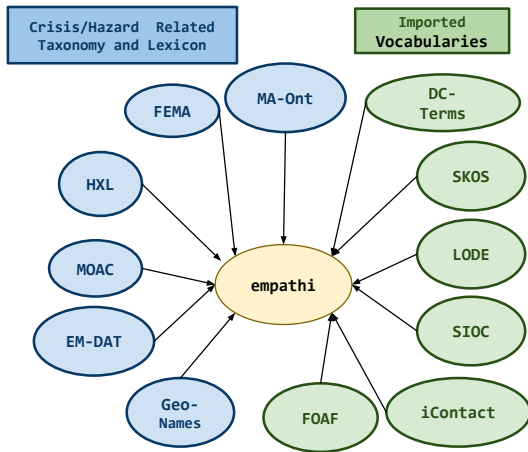


Figure 1: Importing existing ontologies in *empathi*

or not. Then, we list the significant publications utilizing this vocabulary within the sixth column (i.e., **MC for Major Citations**). The seventh and eighth columns represent the number of classes **#C** and relations **#R** specified within the vocabulary. The last column represents the external resources (i.e., **IV for Imported Vocabularies**) imported by their respective vocabulary in the first column. In the following, we shortly describe each vocabulary.

a) **HXL**: HXL stands for Humanitarian eXchange Language. HXL¹ is a standard aiming at information sharing during humanitarian calamity by overcoming the burden of interoperability. HXL ontology has a total of 50 classes and 66 relations. Main concepts contained in HXL are Place, Survey and assessment, Operation, Cash and Finance, Crisis. Furthermore, HXL provides links to UN OCHA vocabularies such as Global Coordination Groups², Disaster Types³, Organization Types⁴, Vulnerable groups and Humanitarian themes⁵. Furthermore, HXL provides a hashtag schema containing related social media tags such as #channel, #crisis, #impact, #event, etc. [1].

b) **MOAC**: MOAC⁶ which is concerned with management of crisis is a vocabulary [40] providing concepts mainly related to crisis management. It was created by the Inter-Agency Standing Committee (IASC)⁷, Emergency Shelter Cluster in Haiti⁸, UN-OCHA 3W Who What Where Contact Database⁹ and Ushahidi Platform¹⁰.

c) **DoRES**: DoRES stands for Document-Report-Event-Situation Ontology. DoRES¹¹ is an ontology sharing information between individuals and organizations using situational

¹<https://github.com/hxl-team/HXL-Vocab/blob/master/Tools/hxl.ttl>

²<https://goo.gl/CD6vHY>

³<https://reliefweb.int/taxonomy-descriptions#disastertype>

⁴<https://goo.gl/Uzy9UA>

⁵<http://vocabulary.unocha.org>

⁶<http://observedchange.com/moac/ns>

⁷<https://goo.gl/ESn99F>

⁸<https://goo.gl/nDoa9F>

⁹<https://goo.gl/jQLnYh>

¹⁰<https://goo.gl/XPSnyG>

¹¹<https://goo.gl/Sw4XGt>

reports for describing the situation [12]. It helps humanitarian organizations to structure their plans.

d) **EDXL-RESCUER Ontology**: EDXL-RESCUER¹² stands for Emergency Data Exchange Language Reliable and Smart Crowdsourcing Solution for Emergency and Crisis Management. An ontology-based on EDXL developed for coordinating and interchanging information with the legacy system [32], [2].

e) **Emergency Fire (EF)**: It is an ontology explicitly designed for fires in the building. It comprises of 131 terms along with definitions created after subjective research. It serves as a protocol for information sharing, analysis, evaluation and comprehension by an organization in the situation of disaster caused by fire [41].

f) **Social Media Emergency Management (SMEM)**: During an unprecedented onset of a natural hazard, social media overflows with textual content about situational information, prayers, weather information, the impact of crisis and events. Of all the available information on the social media, which portion is contextually relevant and action-oriented to personnel in charge of a relief-giver organization. SMEM ontology provides concepts in a hierarchical structure which transforms high-volume of messy content to low-volume of action-oriented information [42].

g) **Disaster Ontology**: It is one of the ontologies listed in Finnish Ontology Library Service ONKI¹³. Disaster Ontology (DO)¹⁴ is comprised of 97 concepts (classes) concerning man-made and natural hazard. This ontology is useful for managing disaster situations but disregards concepts related to social media (e.g., news reports, modality of data, surveillance, prayer and monitoring the status of the services provided by organizations).

III. INTEGRATION OF EXTERNAL VOCABULARIES

In this section, we list the external vocabularies which partially integrated into *empathi*. Not all of them are necessarily related to the domain of hazard or crisis management (we reuse generic concepts from well-known vocabularies, e.g., FOAF). Figure 1 concisely represents an integration aims at reusing the existing vocabularies following ontology design methodologies (Methontology [33] and NeOn [43]) or interlinking *empathi* to other vocabularies which enhance its visibility.

- **Federal Emergency Management Agency (FEMA)**¹⁵ provides a glossary of terms related to disaster preparation and management [37].
- **Emergency Disasters Database (EM-DAT)**¹⁶ provides a precise definition of concepts and furthermore a categorization of disturbance-related events [19].
- **MA-Ont**¹⁷ supports detailed properties describing media files and appropriate metadata mapping [20].

