University at Buffalo
The State University of New York

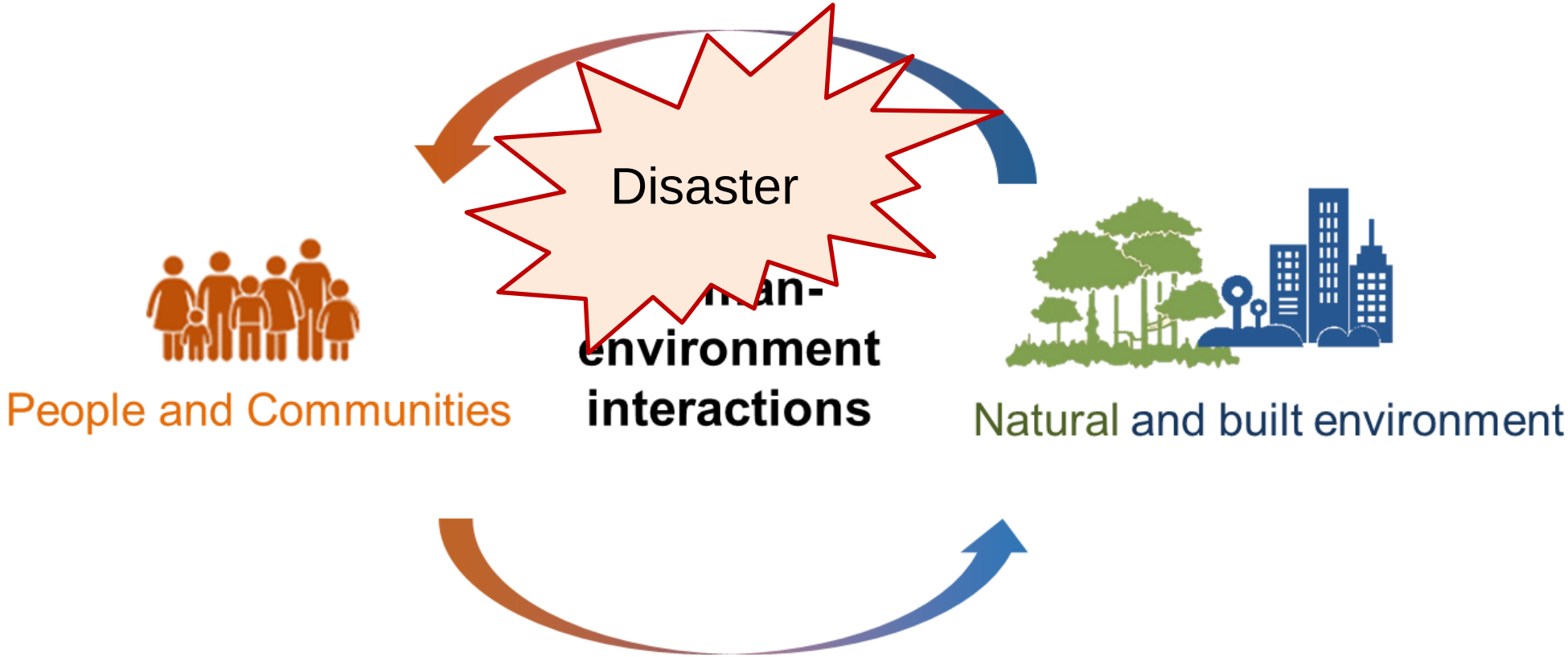
GeoAI
University at Buffalo The State University of New York

Generative AI, GIScience, and Disaster Management

Yingjie Hu

GeoAI Lab, Department of Geography, University at Buffalo, SUNY

Communities under a natural disaster are complex human-environment systems



Disaster management

- An important research topic in GIScience
  - A map uniquely links various types of disaster-related information via location
- Four phases of disaster management
  - Effort in all four phases is important for increasing disaster resilience
  - The unpredicted nature of disasters often make the Response phase the last and most critical line for saving lives and reducing losses

Some information needs for disaster management

- Where are the victims?
  - Who need help and where?
- Where are the damages?
  - Which areas are damaged by the disaster and which areas are OK?
- Where are the potential future affected areas?
  - Which areas will likely be affected in near term?
- All information should ideally be in near real time, have high spatial and temporal resolutions, and meet accuracy requirements

