Street Evidence Model (SEM)

Whitepaper on Neural City's Street-Level Data Model for Outcome-Oriented Urban Measurement

Data Model and Public Disclosure (Whitepaper v0.1)

Version date: 2 Feb 2026 (Asia/Kolkata)

Audience: City officials, researchers, infrastructure practitioners, and data partners

Scope: Street-Level Visual Scoring for Outcome-Oriented Urban Measurement (images and videos)

Table of Contents

Preface

India's urban measurement ecosystem is steadily moving from one-off indices to outcome-oriented, regularly updated datasets that can be audited, compared, and improved over time[1][2].

At the same time, most existing city frameworks rely on administrative reporting, periodic surveys, or aggregate service indicators, which often miss the last-mile reality that residents actually experience on streets and in public spaces[3][4][5].

Neural City's scoring system is designed as a street-level measurement layer that operationalizes outcome frameworks through visual evidence. It is intended to complement, not replace, national and international indicator systems[1][3][4].

This document describes the Neural City data model and scoring approach which translates street-level visual evidence (geo-tagged, timestamped images/videos) into repeatable indicators that can be aggregated to street, ward, roads, and city views.

This data model is designed for two simultaneous needs:

1. Infrastructure meaning (domain lens): indicators reflect real, on-ground conditions that matter for service readiness and daily experience[2].
2. Machine learning readiness (training lens): indicators must be visually distinguishable in diverse Indian street contexts (non-standard designs, colors, materials), so the scoring ladders use 3–5 levels where reliable separation is feasible, and binary presence flags where that is more defensible[2][4].

Neural City treats the scores from visual evidence primarily as trend and relative-comparison signals (within a city over time, and across cities with similar contexts), not as a legal compliance declaration[2].

Abbreviations

AI: Artificial Intelligence
CSCAF: Climate Smart Cities Assessment Framework
EIU: Economist Intelligence Unit
EoL: Ease of Living
GIS: Geographic Information System
ICCC: Integrated Command and Control Centre
ISO: International Organization for Standardization
KPI: Key Performance Indicator
ML: Machine Learning
MoHUA: Ministry of Housing and Urban Affairs, Government of India
NIUA: National Institute of Urban Affairs
NMT: Non-Motorised Transport
QoL: Quality of Life
SDG: Sustainable Development Goals
ULB: Urban Local Body
UOF: Urban Outcomes Framework

1. Introduction

1.1 Why outcomes need street-level measurement

Outcome-oriented approaches such as the Urban Outcomes Framework explicitly aim to strengthen evidence-based policymaking by building comparable, regularly updated urban datasets across sectors[1][2].However, many of the outcomes citizens care about are experienced in the public realm through daily exposure: cleanliness on streets, usability of bins, accessibility of toilets, continuity of walkable space, drain safety, or visible encroachments.

Administrative indicators can confirm whether infrastructure exists or whether services are provisioned, but they often cannot validate whether those services are functional, accessible, and consistent at street level across neighborhoods. This is a known challenge in urban quality-of-life measurement, where composite indices frequently struggle to bridge the gap between high-level indicators and lived experience[6].

Neural City's Street-Level Visual Scoring is built specifically to close this gap by converting street imagery into structured, comparable, outcome-relevant scores.

1.2 Positioning within existing frameworks

Neural City's method draws domain framing and governance intent from established frameworks and standards. It does not claim compliance with any single framework. Instead, it translates outcome domains into observable street conditions using a consistent rubric.

Key reference anchors include:

- Outcome orientation and open, comparable datasets: Urban Outcomes Framework[1][2].
- Standardized service and quality-of-life indicators: ISO 37120[3].
- Composite liveability framing and cross-city benchmarking logic: EIU liveability approach[5].
- India-specific city assessment lineage: NIUA liveability and climate frameworks[4][7].
- Methodological literature on QoL indices and comparability[6].

2. Scope and intended use

2.1 What this methodology measures

Neural City currently measures visible conditions in images that act as proxies for citizen-facing outcomes in the public realm. These include, for example:

- Cleanliness conditions such as garbage and litter presence and maintenance signals.
- Public service usability conditions such as bin usability and overflow.
- Accessibility and upkeep conditions such as toilet accessibility and maintenance.
- Safety-relevant public realm conditions such as open drains or damaged drain covers.

These are structured into a parameter and sub-parameter model described in Section 3.

2.2 What this methodology does not claim today

This methodology is intentionally limited to image-based acquisition. It is not yet a full multi-telemetry city intelligence stack. It does not currently incorporate LiDAR, air quality, heat, traffic telemetry, land use, transit operational feeds, or official geofenced rule layers (for example, formal parking permissions). The method is designed so such layers can be integrated later for cities where data availability and governance arrangements allow.

Scores should primarily be interpreted as:

- Relative comparisons across locations, corridors, and area types within similar contexts
- Trend signals when collected repeatedly
- Diagnostic pointers to "where conditions differ sharply" rather than precise engineering compliance certification

2.3 Transparency and disclosure boundary

This document publishes the indicator catalogue (parameters and sub-parameters), scale types, evidence rules, and representative rubric examples to enable interpretation and external scrutiny. It intentionally does not publish aggregation coefficients, weighting parameters, model confidence thresholds, or training heuristics. This boundary protects system integrity, prevents score gaming, and preserves the ability to evolve the methodology while maintaining comparability over time.

2.4 Representativeness and coverage

Scores reflect the observed condition of the locations captured in the dataset. They should not be interpreted as statistically representative of the entire city unless collection coverage and sampling strategy explicitly support that claim. Neural City therefore reports coverage metadata (where available) and encourages interpretation through relative comparisons, hotspot identification, and time-based trend tracking within comparable corridors and area types.

3. Data model

3.1 Unit of observation: an image as a micro-audit

The atomic unit of scoring is a single image (or frame extracted from a video). Each image is treated as a bounded observation window of the streetscape at a time and location.