¹²<http://www.rescuer-project.org>

¹³<https://onki.fi/en/>

¹⁴<http://onki.fi/en/browser/overview/disaster>

¹⁵<https://goo.gl/QtKzev>

¹⁶<http://www.emdat.be/Glossary.Itisadatabase>

¹⁷<https://www.w3.org/ns/ma-ont>

Vocabulary	Domain Coverage	U	F	D	MC	#C	#R	IV
HXL	Disaster, Geography, Damage, Organization, Humanitarian Response	✓	TTL	✓	[3], [1], [4] [8], [9], [10], [11], [12]	50	66	[5], [6], [7]
MOAC	Impact of Crisis, Recovery and Response Activities, Geo-locations	✗	RDF	✓	[3], [13], [6] [9], [17], [18]	70	30	[14], [15], [16]
SMEM	Social Media and Emergency Management	✗	✗	✓	-	-	-	[19], [20], [21], [22], [6], [7] [23], [24], [1], [25], [26], [27]
DO	Temporal and Spatial Concepts, Impact, Rehabilitation and Facilities	✓	Web-App	✗	-	97	-	-
ERO	Report Specification	✗	✗	✓	[28], [29], [30], [31], [32]	-	-	[33], [34], [35]
DoRES	Events and Reports Specification	✓	RDF	✓	-	96	261	[6], [7] [23], [36]
EF	Fire Disaster Specification, Protocol Design and Planning	✗	✗	✓	-	37	90	-
empathi	Hazard Situational Awareness, Crisis Management, Hazard Events	✓	OWL	✓	-	423	338	[37], [19], [20], [7], [6] [36], [21], [23], [38], [39]

Table I: Comparison of the state-of-the-art Hazard-related Vocabularies, Taxonomy, and Ontologies. U, F, D respectively stand for referenceability of URI, availability of File and Documentation. Furthermore, MC, C, R and IV respectively stand for Major Citations, number of classes, number of relations, and imported vocabularies. ERO: EDXL-RESCUER Ontology, DO: Disaster Ontology, EF: Emergency Fire.

- **iContact**¹⁸ provides conceptual classes for defining international addresses. It is relevant for using GeoNames [36], [21] an ontology for describing places.
- **Friend Of A Friend (FOAF)**¹⁹ was created for describing people, relations, and associated events. Coupled with SIOC [23] and disaster domain model [39], [38], it can describe social media communities formed during the disaster scenarios [6].
- **GeoNames**²⁰ is a part of GeoNames Database providing information about 11 million places (toponyms) covering all the countries. Integration of GeoNames ontology to our ontology adds geospatial semantic information which is critical for actionable hazard response. The ontology contains 150 classes and two relations forming 758 axioms on location dereferencing. Mapping syntax provided by this ontology is compatible with schema.org, DBpedia ontology, LinkedGeoData ontology and INSEE ontology [36].
- **Linked Open Descriptions of Events (LODE)**²¹ defines event as an action which takes place at a certain time and has a specific location. It can be a historical action as well as a scheduled action. Thus, it provides the generic concept of Event along with locational (i.e., *atPlace*), temporal (i.e., *atTime*) aspects and people who play a role (i.e., *involvedAgent*).
- **Simple Knowledge Organization System (SKOS)**²². We utilized this data model to describe the concepts of our domain. It provides a better organization of domain knowledge (i.e., Hazard Crisis) [7].
- **Semantically-Interlinked Online Communities (SIOC)**²³ is a W3C ontological standard to describe information from online communities. It can support a

volunteer or caregiver with actionable information in the realm of social media [23].

IV. CORE CONCEPTS OF EMPATHI

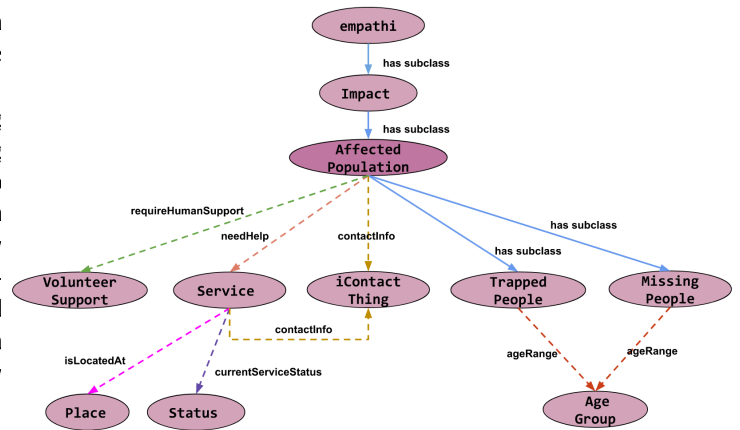


Figure 2: Partial representation of the concept Affected Population in *empathi*

As Table I shows, *empathi* contains 423 classes and 338 relations. In Figure 2, concepts linked to “Affected Population” via solid lines, forms structural concepts (“is-a”/ “subclass”), while the concepts linked via colored dotted line are semantically related concepts to “Affected Population”. Furthermore, in Figure 2, the concept “iContactThing” is taken from the vocabulary *iContact* shown in Figure 2. In the following, we present only the super-classes which imply the generic coverage.

