Natural Language Processing for Analyzing Disaster Recovery Trends Expressed in Large Text Corpora

Lucy H. Lin
Allen School
University of Washington
Seattle, USA
lucylin@cs.washington.edu

Scott B. Miles
Human Centered Design & Eng.
University of Washington
Seattle, USA
milessb@uw.edu

Noah A. Smith
Allen School, University of Washington;
Allen Institute for Artificial Intelligence
Seattle, USA
nasmith@cs.washington.edu

Abstract—We are developing a new natural language processing (NLP) method to facilitate analysis of text corpora that describe long-term recovery. The aim of the method is to allow users to measure the degree that user-specified propositions about potential issues are embodied within the corpora, serving as a proxy for the disaster recovery process. The presented method employs a statistical syntax-based semantic matching model and was trained on a standard, publicly available training dataset. We applied the NLP method to a news story corpus that describes the recovery of Christchurch, New Zealand after the 2010–2011 Canterbury earthquake sequence. We used the model to compute semantic measurements of multiple potential recovery issues as expressed in the Christchurch news corpus that span 2011 to 2016. We evaluated method outputs through a user study involving twenty professional emergency managers. User study results show that the model can be effective when applied to a disaster-related news corpus. 85% of study participants were interested in a way to measure recovery issue propositions in news or other corpora. We are encouraged by the potential for future applications of our NLP method for after-action learning, recovery decision making, and disaster research.

Keywords—*natural language processing, semantic matching, disaster recovery*

I. INTRODUCTION

The literature on empirical research of disaster recovery is growing, as are jurisdictional efforts to do pre-event recovery-based planning. Even so, recovery is one of the least understood disaster topics [1]. In general, recovery research findings are narrowly focused on a single recovery aspect, represent a single point in time, and do not facilitate cross-case comparison [2]. The primary reason for this is likely the scarcity of longitudinal data describing many, if not all, aspects of recovery, which is the result of the expense and difficulty associated with recovery data collection. This means that there is a critical constraint for theory building, hypothesis testing, best practices development, and case-based learning.

An under-utilized and more readily available source of data characterizing disaster recovery is big text data—corpora of documents (e.g., news stories, government reports, transcripts, etc.) that describe aspects of recovery for past or ongoing disasters. Consequently, we follow the example of [3], [4] who used text data from numerous news articles, government documents, and industry reports about four disaster events

to manually characterize the extent and severity of failure interdependencies associated with electric power outages. Such manual content analysis is labor-intensive and slow. Automated natural language processing (NLP) methods to search, compile, characterize, structure, and compare textual descriptions of disaster recovery are needed to increase the appeal of analyzing large text corpora to further research and inform pre- and post-event recovery planning.

To meet this need, we are developing a new NLP method to facilitate exploration and analysis of text corpora that describe disaster recovery. The aim of the method is to allow users to measure the degree that user-specified propositions about potential issues are embodied within corpora, which serves as a proxy for the disaster recovery process. The method employs a statistical syntax-based semantic matching model and was trained on a standard, publicly available training dataset.

We applied the NLP method to a newspaper story corpus that describe the impacts and recovery of Christchurch, New Zealand after the M7.1 September 2010 and M6.3 February 2011 Canterbury earthquakes. Prior to applying the method to explore recovery aspects for this case study, outputs of the method were evaluated through a user study of twenty professional emergency managers.

The following section summarizes literature on measuring disaster recovery and how NLP may be used for this purpose. Our case study—the 2010–2011 Canterbury earthquakes disaster—is presented to provide context for the NLP analysis of newspaper coverage. A non-technical overview of the NLP method we are developing is given, followed by description of the user-based evaluation of method outputs. We describe selected results from applying the NLP method to the news corpus to provide provisional insights into Christchurch's recovery from the earthquakes. From these results, we provide directions for future work.

II. BACKGROUND

A. Measuring Disaster Recovery

Disaster recovery can be quantified or measured in multiple and, at times, opposing ways [2], [5]. The most common way of measuring recovery is to compare post-event states to immediate pre-event states. Less commonly, recovery can be measured by comparing post-event states to some modeled or assumed counterfactual state without a disaster [6]. Disasters

can manifest substantial changes, adaptation, and restructuring so that impacted communities deliberately or unintentionally do not fully return to either a pre-event state or a foreseeable counterfactual state [7]. More realistically, then, recovery can be measured as longitudinal changes to diverse indicators of adaptation or community identity [8]. Most systematic means of measuring recovery in the research literature are quantification approaches that emphasize quantity of supply and speed of recovery. This does not represent qualitative characteristics of recovery such as negative or inequitable impacts of differential recovery across space and time [2], which are arguably more likely to be captured by text-based data (including transcripts).

B. Natural Language Processing for Measuring Disaster Recovery

The use of NLP to understand social, political, and economic processes—aspects of disaster recovery—has become popular with the increase in the volume of data about human communication, including text, audio, and video [8]. Example applications include automatic extraction of international events from political context [9], public opinion measurement from social media posts [10], sense of place [11], and community happiness [12]. There are a growing number of uses of NLP methods to understand topics of disasters [13], [14].

