Landscape Architecture

ARTICLE

Evidence-Deficit Allocation for User-Centred Public-Space Measurement

Kongjian Yu*

College of Architecture and Landscape Architecture, Peking University, Beijing 100871, China

* Correspondence: kjyu@urban.pku.edu.cn

Abstract

Public space analytics has become capable of capturing visual data, text data, behaviour data, environmental data, and administrative data. However, increased opportunities in evidence collection do not necessarily mean balanced evidence on users, since observable actions and simple perception tend to attract significantly more computational power than, for instance, safety, accessibility, climate comfort, universality, and management experience. The objective of this paper is to apply the proposed Evidence-Deficit Allocation approach to a ten-dimensional public space evidence register, featuring 427 dimension assignments and 58 dimension-level machine learning assignments. These ten dimensions include feeling towards place, satisfaction, sensory experience, use and activity, sense of safety, health, climate comfortability, perceived accessibility, universality, and feeling towards management. Evidence-Deficit Allocation involves several components including evidence share, machine learning share, local uptake, positive evidence-to-method deficit, and constraint load, which are then used to calculate a size-sensitive priority score and an under-adoption urgency score. Analysis finds that use and activity and feeling towards place represent 64.17% and 86.21%, respectively, of the total dimension evidence assignments, and machine learning assignments. It can thus be confirmed that there is a pronounced dominance of machine learning in behavioural and affective dimensions. Meanwhile, climate comfortability, universality, and feeling towards management represent 14.29% of dimension evidence assignment, but none of these three has any machine learning assignments. Use and activity achieves a maximum size-sensitive priority score of 100.00, while sense of safety scores 100.00 in urgency. Perceived accessibility, climate comfortability, management perception, and universality are four under-instrumented fields in public space evidence. In a 10,000-run perturbation analysis, the results are found robust against alternative constraint weighting.

Keywords: public space; user-centred assessment; urban analytics; multimodal measurement; machine learning; accessibility; safety; thermal comfort

1. Introduction

Urban public spaces facilitate everyday civic activities. Streets, squares, parks, waterfronts, transit forecourts, playgrounds, and civic plazas enable moving, resting, meeting, recreation, protests, care, commerce, and observing. The quality of urban public spaces cannot be inferred from their physical form alone since the same setting will be experienced differently based on capacity, history, vulnerability, expectations, and entitlement to stay. A shady

Received: 22 August 2024; Revised: 11 December 2024; Accepted: 12 March 2025; Published: 30 June 2025.

doi: 10.70517/la20251152-04

plaza can facilitate comfort but not belonging; an active street can be socially engaging but inaccessible to some users; an immaculate and highly managed square can be legible but restrictive. Public spaces need to be evaluated in terms of use, feeling, accessibility, safety, climate comfort, inclusion, satisfaction, health, sensory experience, and legitimate management.

These dimensions have been discussed in classical urban scholarship, which showed that the value of urban public spaces can be assessed in terms of lived practices rather than geometrical configurations alone. The contributions of Jacobs, Lynch, Whyte, and Gehl indicated that urban public spaces gain in quality through diverse activities, observation, social contact, imageability, environmental legibility, orientation, comfort, and animation [20, 23, 35, 53]. Later work on publicness, sociability, and quality of urban design clarified the role of dimensions like access, control, comfort, animation, ownership, civility, and management, showing that these are interdependent rather than interchangeable aspects of urban public life [9, 11, 37, 50]. Such an approach highlights the importance of critical assessment because visible occupancy should not be treated as sufficient evidence of value. An urban space can be heavily occupied for several reasons: attraction, unavoidable presence, commercial programme, dependency on transit connections, and lack of other possibilities for the users; hence, interpretation is required in order to make activity counts useful for design or policy decisions.

Since the empirical basis of public space evaluation has expanded significantly, a wide range of methods can be employed to evaluate such spaces. Street-view imagery and videos provide an easy-to-scale basis for observation of vegetation, enclosure, sky, benches, sidewalks, surface condition, pedestrian pathways, and visible activity. Online reviews, participatory comments, complaint logs, and texts generated during user interactions with a digital platform are useful for grouping textual content by topic and assessing positive or negative tone. Relational machine learning is used for detecting nonlinear associations between spatial, behavioural, textual, and environmental data streams, and perceived safety, preference, satisfaction, comfort, and use intensity [3, 5, 21, 22, 25, 52]. Computational tools thus help to increase measurement capacity, but they also lead to a new agenda prioritising dimensions that can be easily translated into images, labels, counts, and tones of expression.

Urban perception studies illustrate the power and the pitfall of computational evidence in this context. Large image-based and image-rating approaches showed that the city-wide estimates of perceived safety, beauty, and visual quality were possible. This makes previously qualitative assessments more comparable across various locations [48]. However, the existence of a large labelled corpus does not imply that all important public-space dimensions became easy to measure. Labels used in the perception task typically capture what was inferred by survey respondents rather than what was perceived during repeated use in relation to governance, maintenance, heat, disability, social identity, and fear. The measurement issue in question is thus not just the scalability of the technology, but the coverage of construct space. Even a highly scalable evidence system may fail if the largest data streams are badly calibrated with the dimensions that require attention.

The current aim is to specify which user-centred dimensions should get multimodal measurement priority taking into account the size of evidence base, the current uptake of machine learning, the evidence-to-method gap, and the total weight of constraints associated with each dimension. The scope of this investigation does not include a broad overview of computational urban analytics. Neither does the analysis intend to classify public-space studies according to some universal criteria. Instead, a practical goal is set: the analyst can now assess which dimension out of the list needs to be the next target of measurement. This is done by estimating priority and urgency for each dimension based on evidence share, machine learning share, local uptake, evidence-to-method deficit, and total constraint load.

EVIDA accomplishes this objective by transforming the empirical register into two outputs: a size-sensitive score for dimensions, where additional measurement will affect the evidence register significantly, and an urgency indicator for dimensions, where machine-learning uptake is low relative to constraint severity. While the basic calculation structure is transparent and easily auditable, the key contribution to public-space research lies in the full consideration of output values and in the connection of numerical results to specific dimensions.

An empirical register used in the analysis contains 427 evidence assignments and 58 dimension-level machine-learning assignments across ten user-centred dimensions of urban public spaces: feeling towards place, satisfaction, sensory experience, use and activity, sense of safety, health, climate comfortability, perceived accessibility, universality, and feeling towards management.

2. User experience in public-space assessment

Urban public spaces are evaluated through their visible characteristics, including benches, seats, trees, lighting, cleanliness, walking paths, amenities, and greenery. These features matter, but they are meaningful only in relation to actual use. A bench provides a comfortable seat only if it is placed in a desired location, is comfortable to use, accessible to users, and does not make any group avoid the space due to governance or maintenance practices. An accessible path exists not only when it appears on a map but when it can be recognized and walked by users without excessive effort. Lively public spaces provide good opportunities for civic activities, but they also generate crowding, noise, and other negative externalities. Measurement thus needs to link visible features and conditions with users' interpretations of urban environments.

Publicness in particular is difficult to assess because it emerges as a result of interaction between physical openness, legal access, social norms, control, and behaviour [11, 33, 39, 50]. Urban spaces that seem to be physically open can turn out to be exclusionary due to restrictive rules, constant surveillance, commercial programming, hostile design, and maintenance regime. Conversely, urban spaces that look modest in design but allow free use and encourage repeated participation can prove socially open. The problem of evaluating publicness does not involve merely checking whether the space offers any visible features; it requires measuring whether these features are accessible and usable to various people.

Attachment to the place and environmental significance are other dimensions to consider. Humanistic geography and environmental psychology showed that places acquire emotional significance, meaning, and significance through repeated encounters, bodily orientation, and social interaction [14, 41, 46, 49]. This literature matters for computational assessment since positive tone expressed by a user can correspond to a variety of evaluations: beautiful place, convenient location, enjoyable activity, or positive personal experience. Negative tone also means something else: discomfort, fear, restrictions, access problems, or neglecting maintenance. The construct that needs to be estimated thus has to be specified.