- **Age Group:** classifies people based on their age similarity by providing the following sub-classes (i) Adolescent, (ii) Adult, (iii) Child and (iv) Infant.
- **Event:** defines events along with spatial and temporal constraints happening in any phase of hazard. This concept embodies the following sub-classes (i) Climate Change (ii) Criminal Activity (iii) Emergency Exercises

¹⁸ <http://ontology.eil.utoronto.ca/icontact.html>

¹⁹ <http://xmlns.com/foaf/spec/>

²⁰ <http://www.geonames.org/ontology/documentation.html>

²¹ <http://linkedevents.org/ontology/>

²² <https://www.w3.org/2004/02/skos/>

²³ <http://rdfs.org/sioc/spec/>

Hazard Type	Tweet	Hazard Concept
Flood	1. 188 killed, Airport closed they say that the runway came peeling off. it may take time to resume.	Impact (AP, ID)
	2. @? Hope people get adequate relief and no one is left out.Nation stands with Chennai.	Event (HP)
	3. @? in who can offer places to stay, pls fill the form for volunteers collating info #chennairains	Event (VS)
	4. @? Chennai has been declared disaster zone. Army has been deployed. Army Helpline - +XXX XXX XXXX #ChennaiRains	Service (H)
Hurricane	5. Surveillance video captures several people looting a Houston store #Harvey in the wake of Hurricane	Event (CA)
	6. The aftermath. 10,000 people now homeless because of Hurricane Irma. #WednesdayWisdom #climatechange	Impact (AP), Event(CC)
	7. #HurricaneIrmaRecovery Drive for #Homestead & #FloridaKeys today! Drop off supplies at @? #DJLMS #dontgivebackjustgive.	Service (S)
	8. At least 56 of Florida's 639 nursing homes still have *no* electricity this morning, five days after #HurricaneIrma	Impact (ID)
Blizzard	9. @? "2-3 days, it could take before airlines begin to clear the backlog? @? at on #blizzard2016.	Impact (ID)
	10. @?: I-75 in Kentucky closed due to large number of accidents, state patrol says #blizzard2016	Impact (AP,ID)
	11. @?: The baton is passed. Buoy 50 miles south of Wilmington #Jonas #blizzard2016	Place (L)
	12. @?: Blizzard with "life and death implications" hits Washington, Mid-Atlantic #blizzard2016	Place (L), Impact (AP)
Landslide	13. @?: Be sure to follow @BGSLandslides for lots of up to date information on landslides across the UK #StormFrank	Report (ER)
	14. #StormFrank landslide at Glasscarriga Norman Motte & Bailey in Co Wexford — @? #archaeology #floods	Place (L)
	15. SRI LANKA: At least 73 dead after week of flooding, landslides; 243,000 in temp shelters — TorStar #ExtremeWeather	Service (SH), Impact (AP)
	16. @?: #EcuadorEarthquake - landslides closing down roads & making it challenging for help to reach hardest hit towns	Impact (ID), Hazard Type (*)

Table II: Sample of hazard-related tweets from different hazard types. AP: Affected Population, ID: Infrastructure Damage, HP: Human Prayer, VS: Volunteer Support, H: Helpline, CA: Criminal Activity, S: Supply, L: Location, ER: Expert Report, SH: Shelter, Hazard Type (*): One hazard (Earthquake) causing another hazard (Landslide).

- (iv) Evacuation Plan (v) Humanitarian Event (vi) Recovery Plan (vii) Rescue Operation (viii) State Mitigation Plan (ix) Volunteer Support and (x) Early Warning.
- **Facility:** defines an amenity made accessible for a specific purpose. It attributes to following sub-classes (i) Communication (ii) Electricity (iii) Gas Facility (iv) Water Facility and (v) Education Resource.
- **Hazard Type:** lists different types of hazards that can affect human community. It is an entity type that embodies sub-classes (i) Airburst (ii) Coastal erosion (iii) Drought (iv) Earthquake (v) Explosion (vi) Fire (vii) Flood (viii) Hurricane (ix) Landslide (x) Sandstorm (xi) Storm (xii) Tornado (xiii) Toxic Radioactivity (xiv) Tsunami (xv) Volcano and (xvi) Winterstorm.
- **Hazard Phase:** categorizes different activities carried out by various organizations before, during and after a catastrophic event into three sub-classes (i) During Hazard (ii) Pre-Hazard and (iii) Post-Hazard.
- **Impact:** affects someone or something in an unprecedented manner. This concept embodies sub-classes (i) Affected Population (ii) Animal Loss (iii) Health Issues (iv) Food Shortage (v) Financial Crisis (vi) Contamination (vii) Infrastructure Damage (viii) Severity.
- **Involved Actors:** associates People or Organisation with any catastrophic event. Sub-classes of this concept are (i) Organisation (ii) People.
- **Modality of Data:** conveys or represents Information (raw, structured or semi-structured) using a sequence of text, audio, video or photos. Sub-classes included by this concept are (i) Audio (ii) Photo (iii) Text and (iv) Video.
- **Place:** defines a physical surrounding by longitude, latitude, and area, providing a relative position of the someone or something during a hazard situation. One sub-class of the Place is Location, described by longitude, latitude, and the area of the affected place.
- **Report:** documents evidence of the destruction caused by the natural disaster. All the activities carried out by various governmental and non-governmental organizations (NGOs) are stated in the report. The report is a way to keep people vigilant. Sub-classes encompassed under the concept: Report is (i) Expert Report (ii) Human Sensing Report and (iii) Media Report.
- **Service:** is an act of providing support to someone in a situation of distressing incidents. Core sub-classes of this concept are (i) Financial Care (ii) Healthcare Service (iii) Helpline (iv) Human Remains Management (v) Resource and Information Centre (vi) Supply (vii) Transportation and (viii) Prayer Location.
- **Status:** defines the state of services that are planned during the pre/in/post hazard phases. Associated sub-classes are (i) Available (ii) Offered (iii) Requested and (iv) Unavailable.
- **Surveillance Information:** A systematic, ongoing collection, collation, and analysis of data and the timely dissemination of information to those who need to know so that action can be taken. The surveillance concept in the setting of natural disasters can help to identify the resulting health-related needs which in turn, will lead to