Generative AI

- Generative AI: AI models that can generate new content based on some input

DALL·E 3

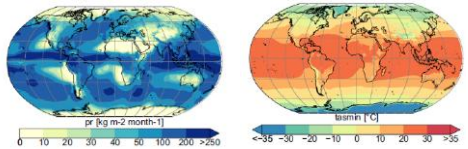
- Today's powerful generative AI models are often foundation models (FM)
  - Trained on a large amount of unlabeled data
  - Can be adapted to various downstream tasks
  - E.g., ChatGPT, DALL-E, NASA+IBM Prithvi, …

NASA IBM

Prithvi

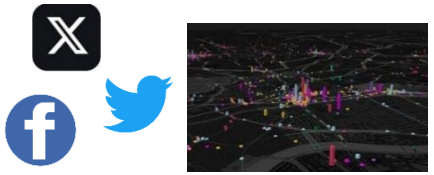
...

Satellite & aerial imagery data



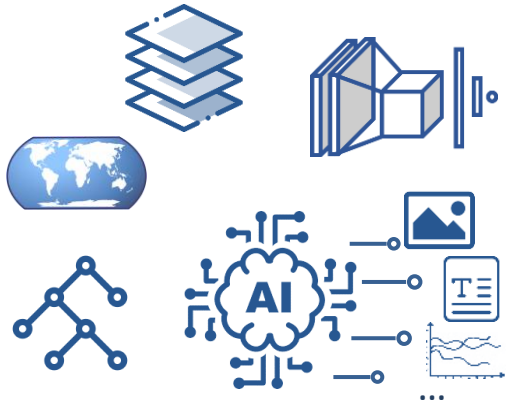
Weather and climate data

Anonymized mobile phone location data



Social media data

...



GIS and GeoAI methods including generative AI

GIS experts

Cyber-infrastructures

GIS methods, experts, and cyberinfrastructures

Finding victims

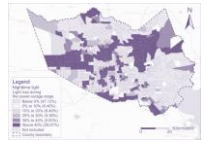
Assessing damages

Predicting future affected areas

Obtaining higher quality data

...

Contributions to disaster management



Disaster impacts

Socioeconomic demographic factors

Resource allocation

Environmental justice

...

Contributions to science and decisions

Case study:

Geo-knowledge-guided GPT models for extracting location descriptions from disaster-related social media messages

Hu, Y., Mai, G., Cundy, C., Choi, K., Lao, N., Liu, W., Lakhanpal, G., Zhou, R.Z., & Joseph, K. (2023): Geo-knowledge-guided GPT models improve the extraction of location descriptions from disaster-related messages. **International Journal of Geographical Information Science**, 37(11), 2289-2318.

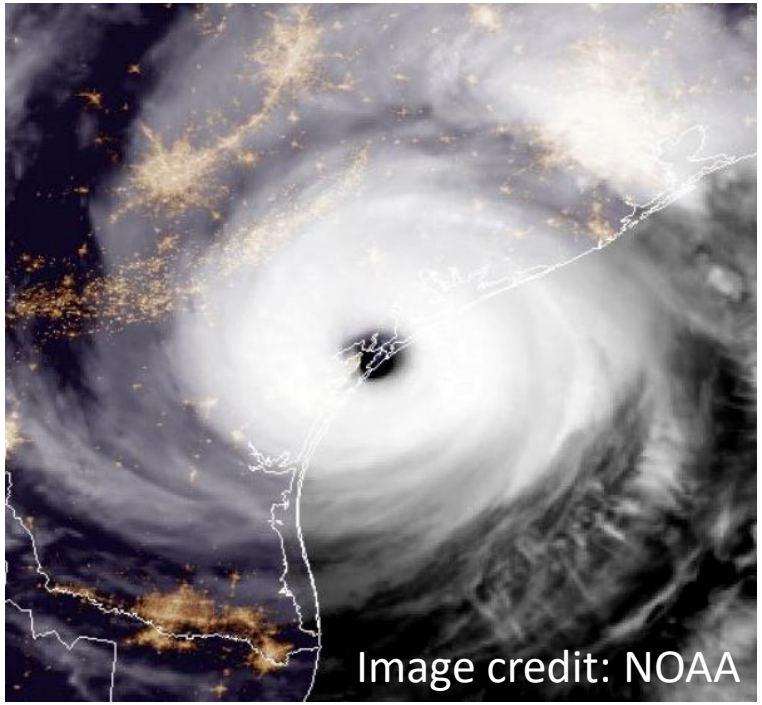
Social media and natural disasters

- Social media platforms, such as Twitter (or X), have been increasingly used by disaster victims to request help and share urgent information

