# Coordinates from Context:
## Using LLMs to Ground Complex Location References

**Tessa Masis**
*they/them*

**Brendan O'Connor**
*he/him*

**University of Massachusetts Amherst, USA**

1

# Using geospatial info in text

Location reference identification from tweets during emergencies: A deep learning approach

Abhinav Kumar*, Jyoti Prakash Singh

Department of Computer Science & Engineering, National Institute of Technology Patna, India

Real-Time Disease Surveillance Using Twitter Data: Demonstration on Flu and Cancer

Kathy Lee          Ankit Agrawal          Alok Choudhary
                   EECS Department
                   Northwestern University
                   Evanston, IL USA
                   choudhar}@eecs.northwestern.edu

Tracking 19th Century Late Blight from Archival Documents using Text Analytics and Geoparsing

Laura Tateosian[a,*], Rachael Guenter[b], Yi-Peng Yang[a], Jean Ristaino[1]

[a] Center for Geospatial Analytics, North Ca...

[b] Department of Entomology and Plant Pathology, N...

From Text to Maps: LLM-Driven Extraction and Geotagging of Epidemiological Data

Karlyn K Harrod*              Prabin Bhandari*†           Antonios Anastasopoulos
Oak Ridge National Laboratory   George Mason University     George Mason University
Oak Ridge, Tennessee, USA       Fairfax, Virginia, USA      Fairfax, Virginia, USA
harrodkk@ornl.gov               pbhanda2@gmu.edu            antonis@gmu.edu

2

# Geocoding: extracting geospatial info

St Queen Ketevan's relics were found in St. Augustine Convent in Old Goa, India in 2005. DNA analysis by the CSIR-Centre for Cellular and Molecular Biology, Hyderabad at the instance of the Archaeological Survey of India confirmed its authenticity. It is believed her relics were brought to Goa in 1627.

https://www.wionews.com/india-news/pm-modi-pays-tributes-to-georgias-saint-queen-ketevan-lauds-goa-for-preserving-her-relics-400676

# Geocoding: extracting geospatial info

St Queen Ketevan's relics were found in St. Augustine Convent in Old Goa, India in 2005. DNA analysis by the CSIR-Centre for Cellular and Molecular Biology, Hyderabad at the instance of the Archaeological Survey of India confirmed its authenticity. It is believed her relics were brought to Goa in 1627.

https://www.wionews.com/india-news/pm-modi-pays-tributes-to-georgias-saint-queen-ketevan-lauds-goa-for-preserving-her-relics-400676

GeoNames ID: 1271146

# Geocoding: compositional locations?

the Order dated 28th December 1962, the result of my Inquiry into the collision between two passenger trains that occurred at 6.1 p.m. on 26th December 1962, on the four tracked electrified line between Winsford Station and Coppenhall Junction signal boxes, just north of Crewe, on the former London and North Western main line of the London Midland Region, British Railways.

https://www.railwaysarchive.co.uk/documents/MoT_Coppenhall1962.pdf

The southern part of Madagascar, the island nation off the southeast coast of Africa, is facing its worst drought in 40 years. The lack of rainfall and other environmental factors have taken a toll on the region's vegetation and agricultural production, leaving many in the region facing severe hunger.

https://science.nasa.gov/earth/earth-observatory/drought-in-madagascar-148636/

# Geocoding: compositional locations?

the Order dated 28th December 1962, the result of my Inquiry into the collision between two passenger trains that occurred at 6.1 p.m. on 26th December 1962, on the four tracked electrified line between Winsford Station and Coppenhall Junction signal boxes, just north of Crewe, on the former London and North Western main line of the London Midland Region, British Railways.

https://www.railwaysarchive.co.uk/documents/MoT_Coppenhall1962.pdf

?

The southern part of Madagascar, the island nation off the southeast coast of Africa, is facing its worst drought in 40 years. The lack of rainfall and other environmental factors have taken a toll on the region's vegetation and agricultural production, leaving many in the region facing severe hunger.

https://science.nasa.gov/earth/earth-observatory/drought-in-madagascar-148636/
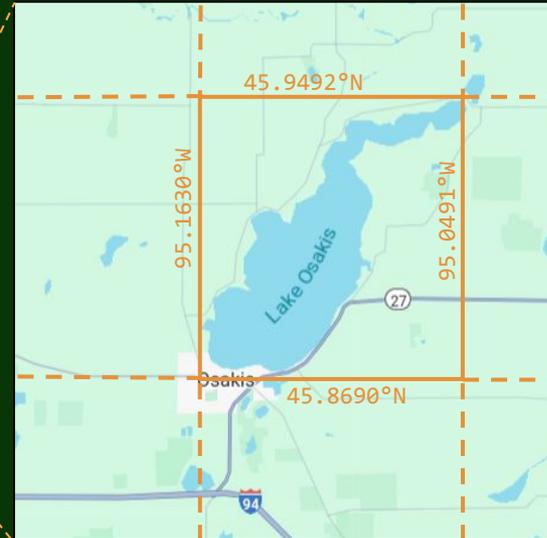
?