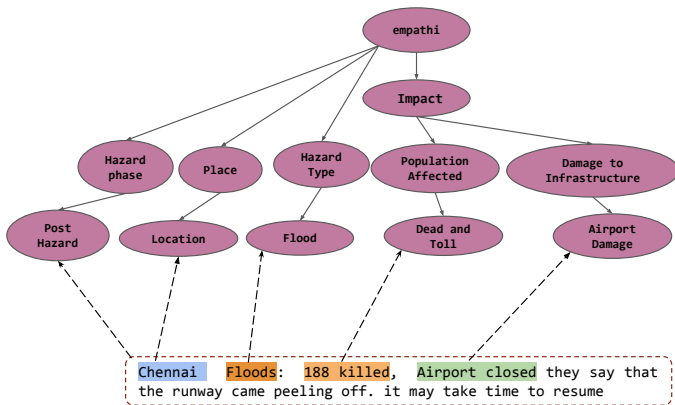


Figure 3: Mapping tweet and its words to *empathi* concepts.

the more rational and effective deployment of resources to affected populations.

V. CASE STUDY

A substantial impact of *empathi* ontology is empowering us to semantically annotate text (e.g., tweets posted during hazard). Figure 3 shows mapping of segments within a given tweet to *empathi* concepts. Thus, from an abstract level, this tweet is related to the flood occurred in Chennai and reports two associated impacts (*Airport Closed*, *188 Killed*). With this respect, in a first experiment, we compiled 53 million tweets from the 30 significant hazards happened in the past²⁴. Such as Hurricane Sandy in 2012 and Irma in 2017, Oklahoma Wildfires in 2017, Chennai Flood in 2005, Alaska Earthquake in 2018, Florida Rains in 2000 and 2016, Houston Floods in 2017, New-Zealand Earthquake in 2016, Typhoon Haima in 2016, Winter Storm Kayla in 2016, and many more. After that, we identified the tweets related to the six central concepts of *empathi* i.e., (i) impact, (ii) modality of data, (iii) hazard type, (iv) place, (v) transportation, (vi) surveillance, and (vii) facility. Table III shows the statistics of identified tweets related to each of the chosen concepts and sub-concepts. The mapping was performed training a word embedding model²⁵ over 53 million tweets and utilize it in generating vectors of ontology classes and tweets. In the mapping process, first, we generate embedding of n-grams of a tweet and aggregate using averaging operation to create a representation of a tweet. To determine the relevant ontology classes to the tweet, we calculate cosine similarity scores. After finding the relevant ontology classes to the tweet using a threshold of 70% similarity, we perform similar process to find similarity between n-grams of the tweet and the sub-concepts within an ontology class. Moreover, considering the dynamic nature of disaster domain, we assume our approach to be a feasible component of ontology evolution and population process. Table II represents samples of these tweets (i.e., column two) along with the mapped *empathi* concepts (i.e., column three) from various hazard type (i.e., column one).

²⁴<http://twitris.knoesis.org/>

²⁵<https://radimrehurek.com/gensim/models/word2vec.html>

Concepts of <i>empathi</i>	#Tweets
Hazard Type	3,034,257
Impact	618,446
Modality of Data	509,645
Facility	258,117
Place	16,397
Transportation	6,694
Surveillance	1,588
Water Facility (Sub-Concept of Facility)	218,968
Gas Facility (Sub-Concept of Facility)	33,047
Involved Organization (Sub-Concept of Involved)	4,249
Severity (Sub-Concept of Impact)	1,344

Table III: Mapping concepts of *empathi* to tweets of hazards.

Extensive coverage by *empathi* provides the capability of extracting structured information from unstructured and sparse content (e.g., Twitter) [44]. For identifying relevant information from unstructured social media text and to classify tweets as relevant or irrelevant to crisis, it is essential to map the words to ontology classes. For instance in Figure 2, the tweet: *Chennai Floods: 188 killed, Airport closed they say that the runway came peeling off. It may take time to resume* is identified as a post-hazard tweet. *Chennai* links to concept *Place*, *Floods* links to concept *Hazard type*, *188 killed* links to concept *Affected population*, and *Airport closed* links to concept *Infrastructure Damage*. Moreover, such a procedure is termed as semantic annotation using expanded concepts (a.k.a. hypernyms). It can improve understanding of social media messages which pose challenges like ill-formed sentences, ambiguous word senses, poor syntactic structure, and implicit referencing. Semantic features formed using *empathi* can enhance supervised and unsupervised learning in crisis domain [45].

VI. EVALUATING QUALITY OF *empathi*

To build up a quality ontology, we followed the principles of ontology methodologies such as NeON [43] and Methontology [33] which encourage the reuse of existing ontologies. However, to quantitatively measure the quality of *empathi*, we designed a user evaluation survey. This survey contained 17 questions concerning with hierarchical, relational and lexical aspects of *empathi* inspired by [46]. Precisely, the participants (e.g., researchers, PhD students) in the survey evaluated the following criteria: (1) the correctness of structure (hierarchy) (2) the correctness of relations between concepts, and (3) lexical evaluation, i.e., quality of annotations associated with both concepts and relations. In the following, we elaborate on these criteria.