III. 2010–2011 CANTERBURY EARTHQUAKES DISASTER

A moderately damaging M7.1 earthquake struck the Canterbury region of New Zealand's South Island on September 4, 2010; the epicenter was located near the town of Darfield, approximately 35 kilometers west of the large city Christchurch. Six months later, on February 22, 2011, another earthquake struck but with the epicenter only 10 kilometers southeast of the central business district of Christchurch. The 2011 earthquake generated more than 7300 felt aftershocks in the first year alone. Measured ground acceleration from the earthquake was the highest ever recorded in New Zealand and one of the highest recorded worldwide at the time of the earthquake.

The February 2011 earthquake had significant impacts on community functioning and well-being in Christchurch, including the death of 185 people [15]–[17]. New Zealand Treasury estimated the capital cost of the Canterbury earthquakes to be around $40 billion. The high shaking intensity, the simultaneous vertical and horizontal ground movement, and the extreme liquefaction of the February 2011 earthquake caused significant damage. Following the 2011 earthquake, access to 45% of the 4,000 downtown buildings was banned for safety reasons, and 1,000 buildings were marked for demolition. Roughly 7,500 houses in Christchurch required demolition and zoning changes to restrict future construction, while almost 100,000 houses needed repairs.

The February 2011 earthquake damaged and disrupted the infrastructure of the city, including roads, water and wastewater networks, and electricity systems. Electric power was restored to 98% of occupied homes in less than two weeks of the earthquake. Roads and bridges were extensively damaged, as were water and wastewater systems. The Christchurch City Council received over 36,000 water and wastewater service requests in the six months following the earthquake. After those six months, around 800 houses remained without wastewater service.

In March 2011, the Canterbury Earthquake Recovery Authority (CERA) was established to lead economic, residential, social, ecological, and cultural recovery for the subsequent five years. SCIRT (Stronger Christchurch Infrastructure Rebuild Team) was formed to rebuild the city's horizontal infrastructure and, similar to CERA, sunset after five years. Funds for the recovery have largely been from a combination of sources: insurance (government and private), central government, local government (including borrowing), and private savings or debt. New Zealand's Earthquake Commission (EQC), a government organization, provides earthquake insurance at a very low rate to residential policyholders. As a result, compared to other earthquake disasters, insurance has played a larger role in recovery. At the time of writing, the recovery process in Christchurch is very much ongoing. A locally-based organization called Regenerate Christchurch was formed to essentially take over the mission of CERA and SCIRT. Significant issues seven years after the February 2011 are completing major public development projects in the central business district and planning for reuse of the extensive land that was formerly residential. Recovery, regardless of definition, for Christchurch is still years away.

IV. A NEW NLP METHOD FOR SEMANTIC MEASUREMENT OF DISASTER RECOVERY

Our work represents a new general application of natural language processing: measuring the extent to which an idea or proposition is expressed in a corpus. Specifically, we are developing a new NLP method to facilitate analysis of text corpora that describe disaster recovery. The aim of the method is to allow users, such as disaster researchers, emergency managers, and recovery planners, to measure the degree that user-specified propositions are matched within a corpus that describes aspects of the disaster recovery process. For example, end-users of a tool built on our research could include local emergency managers evaluating how public sentiment that "government is not doing enough" changes over time or researchers determining whether "lack of construction workers is slowing reconstruction" was more of a factor for one disaster versus another. The frequency of matches between a given proposition and candidate sentences within the corpus can be plotted against time to support understanding disaster recovery processes (Figure 1). (Additionally, results can be aggregated by location or other variables supported by available metadata.)

The presented method employs a statistical syntax-based semantic matching model; we call the desired output of this model "semantic measurement." Finding instances of a proposition in a corpus is similar to information or document retrieval, which is a component in many web-scale question answering systems. The main difference with our method is that it seeks more than a single answer to a query; it seeks *all* matches to the query. Since natural language offers many ways to express any idea, this matching must be more powerful than simple word or phrase-level matching.

Our approach is to create semantic matching models that return a high score if a candidate sentence within the analyzed corpus expresses a similar idea as in the queried proposition.

Proposition query:
"Dealing with authorities is causing stress and anxiety."

↓ query corpus

Matched sentences:
"Relationships Aotearoa... said it is unfamiliar bureaucratic systems which are causing the majority of the stress."

"... those in charge of the earthquake recovery are making moves to appease the growing anger among homeowners."
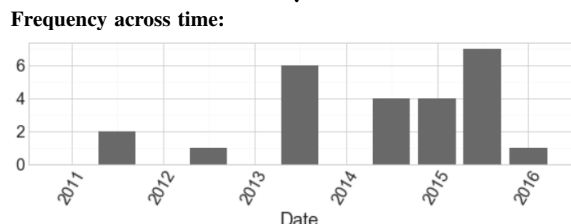
↓ aggregate

Frequency across time:

Fig. 1. Illustration of the new natural language processing method for computing the frequency that user-specified propositions are expressed in a corpus that describes a disaster recovery process. The frequency of these occurrences can be plotted across time.

Our method employs a two-step process to balance speed and precision. First, a faster and coarser semantic matching model is used to filter sentences in the corpus to a subset likely to match the query proposition. Second, a richer and more expensive model is used to score matches and optionally return the sentences in the corpus that express the idea. A conceptual overview is given in the rest of the section; the reader is referred to [18] for formal description of our method.

For the fast filter, we use a simple scoring function inspired by work of [19], [20]. We tested two versions of the filter based on different training data—a larger, more general-purpose corpus of Google News articles [21], and a smaller dataset designed for semantic similarity applications similar to ours [19]. This step provides efficiency, since many sentences in a corpus can be easily determined not to match a provided sentence.