Dimensions of accessibility and safety are complex too since both relate to multiple factors. Accessibility includes not only distance and network availability, but also perceived effort, legibility, continuous surface, traffic exposure, disabilities, care obligations, and confidence to walk or stay [1, 8, 18]. Safety involves lighting, visibility, route availability, pedestrian presence, and traffic exposure, but also harassment, policing, prior victimization, fear among women, age and ethnic groups, and disabilities [26, 30, 44]. Universality of usability cannot be addressed with one single design standard because its inclusiveness depends on who uses the space, who avoids the space, who interacts online, and who is prevented from using the space.

Climate comfortability and sensory experience have acquired additional significance in recent years since cities face challenges related to extreme weather, poor air quality, and unequal distribution of shade. The outdoor environment can influence comfort depending on air temperature, radiation exposure, humidity, wind, surface condition, tree cover, clothing worn, activity level, adaptation, and behavioural adjustments [13, 17, 28, 42]. Comfort varies hourly and seasonally, requiring more comprehensive measurements beyond the visible conditions. Sensor-based and user-generated data are needed for the evaluation of sensory experience since it includes sound, smells, visual glare, crowding, perceived calmness, and contributes to restorative benefit and stress reduction [4, 24, 27]. Inclusion of such variables makes the analysis of public spaces multimodal.

The dimension of management perception should also receive attention because governance practices contribute to comfort and safety. At the same time, governance can restrict spontaneity, encourage surveillance, and exclude some groups who are perceived as disorderly [31, 32, 36, 40]. Administration is not necessarily visible in an image, and it should therefore be covered if one uses computational methods to detect user perceptions of public space. Rules, maintenance practices, programming, cleaning, signs, public information, security measures, and complaint response mechanisms should be incorporated into machine learning analysis if necessary.

In the empirical register, user-centred dimensions include feelings about place, satisfaction, sensory experience, use and activity, sense of safety, health, climate comfortability, perceived accessibility, universality, and feelings about management. The first dimension captures feelings about the place, preference, emotions, memory, identity, and evaluations. Satisfaction reflects users' appraisals of the spaces and fulfilment of expectations. Sensory experience involves various sensations including visual, auditory, tactile, olfactory, and other subjective experiences. User actions like moving, staying, attending events, informally participating, and engaging in other activities fall into the category of use and activity. Sense of safety reflects risk perception, exposure, vulnerability, and situational confidence. Dimensions relating health cover issues associated with physical and mental health, restorative benefits, and social well-being. Climate comfortability includes thermal comfort, shade, wind, humidity, air quality, and extreme weather. Perceived accessibility involves issues related to legibility, entry, movement, connection, and effort. Universality refers to usability across groups, abilities, age, and social position. Finally, feelings about management concern perceptions of governance, maintenance, response to users, cleanliness, and restriction practices.

An efficient allocation method needs to respect these distinctions in order to distinguish popularity from publicness, visual attractiveness from comfort, and governance from social inclusion. Simply summing up all dimensions as a generic quality variable would confuse popularity with inclusion. Assigning equal weight to all evidence types used for different dimensions would make no distinction between what is easy and what is challenging to measure. Poor estimation of safety, universality, and management perception would cause damage to user experience since improper proxy might result in invisibilizing the most vulnerable. Poor estimation of climate comfortability could ignore time-specific measurements, relying only on visible conditions. The point of having a dimension-level register is precisely in keeping the differences in mind.

3. Machine-learning evidence in urban analytics

The need for computational measurement of user experience comes from the fact that experience is spread throughout multiple places, time periods, and data types. Visual audit conducted manually provides high interpretative value but is hard to sustain regularly. Perceptions gathered through interviews or surveys provide rich information, but these methods can hardly be used in long-run evaluation of public spaces. Street-level videos, images, texts, participatory ratings, sensors, and administrative data allow to extend evidence streams. A complete public space evidence system will include all evidence types, but not necessarily in the same quantity.

Image classification and object detection are best developed for the purpose of recognizing visual features and objects. Fully convolutional networks, Faster R-CNN, YOLO-family detectors, and segmentation networks including DeepLab allow researchers to classify pixels and recognize objects on street-view imagery and video [12, 29, 45, 47]. Image segmentation algorithms allow for automatic detection of vegetation, road, sidewalk, enclosure, building, water, and other visible elements. Object detection methods can classify pedestrians, vehicles, bicycle, lamps, benches, steps, litter bins, signs, and other visible features. Action detection methods recognize movements like sitting, standing, walking, waiting, gathering, and other activities. The advantages of this technology lie in its scalability. Its disadvantages include limitations regarding the validity of inference. A tree detected on the picture does not prove shade at the moment of use. Crowd seen on the street does not necessarily indicate inclusion.

Natural language processing can help to extend the evidence base by analyzing verbal accounts. Reviews, comments, complaints, social media posts, and other texts can contain mentions of maintenance, noise, crowding, safety, access problems, amenities, events, comfort, and feelings of belonging. Topic modelling, word embeddings, transformers, and sentiment analyses allow for text corpus exploration [6, 15, 38, 51]. Although natural language processing allows to extract more valuable information, this technique involves bias. Written texts tend to reflect opinions of those users who actively express themselves and may represent exaggerated attitudes. People who feel uncomfortable or excluded may not appear in the dataset. Hence, textual information should be combined with observation and other sources.

Relational machine learning is useful in studying non-linear associations. Random forest, gradient boosting, neural networks, support vector machines, clustering, and principal component analysis help to find complicated patterns between various data streams [7, 19, 34]. Perceived safety can depend on lighting, visibility, presence of pedestrians, traffic presence, time of day, enclosure, and gendered experience. Satisfaction can depend on vegetation, benches, cleanliness, management practices, noise, and social atmosphere. Associations are not additive and should be explored using machine learning methods to better understand the underlying relations.

Some dimensions may receive relatively little attention from computational researchers due to technical difficulties of automation. While places can be visually attractive and active, accessibility requires user perception. Safety is partially measured through lighting, visibility, and pedestrians, but fully measured when including identity, social relations, prior victimization, policing, and timing. Temporally specific microclimate data is required for measuring climate comfortability. Universality requires measurement of inclusion and absence of marginalized users. And finally, governance is also not measured by images since it includes administrative data and user perceptions of management. These dimensions can suffer from delays despite their public significance since computational analysis is constrained by technical issues.

Such a pattern calls for the allocation of resources. Simple counting of studies favours popular dimensions. Simple counting of machine learning applications favours dimensions that are relatively easy to automate. Simple deficit of evidence-to-machine learning ratio does not take into account the size, constraints, and urgency of the particular dimension. EVIDA is designed to tackle the allocation issue. It combines evidence share, machine learning share, uptake in the locality, evidence deficit, and total constraints into a two-component score. The method thus keeps the calculation process visible and allows for inspection, modification, and challenge of conclusions.

In case of public spaces, EVIDA becomes particularly important since researchers and local authorities face resource shortages and need to allocate efforts carefully. Cities cannot continuously gather survey data, image observations, sensor data, route measurements, and administrative information about all public spaces. The choice of dimensions and methods needs to be made according to some strategy. Without the use of the method in question, choices may be guided by routine, preferences, or institutional habit. However, EVIDA asks researchers and city officials to consider which particular dimension needs extra attention given its evidence base, current computational uptake, and constraint burden.

4. Evidence register and EVIDA scoring

Empirical register contains all dimensions along with number of evidence assignments n_i and machine learning assignments m_i. Local machine-learning uptake is calculated as a proportion of the number of machine learning assignments divided by the number of all evidence assignments in each dimension, i.e. $u_i = m_i/n_i$. Dimension evidence share is calculated as $e_i = n_i/\sum_i n_i$. Total constraint code includes three types of constraints.

Table 1. Dimension register.