This choice is deliberate:

- It allows transparent traceability from score back to evidence
- It enables re-audit when methods evolve
- It supports AI training because labels are tied to discrete visual samples

3.2 Parameter and sub-parameter hierarchy

Neural City uses a two-level hierarchy:

- Parameters: outcome-relevant domains that are interpretable for governance and public communication
- Sub-parameters: visually observable elements that can be scored consistently from imagery

The current mapping includes multiple high-level parameters and several dozen sub-parameters.

3.3 Parameter naming convention note

Neural City currently uses parameter names that are legible to broad audiences. For example, "Walkability" is used as an interpretable umbrella even when limited NMT-related signals are included, because:

- Public comprehension is higher
- Current cycling-specific data coverage is smaller compared to pedestrian coverage
- Many NMT features are physically coupled with footpath continuity and design in typical Indian streets

This naming convention will be refined as data coverage expands.

4. Indicator catalogue

Neural City currently uses six parameters, each consisting of visually scored sub-parameters. The catalogue below lists what is scored.

4.1 Parameter and Subparameters

1) Cleanliness and public hygiene

Focus: visible waste, hygiene conditions, and usability of basic public hygiene infrastructure.

Sub-parameters:

- Garbage and litter intensity
- Tobacco spit presence
- Bin usability (functional integrity)
- Bin overflow status
- Bin cleanliness (external maintenance)
- Segregated bin availability
- Open urination presence
- Public toilet accessibility features
- Public toilet maintenance condition
- Drain presence and openness
- Drain waste accumulation
- Manhole or drain cover condition
- Median cleanliness

2) Walkability and pedestrian infrastructure

Focus: continuity, usability, and safety of pedestrian movement space.

Sub-parameters:

- Sidewalk availability
- Obstruction on sidewalk
- Parking on sidewalk
- Street furniture presence
- Street furniture usability

- Tactile paving presence and quality
- Sidewalk surface quality
- Sidewalk effective width
- Sidewalk continuity and termination
- Ramp slope (where ramps exist)
- Ramp build quality
- Bollards presence and spacing logic (where visible)
- Curbs presence and continuity
- Fencing and guardrails (where used)
- Traffic islands (pedestrian refuge)
- Crosswalk presence and maintainability
- Raised crosswalk presence and quality
- Overhead bridge deck condition
- Overhead bridge stairs condition
- Subway or underpass availability and usability
- Signage adequacy for pedestrian movement
- Cycling infrastructure presence and condition (where visible)
- Bus stop build quality
- Bus stop discipline conditions (queuing, spillover, blockage)
- Bus stop restricted accessibility (step-free access cues)
- Bus stop commuter facilities (basic shelter and amenities)
- Electrical safety in pedestrian realm
- Street lighting coverage for pedestrian areas

Note on naming: "Walkability" is used as a public-facing umbrella because most measured signals are pedestrian-first. Cycling infrastructure is included only where it directly shares the same right-of-way and can be reliably observed.

3) Roads and traffic operations (visual layer)

Focus: visible surface and operations that directly affect ride quality and flow.

Sub-parameters:

- Surface quality
- Blacktop quality (finish, patching, pothole cues)
- Roadside gutter presence and condition
- Road-drain connection visibility and functionality cues
- Road markings visibility and maintenance
- Median quality (physical condition)

- Speed breaker or bump quality
- Parking on road presence and severity (visual)
- Traffic signs (captured as visual presence in this phase)
- Taxi queue and lane blocking (captured where observed)

4) Dust and debris

Focus: visible dust loads and construction debris in the public realm.

Sub-parameters:

- Dust level on street surface
- Construction material presence in public realm

5) Encroachment and right-of-way pressure

Focus: observable occupation of public space.

Sub-parameters:

- Encroachment extent
- Occupant type (broad category)
- Structure type (broad category)

6) Streetscape and aesthetics (public realm cues)

Focus: visible maintenance cues that influence perception and comfort.

Sub-parameters:

- Road-facing facade condition
- Street installations condition (public fixtures)
- Street art condition
- Overhead utilities clutter
- Ornamental plants maintenance
- Decorative lighting presence and condition

4.2 Quality-band Examples

The band examples below are illustrative and intended to convey boundary logic. Final band interpretation is governed by Neural City's visual reference and scorer calibration rules, which are periodically updated as new city contexts and edge cases are observed.

Garbage and litter intensity (4-band)

- Band A: large visible dump or concentrated garbage heap
- Band B: severe littering across the visible area
- Band C: slight littering, scattered waste
- Band D: no visible litter or garbage in the frame

Bin usability (3-band)

- Band A: broken or unusable bin body or lid
- Band B: minimally functional bin that meets basic use
- Band C: good quality bin with intact usability cues

Bin overflow (binary)

- Band A: overflow present or garbage spilled around the bin
- Band B: no overflow visible

Public toilet accessibility (4-band)

- Band A: toilet locked or effectively inaccessible
- Band B: no visible accessibility features (step-free access cues absent)
- Band C: accessibility features present but poorly designed or constrained
- Band D: strong step-free accessibility cues and usable design

Public toilet maintenance (5-band)

- Band A: filthy condition, strong visible neglect cues
- Band B: broken fixtures, stains, non-functional cues
- Band C: adequate maintenance, usable but not well-kept
- Band D: good maintenance, clean and usable
- Band E: excellent maintenance, high cleanliness cues

Manhole or drain cover condition (4-band)

- Band A: open chamber or missing cover

- Band B: broken, raised, uneven cover creating hazard cues
- Band C: adequate cover quality, minor unevenness
- Band D: well-aligned cover, smooth integration cues

Walkability examples

Sidewalk availability (3-band)

- Band A: no sidewalk present in the frame
- Band B: local street context with no practical space for sidewalk provision
- Band C: sidewalk present and visually usable as a walking space

Obstruction on sidewalk (4-band)

- Band A: sidewalk physically broken to the extent of being unusable
- Band B: encroached by vendors or vehicles, walking space compromised
- Band C: obstructed by poles, trees, debris, equipment, narrowed passage
- Band D: clear sidewalk with no meaningful obstruction in the frame

Sidewalk continuity (4-band)

- Band A: sidewalk ends into a structure or dead-end barrier
- Band B: sidewalk termination is undefined and unsafe to continue walking
- Band C: sidewalk end is constructed but lacks a usable ramp cue
- Band D: sidewalk ends with a clear ramp cue enabling continuity

Roads examples

Parking on road (3-band)

- Band A: vehicles parked or stopped on the active lane
- Band B: vehicles parked or stopped on the shoulder, partial interference
- Band C: no visible parking on lane or shoulder in the frame

Safety cue

Electrical safety in public reach (3 to 4 band, depending on scene)

- Band A: exposed cables or electrical elements within public reach
- Band B: unsafe connections or equipment in reach without perimeter control
- Band C: equipment perimeter control and separation cues

Published subset source: Neural City rubric mapping.

