A Simple Disaster-Related Knowledge Base for Intelligent Agents

Clark Emmanuel Paulo, Arvin Ken Ramirez, David Clarence Reducindo
Rannie Mark Mateo, Joseph Marvin Imperial
National University
Manila, Philippines
`jrimperial@national-u.edu.ph`

Abstract

In this paper, we describe our efforts in establishing a simple knowledge base by building a semantic network composed of concepts and word relationships in the context of disasters in the Philippines. Our primary source of data is a collection of news articles scraped from various Philippine news websites. Using word embeddings, we extract semantically similar and co-occurring words from an initial seed words list. We arrive at an expanded ontology with a total of 450 word assertions. We let experts from the fields of linguistics, disasters, and weather science evaluate our knowledge base and arrived at an agreeability rate of 64%. We then perform a time-based analysis of the assertions to identify important semantic changes captured by the knowledge base such as the (a) trend of roles played by human entities, (b) memberships of human entities, and (c) common association of disaster-related words. The context-specific knowledge base developed from this study can be adapted by intelligent agents such as chat bots integrated in platforms such as Facebook Messenger for answering disaster-related queries.

1 Introduction

The Philippines is a common ground for natural disasters such as typhoons and flooding. According to the latest statistics of Philippine Atmospheric, Geophysical and Astronomical Services Administration (PAG-ASA)[1], there are more tropical cyclones entering the vicinity of the Philippines than anywhere else in the world. In a year, almost 20 tropical cyclones enter with 70% chance of developing into a full-blown typhoon. As possible result, catastrophic aftermath such as economic failure, destruction of infrastructures, and loss of lives may be imminent without proper information dissemination and preparedness. Thus, directing all research and technological efforts to disaster preparedness and disaster risk reduction to mitigate the tremendous impacts of natural calamities have been prioritized by the country for years.

According to Statista[2], the Philippines has 44 million active Facebook users in 2019 and is predicted to reach approximately 50 million by 2023. Thus, taking in consideration the disaster-prone situation of the country, we note the importance of establishing context-specific knowledge bases where it can be integrated to commonly used digital platforms such as Facebook Messenger to answer context-specific questions on disasters for possible information dissemination and awareness. To make this possible, we present our initial efforts in building a simple knowledge base composed of word concepts joined by semantic word relationships extracted from word embeddings trained from a large online news corpus on disasters.

The overview of the study is as follows: First, we compared and contrasted previous works done in field of ontology learning and mining information from word embeddings for building knowledge

---

bases. Next, we discussed the method of extraction of concepts and entities from our news article dataset as well as the semantic labels used for building concept relationships. In the Results section, we detail the outcome of using word embeddings trained from our collected dataset for extracting semantically similar concepts using an initial concept seedwords list. In addition, we also discussed the integrity of the disaster ontology from expert validation. We wrap up the study by extensively discussing the semantic changes exhibited by concepts present in the knowledge base over time.

2 Previous Work

In this section we highlight significant works in using knowledge bases (KBs) as the main source of intelligence for virtual agents as well as current trends using word embeddings as one way of extracting information and semantic word relationships.

2.1 Building Knowledge-Based Intelligent Agents

Over the years, the use of powerful knowledge bases are seen on a wide variety of intelligent agents such as chat bots, storytelling agents, and recommender systems. The work of (Ong et al., 2018) focused on building a commonsense knowledge base for a storytelling agent for children using assertions extracted from ConceptNet (Speer et al., 2017). ConceptNet is a large semantic network of word relationships that can be adapted by intelligent agents for commonsense reasoning and identifying object relationships. The study developed a knowledge base by filtering out concepts and concept relationships that are not used in the context of children storytelling from the ConceptNet.

Similarly, Han et al. (2015) developed a natural language dialog agent that utilizes a knowledge base to generate diverse but meaningful responses to the user. The system extracts related information from knowledge base, which was adapted from FreeBase (Bollacker et al., 2008), via an information extraction module. The use of a large, external knowledge base allows the dialog system to expand on the information from the users response for a more detailed and interactive reply.

On the other hand, the work of Wang et al.