Dimension	n_i	m_i	u_i (%)	e_i (%)	Code
Feeling towards place	135	29	21.48	31.62	R
Satisfaction	37	3	8.11	8.67	G
Sensory experience	18	2	11.11	4.22	R
Use and activity	139	21	15.11	32.55	R+A
Sense of safety	14	1	7.14	3.28	R+G+A
Health	13	1	7.69	3.04	A
Climate comfortability	6	0	0.00	1.41	R
Perceived accessibility	10	1	10.00	2.34	R+G
Universality	23	0	0.00	5.39	G
Feeling towards management	32	0	0.00	7.49	A

The register in Table 1 demonstrates the empirical imbalance that the computation must take into consideration. Usage and activities represent the largest number of evidence assignments, with 139, and feeling toward place ranks second at 135. The two have most of the corpus covered already, and the two also have most of the machine-learning assignments, with 21 and 29 respectively. On the other hand, management perception has 32 evidence assignments and no machine-learning assignment, universality has 23 and none, and climate comfortability has six and none. An evidence-volume-based ranking would not account for these zero-uptake cases; a local uptake-based ranking would not capture the magnitude of the evidence field covered.

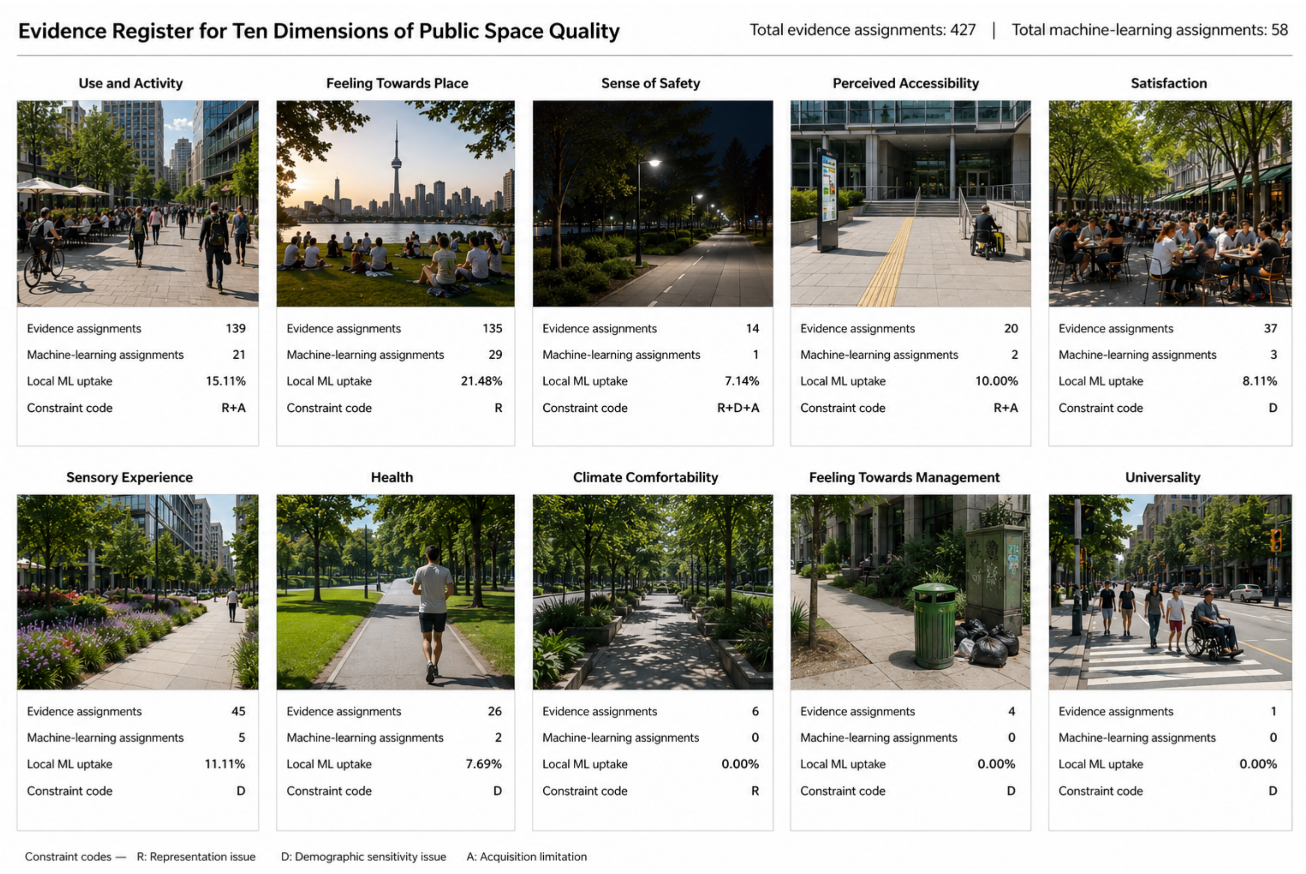


Figure 1. Evidence register.

The visual representation in Figure 1 shows both the numbers and the photographs from the corpus as part of the same visual space. This depiction is functional rather than decorative, since the dominance of the two dimensions can easily be seen even in this form. However, the representation also preserves the rows with no uptake whatsoever, so that one cannot ignore these under-adopted evidence dimensions when the analysis focuses on success stories. This figure highlights the fact that there are no personal records, no image-level annotations, and no user-level response in the dataset.

The information is organized in method groups according to the kinds of evidence they produce, rather than specific algorithms. Visual parsing provides information about compositional and affordance dimensions of a scene. Object and action detection captures human presence and activity in the scene. Affective image analysis provides affective information, but it needs consented validation. Text semantic information identifies recurring themes within texts. Text affect analysis computes evaluative tone. Satisfaction model also computes evaluative tone but does so based on the semantics of a set of experiences. Relational learning links environmental aspects to user-centered outcomes.

The method families in Table 2 clarify the practical meaning of "machine-learning assignment". A machine-learning assignment is not treated as a claim that a dimension has been solved. It only indicates that a computational method has been connected to that dimension. For example, affective image analysis may contribute to feeling towards place, but it cannot by itself validate attachment or belonging. Object detection may describe use and activity, but it cannot determine whether the observed activity is voluntary, comfortable, or inclusive. The table therefore supports a cautious interpretation of the counts: uptake indicates methodological presence, not measurement adequacy.

Table 2. Operational method families.

Family	Algorithmic examples	Measurement role
Visual parsing	SegNet, PSPNet, SegFormer, DeepLabV3+, Canny edge identification	Identifies visible spatial composition, vegetation, enclosure, pedestrian infrastructure, surface condition, disorder proxies, and activity affordances.
Object and action detection	Faster R-CNN, YOLO, VGG-type image classification, action-conditioned models	Detects people, stationary use, movement, sitting, walking, jogging, gathering, waiting, and activity intensity.
Affective image analysis	Face-based emotion services and image emotion classifiers	Provides limited affective signals that require validation against consented and context-aware user feedback.
Text semantics	LDA, Word2Vec, TF-IDF, topic identification	Identifies recurring experiential themes in reviews, captions, complaints, participatory comments, and survey narratives.
Text affect and satisfaction	LSTM-type sentiment models, platform sentiment engines, sentiment classifiers	Estimates positive, negative, and mixed evaluations of comfort, management, safety, accessibility, and place attachment.
Relational learning	Artificial neural networks, random forests, support vector machines, boosted gradient trees	Estimates nonlinear associations between environmental attributes and user-centred outcomes.
Pattern discovery	t-SNE, principal component analysis, k-means clustering	Detects latent spatial, behavioural, and perception clusters without imposing fixed categories.

For each dimension $i \in \{1, \dots, 10\}$, the evidence share is

$$e_i = \frac{n_i}{\sum_{j=1}^{10} n_j}. \tag{1}$$

The machine-learning share is

$$a_i = \frac{m_i}{\sum_{j=1}^{10} m_j}. \tag{2}$$

Local uptake is

$$u_i = \frac{m_i}{n_i}. \tag{3}$$

The evidence-to-method difference is

$$\delta_i = e_i - a_i. \tag{4}$$

A positive δ_i means that the dimension occupies a larger share of the evidence register than of the machine-learning register. A negative δ_i means that machine-learning assignments are concentrated in that dimension beyond its evidence share. The sign of δ_i is interpretive: negative values do not make a dimension unimportant; they show that computational attention is already strong relative to its evidence share.

Each dimension also has a constraint vector $\mathbf{c}_i = (R_i, G_i, A_i)$, where R_i, G_i, and A_i are binary indicators for representation, demographic sensitivity, and acquisition limitations. The constraint load is

$$q_i = 0.4R_i + 0.3G_i + 0.3A_i. \tag{5}$$

Representation receives a slightly higher main-weight because construct interpretation and proxy validity affect both visual and textual measurement. This weighting is transparent rather than hidden; the perturbation analysis later tests whether the main conclusions depend on it.

