GALLOC: A GeoAnnotator for Labeling LOCation Descriptions from Disaster-Related Messages

Kai Sun, Yingjie Hu, Kenneth Joseph, and Ryan Zhenqi Zhou

ABSTRACT: Disaster-related text messages often contain important locations descriptions about victims and accidents. Accurately extracting these descriptions can help disaster responders reach victims more quickly and even save lives. Recent research has shown that these location descriptions are often beyond simple place names, and cannot be extracted using typical named entity recognition (NER) approaches. While new machine learning models can be trained, they require labeled training data that are difficult to create without an effective data annotation tool. To fill this gap, we develop GALLOC, a GeoAnnotator for Labeling LOCation descriptions from disaster-related text messages.

Introduction

During a natural disaster, people often post messages on social media platforms to share urgent information and seek help (Huang & Xiao, 2015; Zou et al., 2022; Suwaileh et al., 2023). Many messages contain location descriptions about victims and accidents. Accurately extracting these location descriptions can help responders reach victims more quickly and even save lives. Existing studies often used an NER approach to extract location entities (Berragan et al., 2022; Wang et al., 2020). However, location descriptions may consist of multiple entities and cannot be extracted using an NER approach. As shown in Figure 1, typical NER approaches extract separate entities while we need the complete address to geo-locate the victims.
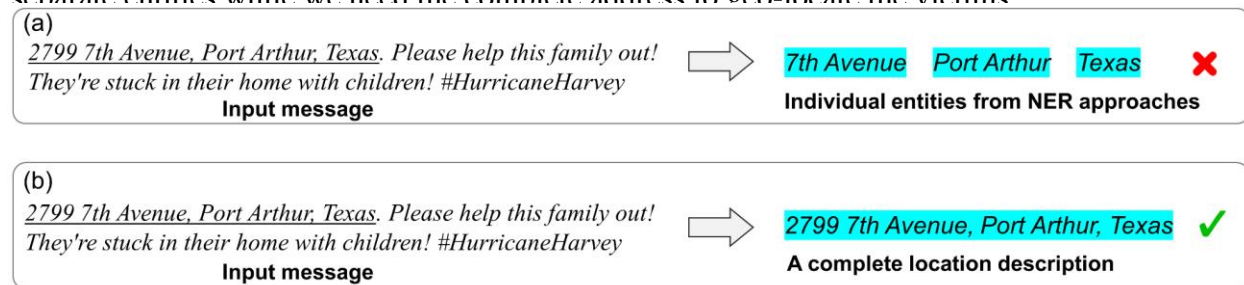

Figure 1. A comparison of (a) separate entities output from typical NER approaches and (b) complete location description needed for geo-locating victims.

Technically, we can overcome the above issue by training new machine learning models. However, there is a lack of datasets labeled with complete location descriptions, and it is difficult to create new datasets without an effective data annotation tool. There exist some data annotation tools, such as WOTR GeoAnnotator (DeLozier et al., 2016), PSU GeoAnnotator (Karimzadeh & MacEachren, 2019), and GeoViz (McDonough et al., 2019). However, they are limited in supporting the annotation of location descriptions and their categories. In this paper, we present GALLOC: a *GeoA*nnotator for *L*abeling *LOC*ation descriptions from disaster-related text messages.