a) **Structural evaluation:** In this evaluation, the hierarchical structure is assessed concerning the correctness of “is-a” relationship as of whether the given concept A is-a particular type of the given concept B. For instance, “Animal Loss” is-a subclass of “Impact” in *empathi*. Such evaluation is necessary to confirm the utility of ontology for classification task [47]. We presented parts of hierarchy in the survey, and asked the participants, how far this hierarchy makes sense to them. They have to rate in the range 1 (fully disagree) - 5 (fully agree).

b) **Semantic relational evaluation:** Ontology is evaluated for holding semantically correct relations between con-

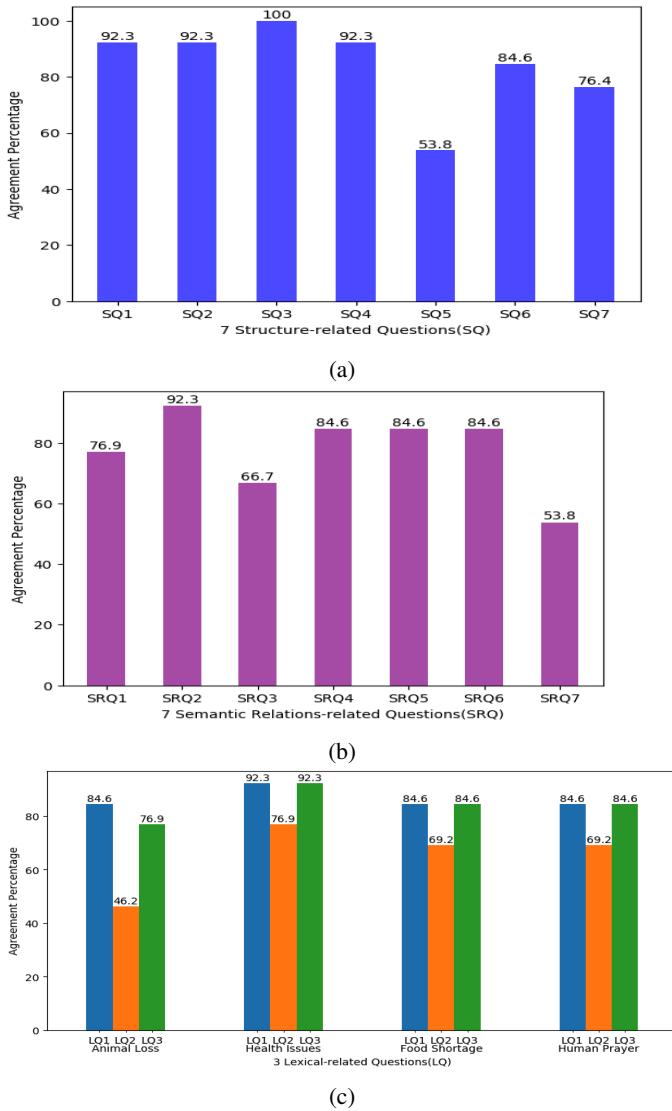


Figure 4: Agreement of 13 evaluators having some understanding of the domain. (a). 7 Questions that evaluate Structure of *empathi*, (b). 7 Questions that evaluate Semantic Relations of *empathi*, and (c). 3 Questions that evaluate Lexical characteristic of *empathi* for 4 concepts: Animal Loss, Health Issues, Food Shortage, Human Prayer. Agreement percentage is calculated as the percentage of evaluators responded “agree” or “yes” in the ontology evaluation form.

cepts. For instance, there is a relation between the concept “Affected Population” and the concept “Service” referring “need for help”. Thus, Affected Population is the domain and Service is the range. Concerning w.r.t. the prior survey [47], having quality relations, higher capability for summarization, subgraph extraction, and contextualization tasks. We represented a number of relations of *empathi* to the participants and asked them whether or not they confirm having such a relation or not.

c) **Lexical evaluation:** This part examines *expressiveness*, *completeness*, and *clarity* of annotations of a given

ontology. Expressiveness states the efficacy of the ontology to identify relevant information using natural language processing techniques. Completeness [48] evaluates whether an illustration of a concept using definition and labels adequately define various scenarios in crisis domain. For instance, w.r.t. the given concept “Human Prayer” the definition “prayer is a message to God from victim’s relatives and family for protecting their lives and health” is provided along with the labels “send prayers”, “heart prayers”, “heart praying”, “join praying”, “love prayers”, “prayers affected”, “temple”, “church”, “prayers city”, “prayers families”, “prayers involved”, etc. Clarity evaluates whether or not the concept name in the ontology is meaningful and easily understandable to human and machine.

d) **Results:** Our survey had 13 participants who understand the *crisis* domain and have published articles in relevant conferences (e.g. ISCRAM). The results recorded through survey have been illustrated in Figure 4.

The structural evaluation section comprised of seven questions expressed as follows; (SQ1) Concerning a hazard situation, is “Mental Stress” and “Physical Stress” two important concepts under “Health Issues”? (SQ2) Does “No Effect”, “Minor”, “Major”, “Hazardous”, “Catastrophic” represents sub-classes of “severity”? (SQ3) Are “Financial Crisis”, “Food Shortage”, and “Contamination” probable impacts of a Hazard? (SQ4) Are following triples meaningful: “Animal Loss is-a Impact”, “Communication Lines Failure is-a Infrastructure Damage”, “Power Outage is-a Infrastructure Damage”, and “Survived People is-a Affected Population”? (SQ5) Do you consider non-government organization’s (NGO) report is an expert report? (SQ6) Do health report, service feedback, and weather report define human sensing? (SQ7) Can “News Agencies Report” and “Social Media Report²⁶” be categorized under “Media Report”, a sub-class of “Report”? Questions SQ1-SQ6 were Yes/No questions, and SQ7 follows Likert Scale. The detailed results are represented in Figure 4a. We observe the average agreement rate above 84.5%.