4.3 Standards anchor and non-compliance statement

Where sub-parameters relate to pedestrian infrastructure, accessibility, roadway elements, and public-realm safety, Neural City anchors its conceptual understanding primarily in Indian Roads Congress (IRC) guidelines, which define accepted practice for design, usability, and safety of urban streets in India.

Relevant IRC documents include, but are not limited to:

- IRC guidelines on pedestrian facilities, footpath design, crossings, and accessibility features
- IRC provisions on road geometry, drainage, surface condition, signage, and safety elements
- IRC guidance on universal access and facilities for persons with reduced mobility, where applicable

These IRC guidelines inform what constitutes a *functionally usable*, *partially degraded*, or *failed* condition in the Indian urban context. Neural City does not assess dimensional compliance or certify adherence to IRC drawings. Instead, it uses IRC intent to interpret whether the visible street condition supports or undermines the intended outcome of safe and usable public infrastructure.

Where IRC guidance is silent, high-level, or not visually verifiable from street imagery, Neural City draws secondary conceptual reference from internationally recognized accessibility and built-environment standards, such as ISO 21542 and EN 17210, to maintain consistency with global best practices in inclusive design. These references are used only to inform qualitative interpretation, not to enforce international compliance thresholds.

Importantly, Neural City scores are not engineering compliance certificates. They are outcome-proxy signals derived from visible evidence, intended to highlight relative condition, functional breakdowns, and spatial patterns in infrastructure performance. Detailed engineering audits, statutory compliance checks, and dimensional verification remain outside the scope of this methodology

4.4 Encroachment: interpretation, categorisation, and intent

Neural City uses the term *encroachment* as a technical shorthand to describe the occupation of public right-of-way or public space that constrains intended movement, access, or service function. The term is widely used in urban governance and legal contexts, but it also carries a negative connotation. Neural City acknowledges this limitation in terminology.

In practice, the on-street reality of Indian cities includes a wide spectrum of public space usage, ranging from informal livelihoods to permanent commercial spillover. To reflect this reality, Neural City deliberately differentiates encroachment by type of occupant and form of structure, rather than treating all occupation as a single negative condition.

The scoring framework therefore distinguishes between, for example:

- roadside vendors operating without fixed structures
- hand carts and mobile vending units
- temporary or semi-permanent sheds (including metal or fabric structures)
- spillover by formal shops onto footpaths or carriageways
- occupation by parked or stationary vehicles

This differentiation is intentional. It allows cities to understand *how* public space is being used, not merely *that* it is being used. The presence of informal vending or livelihood activity is not automatically equated with poor urban outcomes. Instead, scores reflect the extent to which space occupation constrains movement, safety, or accessibility, particularly for pedestrians and vulnerable users.

Neural City's encroachment indicators are therefore not moral or legal judgments. They are spatial diagnostics that surface trade-offs between:

- clearance and continuity of public movement, and
- the lived economic realities of informal and semi-formal employment in cities.

The term *encroachment* is retained in this methodology because alternative phrases such as "public space usage" or "vendor presence" do not fully capture the operational constraint that occupation can impose on sidewalks, carriageways, or service corridors. However, the underlying scoring logic explicitly recognises variation in form, intensity, and impact.

Encroachment scores should be interpreted as signals of pressure on public space, not as prescriptions for eviction or removal. Policy responses may range from redesign and reallocation of space, to formalisation, vending zone planning, or improved management, depending on local context and governance objectives.

5. Visual scoring rubrics

5.1 Conservative evidence rule

Only score what is clearly visible. Do not assume. If uncertain, leave blank and flag for review. This rule is essential for auditability and model integrity.

"Not observed" is a distinct state, not a low score. If the relevant feature is not visible due to obstruction, framing, or lighting, the field is left blank and excluded from aggregation for that sub-parameter. This prevents penalizing locations for missing evidence and protects interpretability.

5.2 Rubric is designed for Model training

The sub-parameter rubric is designed to satisfy two constraints simultaneously:

1. Domain relevance: it reflects how infrastructure and street conditions are described by planners, auditors, and civic frameworks
2. Visual separability: score categories must be distinguishable enough in images to train early ML models reliably

Only include sub-parameters that can be reliably observed from typical street imagery across multiple cities and design contexts, without requiring hidden knowledge (such as underground asset maps or rulebooks not present in the scene).

5.3 Multi-city variability and scale choice

India's on-street design signals are highly non-standardized across cities, corridors, and even within wards. Colors, materials, maintenance patterns, and signage consistency vary sharply.

Because of this, Neural City generally uses short ordinal scales (typically 3 to 5 levels) rather than attempting 7 or 10 level quality scales. This choice reduces false precision while preserving meaningful differentiation that is visible and learnable in images.

5.4 Binary indicators for high-salience behaviors

Certain street behaviors are not "infrastructure quality" but strongly influence public perception and lived experience. For such cases, a binary present or not-present indicator is used instead of an ordinal quality scale (examples include tobacco spit presence and open urination).

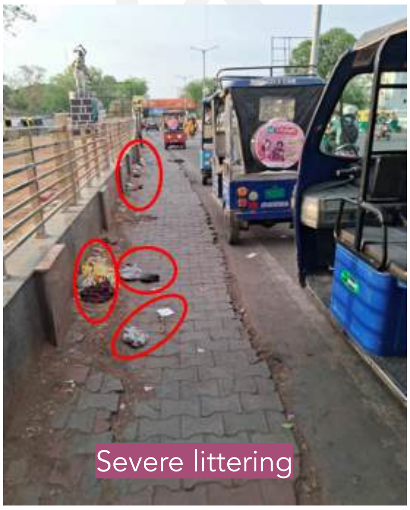
This prevents forced grading where "quality levels" are not conceptually meaningful.