After scoring every sentence in the corpus according to the fast filter, the method then applies a richer semantic matching model to the top n candidates for selection. For this matching function, we start with a syntax-based model introduced by [22]. We use this model because it is simple, interpretable, and suitable for a range of semantic similarity problems. We trained two versions of the model: the original logistic regression version and a new neural network. For training data, we used the Stanford Natural Language Inference corpus (SNLI; [23]). SNLI contains approximately 570,000 pairs of sentences that linguistically express a premise and a hypothesis. Each sentence pair is human-annotated with an entailment, contradiction, or neutral label of the relationship between the two sentences.

V. User Evaluation of Disaster Recovery NLP Method

To evaluate the method described above, we conducted a user study with twenty emergency managers using a convenience and snowball sampling strategy. We first created a

TABLE I. Examples of scored candidate sentences provided to user survey study participants.

	Idea: There is a shortage of construction workers.
Score (1-5)	Example candidate sentence (bold) and preceding and following sentences for context.
1	The data was the latest demand and supply information on the Canterbury rebuild and wider recovery, MBIE said. **The quarterly report for Canterbury included analysis on Greater Christchurch Value of Work, Employment and Accommodation projections.** The forecasts were based on Canterbury Earthquake Recovery Authority projections of work to be done on the residential rebuild and repairs, infrastructure and commercial work.
3	Migrants were now filling most of the rising number of construction jobs but beneficiaries moving into work were also contributing, MBIE's quarterly "job-matching" report said. **The construction sector's workload was expected to peak in the December 2016 quarter at a value of about $1.6 billion.** The residential rebuild would run at "elevated levels" from 2015 until 2018 but commercial work would become increasingly important.
5	The additions to the current workforce of 30,000 will mostly work on commercial projects or infrastructure, the Ministry of Business, Innovation and Employment (MBIE) predicts. **Greater Christchurch's labour supply for the rebuild was tight and was likely to remain that way for the next three years.** Migrants were now filling most of the rising number of construction jobs but beneficiaries moving into work were also contributing, MBIEs quarterly "job-matching" report said.

corpus of 982 earthquake-related articles from New Zealand newspaper websites (stuff.co.nz, nzherald.co.nz) that were published sometime between 2011 and 2016. We composed 20 proposition queries that cover topics like community well-being, infrastructure restoration, decision-making, and public opinion. An example query is: "The council should have consulted residents before making decisions." The output scores for each candidate sentence express the degree of match between each candidate sentence in the corpus and the queried proposition. For each sentence in the corpus, we generated separate scores related to all 20 propositions using four combinations of fast filter and syntax-based matching models; we set the number of sentences passed through the filter to the syntax-based model to be 250. Because it was not feasible to have study participants evaluate every sentence in the corpus, we instead sampled from the following output categories (per query) as follows: 25 random sentences not selected by the filter, 25 random sentences selected by the filter but not matched by the syntax-based model, the 25 highest-scoring sentences from the filter (without applying the syntax-based model), and the 25 highest-scoring sentences matched by the method as a whole.

Using email and a web-based form, we gave each study participant the following prompt: "Given an idea sentence, score each candidate sentence on a 1–5 scale based on how well it expresses the idea. The preceding and following sentences for each candidate are provided for context, but please score the quality of only the bolded candidate sentence." Criteria for each score level were provided to the participants as the following: 1) "...is completely unrelated to the idea sentence", 2) "...is tangentially related to the idea sentence", 3) "...is related to but does not adequately express the idea sentence", 4) "...almost expresses the idea sentence", and
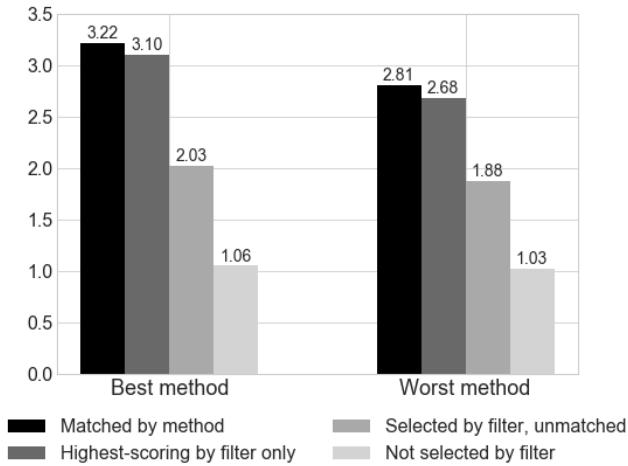
Fig. 2. Example results from the user study for the best and worst performing fast filter and matching method combinations. Decreasing saturation of bars indicate decreasing expectation that the method identifies a match between a candidate sentence and given proposition.

Fig. 3. The computed frequency for six-month intervals that the proposition "Dealing with authorities is causing stress and anxiety" is expressed in a corpus of New Zealand news articles about recovery after the 2010–2011 Canterbury earthquakes.

TABLE II. RANDOMLY-SELECTED SUBSET OF SENTENCES FOUND AS EXPRESSING THE PROPOSITION "DEALING WITH AUTHORITIES IS CAUSING STRESS AND ANXIETY" IN A CORPUS OF NEW ZEALAND NEWS ARTICLES ABOUT RECOVERY AFTER THE 2010–2011 CANTERBURY EARTHQUAKES.