# Geocoding: compositional locations

The location is a lake in Todd and Douglas counties in west-central Minnesota. The town of Osakis is situated on the southwest shore of the lake.

# Geocoding: compositional locations



The location is a lake in Todd and Douglas counties in west-central Minnesota. The town of Osakis is situated on the southwest shore of the lake.

(-95.163, 45.869, -95.049, 45.949)

# Do LLMs know geographic bounding boxes?

## Knowledge

Hotel Astoria, in Denmark

## Knowledge

Hotel Astoria, in Denmark

## Reasoning

The location is a design hotel next to the Central Station in Copenhagen, Denmark. Central Station has a longitude of ... Copenhagen has a longitude of...

# Do LLMs know geographic bounding boxes?

## Knowledge

Hotel Astoria, in Denmark

(55.6734, 12.5631)

## Knowledge

Hotel Astoria, in Denmark

(12.5628, 55.6731, 12.5634, 55.6738)

## Reasoning

The location is a design hotel next to the Central Station in Copenhagen, Denmark. Central Station has a longitude of … Copenhagen has a longitude of…

# Do LLMs know geographic bounding boxes?

## Knowledge

Hotel Astoria, in Denmark

(55.6734, 12.5631)

## Knowledge

Hotel Astoria, in Denmark

(12.5628, 55.6731, 12.5634, 55.6738)

## Reasoning

The location is a design hotel next to the Central Station in Copenhagen, Denmark. Central Station has a longitude of … Copenhagen has a longitude of…

(12.5628, 55.6731, 12.5634, 55.6738)
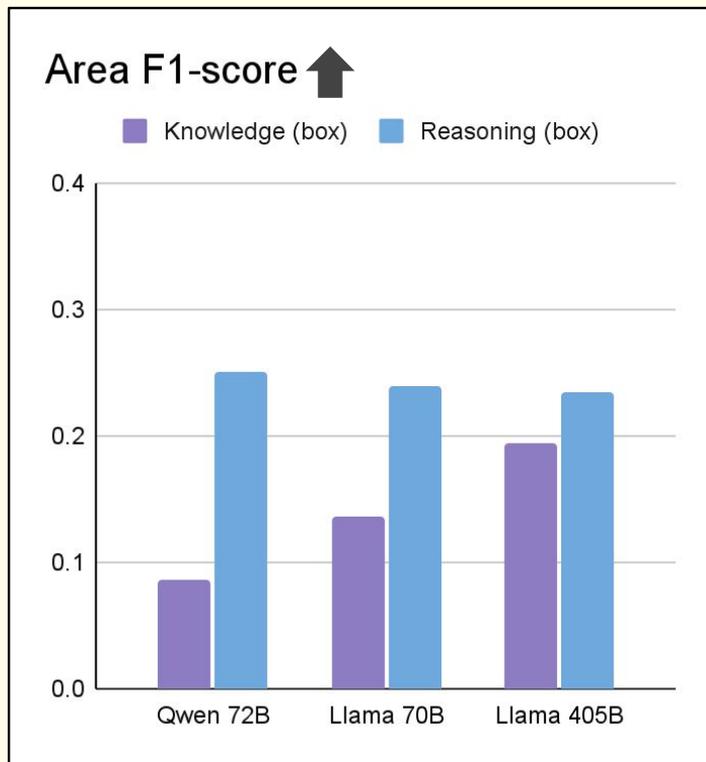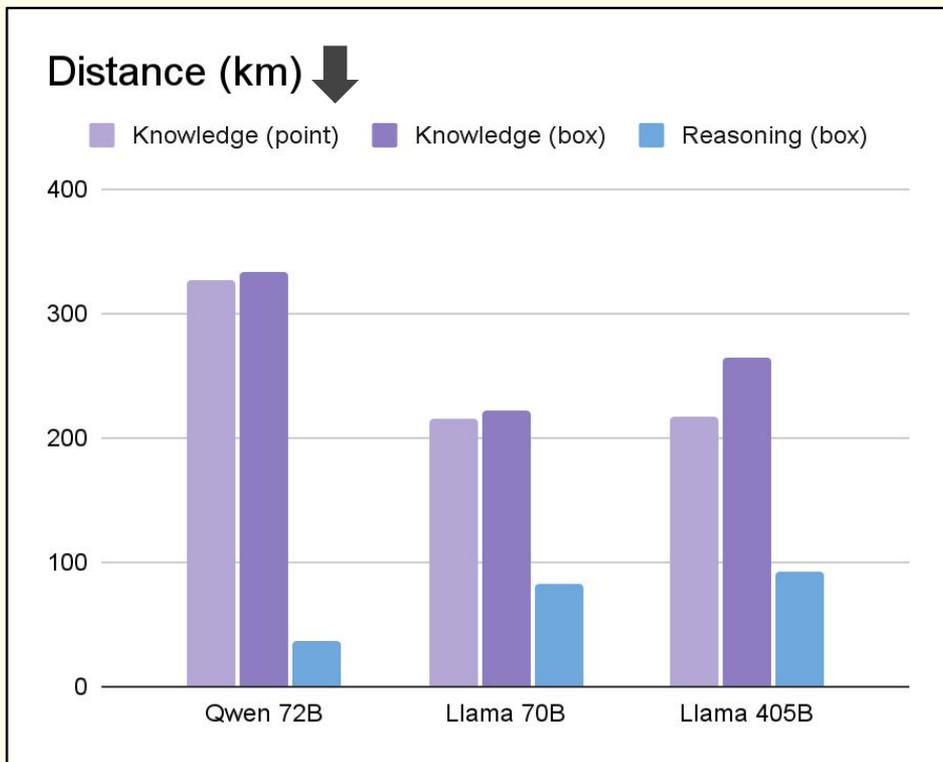
# Do LLMs know geographic bounding boxes?

