International Cartographic Association
Association Cartographique Internationale

From Geospatial Data to Narrative: A GIS-LLM Pipeline for Generating Personalised Outdoor Route Descriptions

Ilya Ilyankou[a, *], James Haworth[a], Tao Cheng[a], Stefano Cavazzi[b]

[a] *UCL SpaceTimeLab, ilya.ilyankou.23@ucl.ac.uk, j.haworth@ucl.ac.uk, tao.cheng@ucl.ac.uk*

[b] *Ordnance Survey, stefano.cavazzi@os.uk*

* Corresponding author

Keywords: Large Language Models (LLMs), route description generation, in-context learning, pedestrian navigation, cartographic communication, personalised navigation, hallucination mitigation

Abstract:

With the growing availability of detailed geospatial data and the rise of generative AI, there is increasing potential to enhance how route-based information is communicated to end users. Current cartographic practices rarely incorporate engaging narrative descriptions, and while generative AI can address this gap, factual inaccuracies (“hallucinations”) present significant challenges. These hallucinations can dangerously mischaracterise terrain difficulty and compromise walkers’ safety (for instance, mistakenly labelling steep or uneven paths as suitable for families with young children or individuals with limited mobility). This work presents a GIS-LLM pipeline that generates human-centred walking route descriptions while highlighting the importance of accuracy and safety.

Our approach bridges traditional cartography with AI-generated narrative content by drawing on open geospatial datasets and using in-context learning with LLMs. While complete factual grounding has not yet been achieved, this work represents an important first step toward that goal. This work presents the conceptual framework and implementation; a subsequent study will address formal validation protocols.

The proposed pipeline processes walking routes input as polylines (e.g., GPX or GeoJSON format) from proprietary recorded hikes via the OS Maps app, as well as public routes from CC-GPX[1] and synthetic, hand-drawn routes. Traditional GIS operations such as overlay, buffering, and nearest neighbour join, are used to extract contextual data by intersecting the route with multiple open geospatial datasets: Overture Places[2] points of interest (POIs), administrative boundaries from the OS and OpenStreetMap, land cover, elevation, and quantitative metadata (length, shape, difficulty). This spatial information is systematically structured into a YAML-formatted schema that forms the foundation for description generation. POI names and categories within 50m of the route are listed in order of appearance, with distances measured from the starting point.

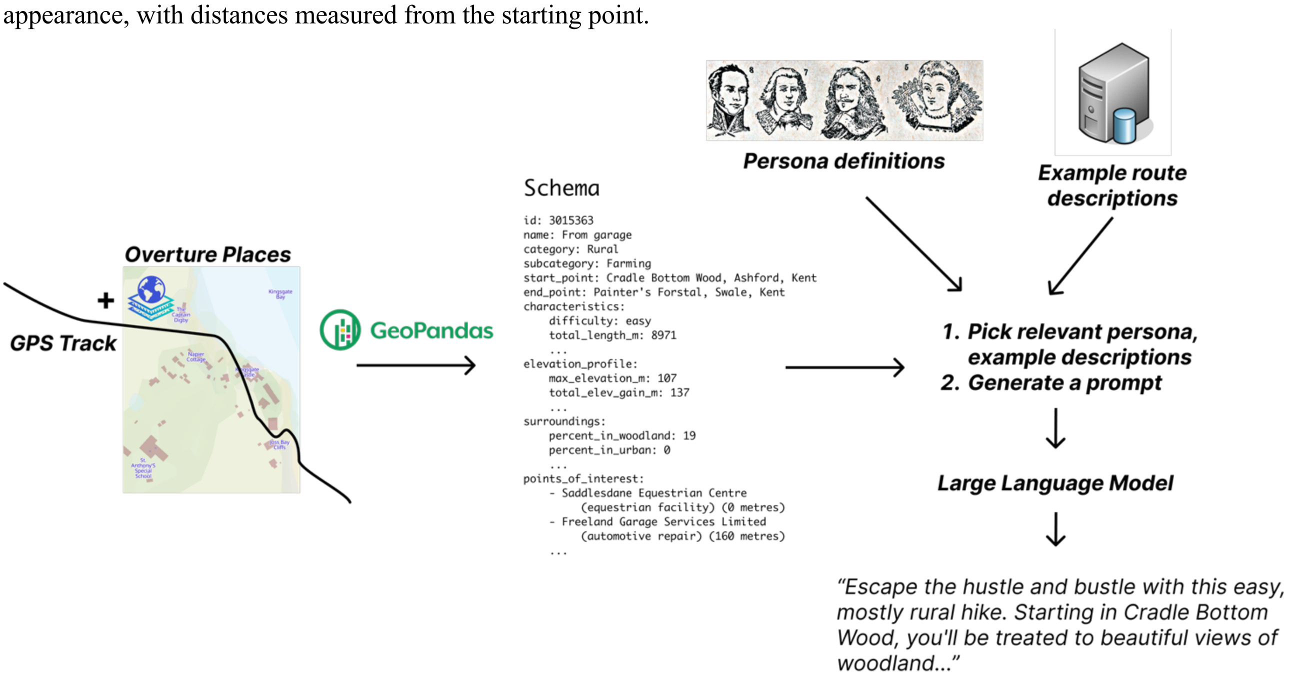

Figure 1: Pipeline overview. Traditional GIS tools and geospatial datasets are used to construct schemas for routes, and an LLM is used to generate engaging, factual, and personalised route descriptions.