1) Stress and anxiety caused by dealing with authorities was "more debilitating" than the quakes, she said.
2) The battle to get back what they lost created huge financial stress.
3) $800m spent on east Christchurch not enough for frustrated residents
4) Most policies have clauses allowing replacement with "materials in common use", frustrating owners of older villas and bungalows.
5) A Canterbury Earthquake Recovery Authority draft document on their psychosocial plan for the city says anxiety and stress will continue to dog the population due to ongoing battles with insurance, land issues, changes to schooling and problems rebuilding homes and businesses.
6) The key findings indicate that the secondary stressors of damaged homes, insurance wrangles, financial challenges and grief over the 'lost Christchurch' are taking its toll.
7) Add to this the growing frustration among the new, youthful leaders of the community who emerged in the wake of the quakes.
8) Mr Hodder accepted that the homeowners involved have been subjected to "great distress" but added that it was a "difficult" situation and it has taken time to get to this point.
9) ICNZ chief executive Tim Grafton said insurers could understand homeowners, whose claim had just reached their insurer, were frustrated.
10) Council anchor projects unit manager Liam Nolan said the council recognised retailers' frustration that the carpark has not come down sooner.

5) "...expresses the idea sentence in its entirety". We also provided participants a sample idea sentence and pre-scored candidate sentences (Table I). For ease of understanding, we used the term "idea sentence" when referring to the proposed sentence instead of the more technical "proposition query."

In total, each participant scored approximately 400 candidate sentences spread across multiple proposition queries, filter-matching method combination (among four), and output category (matched by method; highest-scoring from filter alone; selected by filter but not matched; not selected by filter). To allow calculation of agreement between users, half of the sentences received three judgments. Using the overlapping judgements, we computed Krippendorf's α [24] to be 0.784, which indicates reasonable agreement when users rate the same sentence.

Example results from the user study are shown in Figure 2 for the best and worst performing method combinations. On average, study participants consistently gave low scores (around 1) to candidate sentences that were rejected by the particular fast filter. Lower scores were often given to candidate sentences that only partially expressed the queried proposition—possibly sentences selected by the filter but not matched. Significantly higher scores were given to candidate sentences that were scored high enough by the method to be consider a match. However, the average user-provided score was around 3 on average—based on the provided guidelines, "...related to but does not adequately express the idea sentence."

We conducted a follow-up user study. It was executed similarly to the first one described above, but with proposition queries solicited from users themselves during the first survey. Instead of randomly distributing sentences among the follow-up study participants, we gave each user who participated in the follow-up the output for their own proposition queries. There were 18 idea sentences and seven participants in this study. Each participant scored approximately 400 sentences, which were dr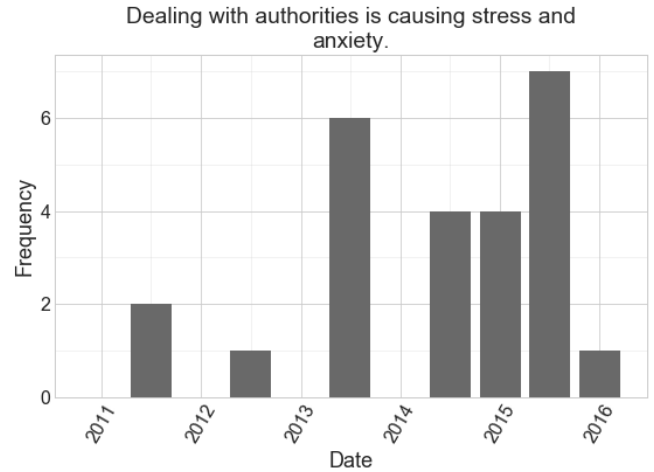awn from different parts of the output (as in the original study). We found that the follow-up study replicated the findings of the first study.

VI. SEMANTIC ANALYSIS OF CANTERBURY DISASTER RECOVERY NEWS CORPUS

We used the best performing method presented in the previous section to explore six-year recovery trends after the 2010–2011 Canterbury earthquakes. We computed trends of the frequency in which the same 20 propositions provided to the user study participants are expressed in the corpus of New
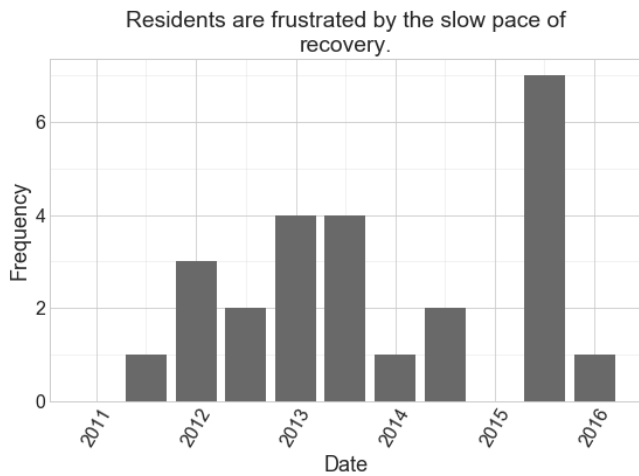
Fig. 4. The computed frequency for six-month intervals that the proposition "Residents are frustrated by the slow pace of recovery" is expressed in a corpus of New Zealand news articles about recovery after the 2010–2011 Canterbury earthquakes.

TABLE III. RANDOMLY-SELECTED SUBSET OF SENTENCES FOUND AS EXPRESSING THE PROPOSITION "RESIDENTS ARE FRUSTRATED BY THE SLOW PACE OF RECOVERY" IN A CORPUS OF NEW ZEALAND NEWS ARTICLES ABOUT RECOVERY AFTER THE 2010–2011 CANTERBURY EARTHQUAKES.

