StreetSense

Using Visual Language Models as Virtual Auditors of Street-Level Pedestrian Experience

Lasith Niroshan and Francesco Pilla
UCD School of Architecture, Planning and Environmental Policy

Spatial Dynamics Lab

The "Blind Spot" of Smart Cities

If the map says you **CAN** walk here, why don't you **WANT** to?

High Walkability Score

Verified Sidewalk Presence

High Intersection Density

Active construction site

Blank Hoarding

Limitations of Manual Auditing

Temporal Friction

Manual surveys provide only a single snapshot, making longitudinal tracking impossible.

Resource Intensity

Surveying thousands of segments manually requires prohibitive labor and financial costs.

Scalability Bottlenecks

Audits are usually restricted to localized pilot studies rather than city-wide deployments.

Spatial Dynamics Lab

Research Objectives

Methodology

Define a scalable system architecture to convert unstructured imagery into granular safety and comfort data.

Subjectivity

Demonstrate VLM capacity for "Chain-of-Thought" reasoning that mirrors a human urban planner.

Proof of Concept

Identify environmental barriers that traditional GIS metrics systematically overlook.

Spatial Dynamics Lab

The New Paradigm

StreetSense deploys Vision-Language Models to operationalize subjectivity at scale.

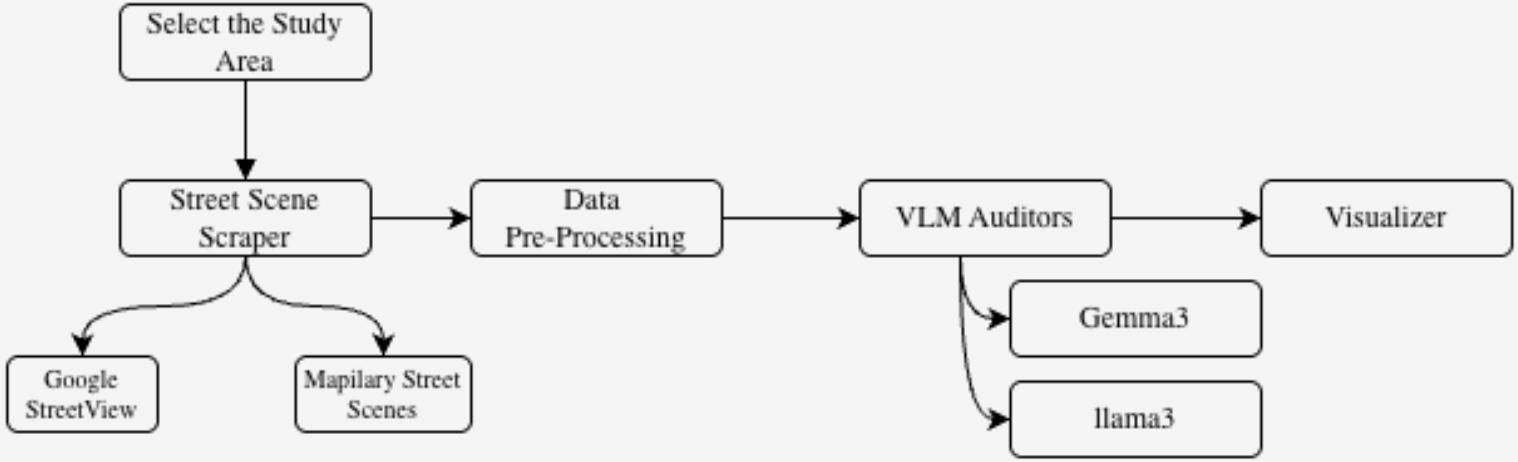
The Virtual Auditor
Bridging the gap between quantitative spatial analysis and qualitative human geography.

Scale & Efficiency
Auditing thousands of street segments for perceived safety and comfort without the prohibitive cost of manual surveys.

Model-Agnostic Architecture
designed for modular integration of next-generation VLMs.