Regarding evaluation of semantic relations, we designed seven questions as follows: SRQ1: Do you think the following triples make sense; “Event occurs in a Place”, “Service is offered in a Place”, “Each Hazard is associated to a couple of Event” and “each Hazard leads to a couple of Services”? SRQ2: Can “Event” and “Service” concepts be linked to the concept “Place” using “isLocationAt” relation? SRQ3: Do you think two different types of Hazard can be related concerning Event, Service, and Place? SRQ4: Is the “currentStatus” relation between Facility and Status semantically correct? SRQ5: Is available, offered, requested and unavailable suitable categories for Status? SRQ6: Can the concepts “Service” and “Organization” be concerning the concept “Hazard Type”? SRQ7: Is the relation “needHelp” correctly links “Affected Population” with “Organization” and “Service”? Questions SRQ2-SRQ7 are Yes/No, and SRQ1 follow Likert Scale. From Figure 4b, an average agreement of 75.4% was concluded in

²⁶<http://www.aidforum.org/topics/disaster-relief/top-100-twitter-accounts>

confirming above facts.

The lexical evaluation of *empathi* was performed by representing definition and synonyms (or labels) describing each concept and asking participants to respond the following questions; (LQ1) Do labels appropriately represent the concept? (LQ2) Are labels complete? (LQ3) Are definitions and labels enough clear? Figure 4c shows the results with the total agreement rate 78.8%. From Figure 4c, we see LQ2, SR7, and SQ5 to have a low agreement, however, the ontology will be improved with additional emerging concepts and relations extracted from social media content.

VII. RELATED WORK

The aftermath of the disaster causes agencies/organization to structure a plan for the thoughtful recovery of the area. An opportunity to hasten this process is a need for a knowledge structure containing concepts, constraints, and links that provide before-hand information for disaster preparedness and act. In [49], the author designed a process model of the aftermath of the disasters using the Petri-Net containing inter-linked concepts for efficiency after-disaster execution and adaptation. Furthermore, the cost associated with the recovery of an area urges the need to have a structured source of events, concepts, and relations that define a distressing incident. In [50], the author defines the utility of an ontology providing seismic risk definition, prediction, and management to reduce damages. There is a recent study on crowdsourced emergency event detection in [51]. There, the authors propose the utilization of emerging knowledge from text using concepts and temporal information for events. In a recent work on crisis management [52], there has been a development of a suite of tools which can leverage our ontology for context-aware response generation. CrowdGeoKG [53] is a framework using the entities in the OpenStreetMap and is enriched by Wikidata.

Events mainly characterize a disaster scenario, and it is essential to understand what is happening in an emergency situation. In [54], a given user can utilize GPS data to create a simulation model for predicting human mobility. We assume that identification of core disaster domain-specific concepts can help in annotating the GPS data similar to OpenStreetMap [55]. In 2012, Hurricane Sandy brought in towering traffic on social sensing sites urging the need of an information filtering mechanism for assisting crisis coordination. A psycholinguistic driven domain-dependent lexicon was created in [56] for assistive crisis response. Moreover, tweets on Hurricane Sandy 2012 identified “power blackout” as one of the implications of the disaster. To identify its associated repercussions, one need a structured domain-model. For instance, crashing out of power affected the medical facility. Hence, there is a need for an ontology to bridge the facilities in emergency situation [39]. In [57], the author created a twitter stream of deceptive and peripheral messages using a knowledge source assisting Public Information Officers (PIOs) to make conscious decisions in an emergency situation.

Table 1 motivates creating *empathi*. Prior ontologies related to crisis management and situational awareness fails to

provide structural, lexical and relational benefits needed for extracting situation-specific information extraction from social media content [47]. It is either because these ontologies are diverse in their subject areas or are incomplete concerning referencing, documentation, and expressibility. On analyzing these state-of-the-art vocabularies, we identified subject areas, vocabularies concerning each subject area and used OWL to provide a formal representation. Since ontologies can be domain-specific (crisis domain) or generic (e.g., FOAF), we incorporate concepts relevant in addressing various issues in crisis management. For instance, `Location` concept was taken from Geonames and iContact, crisis-related concepts from HXL and MOAC. Hence, we addressed all the queries raised in [3]. Human social communication during an emergency event provides real-time insight into various domains such as facility, events, impact, report, surveillance, organization involved and activities carried out during and after the hazard. Extracting actionable information from active social channels is challenging because of 2 reasons: (1) absence of an ontology that map multiple concepts, (2) completeness and expressiveness of the ontology. We provide a utility based case-study (section V) where we used our *empathi* ontology for mapping 53M tweets to any concepts of our ontology. Mapping of social media content to ontology concepts will improve classification and summarization task using state-of-the-art natural language processing and learning techniques.

VIII. CONCLUSION AND FUTURE WORK

We designed *empathi* ontology as a domain archetype that aids crisis management, hazard situational awareness and events during emergency scenarios. In our study, we demonstrated the prowess of ontology by its integration with relevant and state-of-the-art crisis vocabularies. Moreover, its efficacy was assessed through appropriate evaluation of its quality across three dominant criteria: structure, lexical and semantic relations using the human judge. Furthermore, we illustrated its effectiveness concerning social media domain by linking tweet content to ontology concepts. The ontology has been created using semantic web language (OWL) and links to FOAF, SIOC, DC-terms, and LOD. We aim to extend the ontology in following directions: 1) introduce Internet of Things (IoT) for disaster management, 2) improve the ontology with quality tools as recommended by the PerfectO methodology²⁷, and 3) disseminate more the ontologies on ontology catalogs such as Linked Open Vocabularies (LOV) and LOV4IoT. LOV4IoT could be refined and extended to support the environment domain with various use cases such as flooding, fire, earthquake, and tsunami.