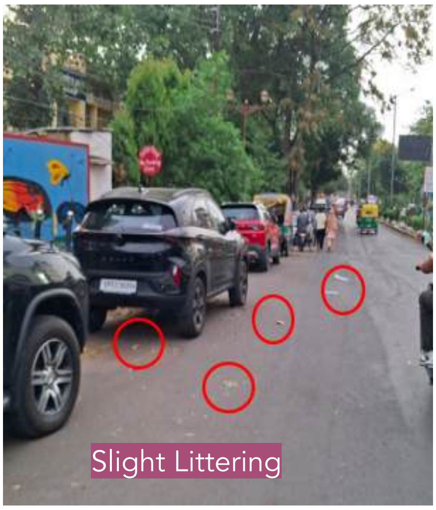
5.5 Reference images as a Visual Standard

Neural City maintains visual references that define sub-parameters with example images and boundary cues. The shared images cover multiple cleanliness and public service elements and show explicit examples for categories such as garbage and litter intensity, bin usability, bin overflow, toilet accessibility and maintenance, and drain coverage states.

5.5.1 Garbage and litter intensity

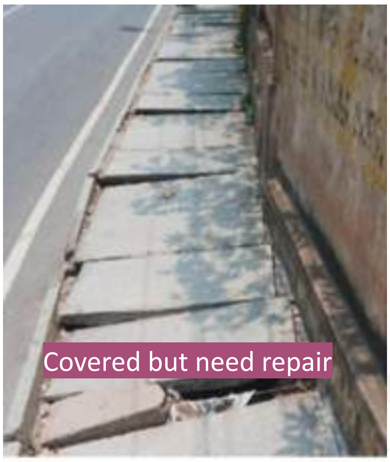

Garbage dump


Severe littering

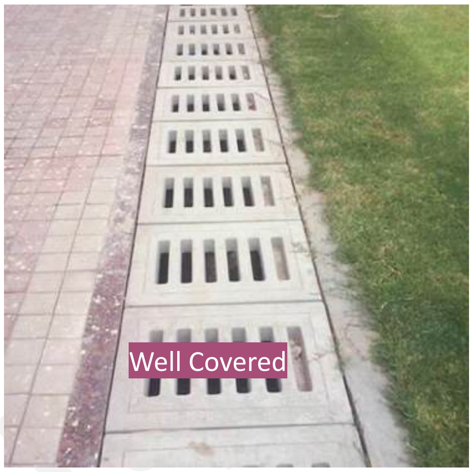

Slight Littering

5.5.2 Drain Cover

Majority Slab Broken,Immediate Hazard

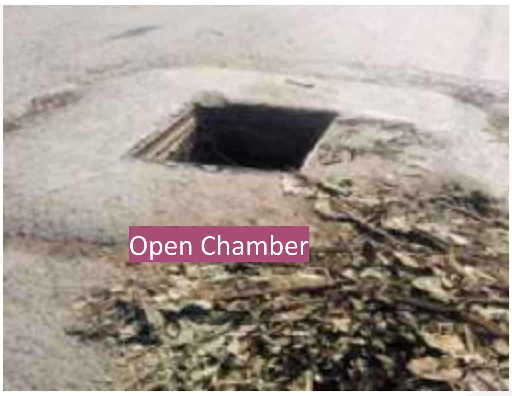


Covered but need repair



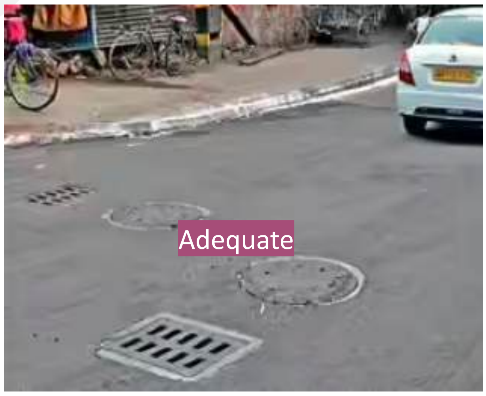
Well Covered

5.5.3 Manhole Cover



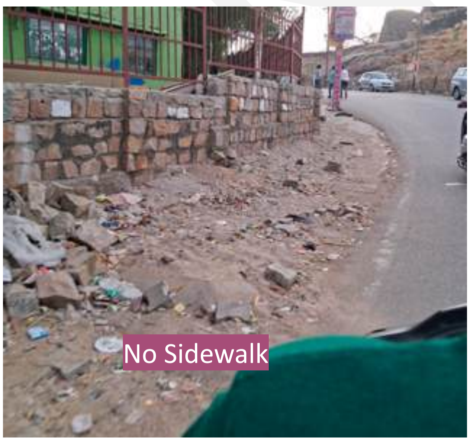
Open Chamber

Broken, Uneven or Raised



Adequate

5.5.4 Footpath/sidewalk Availability

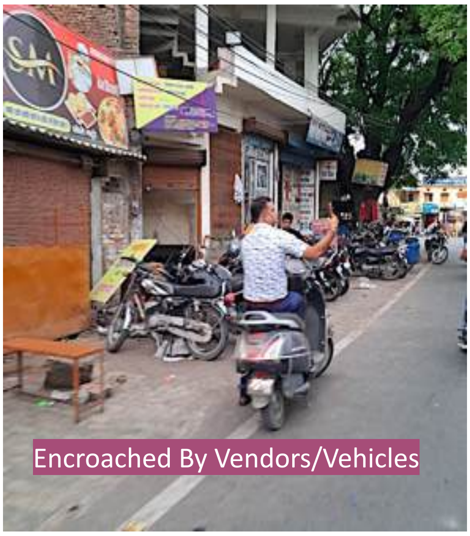


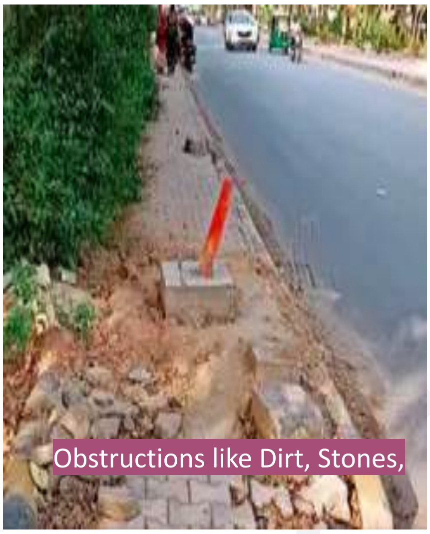
No Sidewalk

Local Street, no space

Sidewalk Available

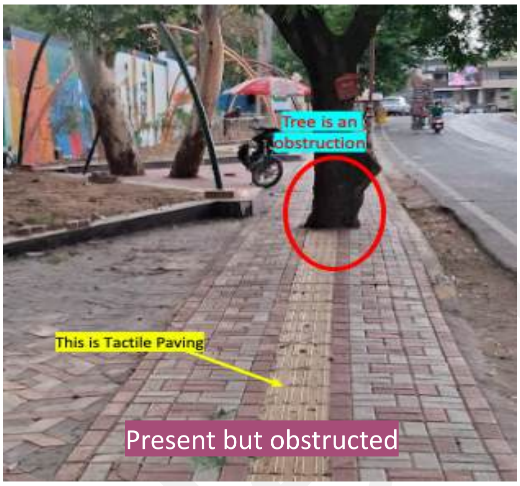
5.5.5 Obstruction on Sidewalk


Encroached By Vendors/Vehicles

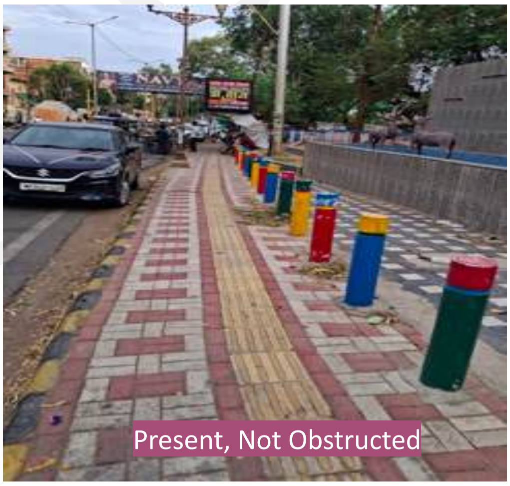

Obstructions like Dirt, Stones,


Clear Sidewalk

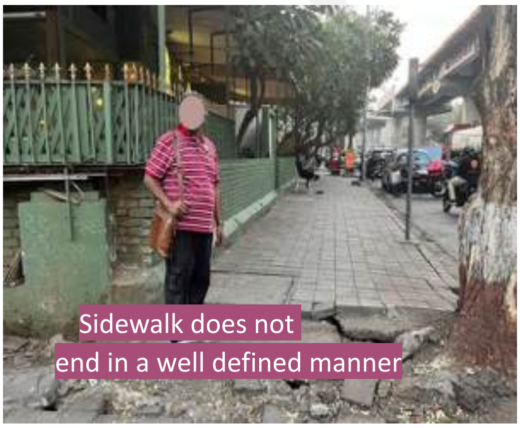
5.5.6 Tactile Paving

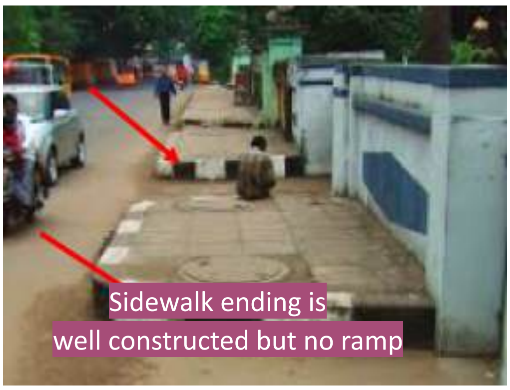


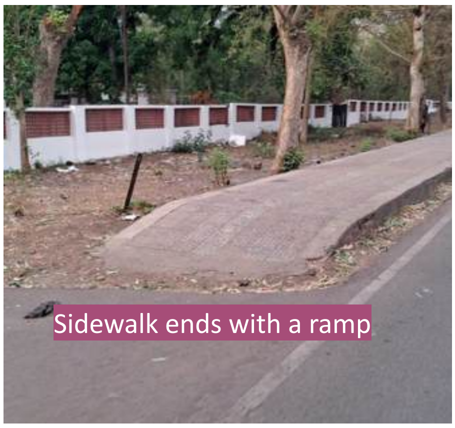
Present but obstructed


Present, Not Obstructed

5.5.7 Sidewalk Continuity


Sidewalk does not end in a well defined manner


Sidewalk ending is well constructed but no ramp


Sidewalk ends with a ramp

6. Scoring Ladder and normalization

Neural City score ladders are designed with "real-world observability" as the primary constraint.

Each ladder is constructed only where conditions can be reliably distinguished from typical street imagery across multiple Indian cities.

Where the on-street variation is visibly graded, multi-band ladders are used. Where conditions are binary or near-binary in images, limited ladders are intentionally retained. This avoids artificial granularity and improves inter-observer consistency, auditability, and replicability across cities.

A sub-parameter is scored only when the relevant feature is clearly visible. When evidence is insufficient, it is recorded as "not observed" and excluded from that sub-parameter's aggregation.

6.1 Why scoring ranges differ across sub-parameters

Not all sub-parameters use the same scale length or the same maximum score. This is deliberate. Some street conditions allow meaningful differentiation (for example, surface quality, continuity, usability). Others have limited observable variation (for example, presence of spit or bin overflow). A uniform ladder across all indicators would misrepresent real-world variability, overstate minor features, and reduce interpretability.

Neural City uses short discrete scales (binary, 3-band, 4-band, 5-band) because visual boundaries for subparameters between fine-grained levels are not stable across Indian cities due to non-standardized materials, colors, and designs. Short scales reduce false precision while preserving meaningful separation for AI training and governance interpretation.