The positive evidence-to-method deficit is

$$\delta_i^+ = \max(0, \delta_i). \tag{6}$$

The size-sensitive priority score is

$$P_i = e_i(1 - u_i)(1 + \delta_i^+)q_i. \tag{7}$$

 doi: 10.70517/la20251152-04

The under-adoption urgency score is

$$U_i = (1 - u_i)(1 + \delta_i^+)q_i. \tag{8}$$

Both scores are rescaled to a 0–100 range:

$$P_i^* = 100\frac{P_i}{\max_j P_j}, \qquad U_i^* = 100\frac{U_i}{\max_j U_j}. \tag{9}$$

The contrast between P_i^* and U_i^* is crucial. P_i^* incorporates the evidence share, thus prioritizing dimensions whose improved measurement could make a meaningful difference to a significant portion of the register. U_i^* lacks the evidence share component, allowing it to emphasize dimensions that, despite small size, face intense under-adoption and high levels of constraint load. As a pair, the scores ensure that one ranking cannot hide the difference between large measurement spaces and strategically under-instrumented spaces.

The analysis consists of three parts. First, evidence shares and machine-learning shares are contrasted to determine levels of concentration and under-adoption. Second, the priority and urgency equations are calculated with the primary constraint weights. Finally, the robustness check is done via 10,000 randomized constraint weight vectors drawn from the uniform Dirichlet distribution among the three constraint components. For each vector, both scores are recalculated and records kept for top-ranking status and top five classification. The perturbation test will not decide on the proper ethical weighting; rather, it will show whether the allocation conclusions hold up in the event of a different constraint weight distribution.

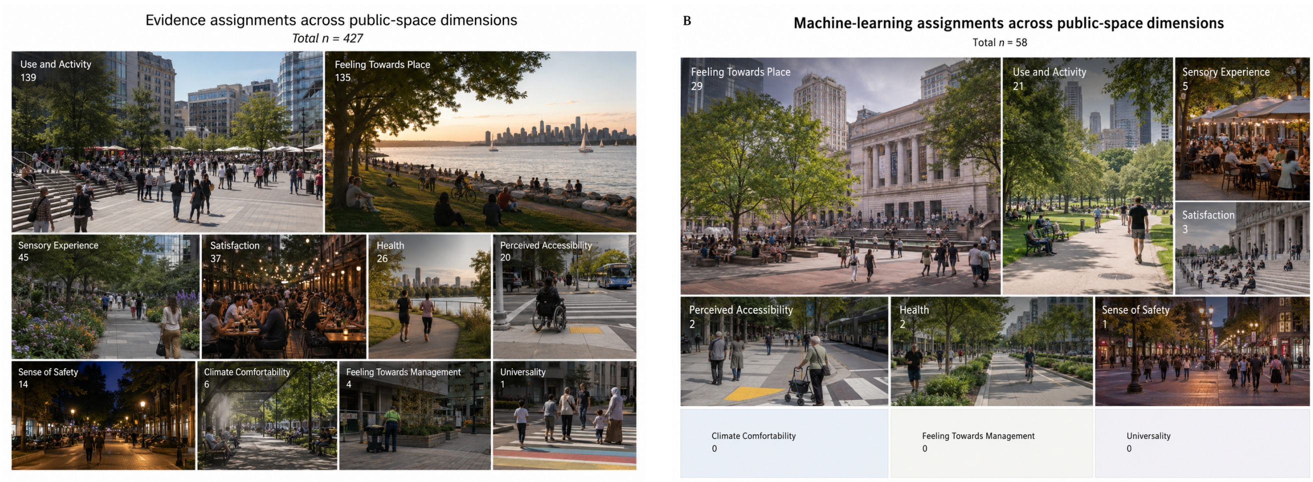
5. Evidence distribution and allocation scores

The evidence register displays high concentration, with two dimensions accounting for the majority of assignment cases. Feeling towards place represents 31.62% of the evidence assignments, and use and activity represents 32.55%. Together, they account for 64.17% of the evidence register. The concentration trend continues in the machine-learning register, where feeling towards place makes up 50.00% of the assignments and use and activity 36.21%. In total, they make up 86.21% of the machine learning register. The difference between 64.17% and 86.21% indicates that the machine learning adoption process intensifies the already high levels of focus on affective and behavioural dimensions.

The evidence and machine-learning concentration visualizations in Figure 2 clearly illustrate this uneven distribution of assignment cases. While the evidence visualization maintains enough breadth to display smaller dimensions, the machine-learning panel visually highlights the prominence of feeling towards place and use and activity. The fields with zero machine-learning assignment cases remain discernible as non-computational blanks, an important detail in understanding the visualization. One might be tempted to interpret the results as a success for technological management; allocation analysis asks us why management perception, climate comfortability, and universality lack local machine-learning representation.

This makes substantive sense as well since feelings towards place, like any other dimension, are consistent with image preferences, reviews, sentiments, and semantics. Similarly, use and activity align well with object recognition, actions, video observation, and counts of people or events. Both, in turn, benefit from abundant and machine-readable streams of evidence. The concern here is that public space analysis could end up being most effective where data is easy, rather than where it is most needed. Issues such as safety, access, thermal comfort, management, and universality demand more purposeful evidence design since their underlying meaning does not translate into an image or sentiment score.

The local adoption rates exhibit the same disparity. The highest machine-learning adoption is seen for feeling towards place (21.48%), followed by use and activity (15.11%), sensory experience (11.11%), perceived accessibility (10.00%), satisfaction (8.11%), health (7.69%), and sense of safety (7.14%). Climate comfortability, universality, and feeling towards management, on the other hand, have zero machine-learning adoption. These percentages reveal the dominance of machine-learning adoption within the two most prevalent rows. Safety and accessibility, key aspects of equitable public space usage, have received merely one assignment each.



(a) Evidence. **(b)** Machine learning.

Figure 2. Evidence and machine-learning concentration.

Evidence-to-method ratio δ_i allows a closer diagnosis of the situation. Feelings towards place has the largest negative value since its machine-learning proportion is nearly twice that of its evidence share (50.00% versus 31.62%). Use and activity is also negatively affected, albeit to a smaller degree, owing to its 36.21% machine learning assignments compared to just 32.55% evidence share. The positive values reveal the dimensions for which evidence share exceeds machine learning share. In order of magnitude, these are feeling towards management (0.075), universality (0.054), satisfaction (0.035), sense of safety (0.016), climate comfortability (0.014), health (0.013), sensory experience (0.008), and perceived accessibility (0.006).

The displacement view in Figure 3, meanwhile, helps distinguish over-concentration from under-adoption without reducing the issue to a simple league table. Feeling towards place lies to the left because of the high number of machine learning assignments, while management perception and universality lie to the right due to their lack of machine learning adoption despite their relatively high evidence shares. The low positive values for accessibility, safety, climate comfortability, and health, meanwhile, deserve attention owing to differences in their constraint load.

The constraint load impacts the interpretation of the difference. The sense of safety dimension is subject to representation difficulties and demographics and acquisition issues; hence its q_i score is 1.00. Use and activity and perception of accessibility are each subject to two issues, hence $q_i = 0.70$. Climate comfortability suffers from representation issue and zero uptake issue. Universality is subject to demographic sensitivity and zero uptake issues. Feeling towards management involves acquisition difficulty and zero uptake issues. It implies that priority is based on a dimension due to inadequacy in measuring it and representing it.

The constraint and scarcity profile illustrated in Figure 4 helps to understand why there is a difference between raw scores and final scores. Low local uptake indicates scarcity, but the constraint load makes the problem more urgent. Climate comfortability, management perception, and universality dimensions have full scarcity owing to $u_i = 0$. The urgency of these dimensions varies according to their scarcity evidence shares and constraint issues. Although the sense of safety dimension has a lower evidence share, it has full scarcity coupled with a high constraint load.