(2010) focused on the development of a system composed of three ontology-based sub-agents for personal knowledge, fuzzy inference, and semantic generation for evaluating a person's health through his/her diet. The system makes use of an ontology with an embedded knowledge base considering the persons health statistics such as BMI, Caloric Difference, Health Diet Status combined with rules laid down by domain experts. Results shows that the proposed system exhibits an intelligent behavior in helping the dietary patterns of users based on their information from the constructed ontology.

2.2 Knowledge from Word Embeddings

The advent of word embeddings as one of the modern approaches in extracting semantic relationships of words has fueled research works to use its potential to build more powerful knowledge bases. Sarkar et al. (2018) used a supervised approach, similar to text classification, for predicting the taxonomic relationship via similarity of two concepts using a Word2Vec embedding (Mikolov et al., 2013b). Results showed that combining the word embedding with an SVM classifier outperformed baseline approaches for taxonomic relationship extraction such as using the Jaccard similarity formula and naive string matching.

The study of Luu et al. (2016) went as far as building a custom neural network architecture with dynamic weighting to significantly increase the performance of statistical and linguistic approaches in extraction word relationships from word embeddings. The neural network considers not only the word relationship such as hypernym and hyponym but also the contextual information between the terms. The proposed approach exhibits generalizability for unseen word pairs and has obtained 9% to 13% additional accuracy score using a general and domain-specific datasets.

Likewise, the work of Pocostales (2016) submitted to the SemEval-2016 Task for Taxonomy Extraction Evaluation (Bordea et al., 2016) focused on using GloVe word embedding model (Pennington et al., 2014) with an offset feature to extract hypernym candidates from a sample word list. Results showed that a vector offset cannot completely capture the hypernym-hyponym relationship of words due to complexity.

2017	Tag	2018	Tag	2019	Tag
family	noun	*act*	verb	*bulletin*	noun
fire	noun	*announced*	verb	*quake*	noun
flood	noun	*ashfall*	noun	*typhoon*	noun
update	verb	*police*	noun	*weakened*	adjective
tricycle	noun	*lava*	noun	*affected*	verb

Table 1: Top 5 seed words per year.

3 Data

For this study, we scraped over 4,500 Filipino disaster-related news articles from years 2017 to 2019 (1,500 articles per year) from Philippine news websites using Octoparse Webscraping Tool as our primary dataset. The corpus covers a wide range of natural disasters that transpired in the Philippines such as typhoons, earthquakes, landslides and also includes statistics from damages and casualty reports. This large collection of news articles will be the groundwork of the knowledge base as it contains disaster-related context words as concepts. We partitioned the dataset into three by year (2017 to 2019) for the word embedding model generation. The purpose of partitioning the dataset will allow us to analyze the temporal changes of sematic relationships of concepts. More on this is discussed in the succeeding sections. To note, all concepts presented in this document are translated to English for the international audience.

4 Building the Knowledge Base

4.1 Extracting Concept Seedwords

Building a knowledge base starts with establishing concepts or words that refer to real world entities such as **nouns**, **adjectives**, and **verbs** that represent everyday objects such as *apple, spoon, cake*, people like *mother, police, mayor*, description of objects such as *beautiful, red, big*, and action words that signify an activity such as *walks, eating* and *jumped* (Ong, 2010).

To identify the grammatical category of each concept to know whether it is a noun, an adjective, or a verb to aid the semantic relationship labelling, we used a Filipino parts-of-speech (POS) tagger[3] developed by Go and Nocon (2017) which is currently

---
[3] github.com/matthewgo/FilipinoStanfordPOSTagger

integrated in the Stanford CoreNLP package (Manning et al., 2014). We extracted the top 50 high-frequency concepts per year from the collected news dataset, having a total of 150 initial seed words. Table 1 shows the top 5 seed words per year from the initial word list. Both common words such as *family*, *tricycle*, *police* and *update* as well as disaster-related words such as *flood*, *ashfall*, *typhoon*, and *quake* to name a few are present. These concepts are then paired with other concepts to form a meaningful representation of knowledge called a **binary assertion** described in the next section.