- Two tweets from 2017 Hurricane Harvey

> *#HurricaneHarvey family needs rescuing at 1280 Grant St. in Cypress, Texas 77249, 2 elderly, one is 90 not steady in her feet*

> *HELP! A pregnant lady is stuck in her car on I-45 between Cypress Hill & Huffmeister exits #harvey*



Hurricane Harvey in 2017

Extracting location descriptions from disaster-related messages

- Existing research often used off-the-shelf named entity recognition (NER) tools for location extraction, such as Stanford NER and SpaCy NER

- An NER tool takes a piece of text as the input, and outputs named entities, such as persons, organizations, and locations

Organization Location

Donations to the Red Cross have provided help for people in Texas impacted by Hurricane Harvey.

Event

Limitations of using typical NER approaches

Input

Tweet 1:

#HurricaneHarvey family needs rescuing at 1280 Grant St. in Cypress, Texas 77249, 2 elderly, one is 90 not steady in her feet

Tweet 2:

HELP! A pregnant lady is stuck in her car on I-45 between Cypress Hill & Huffmeister exits #harvey

Output

Tweet 1:

Grant St., *Cypress*, *Texas* ✗

Tweet 2:

I-45, *Cypress Hill*, *Huffmeister exits* ✗

Separate entities recognized by typical NER approaches

Input

Tweet 1:

#HurricaneHarvey family needs rescuing at 1280 Grant St. in Cypress, Texas 77249, 2 elderly, one is 90 not steady in her feet

Tweet 2:

HELP! A pregnant lady is stuck in her car on I-45 between Cypress Hill & Huffmeister exits #harvey

Output

Tweet 1:

1280 Grant St. in Cypress, Texas 77249 ✓

Tweet 2:

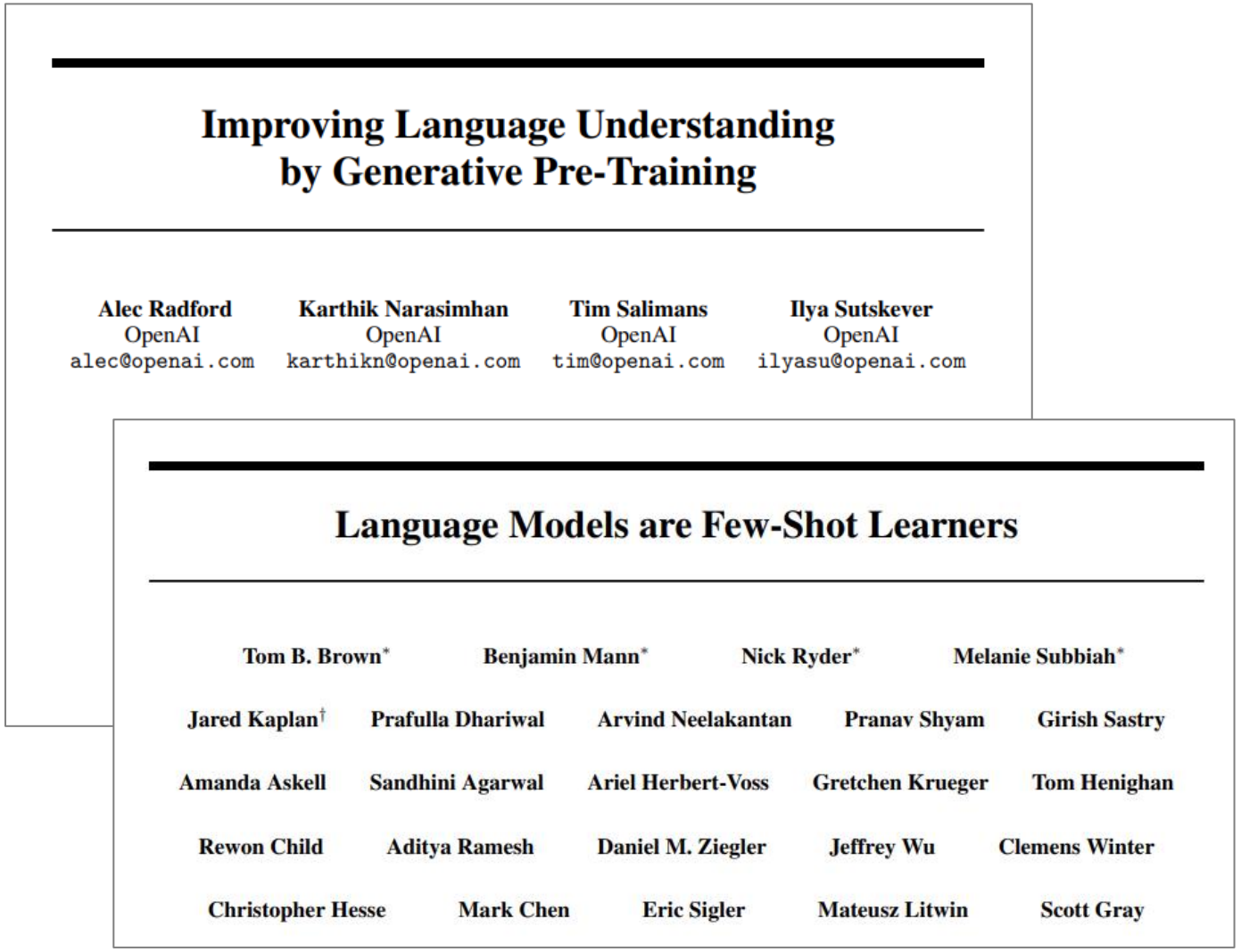
I-45 between Cypress Hill & Huffmeister exits ✓