The computed EVIDA outputs are reported in Table 3. Use and activity reaches the normalized maximum for size-sensitive priority, $P_i^* = 100.00$. Its position is not a simple count effect. It combines the largest evidence share, modest local uptake, and representation plus acquisition constraints. The result indicates that activity measurement remains a large unfinished task: public-space analytics can count people, detect movement, and classify actions, but it must still connect those observations to choice, duration, comfort, inclusion, and social meaning.

The same table shows why size-sensitive priority and urgency serve different allocation functions. Feeling towards place ranks second in priority with $P_i^* = 51.33$ because it is a large evidence field, but its urgency score is only 33.31 because machine-learning uptake is already high relative to other dimensions. Sense of safety has only 14 evidence assignments and therefore a lower priority score of 15.98, but it reaches the urgency maximum of 100.00

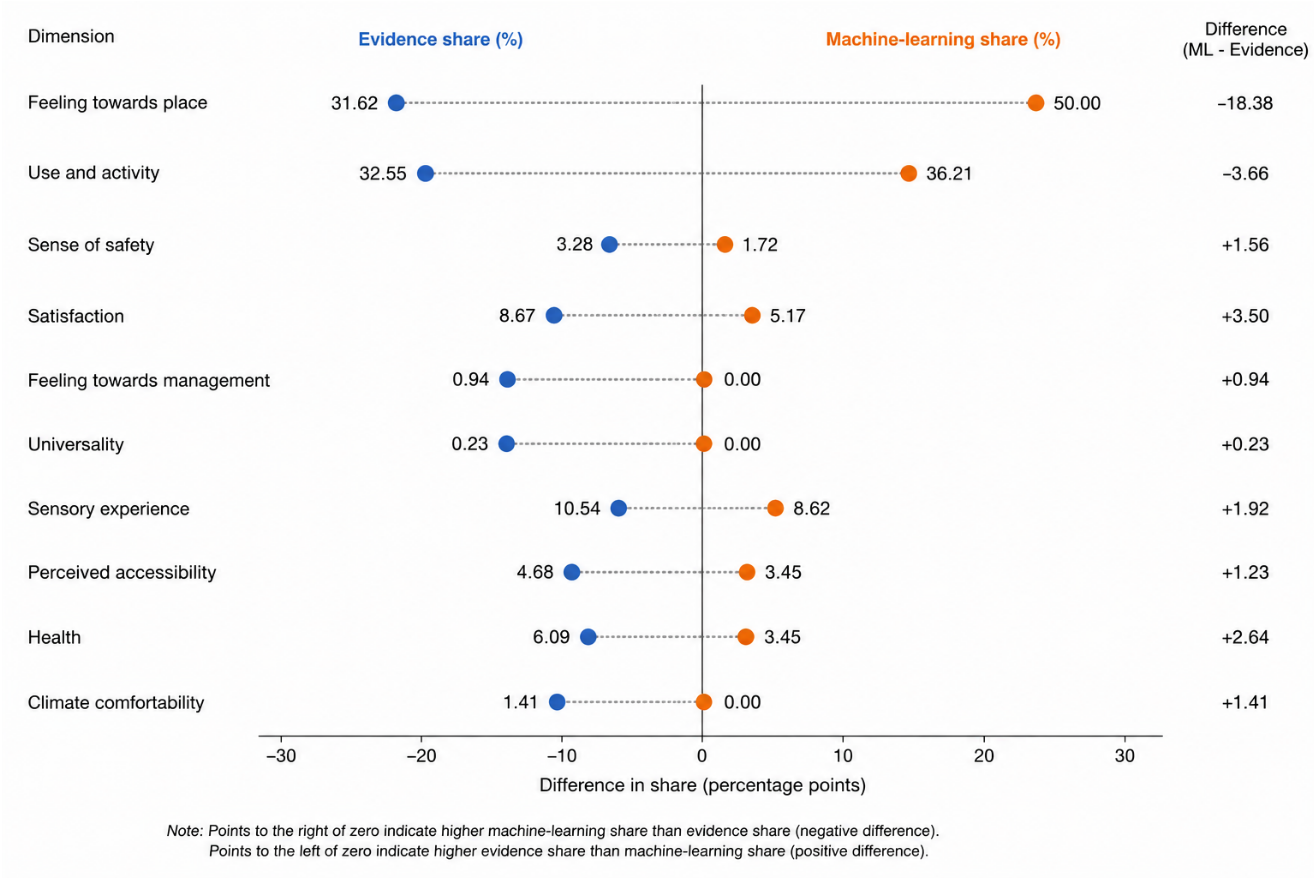


Figure 3. Evidence-to-method displacement.

Table 3. EVIDA scores.

Dimension	δ_i	q_i	$1 - u_i$	P_i^*	U_i^*
Use and activity	-0.037	0.70	0.849	100.00	63.02
Feeling towards place	-0.184	0.40	0.785	51.33	33.31
Sense of safety	0.016	1.00	0.929	15.98	100.00
Satisfaction	0.035	0.30	0.919	12.78	30.25
Feeling towards management	0.075	0.30	1.000	12.49	34.20
Universality	0.054	0.30	1.000	8.80	33.53
Sensory experience	0.008	0.40	0.889	7.81	37.99
Perceived accessibility	0.006	0.70	0.900	7.67	67.22
Health	0.013	0.30	0.923	4.42	29.75
Climate comfortability	0.014	0.40	1.000	2.95	43.01

because uptake is low, scarcity is high, and all three constraint types are present. Perceived accessibility has a small evidence share but a high urgency score of 67.22 because it combines low uptake, scarcity, and representation plus demographic sensitivity.

The score difference in Figure 5 is the essential numerical result here. On the left, it shows the field in which the largest impact of the proposed measurement innovation will take place: use and activity wins hands down, followed closely by feeling towards place. On the right, it shows the dimension where the current under-representation is the most important: sense of safety ranks first, followed by accessibility, and climate comfortability outranks itself in terms of relative importance. The dual output helps to avoid the common mistake of treating either the biggest or the smallest field as self-evidently the most urgent.

Figure 4. Constraint and scarcity profile.

The ranking also suggests the different ways in which the two leaders need to improve their explanatory power. While use and activity should not be left to more automated counting but explained in more depth, feeling towards place requires elaboration of connections between stated feeling, actual condition, frequent use, remembering, and belonging. Such improvements align well with the priority score, since they would enhance the interpretability of a broad spectrum of evidence items.

The interpretation of the two leading dimensions varies. For use and activity, the issue is in the wide range of meanings associated with the observation: voluntary occupancy, unavoidable movement, participation in events, being in waiting lines, going to and from work, absence of alternatives, and constraint. For feeling towards place, the problem is that positivity in texts or images does not necessarily correspond to attachment and ownership. For sense of safety, it is clear that environment does not fully account for fear, confidence, threats of harassment, policing, and time-sensitivity. For perceived accessibility, the challenge is in connecting route distance, legibility, traffic fear, and difficulties related to disabilities.

Climate comfortability scores lowest in the size-sensitive priority ranking but becomes a relatively high scorer in the urgency ranking: 43.01. There is no contradiction. Firstly, there are only six evidence item assignments in this category, hence it cannot possibly win any size-dependent ranking. Secondly, the category has no machine-learning uptake and full representation difficulty, thus becoming an urgent task upon taking into account under-representation. The result is important for cities suffering from heat stress because one's ability to be present in public spaces depends on a variety of factors: shade, radiant exposure, air circulation, humidity, surface characteristics, air pollution, and availability of water. Even if the visual impression conveys the notion of greenery, comfort requires measuring it in practice.

