8. AMEGH P ARUN

NATURAL LANGUAGE PROCESSING FOR ANALYZING DISASTER RECOVERY TRENDS EXPRESSED IN LARGE TEXT CORPORA

Abstract

In this report, a new natural language processing (NLP) method is developed 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.The NLP method is applied to a news story corpus that describes the recovery of Christchurch, New Zealand after the 2010–2011 Canterbury earthquake sequence.The proposed model is used to compute semantic measurements of multiple potential recovery issues as expressed in the Christchurch news corpus that span 2011 to 2016. The 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.This proposed system can be used in future applications of NLP method for after-action learning, recovery decision making and disaster research.