
# The Invisible City: Not All Streets Are Seen Equally

Yu Wang, Joseph Shingleton, Yunus Serhat Bıçakçı,
Ana Basiri, James Ackland

University of Glasgow
Oilthigh Ghlaschu

# Example: Queen Street Station 2020 vs. 2014


University of Glasgow
Oilthigh Ghlaschu

# Background

- **Context:**
  - Street View Imagery (SVI) is widely used in urban planning, health studies, and AI modelling as a proxy for the physical world.
- **The Problem:**
  - Research often treats SVI as static and evenly updated across space
  - Data collection driven by commercial algorithms may lead to temporal bias.
- **Research Question:**
  - Are all urban spaces updated equally? Or are certain communities (e.g., pedestrian zones vs. main roads) systematically "forgotten"?
- **Study Area:**
  - Glasgow, UK (High-resolution 20 metre analysis).

University of Glasgow
Oilthigh Ghlaschu

# Methodology

- **Step 1: Grid Generation (20m Resolution):**
  - Generated a 20m × 20m grid covering the entire Glasgow boundary.
  - Total sample points: ~441k+ grid centres.
- **Step 2: Semantic Integration (OSM):**
  - Used OpenStreetMap data to classify every grid by Road Types:
    - Drivable: Motorways, primary, secondary roads.
    - Non-drivable: Pedestrian paths, cycleways, steps.
    - No-road: Buildings, parks, open land.
- **Step 3: Temporal Metadata Extraction:**
  - Retrieved historical SVI metadata (PanoID, Year, Month).
  - Cleaned and snapped images to the nearest grid point to remove duplicates.
  - Result: A dataset linking spatial infrastructure types with temporal update frequencies.