Universality and feeling towards management belong to the same category of urgency dimensions: they have a low-size-sensitive priority score but full scarcity, positive evidence-to-methods difference, and full zero local machine-learning uptake. In addition to technicality, both dimensions also have conceptual issues: they address

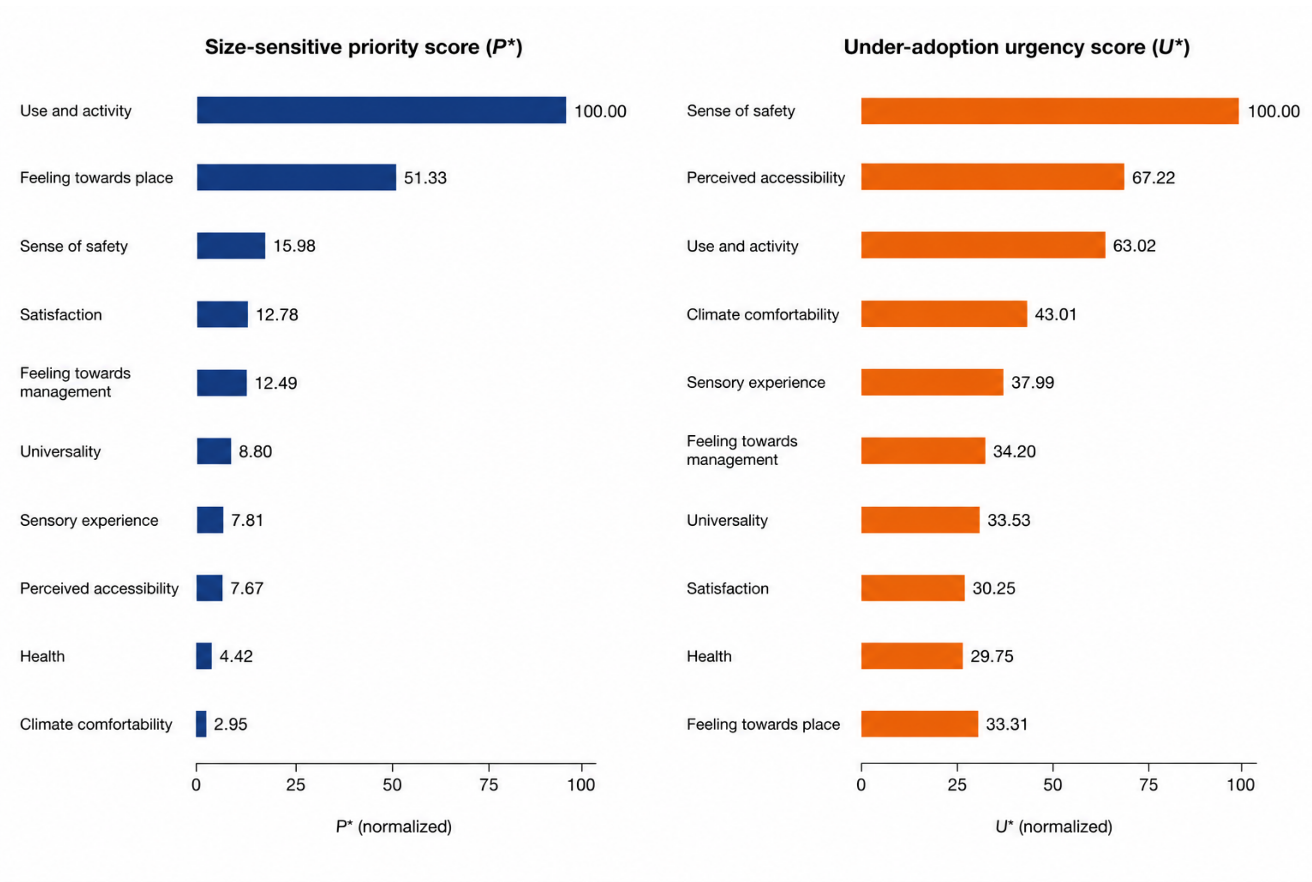


Figure 5. Priority and urgency scores.

questions regarding whose turn it is to use the space and what kind of rulefulness is felt in the place. Evidence related to such dimensions is not readily obtained from the available digital trace and requires active data collection. Thus, they deserve the attention as called for in the urgency score.

Perturbation Analysis Tests the Robustness of Results

The perturbation analysis evaluates whether the results are dependent on the chosen weighting for constraint components. Across 10,000 permutations of constraint weights, use and activity remains the first size-sensitive priority in 94.6% of cases, while satisfaction takes over in 5.5% of cases. More specifically, the latter happens mostly when demographic sensitivity gets higher weights. Sense of safety and use and activity are among the top five in all runs. Feeling towards place appears in the top five in 85.5% of cases, satisfaction in 63.9%, and feeling towards management in 54.9%. These percentages indicate that the priority ranking is stable although not absolute: the leader is robust but the middle ranks are sensitive to the choice of weights.

As for urgency ranking, sense of safety retains its position as number one in 96.9% of runs. Other candidates to replace it are management perception, universality, and climate comfortability, which rise in prominence when acquisition difficulty, demographic sensitivity, or representation difficulty get higher weights. Therefore, the score suggests a balanced approach: safety is the most robust urgency priority while zero-uptake dimensions are not irrelevant either; they become crucial when other criteria receive more weight.

The robustness summary of Figure 6 provides practical meaning for the stability test. The researchers who apply the recommended constraints can conclude that there is nothing accidental about their leaders: use and activity and sense of safety are statistically robust. Still, the probabilities of occurrence of the second and third priority dimensions are significant enough to make the researchers consider other dimensions in a responsible agenda.

The final result is a set of four measurement pathways derived from the existing EVIDA outputs. Activity-perception calibration links action detection and occupation intensity with short prompts, user comments, temporal context, and

 doi: 10.70517/la20251152-04

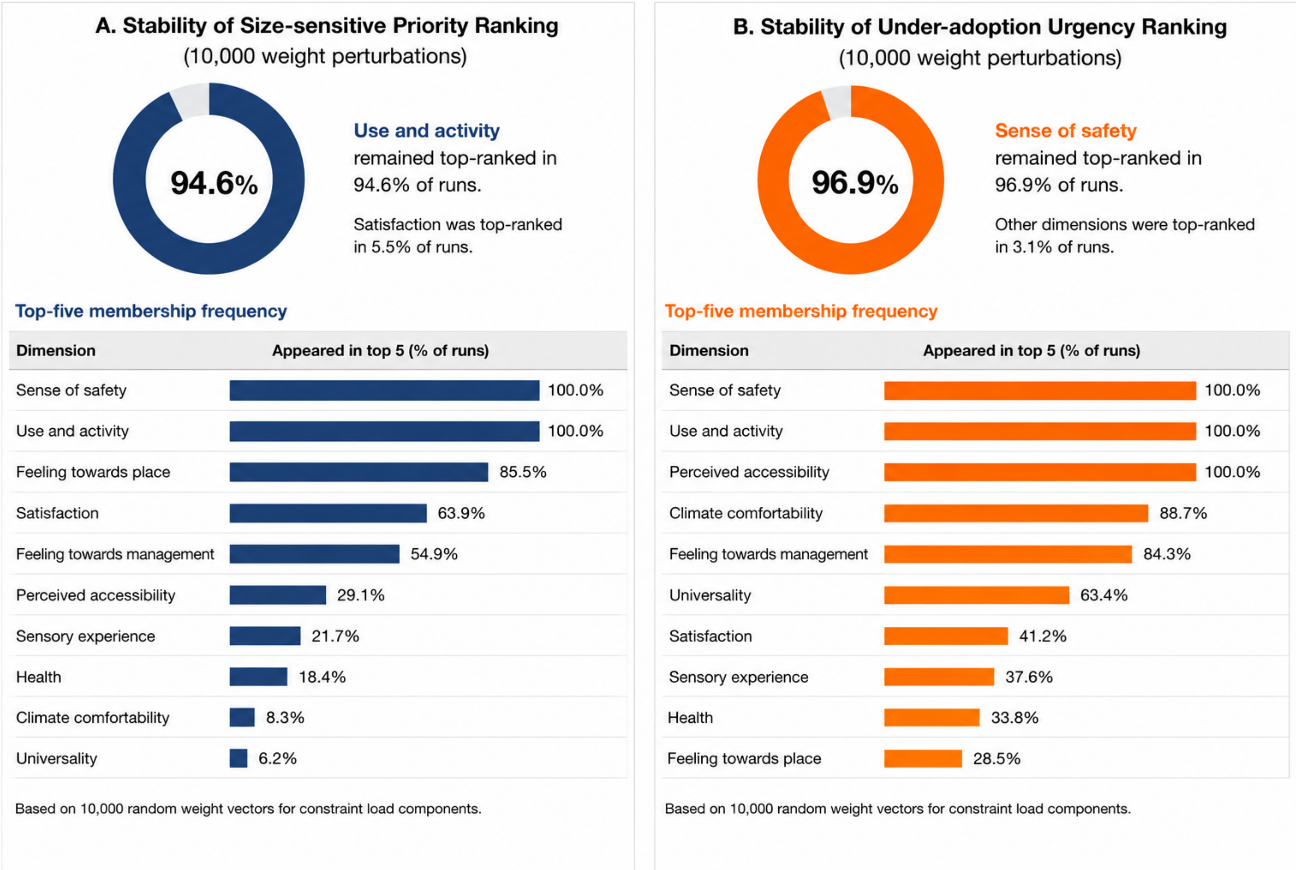


Figure 6. Perturbation stability.

spatial composition. Safety-accessibility triangulation links exposure, lighting, route legibility, traffic conditions, network measures, and consented safety or access feedback. Governance-inclusion observation links complaint topics, maintenance records, rule inventories, event programming, signage, and fairness checks. Thermal-sensory coupling links microclimate sensors, shade and material mapping, air-quality indicators, comfort prompts, and sensory narratives.

