GeoLocator: a location-integrated large multimodal model for inferring geo-privacy

University of Southern California

Yifan Yang

yyang295@usc.edu

Abstract

Geographic privacy or **geo-privacy** refers to the keeping private of one's geographic location, especially the restriction of geographical data maintained by personal electronic devices. Geo-privacy is a crucial aspect of personal security;

With the surge in the use of **Large Multimodal Models (LMMs)**, such as GPT-4, for Open Source Intelligence (OSINT), the potential risks associated with geo-privacy breaches have intensified.

This study develops a location-integrated GPT-4 based model named **GeoLocator** and designs **four-dimensional experiments** to demonstrate its capability in inferring the locational information of input imageries and/or social media contents.

We conclude with the **broader implications** of GeoLocator and our findings for individuals and the community at large, by emphasizing the urgency for enhanced awareness and protective measures against **geo-privacy leakage** in the era of advanced AI and widespread social media usage.


LOCITAL
PRIVACY SENE
ATATA
PRIWVATIAN

Introduction

The Challenges of Geographic Privacy

Definition of Geographic Privacy: The protection and confidentiality of personal geographic information, including real-time whereabouts, historical movement patterns, and specific location data.

Challenges: In the era of ubiquitous smartphones and social media, maintaining geographic privacy becomes more challenging. Navigation, travel ticketing sites, and social media offer convenience but also pose risks to our geographic privacy through potential surveillance, unauthorized data mining, and third-party misuse.

Technology Development and Privacy Leakage: The rapid development of large multimodal models like GPT-4 increases the risk of exposing geographic privacy. These models can extract, interpret, and infer geographic information from published images.

Example Illustration: Publishing a single photo on social media, such as visiting a ballpark, could reveal personal geographic locations

Introduction

GeoLocator Tool and the Future of Geographic Privacy

GeoLocator Tool: A tool that integrates GPT-4 with geolocation functions, capable of inferring locational information from input images and/or social media content.

Experimental Evaluation: A series of experiments based on various datasets, including Google Maps images, daytime/nighttime images, and social media posts, compare the capabilities of regular search engines, GPT-4, and GeoLocator in privacy breaches.

Findings and Risks: GeoLocator can generate specific geographic details with high accuracy, highlighting the risk of users unintentionally exposing geospatial information to the public.

Strategies and Awareness: Emphasizes the urgency of enhancing awareness and protective measures against geographic privacy leakage in the era of advanced AI and widespread social media usage.

Related Work

The emergence of the Transformer architecture has revolutionized large-scale language models.
Key developments: GPT, BERT, and advances in computing power and training techniques. LMM excels at abstract generalization, understanding context, and user intent across different domains.

Integrate with vision-based models:
Transitioning to large multimodal models that integrate text and visual data processing. Breakthroughs in visual understanding, reasoning, and applications in web code generation and image analysis.

GPT-4’s geospatial reasoning capabilities are reflected in location estimation and route planning. GeoLM: Integrating geospatial data with linguistic information to better understand geographic entities. GeoLocator: A novel model that extends LMM to infer geospatial information and enhance geoprivacy.

Workflow to develop the new tool – *GeoLocator*

GeoLocator

You

GeoLocator

Analyzing the image provided, I notice the following key details:

- The prominent sign "The Simpsons Ride" which is a known attraction at Universal Studios theme parks.
- The statue of the character Krusty the Clown from "The Simpsons" television show, which is a part of the themed ride area.
- The overall theme and design of the area, ich matches the cartoonish style of "The Simpsons."