1) Residents are frustrated by the slow pace of recovery and the lack of proposed city council spending in New Brighton and surrounding suburbs.
2) She will also be asking them to put aside an extra $3m for the rejuvenation of New Brighton as part of a package of proposals aimed at addressing residents' concerns about the slow pace of recovery in the east.
3) It has slowed their recovery down hugely and made life for many much harder.
4) The submissions expressed frustration at the slow pace of recovery and the lack of proposed spending in New Brighton and the eastern suburbs.
5) Power has been restored to over 60 per cent of quake-ravaged Christchurch but progress is slow, lines company Orion says.
6) Read more: Christchurch anchor project delays cause frustration.
7) Many Christchurch residents have expressed anger at the slow pace of settling claims for damaged homes.
8) People are also learning to adjust to the pace of the recovery.
9) Worry, despair plague Christchurch residents.
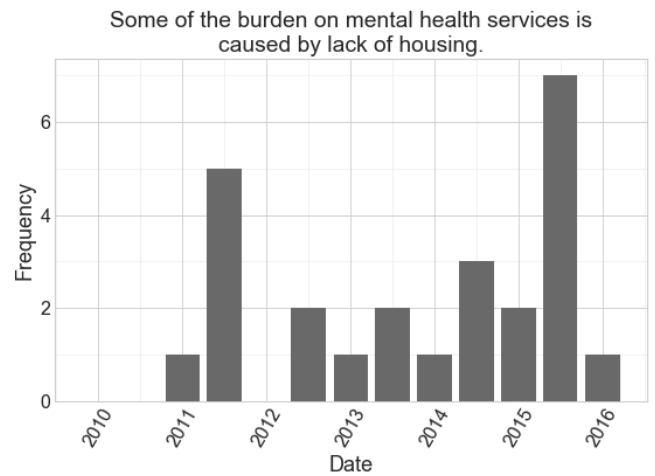10) He said progress in Christchurch is "far too slow".

Fig. 5. The computed frequency for six-month intervals that the proposition "Some of the burden on mental health services is caused by lack of housing" is expressed in a corpus of New Zealand news articles about recovery after the 2010–2011 Canterbury earthquakes.

TABLE IV. RANDOMLY-SELECTED SUBSET OF SENTENCES FOUND AS EXPRESSING THE PROPOSITION "SOME OF THE BURDEN ON MENTAL HEALTH SERVICES IS CAUSED BY LACK OF HOUSING" IN A CORPUS OF NEW ZEALAND NEWS ARTICLES ABOUT RECOVERY AFTER THE 2010–2011 CANTERBURY EARTHQUAKES.

1) "We can see what it is that is causing the mental health issues and we know that if this person could get their housing issues sorted they would be transformed."
2) Four years on we are in the early days of recovery, and the anxiety about aftershocks has given way to stresses about insurance, repairs and relocating offices, schools and homes.
3) Priority has been given to essential service locations [medical facilities, etc] and main streets and thoroughfares."
4) Key urged anxious homeowners to be patient.
5) Some of the burden on mental health services is undoubtedly caused by the lack of housing in Christchurch.
6) Housing security was vital for recovery from mental health illness and for independent living, Duffy said.
7) Generally in Auckland, there will be more demand for residential building because of the lack of housing supply.
8) The pressure on mental health services continued to rise with more homeless mental health patients taking up beds.
9) For the majority of mental health patients cared for in the community, housing stress was a big element in their ill-health.
10) People outside Christchurch also find it hard to understand how mental health could be so affected by the insecurity of housing, or "the importance of your home to your sense of freedom and personal dignity".

Zealand new articles that describe Christchurch's recovery. For each proposition, the 25 sentences with the highest computed match were selected by our method, the publication dates of each sentence's source articles were identified via metadata, and frequencies were calculated and plotted for six-month intervals.

Figure 3 through Figure 8 present histograms of computed semantic matches for six of the 20 propositions. Table II through Table VII include a random subset of matched sentences included in each recovery plot to illustrate specific outputs of our method. For example, Figure 3 is a plot of match frequencies for the proposition "dealing with authorities is causing stress and anxiety." The plot shows a clear upward

trend in expressions about the proposition within the analyzed news corpus. The trend is reasonable considering immediate focus on response after a damaging event, the common "honeymoon" period soon thereafter, and the eventual rise in public frustration as expectations are not met (e.g., due to dealing with bureaucracy or insurance agencies). However, the computed frequency range is not large (1 to 7). Candidate sentences listed in Table II illustrate a subjective range in how well each sentence matches the proposition (e.g., sentence 1 vs. sentence 9). Arguably, topics within the proposition tend to be relatively well expressed in the candidate sentences, but less
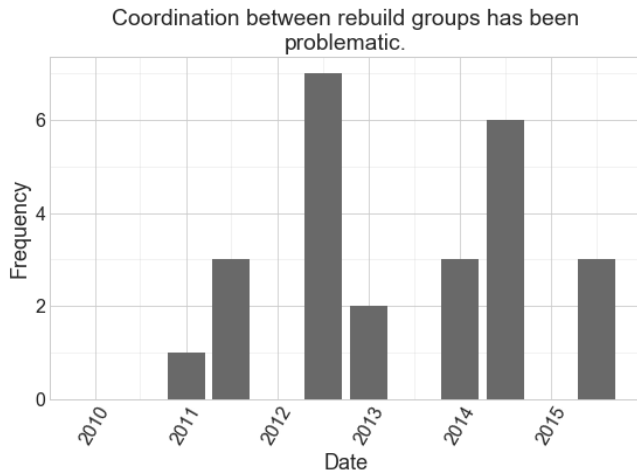
Fig. 6. The computed frequency for six-month intervals that the proposition "Coordination between rebuild groups has been problematic" is expressed in a corpus of New Zealand news articles about recovery after the 2010–2011 Canterbury earthquakes.