ACKNOWLEDGMENT

We are thankful to responders in providing their unbiased evaluation for ontology quality assessment. We acknowledge support from the National Science Foundation (NSF) award EAR 1520870: Hazards SEES: Social and Physical Sensing Enabled Decision Support for Disaster Management and Response. Any opinions, findings, and conclusions/recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the NSF.

²⁷<http://perfectsemanticweb.appspot.com/?p=ontologyValidation>

REFERENCES

- [1] T. Clark, C. Keßler, and H. Purohit, "Feasibility of information interoperability in the humanitarian domain," in *AAAI Symposium*, 2015.
- [2] R. Barros, P. Kislansky, L. Salvador, R. Almeida, M. Breyer, and L. G. Pedraza, "Edxl-rescuer ontology: Conceptual model for semantic integration," in *ISCRAM*, 2015.
- [3] S. Liu, C. Brewster, and D. Shaw, "Ontologies for crisis management: A review of state of the art in ontology design and usability," in *ISCRAM*, 2013.
- [4] G. Poveda, E. Serrano, and M. Garijo, "Designing emergency management services by ontology driven social simulation," *IT CoNvergence PRACTICE*, 2015.
- [5] R. Battle and D. Kolas, "Geosparql: enabling a geospatial semantic web," *Semantic Web Journal*, 2011.
- [6] D. Brickley and L. Miller, "Foaf vocabulary specification 0.91," 2007.
- [7] A. Miles, B. Matthews, M. Wilson, and D. Brickley, "Skos core: simple knowledge organisation for the web," in *International Conference on Dublin Core and Metadata Applications*, 2005.
- [8] H. G. Ceballos, E. Olivarría, and E. A. Fernandez, "Using hxl hashtags for semantic annotation of egovernment tabular data."
- [9] V. Zavarella, H. Tanev, R. Steinberger, and E. Van der Goot, "An ontology-based approach to social media mining for crisis management," in *SSA-SMILE@ ESWC*, 2014.
- [10] S. M. Shah, C. Brewster, and D. Shaw, "Sierra: cooperative request-response for resource management in disasters using semantic web principles," 2013.
- [11] M. Moi and N. Rodehutsors, "Design of an ontology for the use of social media in emergency management," in *International Conferences ICT, WBC, BIGDACL and TPMC*, 2016.
- [12] G. Burel, L. S. Piccolo, K. Meesters, and H. Alani, "Dores—a three-tier ontology for modelling crises in the digital age," in *ISCRAM*, 2017.
- [13] F. Shih, O. Seneviratne, I. Liccardi, E. Patton, P. Meier, and C. Castillo, "Democratizing mobile app development for disaster management," in *Joint Proceedings of the Workshop on AI Problems and Approaches for Intelligent Environments and Workshop on Semantic Cities*, 2013.
- [14] N. Morrow, N. Mock, A. Papendieck, and N. Kocmich, "Independent evaluation of the ushahidi haiti project," *Development Information Systems International*, 2011.
- [15] S. Kar, A.-O. Hussein, K. Thirunarayan, V. Shalin, A. Sheth, and S. Parthasarathy, "D-record: Disaster Response and Relief Coordination Pipeline," in *Proceedings of the SIGSPATIAL International Workshop on Advances in Resilient and Intelligent Cities (ARIC)*. ACM, 2018.
- [16] D. Roman, D. Sukhobok, N. Nikolov, B. Elvesæter, and A. Pultier, "The infrarisk ontology: Enabling semantic interoperability for critical infrastructures at risk from natural hazards," in *OTM*. Springer, 2017.
- [17] A. Aid and I. Rassoul, "Context-aware framework to support situation-awareness for disaster management," *International Journal of Ad Hoc and Ubiquitous Computing*, 2017.
- [18] P. O. C. U. CASE, "Integration of provenance-enabled crowdsourced information with traditional disaster management information using linked open data," 2015.
- [19] S. N. Jonkman, "Global perspectives on loss of human life caused by floods," *Natural hazards*, 2005.
- [20] W. Lee, W. Bailer, T. Bürger, P.-A. Champin, J.-P. Evain, V. Malaisé, T. Michel, F. Sasaki, J. Söderberg, F. Stegmaier *et al.*, "Ontology for media resources 1.0," *W3C recommendation*, 2012.
- [21] S. Nalchigar and M. S. Fox, "An ontology for open 311 data," 2010.
- [22] J. J. Carroll, C. Bizer, P. Hayes, and P. Stickler, "Named graphs," *Web Semantics: Science, Services and Agents on the World Wide Web*, 2005.
- [23] J. G. Breslin, A. Harth, U. Bojars, and S. Decker, "Towards semantically-interlinked online communities," in *ESWC*, 2005.
- [24] D. Brickley, "W3c basic geo vocabulary," 2003.
- [25] W. R. Van Hage, V. Malaisé, R. Segers, L. Hollink, and G. Schreiber, "Design and use of the simple event model (sem)," *Web Semantics: Science, Services and Agents on the World Wide Web*, 2011.
- [26] H.-L. Kim, J. G. Breslin, S.-K. Yang, and H.-G. Kim, "Social semantic cloud of tag: Semantic model for social tagging," in *KES International Symposium on Agent and Multi-Agent Systems: Technologies and Applications*, 2008.