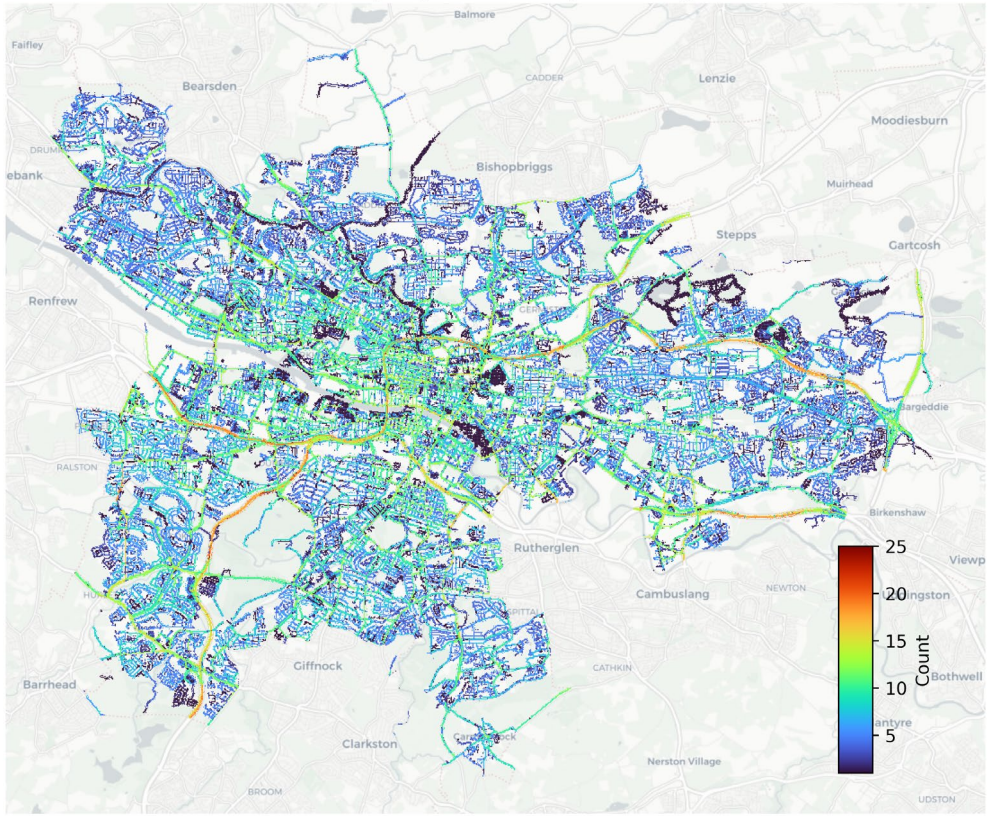
University of Glasgow
Oilthigh Ghlaschu

# Measuring the "Unseen": Key Metrics

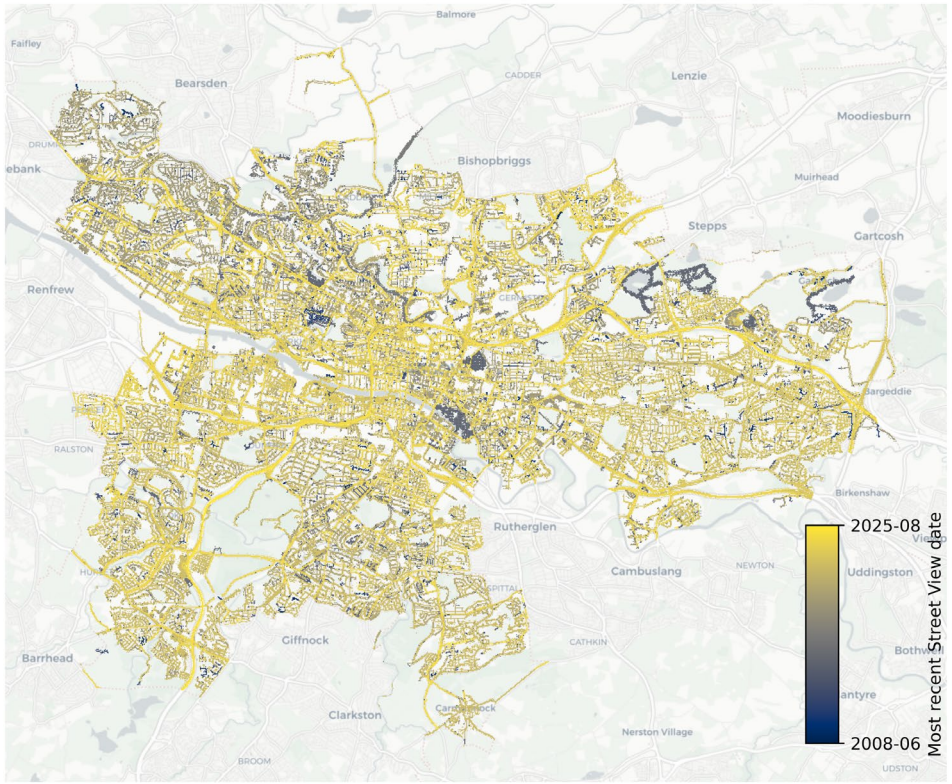
- **Recency (Data Freshness):**
  - Metric: Months elapsed since the most recent image update.
  - Concept: High values indicate the digital representation lags behind physical reality.
- **Consistency (Fragmentation):**
  - Metric: Max Gap Months (Longest time between two consecutive updates).
  - Concept: Identifies "Forgotten" periods where areas went unmonitored for years (e.g., >72 months).
- **Coverage:**
  - Does the grid have any SVI data? (Comparison of coverage rates between road types).