TABLE V.   RANDOMLY-SELECTED SUBSET OF SENTENCES FOUND AS EXPRESSING THE PROPOSITION "COORDINATION BETWEEN REBUILD GROUPS HAS BEEN PROBLEMATIC" IN A CORPUS OF NEW ZEALAND NEWS ARTICLES ABOUT RECOVERY AFTER THE 2010–2011 CANTERBURY EARTHQUAKES.

1) Householders often have different policies and various agencies have to be co-ordinated.
2) There will be lots of co-ordination needed between stakeholders, designers and industry.
3) After more than two years, and tired of fighting insurance companies and bureaucracy for what they believe is a fair deal, the group, described as the "Buggered All Stars Chorus", hope to use the song to raise awareness around New Zealand of their situation.
4) The challenge for the redevelopment of the city is to build demand for commercial, residential and retail space while planning for that redevelopment to occur in a coordinated way that lives up to the vision in the Central City Plan.
5) Appointing commissioners is not the answer to these difficult problems.
6) The Fletcher hubs oversee reconstruction and repair work throughout the rebuild area, where teams of building managers and advisers assess and co-ordinate work.
7) The new authority would pull together all of the resources of central government going into the city and co-ordinate where they went.
8) Calls are flooding into police communications centres around New Zealand, causing problems for staff already trying to deal with the massive earthquake and aftershocks which have hit Christchurch and Canterbury.
9) Another arm's length relationship was created that became a prime problem for the Blueprint [recovery plan].
10) He believed a coordinated response from the various organisations and agents involved with Christchurch's property market was key to dealing with the housing shortage.

so the (characteristics of) relationships within the sentences.

The remainder of the figures and tables are presented without discussion because of space constraints. General observations for these are similar to those for Figure 3 and Table II. The significance of these observations is that, with the current level of performance of our method, the greatest insight
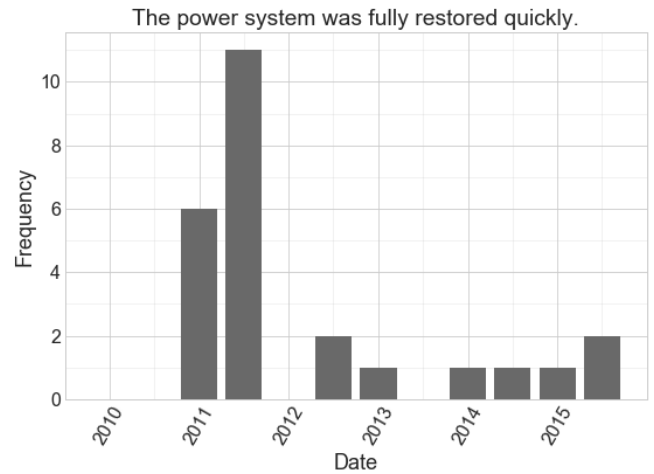

Fig. 7. The computed frequency for six-month intervals that the proposition "The power system was fully restored quickly" is expressed in a corpus of New Zealand news articles about recovery after the 2010–2011 Canterbury earthquakes.

TABLE VI.   RANDOMLY-SELECTED SUBSET OF SENTENCES FOUND AS EXPRESSING THE PROPOSITION "THE POWER SYSTEM WAS FULLY RESTORED QUICKLY" IN A CORPUS OF NEW ZEALAND NEWS ARTICLES ABOUT RECOVERY AFTER THE 2010–2011 CANTERBURY EARTHQUAKES.

1) Canterbury electricity supplier Orion is confident that 90% of Christchurch city will have power restored by nightfall, says spokesman Roger Sutton.
2) Emergency services work frantically to restore water, power and sewerage systems and check for people who might still be trapped in their homes.
3) Then Orion's chief executive, his handling of the power lines company's staff working long hours to restore electricity to the city post-quake impressed many.
4) He had no water but power had been restored in his area.
5) The need to regain some sense of some control over one's life is central to the recovery process.
6) As in Christchurch, electricity, sewerage and chimneys were all down in post-quake Hawke's Bay.
7) Lines company Orion said the last of Christchurch's main power substations to be connected - Brighton substation - was confirmed as working yesterday.
8) City councillors have voted to fully restore the earthquake-damaged Christchurch Town Hall.
9) It had been unable to access the electricity network to restore power and the situation could remain for the next few days.
10) An update from Orion showed power had been restored to 87 per cent of Christchurch yesterday.

about a given recovery proposition is gained through concurrent inspection of time-based visualization and sentence-level inspection of method outputs. Further, these insights should be put in context with attributes of the analyzed corpus.