4.2 Semantic Relationship Labelling

Once the set of context-specific words are obtained from the dataset, in the case of this study, disaster-related concepts, the next process to establish the correct semantic relationship of words. These semantic relationships can be structured in the form of a binary assertion as previously stated. Ong et al. (2018) stated that the binary assertions of concepts are needed by virtual agents such as a story-telling agent or a chatbot to be able to generate responses from a commonsense knowledge base. Semantic relationships can also be used for other tasks such query expansion of words as well as information retrieval (Attia et al., 2016). For this study, we adapt the binary assertion format used by Concept-Net (Speer et al., 2017) shown below:

```
[concept1 semantic-rel concept2]
```

where **semantic-rel** stands for the specified semantic relationship of the two concepts (**concept1** and **concept2**) that contains meaning. We adapted the six semantic relation labels from the CogALex-2016 Shared Task on Corpus-Based Identification of Semantic Relations (Santus et al., 2016) which are **Synonym (SYN) Antonym (ANT)**, **Hypernym (HYP)**, **Membership (PartOf)**,

Relation	Tag	Rule
Synonym	SYN	If `concept1` has the same meaning with `concept2`
Antonym	ANT	If `concept1` has the opposite meaning with `concept2`
Hypernym	HYP	If `concept1` has a broader meaning compared to `concept2`
Performs	DO	If `concept1` is the actor/does of `concept2`
Membership	PartOf	If `concept1` is a member of `concept2`
Adjective	IS	If `concept1` describes `concept2`
Cause	CAUSE	If `concept1` is the cause of event `concept2`
Effect	dueTo	If `concept1` the resulting effect of event `concept2`
Random	RAND	If `concept1` has no direct relationship `concept2`

Table 2: Semantic relations with its corresponding rules and examples.

and **Random (RAND)** as shown in Table 2. In addition, we also added a few semantic labels of our own with respect to the concepts that we will be working on this study which are in the field of disasters. Thus, we added **Performs (DO)** to signify action, **Adjective (IS)** to signify description, and **Cause (CAUSE)** and **Effect (dueTo)** to signify consequences of events in a disaster setting.

4.3 Concept Expansion using Word Embeddings

Word embeddings are representations of an entire document vocabulary in a mathematical vector space. Each word is represented with a set of real-valued numbers given a specific set of dimensions. Word embedding architectures such as **Word2Vec** (Mikolov et al., 2013a; Mikolov et al., 2013b) and **Global Vectors** or **GloVe** (Pennington et al., 2014) capture various relationships of words in a corpus such as the **semantic similarity**, **syntactic similarity**, and **word co-occurrence** to name a few. Thus, if two words are commonly used together in the same context, such as in disaster-related articles, we can expect them to be close together when represented in the vector space. For example, finding the most similar terms using the query word *warning* from a word embedding model trained on a disaster-related dataset will output words such as *typhoon*, *flooding*, and *signal* since the word *warning* is commonly used in texts to notify people of possible natural disasters.

For this study, we generated a word embedding model for each year of partitioned news article dataset from 2017 to 2019. A total of three mod-

els were generated using the Word2Vec architecture (Mikolov et al., 2013b). In addition, we used the initial seedwords list which contains 150 concepts (50 for each year) as query words to extract semantically similar terms which occur in the context of disaster.

5 Results

Table 3 shows the expanded ontology by querying four sample seedwords from the word embedding model. We only filtered the top three semantically similar resulting words from each query word that falls under a manually-annotated and qualified semantic label discussed in Section 4.2. From this, we arrived at an expanded ontology composed of 450 assertions for our disaster-related knowledge base.

From the resulting expanded ontology, it already provides us the *knowledge* that an intelligent agent can piece together or form when asked about something in the context of disasters. Take in, for example, the word *tremor*. Obtaining the top three assertions with semantic labels and expanded using word embeddings informs us that the word *tremor* is synonymous and interchangeable with the word *af-*
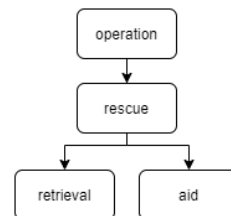

Figure 1: Hierarchy of hypernyms formed from the expanded ontology.

Seedword	Semantic Label	Assertions
police	DO	[police DO risk]
	IS	[police IS armed]
	DO	[police DO commitment]
tremor	SYN	[tremor SYN aftershock]
	dueTo	[tremor dueTo quake]
	dueTo	[tremor dueTo earthquake]
rescue	partOf	[rescue partOf operations]
	HYP	[rescue HYP retrieval]
	HYP	[rescue HYP aid]
experts	IS	[experts IS supporting]
	DO	[experts DO recommend]
	DO	[experts DO impose]

Table 3: Expansion of knowledge base using word embeddings.

tershock which scientifically means the involuntary movement of the surface due to breaking of underground rocks. Likewise, the words *quake* and *earthquake* are annotated with the semantic label *dueTo* which denotes effect since earthquakes is root cause of tremors according to scientific definition (Yose, 2013).