Spatial Dynamics Lab

System Architecture



Spatial Dynamics Lab

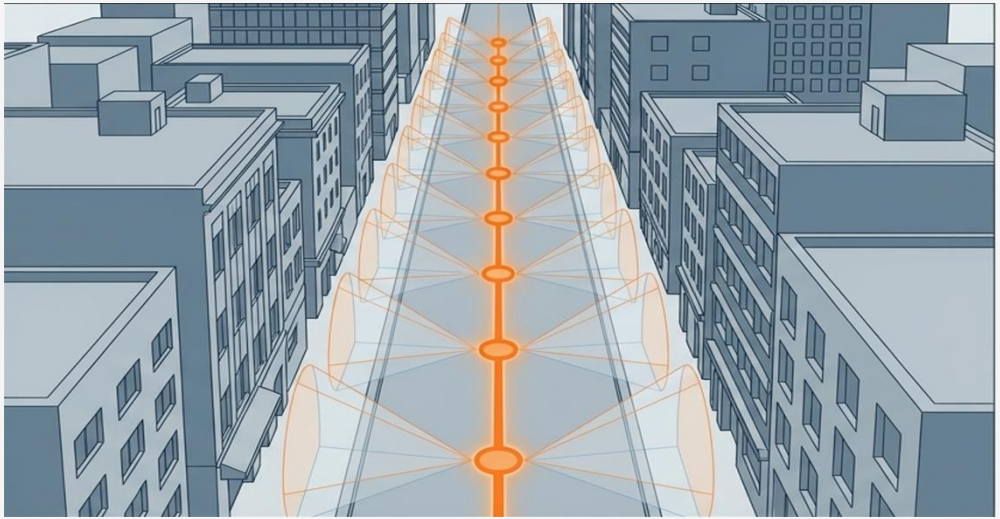
Moving from geometric proxies to perceptual urban analytics

	Traditional Geometric Proxies (e.g., WalkScore®)	Perceptual Analytics (StreetSense)
Core Metric	Measures Physical Access	Measures Human Experience
Analytical Viewpoint	Top-down, Bird's-eye (GIS)	Street-level, Human-eye (VLM)
Detection Limits	Infrastructure Presence (sidewalks, intersections)	Environmental Friction Points (litter, neglect, lack of shade)
Output Type	Opaque Numerical Index	Granular, Explainable Diagnostics

Spatial Dynamics Lab

Data Collection

Systematic, multi-source image acquisition captures the continuous pedestrian experience.



Multi-Source Aggregation

Pulling from Google Street View API (standardized temporal coverage) and Mapillary API (crowdsourced, diverse angles).

Stratified Sampling

Capturing diverse urban typologies.

Continuous Audit Trail

Systematic coordinate sampling along street centerlines at strict **50m** intervals.

Spatial Dynamics Lab

The Chain-of-Thought Pipeline

Standard computer vision relies on strict object detection. StreetSense enforces a structured Chain-of-Thought (CoT) prompting routine, forcing the model to print intermediate findings textually, ensuring an auditable logic trail.

Computer Vision Model

```
Object Detected:
Construction (1)
Pavement (2)
```

Transparent CoT Model

Contextual Understanding:
A functional urban env. with active construction site near to the pavement.

Spatial Dynamics Lab

How the Machine Reasons

1. Inventory

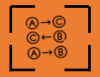
The model first lists objective physical features.

Example: Wide pavement, Overgrown weeds, Graffiti

2. Contextualization

It evaluates the condition of these features in relation to the pedestrian.

Example: The pavement exists but is a tripping hazard due to root uplift



3. Synthesis & Scoring

It assigns a "Pedestrian Comfort Score" (1-10) and provides a **text-based justification.**

Example: Pedestrian Comfort Score: 4/10 (Hostile and unsafe atmosphere)