In this case, Christchurch's disaster recovery process is represented by 982 earthquake-related articles from New Zealand newspaper websites that were published between 2011 and 2016. Computed trends and visualizations of the trends is likely influenced by the size of the corpus. How and whether certain propositions are expressed in the corpus for a computed time period after the earthquake depend on the content of news
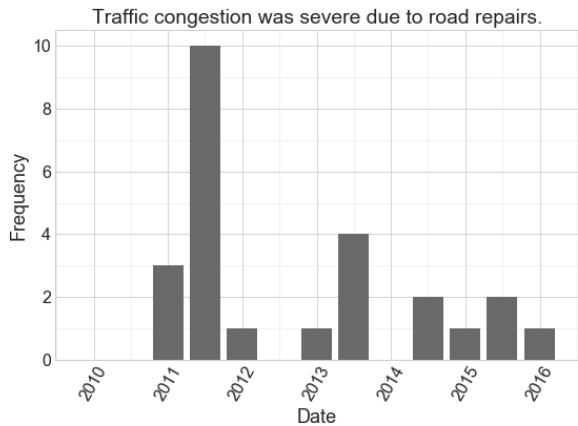
Fig. 8. The computed frequency for six-month intervals that the proposition "Traffic congestion was severe due to road repairs" is expressed in a corpus of New Zealand news articles about recovery after the 2010–2011 Canterbury earthquakes.

TABLE VII.   RANDOMLY-SELECTED SUBSET OF SENTENCES FOUND AS EXPRESSING THE PROPOSITION "TRAFFIC CONGESTION WAS SEVERE DUE TO ROAD REPAIRS" IN A CORPUS OF NEW ZEALAND NEWS ARTICLES ABOUT RECOVERY AFTER THE 2010–2011 CANTERBURY EARTHQUAKES.

1) Karen Atkinson just up his road was on the verge of agreeing to a $280,000 repair job on her home. She called a rapid halt after her insurer told her that, if she wanted her house lifted, she would have to pay for it herself.

2) More than 1.2 million homes were without water, traffic was congested and fires started from gas mains rupturing could not be fought.

3) There was significant traffic congestion and drivers were told to avoid Ferry Road.

4) Congestion was making the road repairs more difficult and was holding up the delivery and servicing of portable toilets, the movement of trucks removing silt and numerous other important services.

5) Campbell-Reid says the council has developed strong criteria around the performance of new buildings, and it is looking at converting road space into green space so the central city streets become about people rather than rush-hour traffic.

6) Emergency service lines are working but remained heavily congested.

7) Civil Defence is asking people not to move road barriers and other signs on Christchurch city streets.

8) People may have been crushed or trapped by collapsing buildings, and brick facades.

9) Three years on from the February 2011 earthquake, the city's road network continues to be disrupted with lines of traffic cones, constant road closures and diversions.

10) People were being asked to stay off the roads to reduce congestion.

coverage and writing style of the individual articles, as well as the actual prevalence of the issue. Different or combined corpora, which could include government documents, firsthand accounts, or scholarly publications, will likely influence semantic matching results—something not investigated to date.

Beyond attributes of the analyzed corpus, our method does not take multiple sentences into account at the same time when determining sentence matches. (For the user study, sentences bracketing the scored candidate sentence were provided as context only.) In some cases, this leads to our method finding a match at the sentence level when it would otherwise be invalid from context; in others, a potential match is spread across a sentence boundary. Furthermore, our method is dependent on general-purpose training datasets (like SNLI). Disaster-specific entities like government agencies and insurance companies may be central to texts describing the disaster, but these general training sets may not contain such unique entities.

VII.   CONCLUSION

This paper presents a new NLP method to facilitate analysis of text corpora that describe disaster recovery. The aim of the method is to allow users to measure the degree that user-specified propositions are expressed within a corpus that describe a disaster recovery process. We applied our method to a news story corpus that describes the recovery of Christchurch, New Zealand from the 2010–2011 Canterbury earthquake sequence. The semantic measurements output by the method were evaluated through a user study of twenty professional emergency managers. The user study results show that our method can be useful for gaining insight into disaster recovery from a specific event when the method is applied to an appropriate news article corpus and, potentially, large corpora in general.

As part of the first user survey described above, we asked study participants whether they are interested in natural language processing as a way to measure recovery-related propositions in news or other text corpora; 85% responded positively. Thus, we are encouraged by the potential for future method improvements and its application for after-action learning, recovery decision making, and disaster research. NLP, and specifically semantic matching of large text data, may facilitate unique and novel understandings of disaster recovery; for example, we envision that disaster management researchers and practitioners could use qualitative text data from past disasters to inform pre-event long-term recovery planning or perform exploratory research for generating new hypotheses.

As mentioned in Section VI, improvements to our method might include accounting for larger and smaller passages (not only sentences), especially if this is coupled with more preprocessing, such as finding all expressions that refer to the same entity. While the richer semantic matching models we investigated offer small improvements over the results from just the fast filter (Figure 2), our method should be improved through detecting more than sentence-level entailment. In addition, a more involved user study that includes more significant user training and engagement time is warranted. This study could leverage their assistance in tuning the threshold for determining what candidate sentence score warrants being considered a match, either in a query-specific or general way. (Because semantic matches to some of the proposition queries are sparse, sensitivity to this threshold might be very high.) Lastly, to ensure utility and credibility of method outputs, it is important to investigate the effectiveness of visualizations, particularly those that integrate quantitative and qualitative representations interactively.

REFERENCES

[1] S. E. Chang, "Urban disaster recovery: A measurement framework and its application to the 1995 Kobe earthquake," *Disasters*, vol. 34, no. 2, pp. 303–327, 2010.