6.2 Raw rubric levels

Sub-parameters are scored using discrete rubric levels. The rubric contains a mixture of:

- Ordinal quality scales (typically 3 to 5 levels depending on the item)

- Binary presence indicators for select features

The definitions and mapping of sub-parameters to possible score levels are maintained in the scoring rubric dataset.

Where visual differentiation is limited, we use calibrated spacing between bands rather than adding more rungs. This ensures that severe failures are not diluted and edge cases do not manipulate the median.

6.3 Reporting as a 0–100 index

To make results legible to non-technical stakeholders, Neural City reports scores on a 0–100 style scale.

In general, ordinal rubric scores are converted into a 20–100 ladder (multiplying by 20 for a 1–5 scale)

This creates an intentionally high visual separation between materially different states (for example, a location that is meaningfully usable versus one that is meaningfully degraded).

A key interpretive note:
The expanded 0–100 style scale is for communication clarity. It does not imply interval precision between adjacent ordinal levels.

Where the internal rubric uses non-uniform rung values (for example, capped maxima or calibrated spacing), the public display index reports the mapped result while retaining the underlying ordinal meaning.

6.4 Why minimum scores are conservative

In many global quality scoring approaches, the lowest categories are harsher, often reflecting strict compliance thresholds rather than lived-function thresholds. Neural City intentionally uses conservative minimum scoring so that the system remains stable during early-stage AI training and cross-city generalization, and so that extremely poor outcomes are not over-assigned due to ambiguity in imagery[6].

This does not mean the system underplays deficiencies. It means the system prefers "unknown" over "overconfidently poor" when evidence is partial.

Where a "0" exists in a ladder, it is reserved for clear functional failure (for example, missing, unusable, or hazardous states), not for mild degradation. For all other cases, the system prefers “not observed” over forcing a low score when evidence is incomplete.

6.5 How to read a score

Scores should be read as “what the image shows”, not as a blanket judgment of the whole neighborhood.

Example 1: Drain condition vs median cleanliness
A frame shows a clean, painted median and swept road edge, but the drain has a few covers open and misaligned.

- Drain: scores at the failure or lowest band (because the drain is not functioning)
- Median cleanliness: scores high (because that element is visibly maintained)
  How to read it: a clean-looking street can still have a critical service failure. Cleanliness cues do not cancel a clogged drain.

Example 2: Sidewalk continuity vs decorative lighting
A frame shows decorative lighting and street art, but the sidewalk ends abruptly into a wall with no ramp, forcing pedestrians onto the carriageway.

- Sidewalk continuity: low band (because walking cannot continue safely)
- Decorative lighting: high or present (because it exists)
  How to read it: enhancement features may be positive, but they do not offset a core walkability break.

Example 3: Not observed is not a low score
A frame shows a crowded market street. The sidewalk edge is fully occluded by people and parked scooters, so curb presence cannot be verified.

- Sidewalk curbs: left blank (not observed), not marked as low
  How to read it: blank means “insufficient evidence in this frame”, not “bad condition”.

7. Scoring workflow

7.1 Ingestion and preparation

Images are ingested with metadata when available (time, approximate location, capture source). The system is designed to operate even when metadata is incomplete, but richer metadata improves downstream analysis.

7.2 Annotation and scoring steps

For each image:

1. Determine which sub-parameters are visible enough to score
2. Assign rubric level only where evidence is clear
3. Leave blanks for not-visible or uncertain elements
4. Flag ambiguous images for secondary review

7.3 Audit trail and traceability

Neural City maintains traceability from:

- Aggregate scores
- Back to sub-parameter scores
- Back to image-level evidence

This principle is essential for governance adoption because it enables challenge-resolution. It also supports iterative rubric upgrades without losing interpretability.

8. Aggregation Philosophy

Neural City aggregates from image-level observations upward while minimizing distortion from missing data.

Neural City aggregates image-level observations into parameter-level scores using a predefined internal model. The public methodology follows three principles:

- **Essentials-first weighting:** core usability and service readiness conditions have greater influence than comfort or enhancement signals.
- **Non-compensatory logic for failures:** where an essential sub-parameter indicates functional failure, cosmetic or enhancement positives do not offset that failure in the aggregate.
- **Diagnostic interpretability:** aggregation preserves parameter-level interpretability to support diagnosis rather than collapsing diverse issues into a single opaque index.

Example 1: Walkability
- *Aesthetic positives: decorative lighting present, some street art, ornamental plants maintained*
- *Core failures: sidewalk missing, sidewalk blocked, no safe crossing*

8.1 Essential vs non-essential signals

Sub-parameters are designed with an "essentials-first" philosophy. Core usability conditions such as continuity, obstruction, and functional access carry greater influence than comfort features. Comfort elements (for example, certain street furniture cues) are treated as enhancers and are intentionally capped so they cannot mask core failures.

8.2 Sub-parameter aggregation

Where multiple images cover the same micro-area or corridor, sub-parameter scores can be aggregated to produce stable estimates. Missing values are treated as "not observed" rather than "poor", consistent with the conservative evidence rule.

8.3 Parameter aggregation

Parameter scores are computed from constituent sub-parameters:

- Parameters roll up from sub-parameters, not from opinions
- Binary indicators are treated distinctly to avoid inflating or suppressing continuous quality domains
- Aggregation is designed to preserve signal where conditions differ sharply within a city, rather than hiding it through citywide averaging

8.4 Interpretation boundary

Scores are best used for:

- Relative comparisons across corridors and zones
- Identifying priority gaps where conditions fall into the bottom bands
- Monitoring improvement across repeated collection cycles

Scores should not be used as sole proof of compliance with engineering standards.

9. Model training intent and performance governance

9.1 Training intent

The rubric and the reference images are structured specifically to enable ML training. The method balances domain meaning with visual separability so that early models learn robust boundaries.

9.2 Performance variation and human review

Visual ambiguity varies strongly across sub-parameters. Some categories are visually crisp and train well. Others are affected by camera angle, occlusion, lighting, and non-standard designs, and require additional data or human confirmation.