Method

We design the overall architecture of GALLOC as in Figure 2. It consists of three major modules: *User Module*, *Project Module*, and *Annotation and Resolution Module*. Together, the three modules support the management of users and projects, and provide various functions for facilitating data annotation.
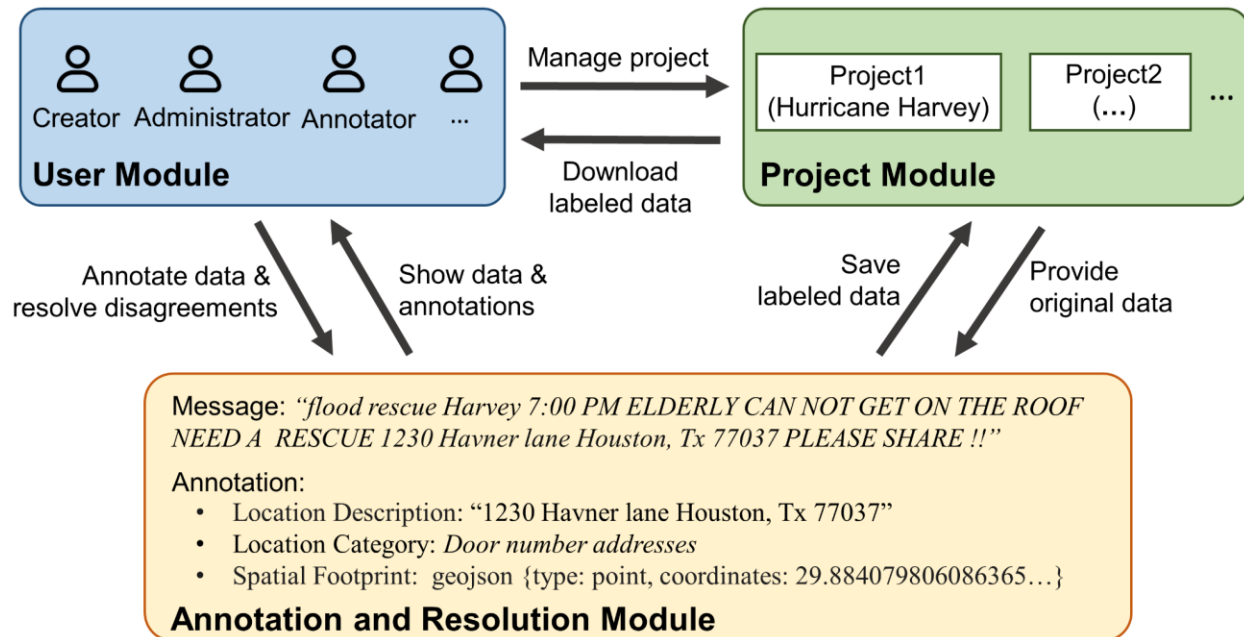

Figure 2. The overall architecture of GALLOC.

Results

GALLOC is designed and developed as an open-source and Web-based geo-annotation tool (Figure 3). It provides seven major functions: 1) annotating complete location descriptions; 2) annotating location description categories; 3) providing automatic pre-annotation; 4) providing automatic identification of spatial footprints; 5) supporting customized drawing of spatial footprints; 6) supporting multi-user collaborative annotation; and 7) supporting multilingual data annotation. An online demo is available at: https://geoai.geog.buffalo.edu/GALLOC.
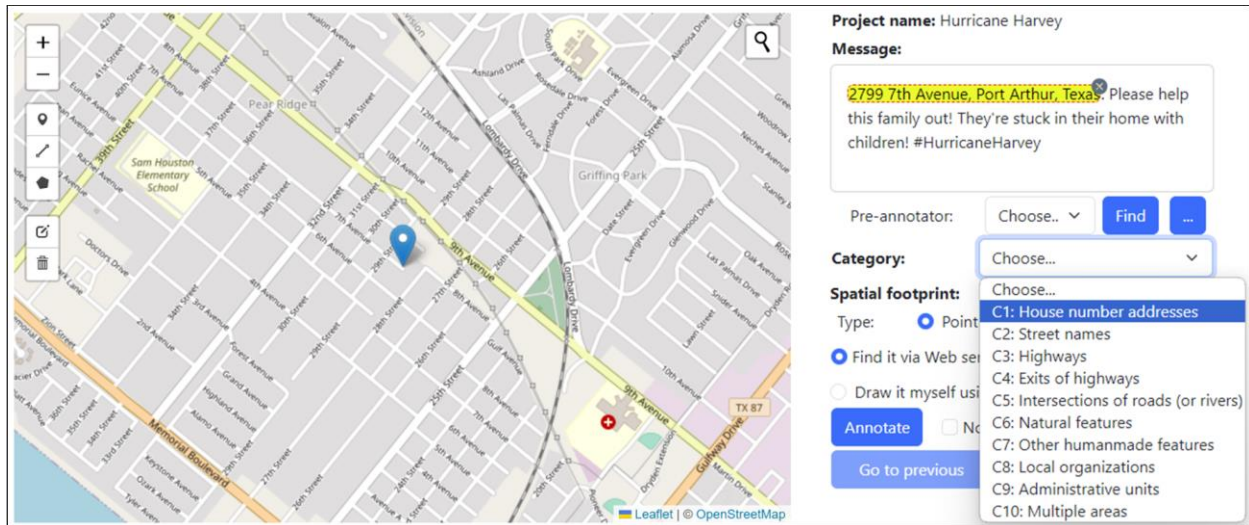
Figure 3. A screenshot of GALLOC.