Another observation from the expanded ontology is the knowledge of understanding general actions to more specific ones. In the example, the action word *rescue* is a **hypernym**, which means it is a general word that can be possibly specified further, in this case, it is the hypernym of the word *aid*. Consequently, the word *operation* is a hypernym of *rescue*. Combining the words altogether, we can interpret the three assertions as a string of successive actions as shown in Figure 1 where in an *operation* involves a *rescue* and the meaning of *rescue* may vary such that it can be some form of (a) *retrieval* of missing people or (b) *aid* for the wounded or stranded people.

Lastly, the word embedding model was able to produce words that imply *responsibilities* by concepts in the form of human entities. In the two examples of human entities shown in Table 3, *experts* and *police*, most of the semantic labels are **DO** which denotes **action** and **IS** which denotes **description**. For *police*, common known actions done in the context of disasters are *risk* and *commitment* while being *armed*. For *experts*, the responsibilities are in the line of *supporting* a claim as well as *recommend-*

ing and *imposing* future actions based on scientific knowledge.

We observe the significant potential of using word embeddings for expanding knowledge bases by capturing various levels of information such as (a) understanding interchangeable context-specific words, in the case of this study on disaster-related words, and (b) understanding roles played by human entities described in this section. When used by an intelligent agent, it will have an idea of what response can be produced when queried with questions such as *"What does a police officer do?"* or *"What happens after an earthquake?"*.

5.1 Expert Validation

To properly gauge the effectivity of using word embeddings for the expansion of knowledge bases, we performed a validation process by inviting three experts in the fields of linguistics, disaster response, and meteorology to evaluate the assertions of the knowledge base. Each assertion is evaluated using a **two-scale metric**: **Agree** if the expert deems that the assertion observes a correct relationship and semantic labelling ([*doctor* SYN *medical person*]) or **Disagree** if the assertion is not properly labelled to form a correct relationship ([*flood* SYN *typhoon*]).

Table 4 shows the averaged agreeability rate of the expert validation process. Results show that the most accurate model with the highest rating is the Word2Vec model trained on disaster-related news articles collected in the year of 2019. This is fol-

Model	Aggreability Rate
Word2Vec 2019	0.64
Word2Vec 2018	0.52
Word2Vec 2017	0.49

Table 4: Expert validation for the knowledge base.

lowed by word embedding models using the 2018 and 2017 dataset. We attribute the semi-low agreeability scores due to the variation of word usage of the experts in their corresponding fields of study. For example, the assertion [*tremor* SYN *earthquake*] was evaluated by the linguist and disaster experts as **Agree** while the meteorologist contested. The expert on meteorology swears by the scientific definition of which tremors are very different with earthquakes in a way that tremors are *caused* by earthquakes and not an interchangeable term as perceived by the other two non-technical experts. This pattern is also observed with other assertions using the synonym labelling such as [*storm* SYN *typhoon*] and [*lava* SYN *magma*].

6 Discussion

In this section, we perform an even more in-depth analysis by considering the changes in semantic meaning or semantic information of the assertions of the knowledge base over time. We break the discussion into three categories we observed from the analysis.

6.1 Roles Played by Human Entities

We observe the changing roles played by three essential human entities, *governor*, *experts*, and *police* over time in the setting of a natural disaster as written in news articles. These entities are expected to be on full alert and their responsibilities are crucial towards mitigating the consequences of disasters.

The entity *governor*, or the highest commanding individual of Philippine province, performs various roles over time. As seen in Table 5, the entity is mostly connected by the semantic label **DO** to denote action with concepts such as *assure*, *explain*, *oversee*, *develop*, and *declare*. There are also a few description words connected by the label **IS** such as *political* and *mandatory* which is obvious since holding a gubernatorial position is indeed *political*

and resolutions filed in a governor's office is in its essence *mandatory*.