Neural City therefore uses a tiered governance approach:

- High-confidence sub-parameters can be increasingly automated
- Low-confidence or high-ambiguity sub-parameters remain human-reviewed

9.3 Validation and calibration

We perform periodic scorer calibration and targeted re-audits to reduce drift over time. For model-supported sub-parameters, validation includes held-out samples from multiple cities and error review focused on visually ambiguous conditions (occlusion, nighttime lighting, dense crowds, and non-standard design cues). Sub-parameters with persistent ambiguity remain human-reviewed until dataset coverage and model stability improve.

10. Known limitations and planned integrations

Neural City distinguishes between two different street realities that require different scoring treatment.

1. Functional failure: the asset exists but does not perform its intended function (for example, a drain that is fully clogged, a toilet that is unusable, or a pedestrian path that is effectively missing).
2. Quality degradation: the asset performs partially, but with reduced usability (for example, partial blockage, partial obstructions, or inconsistent continuity).This distinction is used to clearly signal system breakdowns rather than only documenting cosmetic conditions.

10.1 Images are a partial window

Images do not capture everything. They can miss:

- Underground conditions
- Functional performance not visible in a still frame
- Rule context (for example, whether parking is permitted)

10.2 Rule layers and georeferenced permissions

Some observed "negative" signals may be legally permitted in specific contexts (for example, marked parking bays). Neural City does not yet integrate official georeferenced permissions such as parking zones or designated vending zones, which means the current scoring errs on the side of street experience rather than legal intent.

Planned integration approach:
For cities with digital rule layers, Neural City can condition interpretations of select sub-parameters on these layers, improving fairness and actionability without weakening the visual evidence core.

10.3 From scoring to a city intelligence stack

Scoring is the first layer. The longer-term system vision is to integrate street-condition scores with complementary outcome layers such as:

- Road safety outcomes

- Transit usage patterns
- Land use and activity density
- Vendor zones and encroachment governance
- Green cover, air quality, and surface temperature
- Waste generation and service response cycles

This is aligned with the broader direction of outcome and data-driven governance frameworks.

12. From Measurement to Urban Outcomes

Neural City's street-level data model is designed as a foundational measurement layer for outcome-oriented urban governance. By translating established outcome frameworks into observable public-realm conditions, it enables cities to move from abstract indicators to evidence that can be seen, audited, and acted upon at the scale where citizens actually experience the city.

The immediate objective of this methodology is diagnostic clarity: to help cities identify where essential services and public infrastructure are functionally failing, where quality is degrading, and where conditions are improving over time. In doing so, the system supports more targeted prioritisation, more grounded discussions between departments, and more realistic assessments of on-ground impact.

At the same time, this methodology should be understood as an evolving system, not a finished or absolute representation of urban quality. Visual evidence has inherent limitations, and street conditions are shaped by local design norms, informal practices, and contextual trade-offs. Neural City therefore adopts a conservative evidence rule, prioritises relative comparison and trend analysis over absolute judgments, and treats "not observed" as a valid state rather than forcing conclusions.

The longer-term goal is maturity, not perfection. As coverage improves, datasets expand across cities, and additional data layers are integrated, this street-level evidence can progressively be linked with:

- formal infrastructure inventories and rule layers
- land use and right-of-way allocations
- transit, safety, and environmental outcomes
- service response and maintenance cycles

In this sense, street-level scoring is the first step toward a more comprehensive city intelligence stack, not an endpoint.

Neural City is intentionally open to critique, feedback, and iteration. Scoring rubrics, indicator definitions, and aggregation approaches are expected to improve as cities, practitioners, and researchers engage with the data and highlight edge cases, blind spots, and opportunities for refinement. This openness is essential if outcome-oriented measurement is to remain credible, trusted, and useful over time.

Ultimately, the value of this methodology lies not in the score itself, but in what it enables: clearer conversations about trade-offs, earlier detection of systemic failures, and more informed choices about how limited urban space and resources are allocated.

Street-level evidence, when used responsibly, can help cities move closer to the outcomes they seek, such as safer movement, cleaner environments, and more inclusive public spaces, while remaining grounded in the realities of how cities actually function.

Sources

1. Urban Outcomes Framework Part 1, MoHUA / Smart Cities Mission
2. Urban Outcomes Framework Part 2
3. ISO 37120: Indicators for city services and quality of life
4. NIUA Liveability Standards
5. EIU Global Liveability Index Summary Report
6. Przybyłowski et al., "Quest for a Tool Measuring Urban Quality of Life: ISO 3712
7. NIUA ClimateSmart Cities Assessment Framework
8. EIU-The-Global-Liveability-Index

Frequently Asked Questions

1. Is the Street Evidence Model a compliance or certification score?

No. The Street Evidence Model is not a legal, engineering, or statutory compliance assessment. It is an outcome-proxy measurement based on visible street-level evidence, intended to surface relative conditions, functional failures, and trends. Engineering audits, dimensional checks, and statutory certifications remain outside its scope.

2. Does a low score mean the city is performing poorly overall?

Not necessarily. Scores reflect the condition of observed locations and specific parameters, not a blanket judgment of an entire area or city. The model is designed to highlight spatial variation and hotspots, and to support comparison over time or across similar contexts, rather than to label places as "good" or "bad."

3. Can scores be influenced or "gamed" by cosmetic improvements?

The methodology is designed to limit this risk. Essential conditions and functional failures have greater influence than cosmetic enhancements, and failures cannot be fully offset by superficial improvements. Because scoring is based on distributed, repeat observations rather than a single inspection, short-term cosmetic changes have limited impact on overall trends.