Complete location descriptions needed for geo-locating the victims

Limitations:

- Results are separate entities not complete descriptions
- Not able to identify categories of location descriptions

Harnessing the power of large language models (LLMs)



Improving Language Understanding by Generative Pre-Training

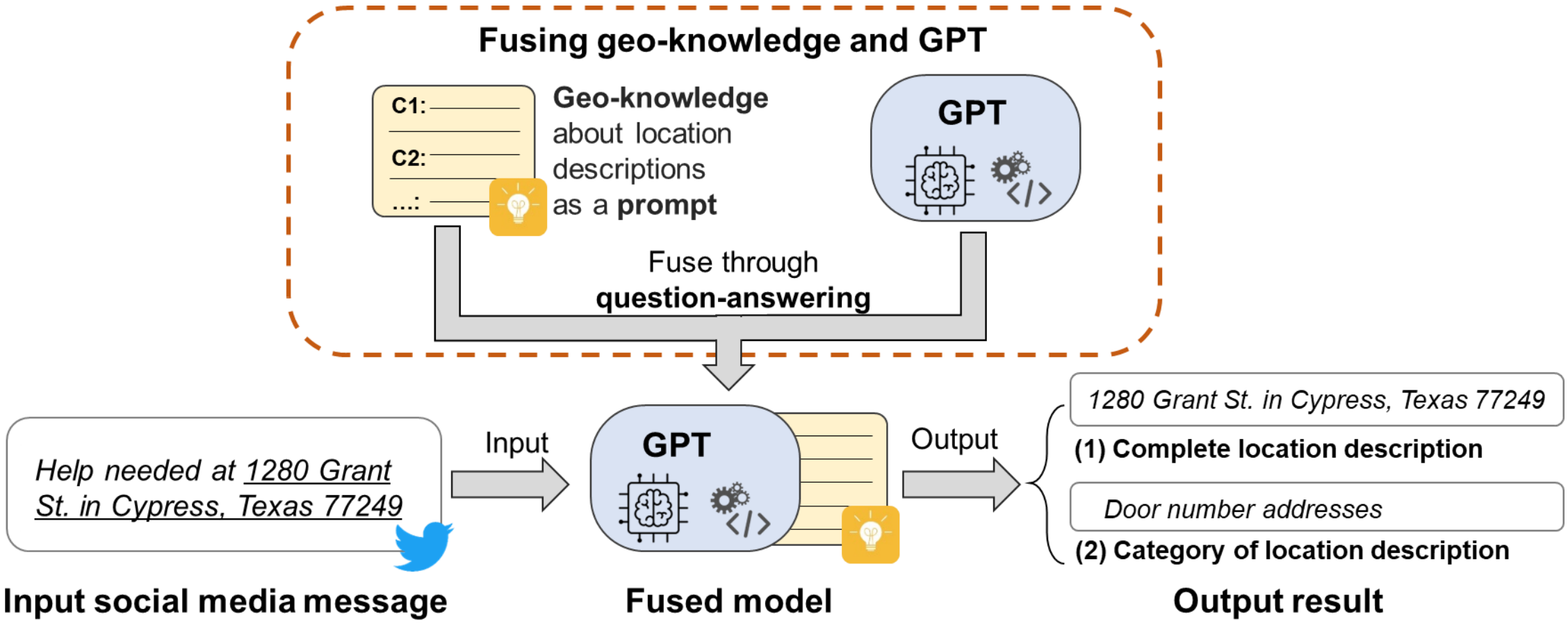
Alec Radford, OpenAI, alec@openai.com
Karthik Narasimhan, OpenAI, karthikn@openai.com
Tim Salimans, OpenAI, tim@openai.com
Ilya Sutskever, OpenAI, ilyasu@openai.com