The location is a lake in Todd and Douglas counties in west-central Minnesota. The town of Osakis is situated on the southwest shore of the lake.

(-95.163, 45.869, -95.049, 45.949)

**GeoCoDe dataset (Laparra & Bethard, 2020)**
1,000 compositional + 350k simple location descriptions from Wikipedia articles; we link each location to its bounding box, using the OpenStreetMap geodatabase

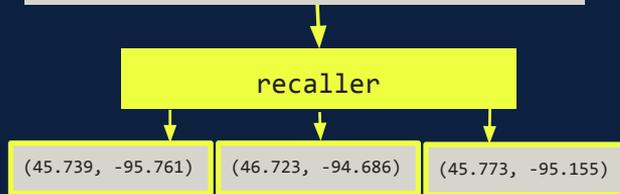# Do LLMs know geographic bounding boxes?

# Tool-augmented geocoding

The location is a lake in Todd and Douglas counties in west-central Minnesota. The town of Osakis is situated on the southwest shore of the lake.

# Tool-augmented geocoding

The location is a lake in Todd and Douglas counties in west-central Minnesota. The town of Osakis is situated on the southwest shore of the lake.

↓

recaller

↓         ↓         ↓

(45.739, -95.761)   (46.723, -94.686)   (45.773, -95.155)

The **recaller**, a traditional geoparsing tool (e.g. Google Maps Geoparser), retrieves mentioned locations' center coordinates

# Tool-augmented geocoding

The location is a lake in Todd and Douglas counties in west-central Minnesota. The town of Osakis is situated on the southwest shore of the lake.
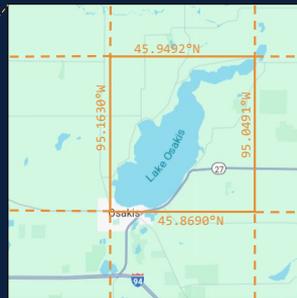
recaller