Similarly, the entity *experts* is also expected to have various changing roles. For 2017 as seen in the Table, common action words associated are *checking* and *verifying* which provides one of the most important roles of experts in the field of disasters: validating the integrity of information being publicized. For 2018, experts play more of an information dissemination role with the concept *announce* as the connecting action word. In 2019, experts assumed a more stricter role as observed with the connected concepts such as *recommend* and *impose*.

For the entity *police*, there is consistent associated descriptor across the years regarding their responsibility: *risk*. This provides us a concrete idea of a consequence when assuming the role of a police. In 2017 and 2018, strong action words such as *control*, *damage*, and *armed* are tied with a policeman's job. In 2019, however, the entity *police* assumed a more passive role as more of an informant with the connected action words being *study* and *alerts*.

6.2 Memberships of Human Entities

We also observed changes in memberships of human entities in context of disasters over time as seen in Table 6. Memberships are denoted by the semantic labels **HYP** for **hypernyms** or general words and **partOf** for **hyponyms** more specific words. We observe memberships played of three entities: *mayor*, *student*, *teacher*.

For the entity *mayor*, memberships are more specific in 2017 and 2018. The entity is expected to partner with small groups in a municipality such as a *sitio* or a barangay cite as well as with large government agencies such as Division on the Welfare of the Urban Poor (DWUP) and Department of Environment and Natural Resources (DENR) in times of disasters. There are also general memberships such as *administration* and *government* to which the entity is an obvious member.

In the case of the entity *student*, there is a mix of specific and generalized memberships. Certain specific universities such as *UP* or the University of the Philippines and *PUP* or Polytechnic University of the Philippines were classified as institutions where a student may belong. General and obvious concepts such as *elementary*, *school*, *organization* and *Uni-*

Seed	2017	2018	2019
governor	[governor DO assure] [governor DO giving] [governor DO explain]	[governor IS political] [governor DO oversees] [governor IS mandatory]	[governor DO declare] [governor DO resolution] [governor DO develop]
experts	[experts SYN representative] [experts DO checking] [experts DO verify]	[experts DO work] [experts IS frantic] [experts DO announce]	[experts IS supporting] [experts DO recommend] [experts DO impose]
police	[police DO risk] [police DO control] [police DO damage]	[police DO risk] [police IS armed] [police DO commitment]	[police DO risk] [police DO study] [police DO alerts]

Table 5: Roles played by human entities.

Seed	2017	2018	2019
mayor	[mayor partOf hall] [mayor partOf DWUP] [mayor partOf sitio]	[mayor partOf administration] [mayor DO communication] [mayor partOf DENR]	[mayor RAND workers] [mayor partOf government] [mayor RAND employers]
student	[student partOf PUP] [student IS victim] [student IS resident]	[student partOf elementary] [student partOf school] [student IS minor]	[student partOf organization] [student partOf University] [student partOf UP]
teacher	[teacher DO education] [teacher partOf DepEd] [teacher partOf school]	[teacher DO research] [teacher DO education] [teacher partOf school]	[teacher partOf house] [teacher SYN employee] [teacher partOf school]

Table 6: Memberships of human entities.

versity contribute to the commonsense information of the knowledge base.

For the entity *teacher*, general concepts of memberships are more prominent compared to the other two entities. The concept *school* is consistent for all years. The term *DepEd* which means Department of Education, the government agency responsible in shaping the educational landscape of the Philippines, is connected to the entity. DepEd oversees elementary and intermediate level schools which may mean that the term *teacher* is commonly tied with educators from these levels. Interestingly, the concept *house* is also tied with the entity *teacher*. Although it may already be obvious that a teacher assumes a different role inside the house maybe as a mother or a breadwinner.

6.3 Common Word Association of Disaster-Related Terms

For the last category, we observe change in association of disaster-related words over time as seen in Table 7. We highlight the importance of this analysis to understand how a knowledge base using information extracted from word embeddings interchange co-occurring and similar words in the context of disasters. For this category, we analyze the four frequently occurring natural disasters in the Philippines which are **earthquakes**, **eruption**, **landslide**, and **typhoon**.

The first disaster-related word is *earthquake*. From the formed assertions for all years, *earthquakes* are synonymous with the term *quakes* which denote a shortened version of the word. In addition, the term *magnitude* is also a consistent term connected to the seedword *earthquake* which tells us that earthquakes have their own corresponding *magnitudes* quantified by some number.