Language Models are Few-Shot Learners

Tom B. Brown*, Benjamin Mann*, Nick Ryder*, Melanie Subbiah*
Jared Kaplan†, Prafulla Dhariwal, Arvind Neelakantan, Pranav Shyam, Girish Sastry
Amanda Askell, Sandhini Agarwal, Ariel Herbert-Voss, Gretchen Krueger, Tom Henighan
Rewon Child, Aditya Ramesh, Daniel M. Ziegler, Jeffrey Wu, Clemens Winter
Christopher Hesse, Mark Chen, Eric Sigler, Mateusz Litwin, Scott Gray

- Pre-trained on large amount of data
- Can be adapted to a particular task via few-shot learning
- However, the default output of the model can be incorrect or inaccurate

Geo-knowledge-guided GPT models for location extraction



Geo-knowledge about how people describe locations during natural disasters

Category	*Location description examples*
C1: Door number addresses	- *"#HurricaneHarvey family needs rescuing at 1280 Grant St. in Cypress, Texas 77249, 2 elderly, one is 90 not steady in her feet"* - *"Papa stranded in home. Water rising above waist. HELP 812 Wood Ln, 77828 #houstonflood"*
C2: Street names	- *"#Harvey LIVE from San Antonio, TX. Fatal car accident at Ingram Rd., Strong winds."* - *"Allen Parkway, Memorial, Waugh overpass, Spotts park and Buffalo Bayou park completely under water"*
C3: Highways	- *"9:00AM update video from Hogan St over White Oak Bayou, I-10, I-45: water down about 4' since last night. . . ""* - *"Left Corpus bout to be in San Angelo #HurricaneHarvey Y'all be safe Avoided highway 37 Took the back road"*
C4: Exits of highways	- *"Need trailers/trucks to move dogs from Park Location: Whites Park Pavillion off I-10 exit 61 Anahuac TX"* - *"Townsend exit, Sorters road and Hamblen road is flooded coming from 59 southbound #HurricaneHarvery #Harvey2017"*
C5: Intersections of roads (rivers)	- *"Guys, this is I-45 @ Main Street in Houston. Crazy. #hurricane #harvey. . ."* - *"Major flooding at Clay Rd & Queenston in west Houston. Lots of rescues going on for ppl trapped…"*
C6: Natural features	- *"Frontage Rd at the river #hurricaneHarvey #hurricaneharvey @ San Jacinto River"* - *"Buffalo Bayou holding steady at 10,000 cfs at the gage near Terry Hershey Park"*
C7: Other human-made features	- *"If you need a place to escape #HurricaneHarvey, The Willie De Leon Civic Center: 300 E. Main St in Uvalde is open as a shelter"* - *"Houston's Buffalo Bayou Park - always among the first to flood. #Harvey"*
C8: Local organizations	- *"Cleaning supply drive is underway. 9-11 am today at Preston Hollow Presbyterian Church"* - *"#Harvey does anyone know about the flooding conditions around Cypress Ridge High School?! #HurricaneHarvey"*
C9: Administrative units	- *"#HurricaneHarvey INTENSE eye wall of category 4 Hurricane Harvey from Rockport, TX"* - *"Pictures of downed trees and damaged apartment building on Airline Road in Corpus Christi"*
C10: Multiple areas	- *"Anyone doing high water rescues in the Pasadena/Deer Park area? My daughter has been stranded in a parking lot all night"* - *"FYI to any of you in NW Houston/Lakewood Forest, Projections are showing Cypress Creek overflowing at Grant Rd"*
C11: Road segments	- *"HELP! A pregnant lady is stuck in her car on I-45 between Cypress Hill & Huffmeister exits #harvey"* - *"Streets Flooded: Almeda Genoa Rd. from Windmill Lakes Blvd. to Rowlett Rd. #HurricaneHarvey #Houston"*