To further evaluate GALLOC, we compare its functions with those of other existing geo-annotation tools (Table 1).

Table 1. A comparison between GALLOC and other tools.

Main Functions	GALLOC	WOTR Geoannotator	PSU Geoannotator	Geoviz
Annotating complete location descriptions not separate entities	✓	✓		
Annotating categories of location descriptions based on a user-defined scheme	✓			
Providing pre-annotation of location descriptions from text	✓ (AI-powered pre-annotation; can be updated with new technologies)		✓ (using a combination of six NER tools; cannot be updated)	✓ (using Edinburgh Geoparser; cannot be updated)
Identifying spatial footprints automatically	✓ (using Nominatim and Google Maps to geo-locate descriptions)		✓ (based on place name matching using GeoNames)	✓ (based on place name matching using GeoNames)
Drawing customized spatial footprints	✓	✓		
Supporting multi-user collaborative annotation	✓		✓	
Annotating multilingual datasets	✓			

Conclusions

This work develops GALLOC as a geoannotation tool to support the labeling of location descriptions from disaster-related text messages.

References:

Berragan, C., Singleton, A., Calafiore, A., & Morley, J. (2022). Transformer based named entity recognition for place name extraction from unstructured text. *International Journal of Geographical Information Science*, 1–20.

DeLozier, G., Wing, B., Baldridge, J., & Nesbit, S. (2016). Creating a novel geolocation corpus from historical texts. *Proceedings of the 10th Linguistic Annotation Workshop Held in Conjunction with ACL 2016 (LAW-X 2016)*, 188–198.

Huang, Q., & Xiao, Y. (2015). Geographic situational awareness: Mining tweets for disaster preparedness, emergency response, impact, and recovery. *ISPRS International Journal of Geo-Information*, *4*(3), 1549–1568.

Karimzadeh, M., & MacEachren, A. M. (2019). GeoAnnotator: A collaborative semi-automatic platform for constructing geo-annotated text corpora. *ISPRS International Journal of Geo-Information*, *8*(4), 161.

McDonough, K., Moncla, L., & Van de Camp, M. (2019). Named entity recognition goes to old regime France: Geographic text analysis for early modern French corpora. *International Journal of Geographical Information Science*, *33*(12), 2498–2522.

Suwaileh, R., Elsayed, T., & Imran, M. (2023). IDRISI-RE: A generalizable dataset with benchmarks for location mention recognition on disaster tweets. *Information Processing & Management*, *60*(3), 103340.

Wang, J., Hu, Y., & Joseph, K. (2020). NeuroTPR: A neuro-net toponym recognition model for extracting locations from social media messages. *Transactions in GIS*, *24*(3), 719–735.

Zou, L., Liao, D., Lam, N. S., Meyer, M. A., Gharaibeh, N. G., Cai, H., Zhou, B., & Li, D. (2022). Social media for emergency rescue: An analysis of rescue requests on Twitter during Hurricane Harvey. *International Journal of Disaster Risk Reduction*, 103513.

Kai Sun, Postdoc Associate, GeoAI Lab, Department of Geography,
University at Buffalo, Buffalo, NY, 14261

Yingjie Hu, Associate Professor, GeoAI Lab, Department of Geography,
University at Buffalo, Buffalo, NY, 14261

Kenneth Joseph, Assistant Professor, Department of Computer Science and Engineering,
University at Buffalo, Buffalo, NY, 14260

Ryan Zhenqi Zhou, PhD Candidate, GeoAI Lab, Department of Geography,
University at Buffalo, Buffalo, NY, 14261