Table of pathways converts scores into measurement strategies. Its main implication is the differentiation between high-volume refinement and blind-spot correction. Activity-perception calibration is effective when there is much evidence about use and activity. Safety-accessibility triangulation is needed when instruments do not reflect vulnerability and group-sensitive access. Governance-inclusion observation brings evidence from administration and fairness into analysis of public spaces. Thermal-sensory coupling solves issues of temporally variable and invisibilized conditions. All four pathways preserve the current method and provide guidance on using the existing results empirically.

Another implication is that each pathway implies a specific validity challenge. Activity-perception calibration requires testing for correspondence between detected activity and user satisfaction, choice and belonging. Safety-accessibility triangulation requires testing for correspondence between affordance and confidence in public space. Governance-inclusion observation requires testing for correspondence between administration responses and user experiences, and evenness of rule application and enforcement. Finally, thermal-sensory coupling requires testing for correspondence between exposure and comfort or discomfort. In all four cases, the computational dimension remains a fraction of the evidence base, and proper interpretation requires complementing it with user-centred validation.

 doi: 10.70517/la20251152-04

Table 4. Measurement pathways.

Pathway	Target dimensions	Measurement design
Activity-perception calibration	Use and activity; feeling towards place	Combine action detection from video or street-level imagery with topic and sentiment identification from user comments and short field prompts; test whether activity intensity corresponds to expressed experience after accounting for visual composition and temporal conditions.
Safety-accessibility triangulation	Sense of safety; perceived accessibility	Combine semantic segmentation, route-network measures, lighting, enclosure, traffic exposure, and consented perception prompts; test whether safety and access perceptions diverge from measured affordances across groups and times.
Governance-inclusion observation	Feeling towards management; universality; satisfaction	Combine complaint-topic modelling, maintenance and event records, signage inventories, rule audits, and demographic fairness checks; test whether management practices distribute usability and satisfaction unevenly.
Thermal-sensory coupling	Climate comfortability; sensory experience; health	Combine microclimate sensors, shade and material mapping, air-quality indicators, perceived comfort prompts, and text affect analysis; test whether exposure conditions correspond to comfort, avoidance, or restorative use.

6. Measurement implications for public-space assessment

The study shows that public-space analytics is determined by the fit between constructs and data streams. Use and activity and feeling toward place are dominant because they are easily observed, easily articulated, and easily processed computationally. This does not mean that they are studied to exhaustion; activity and place feeling still need much better interpretation. However, the uneven development of the computational approach to public space implies that it focuses on dimensions with many possible computational applications. On the other hand, some dimensions have only minimal coverage in computational studies – due to lack of machine learning applications or due to reliance on quantitative counts, qualitative judgements, administrative evidence, or field observation.

The aim to find evidence allocation leads to another twist in interpreting the result. If this were a purely descriptive analysis, it would imply that the high counts for use and activity and place feeling prove the primacy of these dimensions in public space evaluation. In the context of the EVIDA calculation, this conclusion is less obvious. Use and activity and place feeling are dominant because they constitute much of the evidence register, but not equally; use and activity need interpretation based on matching observations and users' intentions, and place feeling needs validation through affective and belonging evidence. Hence additional efforts to enhance the understanding of the two dominant dimensions should concentrate on the calibration and validation tasks.

Safety is also unique in its characteristics. As opposed to use and activity and place feeling, it has relatively few evidence samples, but it dominates in the urgency score. This is an inherent characteristic of safety as a public-space dimension. It is impossible to derive a prediction of people's sense of safety based on their visibility, activity, distances, and other objective variables because it includes physical vulnerability, social interaction, prior experience, personal identity, time of day, and trust. The urgency score, in particular, highlights the importance of combining environmental features with voluntary user reports, taking the time of day into account and being group sensitive.

Accessibility is another unique public-space dimension. Like safety, it has only minimal evidence presence but a relatively high urgency score. This outcome is important to interpret properly as it suggests going beyond mere distance in assessing public space accessibility. Route length, proximity to public transport stops, and network centrality matter, of course, but there is a long list of other considerations related to accessibility perception – including wayfinding convenience, surface quality, comfort, traffic risks, disability and gendered safety, age, and care giving responsibilities. Thus, measurement of accessibility is best performed when it combines spatial network

 doi: 10.70517/la20251152-04

and route information with user feedback.

The zero-uptake dimensions are the clearest evidence for avoiding method-led analysis. Climate comfortability, universality, and management perception cannot be discarded just because they have no machine learning assignments in the register. Their absence from the machine learning stream means that the current methodological tool kit is not well developed for thermal, inclusive, and governance aspects of the public spaces. Climate comfortability requires temperature sensors and periodic prompting of users; universality requires measuring absence, barriers, and usage by different groups of people; management perception requires analyzing administrative evidence and user trust.

Figure 7. Shared field setting.

The field setting in Figure 7 helps explain why multimodal evidence is necessary. A single plaza or park edge can contain visible activity, seating, planting, lighting, signage, surfaces, management traces, informal conversation, route choices, and environmental exposure. None of these elements alone can represent public-space quality. Visual observation can record the presence of people and amenities, but it cannot determine whether users feel welcome or safe. Text can record evaluation, but it may exclude non-posting users. Sensors can record heat or air quality, but they cannot determine how exposure is experienced. The figure therefore supports the paper's main allocation logic: the next measurement step should be selected according to the dimension that needs evidence, not according to the data stream that happens to be easiest to collect.

The protocols depicted in the panels of Figure 8 illustrate how the numerical outcomes are put into practice. The activity-perception panel exemplifies the need for comparing observed occupancy with reported experience. The safety-accessibility panel exemplifies route-specific triangulation instead of coarser access mapping. The governance-inclusion panel includes evidence on institutional practice by means of records, signs, rules, and complaints. The thermal-sensory panel reflects the fact that environmental factors must be assessed together with the perceptual experience, since temperature, humidity, shade, and air quality influence usage through physiological exposure and adaptive cognition.

The literature connection becomes much stronger when we understand why EVIDA should be interpreted as a tool

 doi: 10.70517/la20251152-04

(a) Activity and perception.

(b) Safety and access.

(c) Governance and inclusion.

(d) Thermal and sensory evidence.

Figure 8. Field protocols.

for resource allocation rather than a quality index applicable to all urban settings. Public spaces literature highlights several value criteria, including sociability, accessibility, freedom, comfort, control, remembrance, environmental quality, and civic legitimacy [10, 11, 36, 37, 43]. The EVIDA scores do not replace these values. Instead, they suggest in which of these aspects the current evidence base is insufficient. This point is essential because a numeric score might be wrongly used as an ultimate evaluation of place quality. In this paper, scores only reveal the gaps in measurement capacity, but they do not provide a judgment on whether any specific park, street, or plaza is good or bad.

In this respect, the proposed approach facilitates responsible use of machine learning. Data biases due to platform or sensor design can perpetuate urban selection, proxy, and spatial inequalities [2, 16, 25]. EVIDA does not solve these issues. However, they become evident with constraint loads, which identify problematic constructs that cannot be proxied with algorithms, groups with differential interpretations of certain dimensions, and dimensions with poor availability, continuity, or accessibility of data. This coding encourages asking whether each unmeasured dimension is left aside for some methodological reason or because it is actually less important.