https://chat.openai.com/g/g-qxqvMb6YJ-geolocator

Workflow to develop the new tool – *GeoLocator*

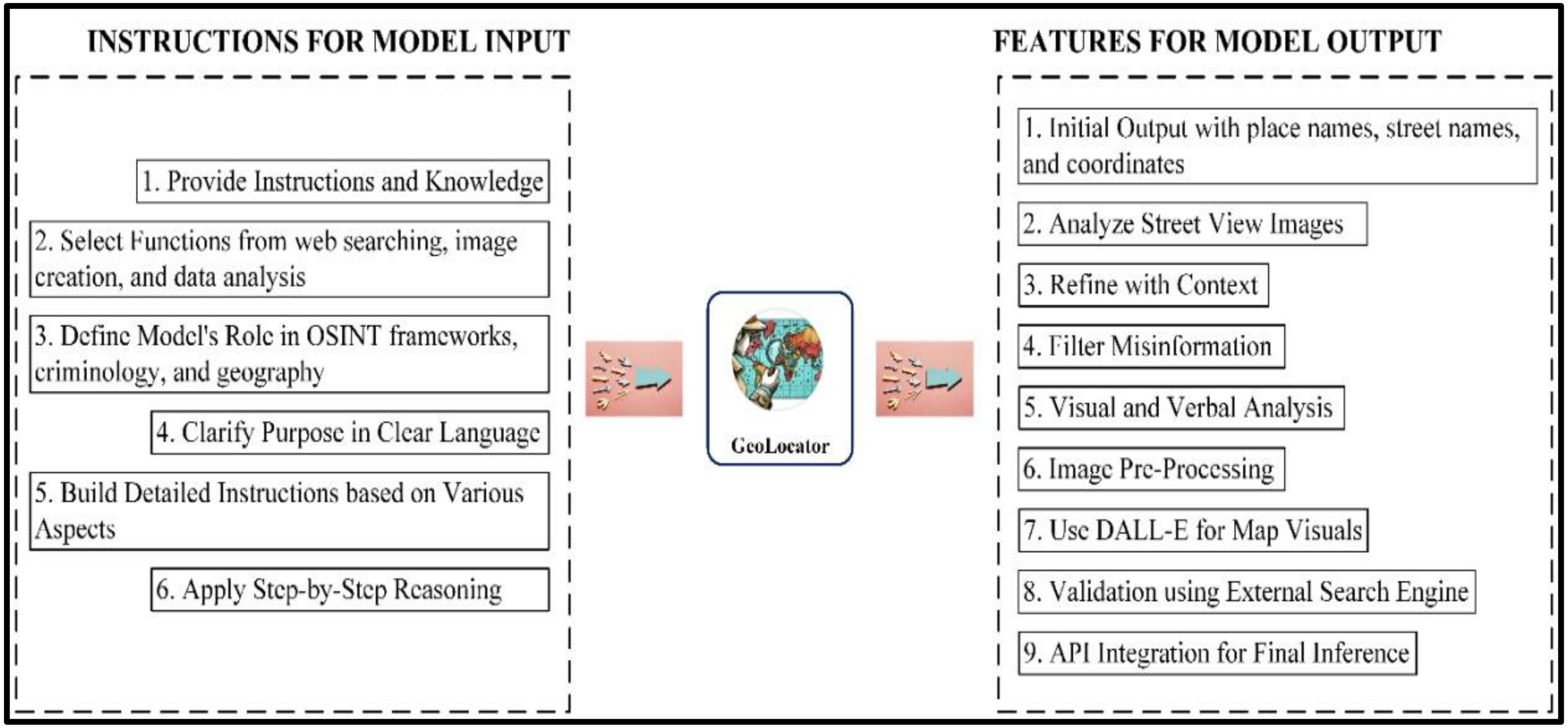


Figure 1. *GeoLocator* Instructions and Features

Workflow to develop the new tool – *GeoLocator*

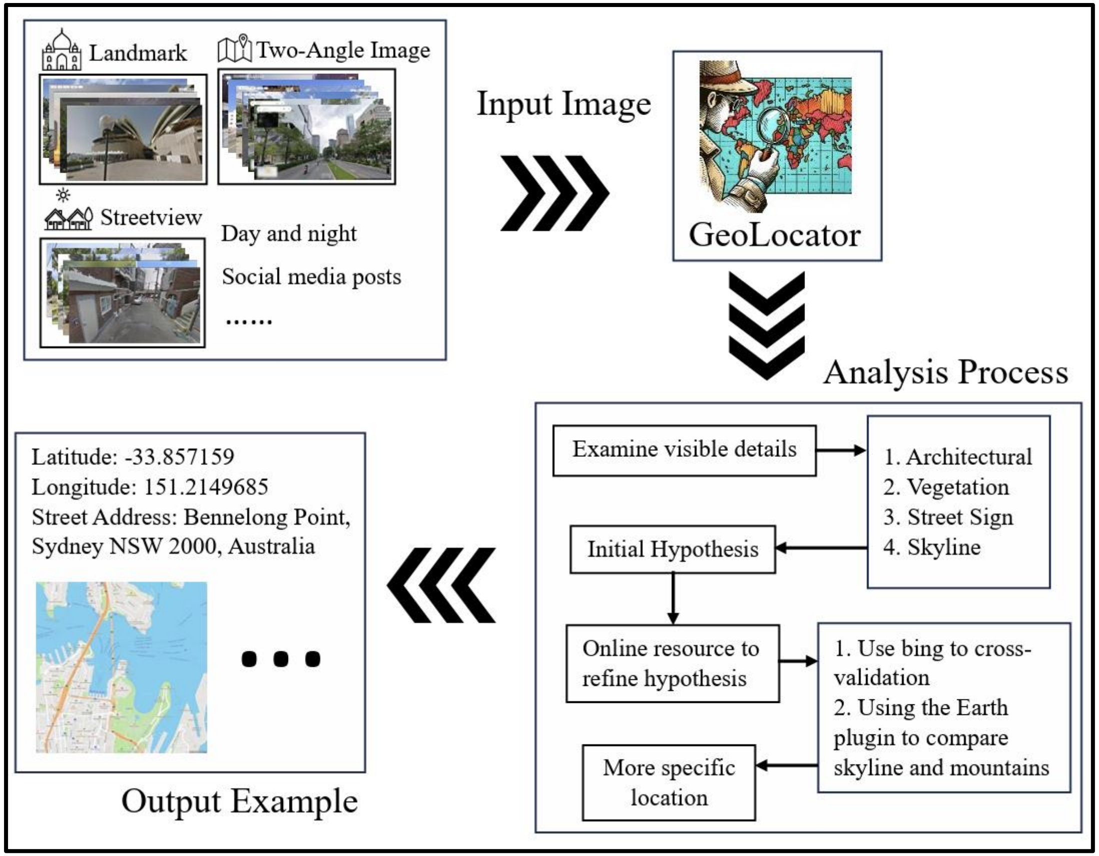


Figure 2. Geolocator working flowchart

Experimental design to test out *GeoLocator*

To evaluate the effectiveness of the Google search engine, GPT-4, and *Geolocator* across various image types, we gathered a diverse set of data sources. This included images from Google Maps, photographs taken by our research team, Google Images, and posts from social media.

Table 1. Data source and description

Data Source	Description	Number of images
Google Maps	Iconic landmarks	50 images
	Street view without obvious landmarks	50 images
	Images of 20 locations from two different angels	20 sets (40 images)
Taken by research team	Images of 20 locations at two times slots (i.e., day and night time)	[illegible] sets (40 images)
Google Images	Images of 10 locations from China to assess the impact of language input	10 images
Posts from social media	Social media posts sent by research team members	3 posts

Experimental design to test out *GeoLocator*

Compare images' location inference ability among Google search engines, GPT-4, and GeoLocator

uploaded the same images to Google search engine, GPT-4, and GeoLocator, and to judge the performance of the tools above based on their inference precision.

Compare the location inference based on image and text instruction

compared GeoLocator's prediction results before and after applying these additional images or textual instructions.

Examine the impact of languages on inference results

provided the model with images containing text prompts in different languages to GeoLocator and observed its predictions for the geographic location of the images.

Evaluate the GeoLocator's performance on social media posts

explored whether finely tuned LMMs like GeoLocator could now infer the location of a place that social media posts show or describe.

Results