In the case of *eruption*, most assertions formed are used with the semantic label **IS** to denote **description** or **property**. Across the years, most of the descriptive words associated with the term are negative such as *amplifying*, *hazardous*, *destruction*, *explo-*

Seed	2017	2018	2019
earthquake	[earthquake IS magnitude] [earthquake SYN quake] [earthquake RAND drill]	[earthquake IS magnitude] [earthquake SYN quake] [earthquake IS intensity]	[earthquake IS magnitude] [earthquake SYN quake] [earthquake RAND signal]
eruption	[eruption IS amplifying] [eruption IS fast] [eruption IS confirmed]	[eruption IS happening] [eruption IS explosive] [eruption CAUSE destruction]	[eruption IS hazardous] [eruption IS magmatic] [eruption IS happening]
landslide	[landslide dueTo flood] [landslide HYP mudslide] [landslide RAND area]	[landslide dueTo rain] [landslide IS torrential] [landslide HYP mudslide]	[landslide RAND mountain] [landslide dueTo rains] [landslide IS widespread]
typhoon	[typhoon IS expected] [typhoon partOf calamity] [typhoon SYN onslaught]	[typhoon SYN storm] [typhoon IS super] [typhoon IS powerful]	[typhoon dueTo Amihan] [typhoon SYN hurricane] [typhoon SYN cyclone]

Table 7: Disaster-related terms and its associated words.

sive, and *magmatic*. These word may denote a sense of urgency compared to the other natural disasters when it happens. The term *happening* has occurred both in 2018 and 2019 due to two of the most active volcanoes in the Philippines, Mount Mayon and Mount Taal, had shown activity[4] by erupting successively and spewing ash. The eruption caused mass postponement of public activities and over 48,000 evacuated locals.

For the concept *landslide*, the common associated word is *mudslide*. Although by scientific definition, mudslides are as specific type landslides (also called *debris flow*). Thus, the *landslide* term conforms the the semantic label **HYP** for hypernyms. The cause of landslides can be tied to *floods*, which in turn, caused by *rains* as joined by the semantic label **dueTo** denoting a **consequence** or an **effect**. In addition, landslides are also often described using the words *widespread* and *torrential* which gives us a quantifiable idea of the magnitude of landslides that occur in the Philippines.

Lastly, we have the word *typhoon*. This disaster-related concept assumes many interchangeable and synonymous terms such as *storm*, *hurricane*, *cyclone*, and *onslaught*. We note the frequent association and interchangeability of these words due to geographic locations (Khadka, 2018) but may mean the same thing—they are all **tropical storms**. Likewise, description words tied to the concept *typhoon*

are *expected*, *super*, and *powerful* which also provides us an idea of the magnitude of typhoons occurring in the country similar to *landslides*.

7 Conclusion

As a disaster-prone country, the Philippines should increase its efforts in mitigating the effects of natural calamities with the help of technology. One way to do this is to consider the potential of intelligent agents such as chatbots as tools for disaster preparedness and information dissemination. In this study, we established a simple context-specific knowledge base by doing three important processes: (a) extracting disaster-related concepts from a collected news article dataset, (b) building a network of binary assertions from a curated list of semantic labels, and (c) expanding the ontology by querying an initial seedwords list from word embeddings generated from the original news dataset. Results show that using word embeddings captured various levels of information that may be useful for intelligent agents to produce responses such as information on roles of human entities, generalization and specification of terms, and common word association when asked in the topic of disasters. Future directions of the study include collection of even more dataset that covers not only news articles but also other media for finer-grained assertions. In addition, the study will also benefit from efforts in testing the capability of the knowledge base in practical applications.

---

[4]Volcano Bulletin: phivolcs.dost.gov.ph/index.php/volcano-hazard/volcano-bulletins3

References

Mohammed Attia, Suraj Maharjan, Younes Samih, Laura Kallmeyer, and Thamar Solorio. 2016. Cogalex-v shared task: Ghhh-detecting semantic relations via word embeddings. In *Proceedings of the 5th Workshop on Cognitive Aspects of the Lexicon (CogALex-V)*, pages 86–91.

Kurt Bollacker, Colin Evans, Praveen Paritosh, Tim Sturge, and Jamie Taylor. 2008. Freebase: a collaboratively created graph database for structuring human knowledge. In *Proceedings of the 2008 ACM SIGMOD international conference on Management of data*, pages 1247–1250.