Using geo-knowledge to guide GPT models

This is a set of location description recognition problems.
The `Sentence` is a sentence containing location descriptions.
The goal is to infer which parts of the sentence represent location descriptions and the categories of the location descriptions. Split different location descriptions with `;`.
--

--
Sentence: Papa stranded in home. Water rising above waist. HELP 812 Wood Ln, 77828 #houstonflood
Q: Which parts of this sentence represent location descriptions?
A: C1: 812 Wood Ln, 77828
--

--
Sentence: Anyone doing high water rescues in the Pasadena/Deer Park area? My daughter has been stranded in a parking lot all night
Q: Which parts of this sentence represent location descriptions?
A: C10: Pasadena/Deer Park

Evaluation Experiments

- Test dataset:
  - 1,000 tweets from Hurricane Harvey; manually labeled with location descriptions and location categories; 22 tweets are included in the prompt, and 978 are used for performance evaluation

- Evaluation metrics:

$$Precision = \frac{tp}{tp + fp}$$

$$Recall = \frac{tp}{tp + fn}$$

$$F\text{-}score = 2 \cdot \frac{Precision \times Recall}{Precision + Recall}$$

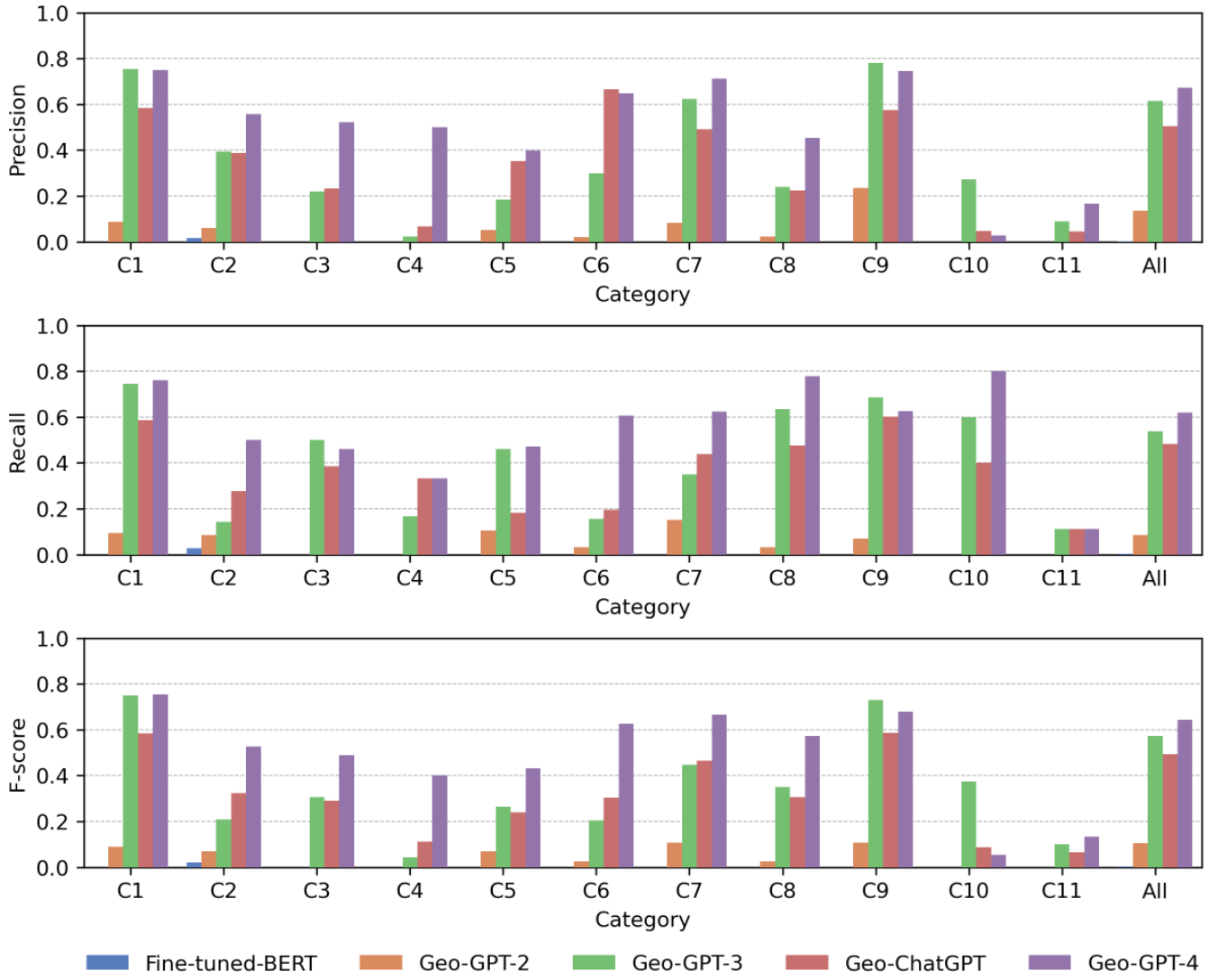
- Test models
  - Default GPT 2, GPT 3, ChatGPT, GPT 4
  - Geo-knowledge-guided GPT models
  - Fine-tuned BERT
  - Stanford NER (narrow and broad versions)
  - SpaCy NER (narrow and broad versions)

