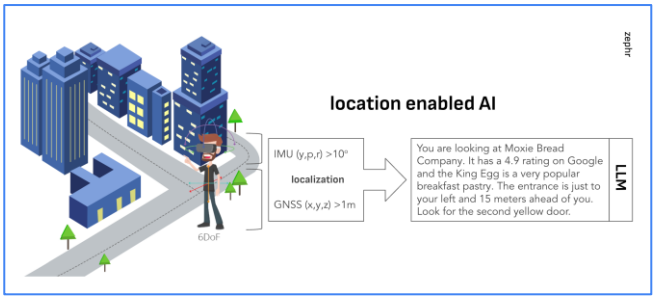
Teaching AI to Understand Place and Context

Zephr.xyz is developing a system that links Large Language Models with geospatial data to make AI assistants spatially aware. Using GNSS, sensors and open map data, the platform enables location-aware AI interactions for AR, mobility, and hands-free applications.

$$\begin{bmatrix} a^o \cap b^o \neq \emptyset & a^o \cap \partial b \neq \emptyset & a^o \cap b^e \neq \emptyset \\ \partial a \cap b^o \neq \emptyset & \partial a \cap \partial b \neq \emptyset & \partial a \cap b^e \neq \emptyset \\ a^e \cap b^o \neq \emptyset & a^e \cap \partial b \neq \emptyset & a^e \cap b^e \neq \emptyset \end{bmatrix}$$

spatial data <> vector embeddings
images <> language



Applications

- **Augmented Reality & Wearables:** Powers audio assistants that know what users are looking at without needing cameras.
- **Mobility & Accessibility Tools:** Enhances audio-based navigation and situational awareness for pedestrians and cyclists.
- **AI-Enhanced Fieldwork:** Supports field researchers and technicians with context-aware, place-linked knowledge delivery.

Key Features

- **Camera-Free Location Awareness:** Uses positioning and sensor data, not visual input, for spatial AI reasoning.
- **Real-Time Spatial Intelligence:** Enables natural, contextual interactions grounded in real-world orientation and location.
- **Open Mapping Integration:** Leverages the Overture Maps and Mapillary database for place recognition and linking.

Teaching AI Spatial Context

Detects geographic entities in a user's view and connects them to LLMs by adding geospatial reasoning for richer, place-based understanding.

Enabling Location-Aware Interactions

Combines GNSS, orientation data, and Overture Maps and Mapillary to deliver real-time, spatially intelligent assistant responses.

Team:
Zephr.xyz Inc. (Sean Gorman, Pramukta Rao, Kostas Stamitou, Joseph Strus, Ovidiu Voda)

Technical Challenges

Large language models (LLMs) can struggle with understanding geographic relationships with 'place'.

"LLMs have little understanding of geography or the relationships between geographical entities within the study area. This makes it hard for them to interpret questions that involve reference to specific spatial features or areas"

— **The Turing Institute**

As we move towards AI hardware devices that need to interact with the real world it is important to solve this disconnect between LLMs and the geography of the real world. We need a fundamental understanding of a user's geographic pose and how geospatial data in turn relates to their view. To accomplish this the AI needs to understand location and align the geographic data that encodes our knowledge of 'place' to how a user views their immediate geography.

Observations

- Generative AI provides an extremely engaging interface to the traditional map on a mobile device. It also helps fuse disparate and messy multi-modal data sources that power the map. These capabilities can anchor the creation of novel systems for screenless audio-based AI, perfect for emerging devices.
- The team had expected the technical challenge to be centered around encoding spatial information in a way that an LLM can use it. However, it was found that current generation LLMs not only struggle with real-world spatial relationships on their own but also flounder at translating mathematical/quantitative spatial descriptions into plain language.
- This challenge with spatial information led the team to using LLMs that excel at tool-calling. However, the current crop of LLMs have trouble composing basic GIS-style spatial relations reliably for human spatial awareness in an AI assistant. They require higher level tooling to be successful.
- A missing, more fundamental, gap in standard geospatial tooling is that human spatial awareness benefits from understanding a person's dynamic geographic pose in relation to mapped geometries, rather than simple static geometry relationships.
- Accuracy of both the user's position and orientation as well as map features (viz. POIs) are extremely important.

Next Steps

- Building infrastructure to manage latency and cost, leveraging caching and indexing in vector stores/databases.
- Adding additional spatial awareness relationships, and visual context to responses.
- Adding the ability for upstream human observations to provide continuous map improvement.