[1] Ilyankou et al., 2024, Extracting High-Quality Annotated Geospatial Data from Common Crawl (https://dl.acm.org/doi/10.1145/3678717.3691215; on GitHub: https://github.com/ilyankou/cc-gpx)

[2] https://docs.overturemaps.org/guides/places/

The methodology employs in-context learning[3] with LLMs rather than model fine-tuning. This approach is resource efficient as it does not require extensive computational resources or a large training dataset of 'schema—description' pairs. In-context learning allows persona customisation, schema flexibility, and rapid iteration via simple prompt adjustments[4]. Additionally, it provides greater transparency for identifying and addressing hallucination patterns through tweaking examples and input formats—a critical consideration given the safety implications of route descriptions. While fine-tuning remains a potential future direction, for this exploratory phase of research in-context learning presents an optimal balance of efficiency and adaptability.

Each prompt is supplied between 1-3 schema—description pairs, from which the LLM extrapolated new route narratives. The prompt incorporates a persona parameter that enables targeted content generation for specific audiences (e.g., families with children, cultural heritage tourists), which results in emphasising different aspects of routes.

A simplified schema structure and the resulting description is illustrated in Figure 2.

```
length: 8.4 mi
difficulty: medium
start: Folkestone, Kent
end: Dover, Kent
is_coastal: True
is_circular: False
points_of_interest:
    - Folkestone Harbour
    - Folkestone Warren Nature Reserve
    - Capel-le Ferne Cliffs
    - Cliff Top Cafe
    - The Royal Oak (pub)
    - Abbot's Cliff Sound Mirror
    - Dover Marina
```

Embark on a scenic coastal journey from Folkestone to Dover, a walk brimming with captivating views and historical intrigue. Starting at the bustling Folkestone Harbour, you'll delve into the natural beauty of the Folkestone Warren Nature Reserve, where chalk cliffs and diverse wildlife await.

As you ascend the dramatic Capel-le Ferne Cliffs, treat yourself to breathtaking panoramic vistas and a well-deserved rest at the Cliff Top Cafe. Continue along the clifftops, passing the charming Royal Oak pub – a perfect pit stop for a traditional English pub experience. Keep an eye out for the fascinating Abbot's Cliff Sound Mirror, a remnant of World War II, before descending to Dover Marina, marking the end of your adventure.

This medium-difficulty walk offers a captivating blend of history, nature, and coastal charm.

Figure 2: Schema (simplified) to description transformation using Gemini 1.5. Underlined are route "facts" that can be traced back to the schema. Highlighted are LLM knowledge and/or "hallucinations".

LLMs exhibit factual hallucinations when generating route descriptions, manifesting as non-existent POIs, spatial misrepresentations, and unverifiable attribute assertions. In Figure 2, the model fabricates a "traditional English pub experience" at the Royal Oak and erroneously dates the Abbot's Cliff Sound Mirror[5] to World War II rather than its actual inter-war construction period. Prompt engineering and retrieval-augmented generation (RAG) are two key strategies for reducing hallucinations in LLM outputs that were used in our work. Other mitigation approaches—multi-agent pipelines that pair a generation agent with a dedicated fact-checking agent, system-level post-editing, and human-in-the-loop reviews—are also promising for this task and should be explored further.

The risk severity of these hallucinations varies significantly—from benign descriptive embellishments to potentially dangerous safety misinformation (e.g., characterising hazardous terrain as child-friendly). This risk differential requires a formal classification framework and targeted factual grounding methods that prioritise safety-critical information, particularly for outdoor navigation contexts where erroneous guidance could lead to physical harm.

Another key consideration is schema design. Drawing inspiration from prior work in semantic route modelling and Schema.org extensions for trails[6], this research implements a flexible yet expressive schema structure that balances detail with LLM readability. The proposed schema architecture prioritises both information density and compatibility with LLM comprehension capabilities.

The proposed approach contributes to the broader field of AI-supported cartography by enabling scalable, automated production of route narratives tailored to different audiences. These descriptions can enhance user engagement in walking apps, improve accessibility for non-map readers, and support personalised cartographic storytelling.

Future work will focus on three areas: (1) developing a risk assessment framework to categorise hallucinations by potential consequences; (2) implementing evaluation protocols combining automated metrics with human participant studies; (3) methods to mitigate hallucinations through improved factual grounding, including by employing multi-agent architectures that separate generation from verification. These directions aim to address the critical challenges of reliability and safety while preserving the engaging qualities of AI-generated route descriptions.

[3] Brown et al., 2020, Language Models are Few-Shot Learners (https://arxiv.org/pdf/2005.14165)
[4] Min et al., 2022, Rethinking the Role of Demonstrations: What Makes In-Context Learning Work? (https://arxiv.org/pdf/2202.12837)
[5] https://www.discoveringbritain.org/activities/south-east-england/viewpoints/folkestone-sound-mirrors-viewpoint.html
[6] Calbimonte et al., 2020, Semantic Data Models for Hiking Trail Difficulty Assessment