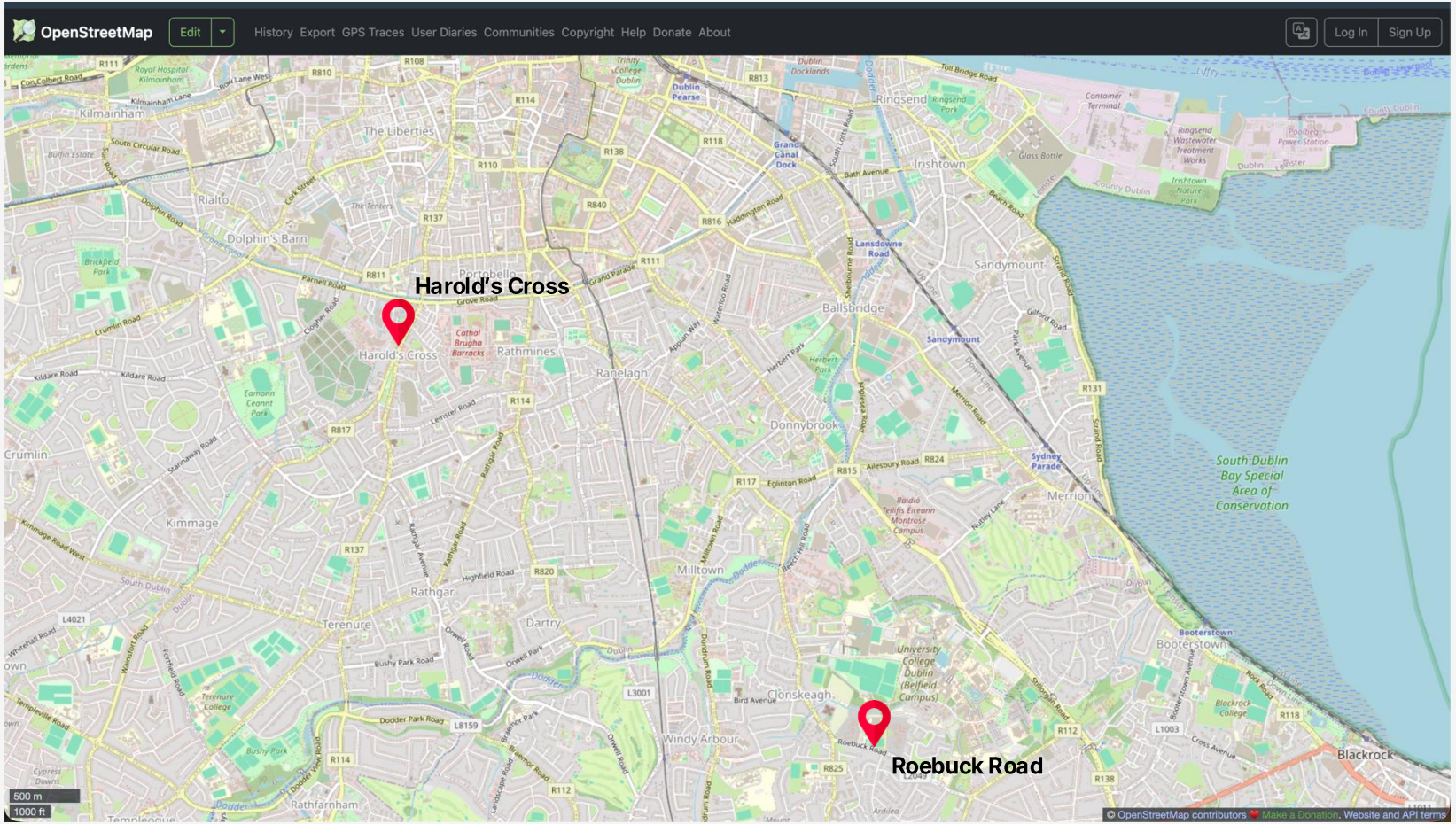
Spatial Dynamics Lab

Actionable Logic Trails

Traditional GIS Metrics

StreetSense Analysis

Spatial Dynamics Lab



Prototype Testing

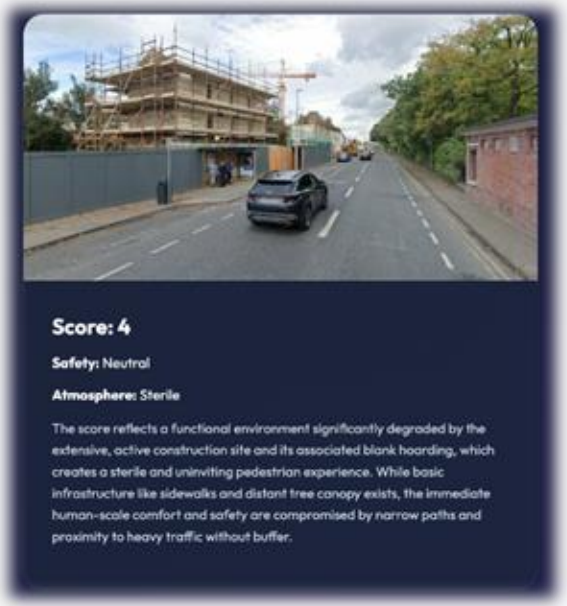
Harold Cross Case Study

Score: 7

Safety: Safe

Atmosphere: Established Commercial

The street benefits from active, human-scale storefronts and a pleasant public park with mature trees, fostering a sense of community and visual interest. However, the multi-lane road design and its implied traffic speed reduce pedestrian comfort and overall walkability, preventing a higher score.