[2] S. E. Chang, T. McDaniels, J. Fox, R. Dhariwal, and H. Longstaff, "Toward disaster-resilient cities: Characterizing resilience of infrastructure systems with expert judgments," *Risk Analysis*, vol. 34, no. 3, pp. 416–434, 2014.

[3] T. McDaniels, S. E. Chang, K. Peterson, J. Mikawoz, and D. Reed, "Empirical framework for characterizing infrastructure failure interdependencies," *Journal of Infrastructure Systems*, vol. 13, no. 3, pp. 175–184, 2007.

[4] S. B. Miles and S. E. Chang, "Modeling community recovery from earthquakes," *Earthquake Spectra*, vol. 22, no. 2, pp. 439–458, 2006.

[5] W. duPont and I. Noy, "What happened to Kobe? A reassessment of the impact of the 1995 earthquake in Japan," *Economic Development and Cultural Change*, vol. 63, no. 4, pp. 777–812, 2015.

[6] J. Kennedy, J. Ashmore, E. Babister, and I. Kelman, "The meaning of 'build back better': Evidence from post-tsunami Aceh and Sri Lanka," *Journal of Contingencies and Crisis Management*, vol. 16, no. 1, pp. 24–36, 2008.

[7] S. B. Miles, "Foundations of community disaster resilience: Well-being, identity, services, and capitals," *Environmental Hazards*, vol. 14, no. 2, pp. 103–121, 2015.

[8] B. O'Connor, D. Bamman, and N. A. Smith, "Computational text analysis for social science: Model assumptions and complexity," in *Proceedings of the NIPS Workshop on Computational Social Science and the Wisdom of Crowds*, 2011.

[9] B. O'Connor, B. M. Stewart, and N. A. Smith, "Learning to extract international relations from political context," in *Proceedings of the 51st Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, 2013.

[10] B. O'Connor, R. Balasubramanyan, B. R. Routledge, and N. A. Smith, "From tweets to polls: Linking text sentiment to public opinion time series," in *Proceedings of the 4th International AAAI Conference on Weblogs and Social Media*, 2010.

[11] B. Adams and M. Raubal, "Identifying salient topics for personalized place similarity," *Research@Locate*, 2014.

[12] N. Ramírez-Esparza, C. K. Chung, G. Sierra-Otero, and J. W. Pennebaker, "Cross-cultural constructions of self-schemas: Americans and Mexicans," *Journal of Cross-Cultural Psychology*, vol. 43, no. 2, pp. 233–250, 2012.

[13] Y.-R. Lin and D. Margolin, "The ripple of fear, sympathy and solidarity during the Boston bombings," *EPJ Data Science*, vol. 3, no. 1, p. 31, 2014.

[14] M. A. Cohn, M. R. Mehl, and J. W. Pennebaker, "Linguistic markers of psychological change surrounding September 11, 2001," *Psychological Science*, vol. 15, no. 10, pp. 687–693, 2004.

[15] S. B. Miles, D. Brechwald, R. Davidson, K. Demeter, D. Johnston, S. Pampanin, and S. Wilkinson, "Building back better–case study of the 2010–2011 Canterbury, New Zealand earthquake sequence," *New Zealand Earthquake Sequence (A Learning from Earthquakes Report)*, 2014.

[16] S. E. Chang, J. E. Taylor, K. J. Elwood, E. Seville, D. Brunsdon, and M. Gartner, "Urban disaster recovery in Christchurch: The Central Business District Cordon and other critical decisions," *Earthquake Spectra*, vol. 30, no. 1, pp. 513–532, 2014.

[17] J. Morgan, A. Begg, S. Beaven, P. Schluter, K. Jamieson, S. Johal, D. Johnston, and M. Sparrow, "Monitoring wellbeing during recovery from the 2010–2011 Canterbury earthquakes: The CERA Wellbeing Survey," *International Journal of Disaster Risk Reduction*, vol. 14, pp. 96–103, 2015.

[18] L. H. Lin, S. Miles, and N. A. Smith, "Semantic matching against a corpus: New applications and methods," 2018, arXiv:1808.09502 [cs.CL].

[19] J. Wieting, M. Bansal, K. Gimpel, and K. Livescu, "Towards universal paraphrastic sentence embeddings," in *International Conference on Learning Representations*, 2016.

[20] M. Iyyer, V. Manjunatha, J. Boyd-Graber, and H. Daumé III, "Deep unordered composition rivals syntactic methods for text classification," in *Proceedings of the 53rd Annual Meeting of the Association for Computational Linguistics and the 7th International Joint Conference on Natural Language Processing (Volume 1: Long Papers)*, 2015.

[21] T. Mikolov, I. Sutskever, K. Chen, G. Corrado, and J. Dean, "Distributed representations of words and phrases and their compositionality," in *Advances in Neural Information Processing Systems*, 2013.

[22] M. Heilman and N. A. Smith, "Tree edit models for recognizing textual entailments, paraphrases, and answers to questions," in *Human Language Technologies: The 2010 Annual Conference of the North American Chapter of the Association for Computational Linguistics*, 2010.

[23] S. R. Bowman, G. Angeli, C. Potts, and C. D. Manning, "A large annotated corpus for learning natural language inference," in *Proceedings of the 2015 Conference on Empirical Methods in Natural Language Processing*, 2015.

[24] K. Krippendorff, *Content analysis: An introduction to its methodology*, 2012.