- [27] D. Shotton, "Cito, the citation typing ontology," in *Journal of biomedical semantics*, 2010.
- [28] R. Barros¹², P. Kislansky, L. Salvador¹², R. Almeida, M. Breyer, L. G. Pedraza, and V. Vieira¹², "Edxl-rescuer ontology: an update based on faceted taxonomy approach," 2015.
- [29] K. Bitencourt, F. Durão, and M. Mendonça, "Emergencyfire: An ontology for fire emergency situations," in *Proceedings of the 21st Brazilian Symposium on Multimedia and the Web*, 2015.
- [30] F. Simas, R. Barros, L. Salvador, M. Weber, and S. Amorim, "A data exchange tool based on ontology for emergency response systems," in *Research Conference on Metadata and Semantics Research*, 2017.
- [31] M. Snaprud, J. Radianti, and D. Svindseth, "Better access to terminology for crisis communications," in *International Conference on Information Technology in Disaster Risk Reduction*, 2016.
- [32] K. Villela, V. Vieira, M. Mendonça, J. Torres, and S. Graffy, "Rescuer-dow-reliable and smart crowdsourcing solution for emergency and crisis management," *Munich, Germany*, 2013.
- [33] M. Fernández-López, A. Gómez-Pérez, and N. Juristo, "Methontology: from ontological art towards ontological engineering," 1997.
- [34] R. Prieto-Díaz, "A faceted approach to building ontologies," in *IEEE IRI*, 2003.
- [35] M. Grüniger and M. S. Fox, "Methodology for the design and evaluation of ontologies," 1995.
- [36] M. Wick, B. Vatant, and B. Christophe, "Geonames ontology," *URL http://www.geonames.org/ontology*, 2015.
- [37] C. Anderson, *The Federal Emergency Management Agency*, 2002.
- [38] H. Purohit, A. Hampton, S. Bhatt, V. L. Shalin, A. P. Sheth, and J. M. Flach, "Identifying seekers and suppliers in social media communities to support crisis coordination," *CSCW*, 2014.
- [39] S. P. Bhatt, H. Purohit, A. Hampton, V. Shalin, A. Sheth, and J. Flach, "Assisting coordination during crisis: a domain ontology based approach to infer resource needs from tweets," in *Web science*, 2014.
- [40] M. Limbu, "Management of a crisis (moac) vocabulary specification. observedchange, january 2012," 2014.
- [41] K. Bitencourt, F. Durão, and M. G. Mendonça, "Emergencyfire: An ontology for fire emergency situations," in *WebMedia*, 2015.
- [42] M. Moi, N. Rodehutsors, and R. Koch, "An ontology for the use of quality evaluated social media data in emergencies," *IADIS International Journal on WWW/Internet*, 2016.
- [43] M. C. Suarez-Figueroa, A. Gomez-Perez, and M. Fernandez-Lopez, "The NeOn Methodology for Ontology Engineering," in *Ontology engineering in a networked world*, 2012.
- [44] L. Derczynski, A. Ritter, S. Clark, and K. Bontcheva, "Twitter part-of-speech tagging for all: Overcoming sparse and noisy data," in *Proceedings of the International Conference Recent Advances in Natural Language Processing RANLP 2013*, 2013.
- [45] P. Khare, G. Burel, and H. Alani, "Classifying crises-information relevancy with semantics," 2018.
- [46] H. Hlomani and D. Stacey, "Approaches, methods, metrics, measures, and subjectivity in ontology evaluation: A survey," *Semantic Web Journal*, 2014.
- [47] M. Imran, C. Castillo, F. Diaz, and S. Vieweg, "Processing social media messages in mass emergency: A survey," *ACM CSUR*, 2015.
- [48] V. Cordt and V. Mascardi, "Checking the completeness of ontologies: a case study from the semantic web," in *CILC Workshop*, 2004.
- [49] D. Fahland and H. Woith, "Towards process models for disaster response," in *BPM Conference*, 2008.
- [50] B. Murgante, G. Scardaccione, and G. Las Casas, "Building ontologies for disaster management: seismic risk domain," *Urban and Regional Data Management*, 2009.
- [51] K. Grolinger, M. A. Capretz, and E. Mezghani, "Knowledge as a service framework for disaster data management," in *IEEE WETICE*, 2013.
- [52] F. Benaben, A. Montarnal, S. Truptil, M. Luras, A. Fertier, N. Salatge, and S. Rebiere, "A conceptual framework and a suite of tools to support crisis management," 2017.
- [53] J. Chen, S. Deng, and H. Chen, "Crowdgeoqk: Crowdsourced geo-knowledge graph," in *CCKS*, 2017.
- [54] X. Song, Q. Zhang, Y. Sekimoto, R. Shibasaki, N. J. Yuan, and X. Xie, "A simulator of human emergency mobility following disasters: Knowledge transfer from big disaster data," in *AAAI*, 2015.
- [55] S. Auer, J. Lehmann, and S. Hellmann, "Linkedgeodata: Adding a spatial dimension to the web of data," in *ISWC*, 2009.
- [56] H. Purohit, "Crisis response coordination in online communities," 2013.
- [57] A. J. Hampton, S. Bhatt, A. Smith, J. Brunn, H. Purohit, V. L. Shalin, J. M. Flach, and A. P. Sheth, "Constructing synthetic social media stimuli for an emergency preparedness functional exercise," 2017.