Score: 4

Safety: Neutral

Atmosphere: Sterile

The score reflects a functional environment significantly degraded by the extensive, active construction site and its associated blank hoarding, which creates a sterile and uninviting pedestrian experience. While basic infrastructure like sidewalks and distant tree canopy exists, the immediate human-scale comfort and safety are compromised by narrow paths and proximity to heavy traffic without buffer.

- Active, human-scale storefronts and a pleasant public park.
- Blank hoarding generating a sterile and uninviting pedestrian experience.

Spatial Dynamics Lab

Note: All operational results and examples shown were generated using the local llama3:2 (2B) model deployment, orchestrated via the Ollama inference framework

Roebuck Road Case Study

Score: 8

Safety: Safe

Atmosphere: Vibrant

The street exhibits high-quality elements like wide sidewalks and ample tree canopy, fostering a sense of safety and vibrancy. The minor issue of litter does not substantially detract from the overall positive perception.

Wide sidewalks and ample tree canopy.

Score: 2

Safety: Unsafe

Atmosphere: Neglected

The environment lacks basic infrastructure like sidewalks and streetlights, and negative maintenance cues such as litter and overgrown vegetation dominate, creating a hostile and unsafe atmosphere.

Lacks basic infrastructure like sidewalks and streetlights, forced onto roadway.

Spatial Dynamics Lab

The **False Positives** of Walkability

Next Steps

Calibrating the virtual auditor for equitable and temporal urban analysis.

Human-in-the-Loop Validation

Deploying large-scale crowdsourcing to collect "ground truth" scores, fine-tuning VLM weights to strictly align with human psychological perception.

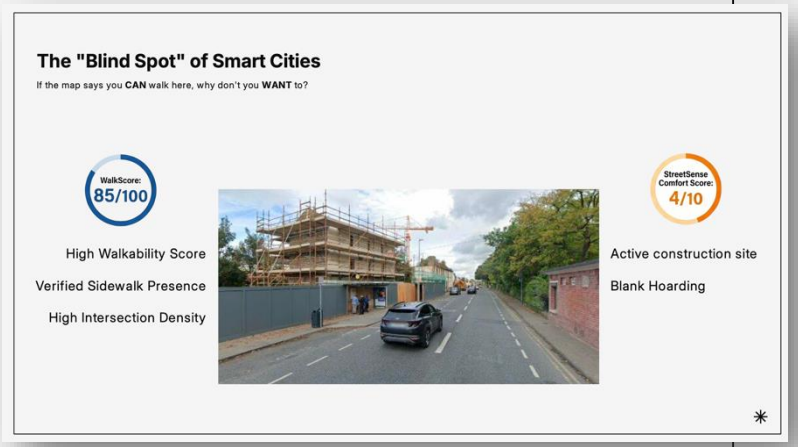
Bias Mitigation & Prompt Optimization

Future work will evaluate how the VLM's training data affects assessments across diverse architectural styles, maintenance conditions, and demographics. Additionally, alternative prompt engineering strategies will be tested to minimize model hallucinations and guarantee equitable evaluation metrics across varying urban socioeconomic contexts.

Longitudinal Analysis

Applying the pipeline to historical imagery to track the degradation or improvement of pedestrian experience over time.

Spatial Dynamics Lab



StreetSense transforms raw, unstructured imagery into auditable diagnostic logs. It provides city planners with a highly accessible methodology to visualize not just where infrastructure dictates citizens can walk, but where localized design choices determine they want to walk.

Spatial Dynamics Lab

Thank You

lasith.kottawahewamanage@ucd.ie

Spatial Dynamics Lab

This research was conducted as part of the RURACTIVE project, funded by the European Union's Horizon Europe programme.