Compare images' location inference ability among Google search engines, GPT-4, and GeoLocator

Table 2. Results of image inferring accuracy

Image Type	Sample size	Google search engine	GPT-4	GeoLocator
Iconic landmark	50	88%	60%	94%
Street view	50	16%	18%	54%
Daytime image	20	25%	40%	70%
Nighttime image	20	10%	15%	35%

Images	Data Type	Google Search Engine	GPT-4	GeoLocator	Distance* (miles)
	Street View	State	City/ Town	Street	*0.0034*
	Street View	City/ Town	Unknown**	City/ Town	*10*
	Street View	State	Unknown**	Country	*32.49*
	Street View	Country	Unknown**	Country	*126.42*
	Landmark	Street	Unknown**	Street	*0.0044*

Image Perspective Enhancement

Tested with different angles of Taipei 101 images, demonstrating significant accuracy improvements with additional images.

Textual Prompt Enhancement

Analyzed using USC University Park Campus case, where textual prompts enabled more precise street-level location inference.

Impact of Languages

Explored the model's reasoning process and result variations with text prompts in Chinese and English, revealing sensitivity to cultural and geographical information inherent in language.

Social Media Post Analysis

Tested the capability to infer locations in complex real-world scenarios from social media posts, accurately pinpointing city and street levels, and building detailed personal profiles of the posters. GeoLocator can process complex information in social media posts, accurately inferring locations and providing detailed personal profiles.

Key Discoveries from GeoLocator Development

Exceptional Performance: GeoLocator surpasses Google search and GPT-4 in inferring specific locations, especially in street views. **Enhanced with Multimodal Inputs**: Accuracy improves significantly when provided with additional images or textual prompts. **Influence of Language**: Input language affects reasoning processes, showing no significant difference in accuracy but variations in focus. **Social Media Proficiency**: Demonstrates strong capabilities in inferring locations from complex and diverse data in social media posts.