(45.739, -95.761)  (46.723, -94.686)  (45.773, -95.155)
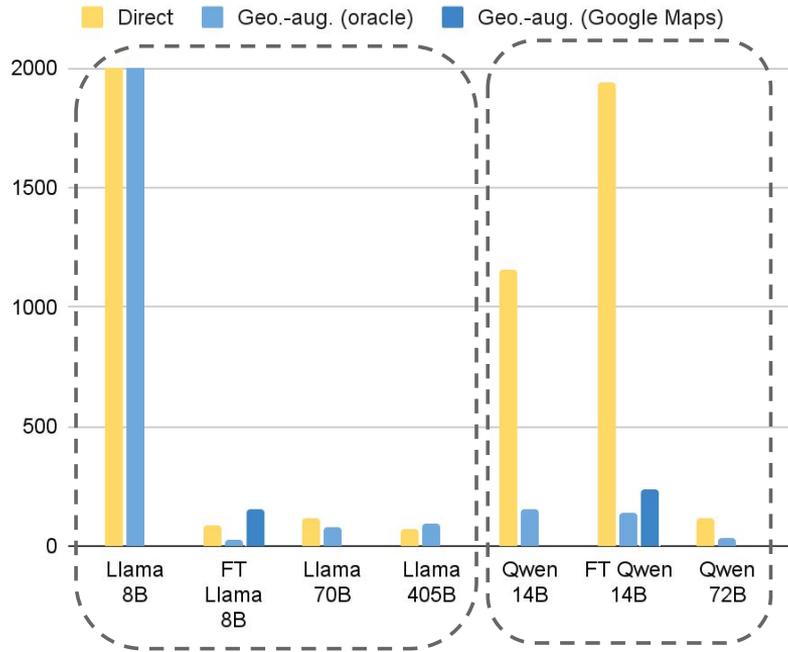
reasoner

(-95.163, 45.869, -95.049, 45.949)

The recaller, a traditional geoparsing tool (e.g. Google Maps Geoparser), retrieves mentioned locations' center coordinates
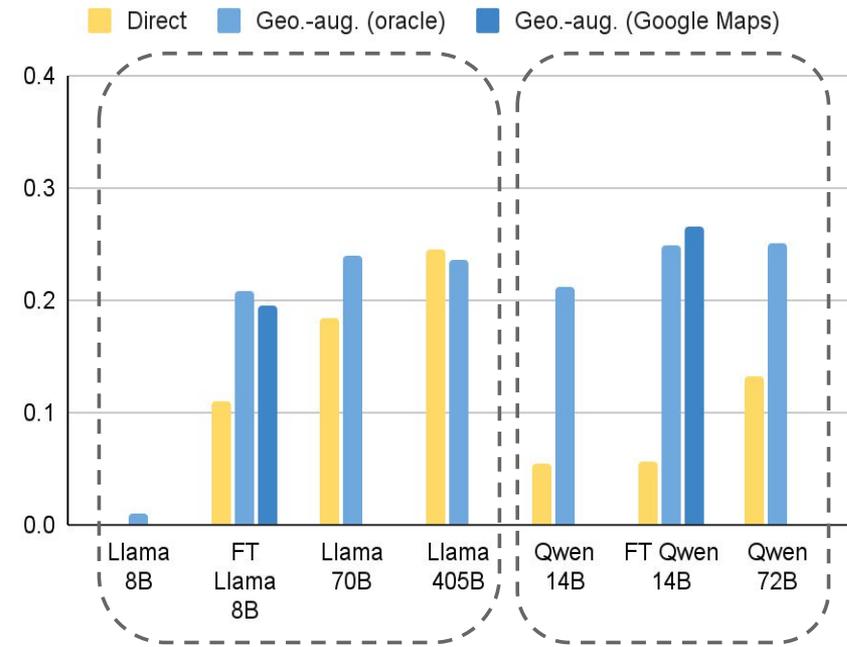
The reasoner, an LLM, uses the location description + mentioned locations' coordinates to generate the described location's bounding box

# Tool-augmented geocoding

# Tool-augmented geocoding

| Approach | reasoner | Coverage (%) ↑ | Distance (km) ↓ | F1-score ↑ |
|---|---|---|---|---|
| Geo.-aug. | FT Qwen 14B | 90.8 | **240.9** | **.266** |
| End-to-end LLM | FT Qwen 14B | **95.9** | 627.5 | .167 |
| GBSP | Semantic parser | 52.8 | – | .240 |

# Summary

- We found that off-the-shelf LLMs have stronger geospatial reasoning skills than geospatial knowledge

- Proposed the first end-to-end methods for geocoding compositional location references

- Novel use of bounding boxes to link locations, providing a simple and informative way to ground locations that don't exist in a geographic database

- Our methods achieve SOTA performance, and a relatively small fine-tuned LLM can achieve comparable performance with much larger off-the-shelf models

# Thank you!

## Slides, poster, and paper available at tmasis.github.io/

**Tessa Masis**
tmasis@cs.umass.edu

**Brendan O'Connor**
brenocon@cs.umass.edu