Results– Overall performance without differentiating categories

Models	Precision	Recall	F-score
Models that cannot identify location categories			
Stanford NER (narrow)	0.621	0.402	0.488
Stanford NER (broad)	0.564	0.440	0.495
SpaCy NER (narrow)	0.643	0.224	0.332
SpaCy NER (broad)	0.352	0.340	0.346
GPT-2	0.012	0.004	0.005
GPT-3	0.255	0.256	0.255
ChatGPT	0.416	0.370	0.392
GPT-4	0.404	0.385	0.394
Models that can identify location categories			
Geo-GPT-2	0.380	0.404	0.391
Geo-GPT-3	**0.693**	0.694	0.693
Geo-ChatGPT	0.633	0.673	0.653
Geo-GPT-4	0.687	**0.704**	**0.695**
Fine-tuned-BERT	0.150	0.242	0.185

- Geo-knowledge-guided GPT models can recognize location descriptions more accurately than typical NER approaches (about 40% better)
- Geo-knowledge-guided GPT models can recognize location descriptions more accurately than default GPT models (about 76% better)

Results– Performance across different categories



- Geo-GPT 3, 3.5 and 4 overall perform better than the two other models across all categories
- While Geo-GPT-4 performs reasonably well in most categories, its performance is still limited in some categories (e.g., C4, C10, and C11)

Conclusions (of this case study)

- Messages posted on social media platforms during natural disasters often contain important location descriptions about victims and accidents
- Typical NER approaches cannot effectively extract complete location descriptions and their categories
- Geo-knowledge-guided GPT models use geo-knowledge about disaster-related location descriptions to guide GPT models for this task; achieves a performance 40% better than default NER tools and 76% better than default GPT models
- The extracted location descriptions can help disaster responders reach victims more quickly and potentially save lives

Other possible applications of generative AI in disaster management

- Mapping damages: estimating areas burnt by a wildfire, identifying buildings damaged by an earthquake, assessing flood depth, …
- Predicting future affected areas: where will the current wildfire spread? What could be the potential trajectories of a hurricane, …
- Increasing information quality: can we increase the spatial resolution of wind speed data from 9 km to 1 km? …
- Community engagement: how would this neighborhood look if it is affected by 1 meter of floodwater? …
- …

Summary

- Communities under natural disasters are complex human-environment systems
- GIScience plays an important role in disaster management. Generative AI and GeoAI provide new tools for studying disasters and increasing community resilience
- Some existing and possible applications: Identifying victims, mapping damages, predicting future affected areas, increasing data quality, community engagement, …
- Important considerations when using generative AI in disaster management:
  - Quality: Does generative AI extract the right information that we need?
  - Sensitivity: Does the output of generative AI vary substantially if given slightly different input?
  - …

University at Buffalo
The State University of New York

GeoAI
University at Buffalo The State University of New York

Thank you!

Yingjie Hu

Email: yhu42@buffalo.edu
GeoAI@UB: https://geoai.geog.buffalo.edu

Acknowledgement:
This work is supported by the National Science Foundation under Grant No. BCS-2117771