Contributions and Advancements:

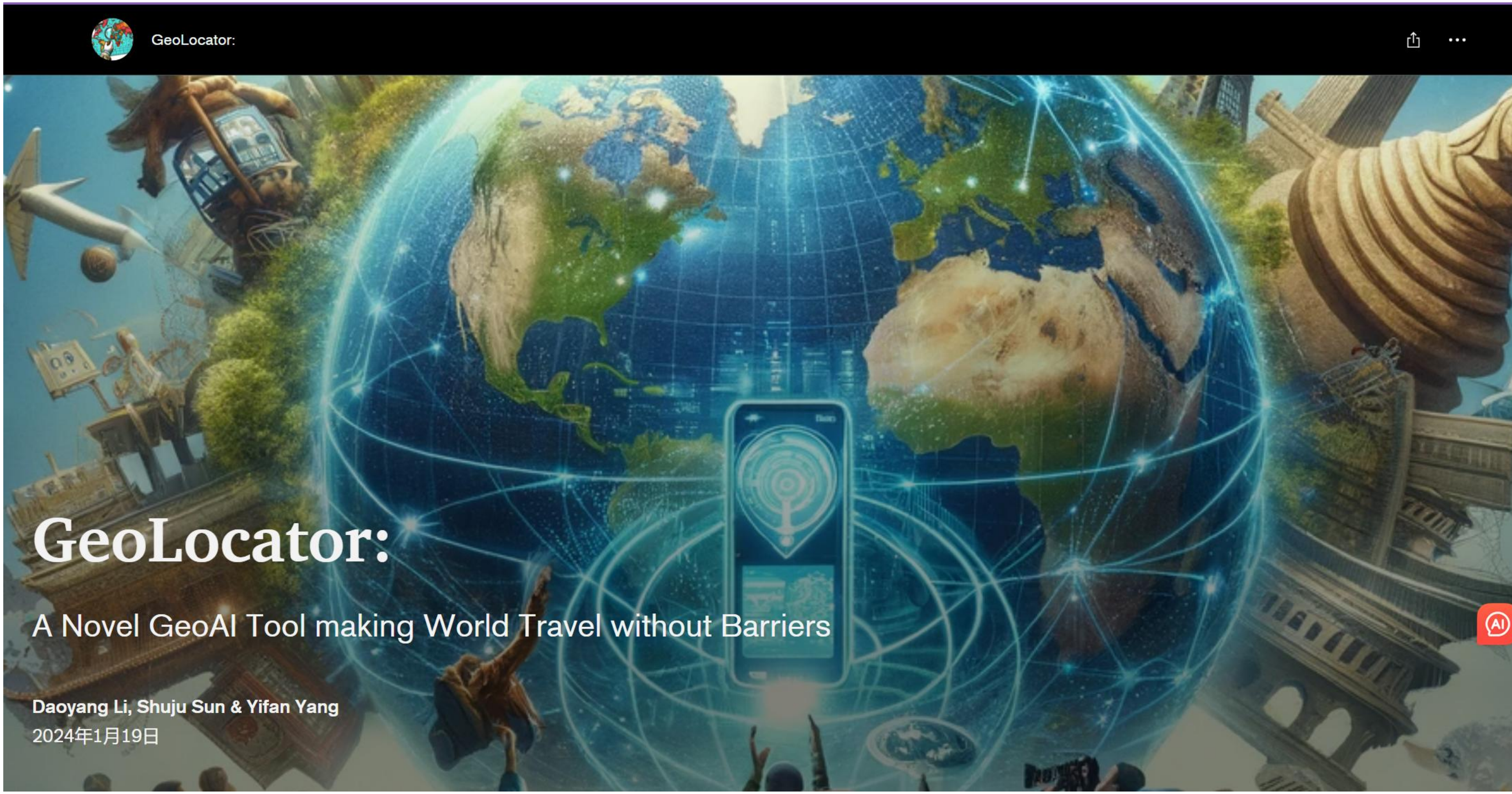
Extends LMM applications in geography, offering granular street-level location inference. Showcases the effectiveness of multimodal analysis in real-world scenarios.

Policy Implications:

Potential tool for privacy regulation enforcement and urban planning.
Useful in public safety and natural disaster management for rapid assessment.

Limitations and Concerns:

Challenges with processing lengthy instructions and stabilizing outputs.
Risks of data exposure and the need for data confidentiality in sensitive geospatial information.



https://storymaps.arcgis.com/stories/a34ffb2aacf74afdb3881b77be65064a