Georgeta Bordea, Els Lefever, and Paul Buitelaar. 2016. Semeval-2016 task 13: Taxonomy extraction evaluation (texeval-2). In *Proceedings of the 10th International Workshop on Semantic Evaluation (SemEval-2016)*, pages 1081–1091.

Matthew Phillip Go and Nicco Nocon. 2017. Using Stanford part-of-speech tagger for the morphologically-rich Filipino language. In *Proceedings of the 31st Pacific Asia Conference on Language, Information and Computation*, pages 81–88.

Sangdo Han, Jeesoo Bang, Seonghan Ryu, and Gary Geunbae Lee. 2015. Exploiting knowledge base to generate responses for natural language dialog listening agents. In *Proceedings of the 16th Annual Meeting of the Special Interest Group on Discourse and Dialogue*, pages 129–133.

Navin Singh Khadka. 2018. Hurricanes, typhoons and cyclones: What's the difference?, Sep.

Anh Tuan Luu, Yi Tay, Siu Cheung Hui, and See Kiong Ng. 2016. Learning term embeddings for taxonomic relation identification using dynamic weighting neural network. In *Proceedings of the 2016 Conference on Empirical Methods in Natural Language Processing*, pages 403–413.

Christopher D Manning, Mihai Surdeanu, John Bauer, Jenny Rose Finkel, Steven Bethard, and David McClosky. 2014. The stanford corenlp natural language processing toolkit. In *Proceedings of 52nd annual meeting of the association for computational linguistics: system demonstrations*, pages 55–60.

Tomas Mikolov, Kai Chen, Greg Corrado, and Jeffrey Dean. 2013a. Efficient estimation of word representations in vector space. *arXiv preprint arXiv:1301.3781*.

Tomas Mikolov, Ilya Sutskever, Kai Chen, Greg S Corrado, and Jeff Dean. 2013b. Distributed representations of words and phrases and their compositionality. In *Advances in neural information processing systems*, pages 3111–3119.

Dionne Tiffany Ong, Christine Rachel De Jesus, Luisa Katherine Gilig, Junlyn Bryan Alburo, and Ethel Ong. 2018. Building a commonsense knowledge base for a collaborative storytelling agent. In *Pacific Rim Knowledge Acquisition Workshop*, pages 1–15. Springer.

Ethel ChuaJoy Ong. 2010. A commonsense knowledge base for generating children's stories. In *2010 AAAI Fall Symposium Series*.

Jeffrey Pennington, Richard Socher, and Christopher D Manning. 2014. Glove: Global vectors for word representation. In *Proceedings of the 2014 conference on empirical methods in natural language processing (EMNLP)*, pages 1532–1543.

Joel Pocostales. 2016. Nuig-unlp at semeval-2016 task 13: A simple word embedding-based approach for taxonomy extraction. In *Proceedings of the 10th International Workshop on Semantic Evaluation (SemEval-2016)*, pages 1298–1302.

Enrico Santus, Anna Gladkova, Stefan Evert, and Alessandro Lenci. 2016. The CogALex-v shared task on the corpus-based identification of semantic relations. In *Proceedings of the 5th Workshop on Cognitive Aspects of the Lexicon (CogALex - V)*, pages 69–79, Osaka, Japan, December. The COLING 2016 Organizing Committee.

Rajdeep Sarkar, John Philip McCrae, and Paul Buitelaar. 2018. A supervised approach to taxonomy extraction using word embeddings. In *Proceedings of the Eleventh International Conference on Language Resources and Evaluation (LREC 2018)*.

Robyn Speer, Joshua Chin, and Catherine Havasi. 2017. Conceptnet 5.5: An open multilingual graph of general knowledge. In *Thirty-First AAAI Conference on Artificial Intelligence*.

Mei-Hui Wang, Chang-Shing Lee, Kuang-Liang Hsieh, Chin-Yuan Hsu, Giovanni Acampora, and Chong-Ching Chang. 2010. Ontology-based multi-agents for intelligent healthcare applications. *Journal of Ambient Intelligence and Humanized Computing*, 1(2):111–131.

Joash Yose. 2013. Earthquake or tremor: What's the difference?, Jul.