4. Why do some sub-parameters use binary scoring while others use multiple quality bands?

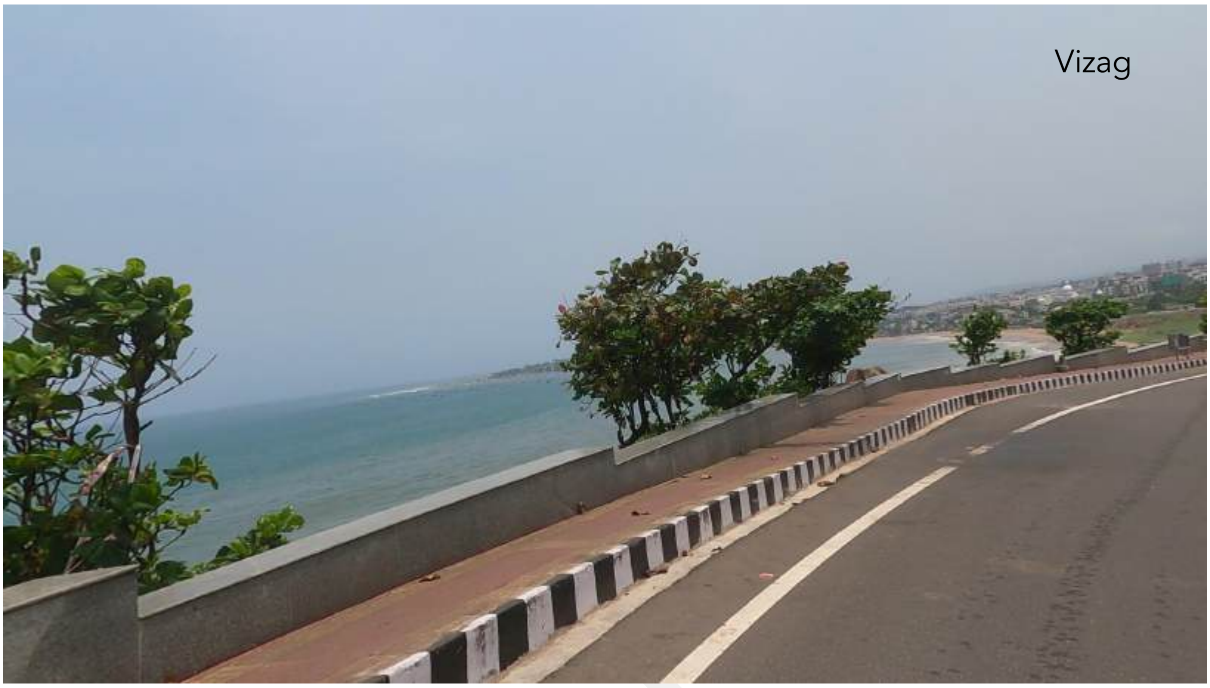
Scoring scales are chosen based on what can be reliably distinguished from street imagery across diverse urban contexts. Some conditions are effectively presence-based in visual data, while others show meaningful gradation. The model avoids artificial precision where visual separation is not dependable.

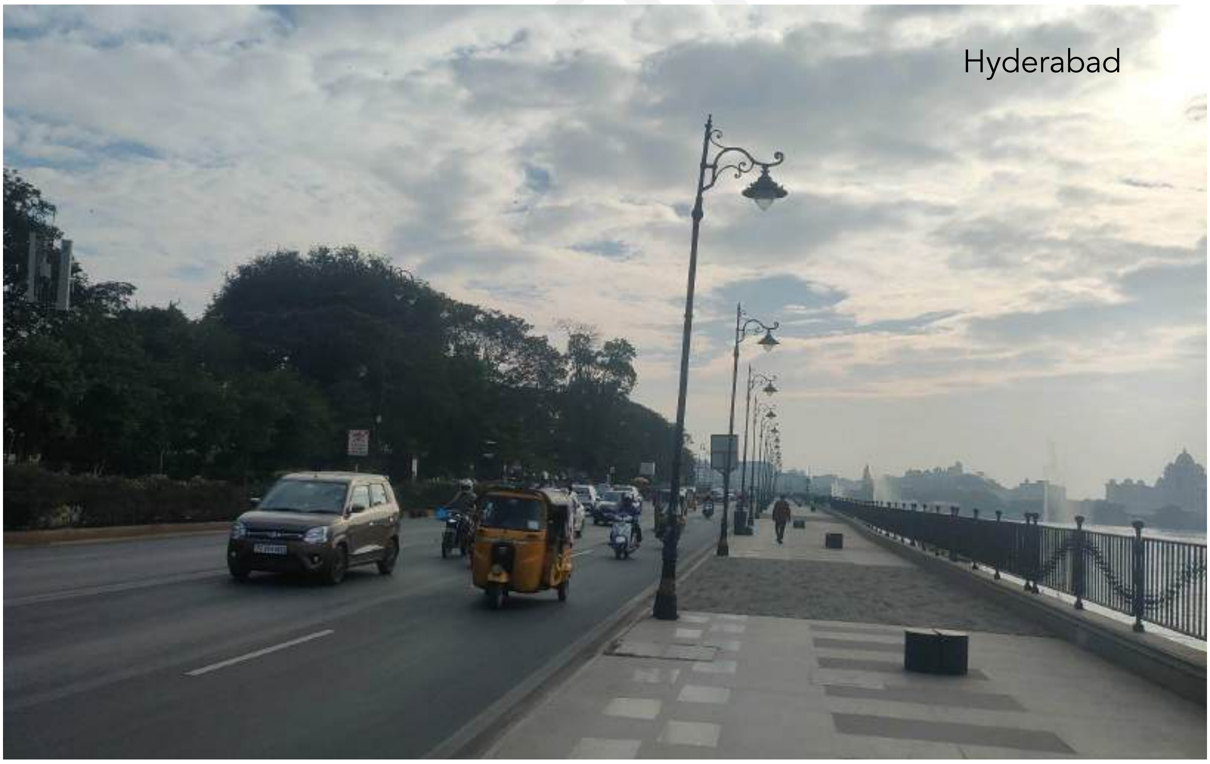
5. How should encroachment or informal vending scores be interpreted?

Encroachment scores are diagnostic, not normative. They differentiate between types and forms of space occupation to reflect how public space is constrained in practice. The intent is to surface spatial pressure and trade-offs, not to prescribe eviction or removal. Policy responses may include redesign, reallocation, or formalisation, depending on local objectives.

6. Will the methodology evolve over time?

Yes. The Street Evidence Model is expected to mature as data coverage expands, additional evidence layers are integrated, and feedback is received from practitioners and cities. Updates are versioned to preserve comparability over time, and changes are guided by the same principles of conservative evidence, interpretability, and outcome relevance.


Vizag


Hyderabad