# Left: Available SVIs per grid. & Right: Recency per grid.

Count of unique year-month per 20m grid cell

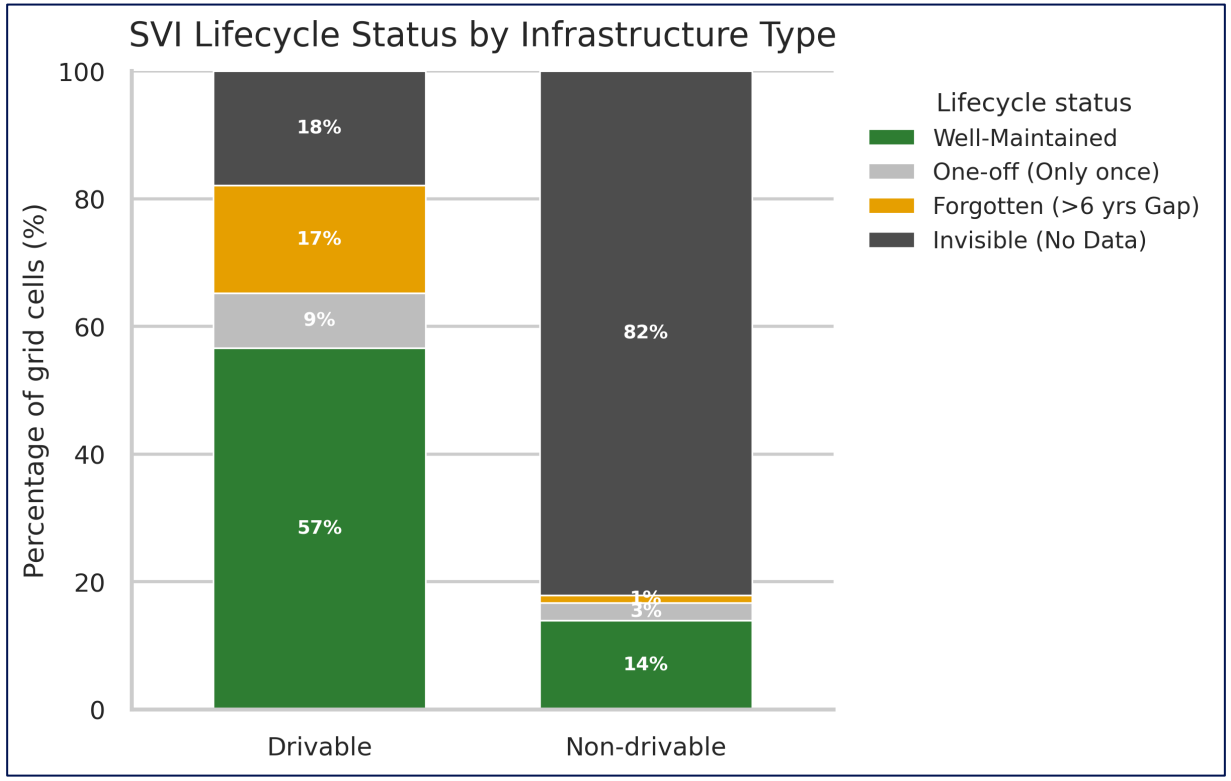



Most recent available Street View date per 20m grid cell

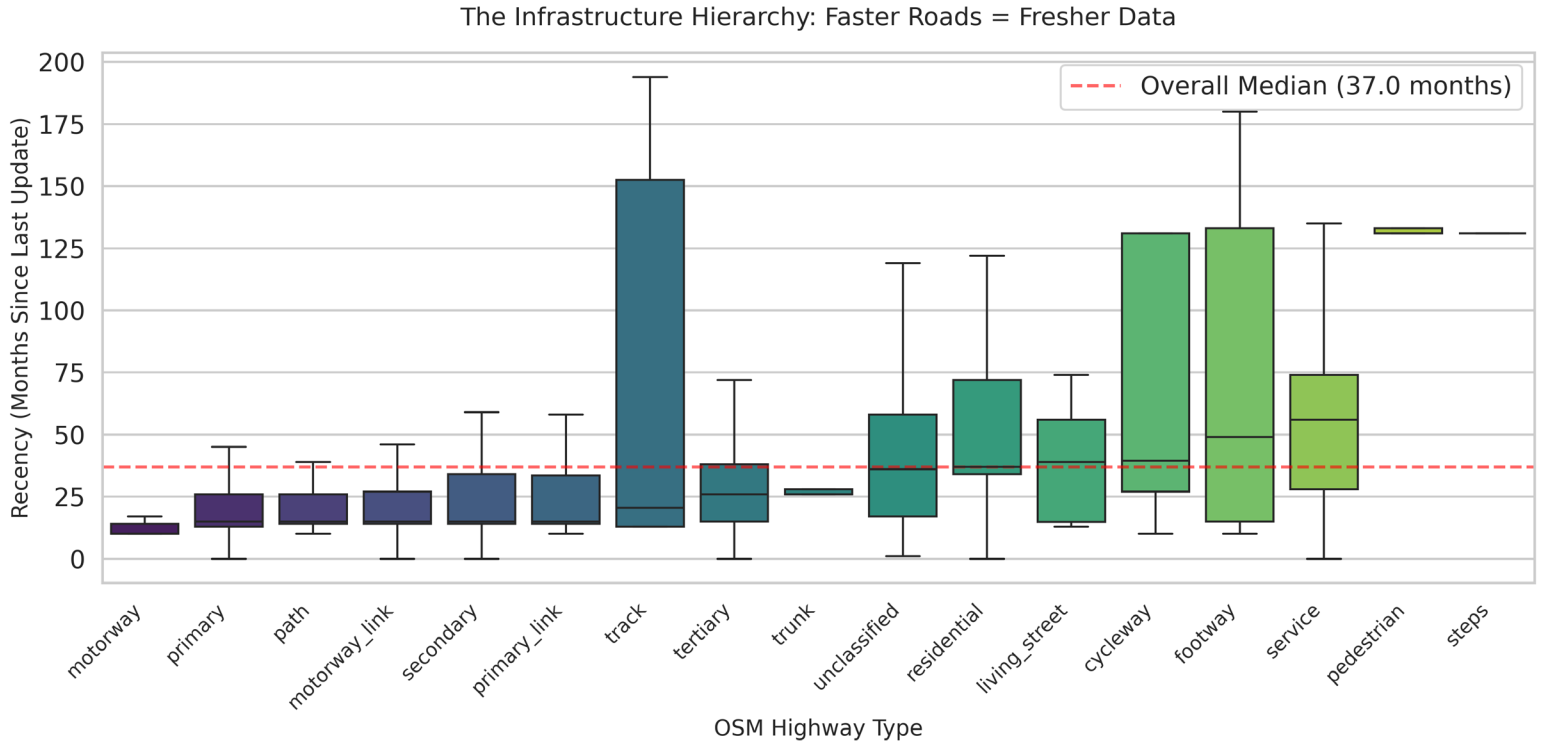


# Lifecycle composition of SVI across drivable and non-drivable grids.

- **Well-maintained:** *Regular updates with no gaps exceeding six years.*
- **One-off:** *A single snapshot, frozen in time.*
- **Forgotten:** *Discontinued monitoring with a data gap of over 6 years (unstable maintenance).*
- **Invisible***: no collection at all.*



# Distribution of SVI recency across different road types, measured as months since the most recent update (2008–2025).




1. **High-Priority (Motorway, Primary) Leads to Rapid Refresh:**
   - Critical infrastructure is updated almost annually (Median ~12 months).
   - Logic: High traffic volume drives frequent data collection.
   - Result: A "Live" digital twin exists only for major highways.

2. **The Long Tail (Residential, Service, & Minor)**
   - Systematic Decay: As road importance drops, data ages rapidly.
   - Residential roads and service lanes lag significantly (3-5+ years old).

# Thanks!