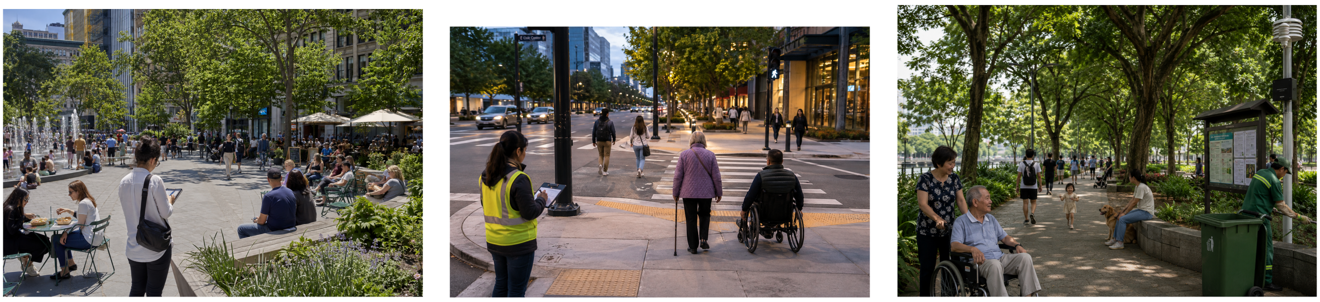
The practical application will vary across urban environments. For example, a city with abundant street view coverage and little user feedback might expect high potential for visual parsing and low capacity for detecting safety, access, and management perception. Another city, equipped with a complaint system, would possess rich governance texts, but would lack field evidence. Climate-related cities might have to allocate more weight to the climate comfortability dimension, whereas in thermal comfort is a less pressing concern. The EVIDA formula allows accommodating these city needs because the weighting parameters are transparent and flexible.

In the realm of design practice, the findings warn against considering images as a measure of public space quality. Rendering images, photographs, and even street views can communicate greenness, activity, and visible amenities; however, they fail to demonstrate whether these places are exclusive, dangerous, heat-exposed, poorly maintained, and poorly connected with the rest of the city. The priority score points at areas in which evidence fields need further development in terms of activity and place feeling dimensions. The urgency score points at under-instrumented dimensions that require focused assessment, such as safety and access perception. Together, they lead towards a measurement sequence in which wide-ranging visual assessment is followed by deep field confirmation.

The findings suggest a multi-step monitoring procedure as well. Visual or textual screening can detect candidates for assessment, but the urgency score will show in which of those places safety, accessibility, thermal conditions, and governance issues need to be examined further. This sequential process will allow reducing the measurement burden without limiting the assessment to the most readily observable dimensions. Such an approach is especially relevant for public authorities with limited budgets and personnel. Evidence-based allocation can provide justifiable reasons for paying more attention to certain dimensions than to others.

The calculated value also provides an opportunity to discuss negative space in academic literature. Absences usually cannot be reported because their absence prevents building an estimation model, computing a coefficient, or plotting a visualization. EVIDA translates absences into numbers by tying them to evidence shares, constraints, local uptake, and scarcity. This is why climate comfortability, universality, and management perception are still meaningful, despite having zero machine learning assignments locally. Their absence is part of the evidence rather than being hidden behind the successes of more measurable dimensions.

The above interpretation has several limitations. First, the evidence register operates at the level of dimensions instead of papers or articles; therefore, it does not evaluate their quality, geographical focus, sampling, choice of venues, or context of implementation in specific cities. Second, the constraint codes are concise. More complex registers can identify different categories of representation problems, demographically sensitive dimensions, and acquisition challenges. Third, the weights assigned to the main constraints are normative choices. Perturbation analysis reveals their robustness but cannot settle ethical questions regarding correct weight assignments per city or research project. Fourth, the visual evidence in Figs. 9 is interpretive in nature because it presents aggregates. Fifth, EVIDA only indicates the dimensions that require further measurement and field verification without validating any specific data collection methodology.



(a) High-volume refinement. **(b)** Blind-spot correction. **(c)** Zero-uptake expansion.

Figure 9. Monitoring portfolio.

The portfolio of monitoring activities in Figure 9 reflects the action logic in its current form without altering the computation. High volume refinement covers the dimensions of large domains in which the existing machine learning coverage requires interpretive calibration. Blind spot correction covers safety and accessibility dimensions in which low uptake accompanies a high level of constraint load. Zero uptake expansion covers climate comfortability, universality, and management perception dimensions in which the absence of local machine-learning assignments implies the deliberate production of additional evidence.

7. Conclusions

This allocation problem finds the user-centred dimensions of public spaces whose multimodal evidence needs to be collected based on the combined evaluation of evidence volume, machine-learning uptake, evidence-to-method deficit, and constraint load. There are two answers. In terms of refinement, the effort should cover use and activity, followed by feeling towards place, because these dimensions cover 64.17% of all evidence and represent 86.21% of all machine-learning assignments and thus require careful calibration of interpretation beyond visible activity and textual positive assessments. In terms of blind-spot correction, effort should focus on sense of safety and perceived accessibility, with special emphasis placed on climate comfortability, universality, and management perception, because these dimensions cover 14.29% of evidence while showing low or zero machine-learning uptake and constraint-loaded representation.

This numerical result is confirmed immediately. Use and activity and feeling towards place represent 64.17% of the evidence register and 86.21% of all machine learning assignments. Climate comfortability, universality, and feeling towards management cover 14.29% of the evidence register but do not contain any machine learning assignments locally. Use and activity has the highest size-sensitive priority score, and sense of safety has the highest urgency score. Perturbation analysis verifies this result; use and activity maintains its top rank for priority in 94.6% of simulations, and sense of safety maintains its top rank for urgency in 96.9% of simulations.

The substantive conclusion drawn from this result is that public space analytics must allocate measurement efforts through dimension need instead of data availability. Machine learning work has proven effective at measuring dimensions of use and activity, as well as positive place feeling through image and text processing. However, other dimensions are harder to measure, requiring multimodal data collection. This means that safety, accessibility, thermal comfort, universality, and management perception each require the pairing of visual, textual, sensor, administrative, and user-governed evidence. EVIDA shows clearly how these priorities can be identified without losing the multiplicity of dimensions relevant to public-space quality.

References

[1] Alfonzo, M. A. (2005). To walk or not to walk? The hierarchy of walking needs. Environment and behavior, 37(6), 808-836.

[2] Barocas, S., Hardt, M., & Narayanan, A. (2023). Fairness and machine learning: Limitations and opportunities. MIT press.

[3] Batty, M. (2013). Big data, smart cities and city planning. Dialogues in human geography, 3(3), 274-279.

[4] Berman, M. G., Jonides, J., & Kaplan, S. (2008). The cognitive benefits of interacting with nature. Psychological science, 19(12), 1207-1212.

[5] Biljecki, F., & Ito, K. (2021). Street view imagery in urban analytics and GIS: A review. Landscape and urban planning, 215, 104217.

[6] Blei, D. M., Ng, A. Y., & Jordan, M. I. (2003). Latent dirichlet allocation. Journal of machine Learning research, 3(Jan), 993-1022.

[7] Breiman, L. (2001). Random forests. Machine learning, 45(1), 5-32.

[8] Brownson, R. C., Hoehner, C. M., Day, K., Forsyth, A., & Sallis, J. F. (2009). Measuring the built environment for physical activity: state of the science. American journal of preventive medicine, 36(4), S99-S123.

[9] Carmona, M. (2010). Contemporary public space: Critique and classification, part one: Critique. Journal of urban design, 15(1), 123-148.

[10] Carmona, M. (2019). Principles for public space design, planning to do better. Urban design international, 24(1), 47-59.

[11] Carr, S. (1992). Public space. Cambridge University Press.

[12] Chen, L. C., Zhu, Y., Papandreou, G., Schroff, F., & Adam, H. (2018). Encoder-decoder with atrous separable convolution for semantic image segmentation. In Proceedings of the European conference on computer vision (ECCV) (pp. 801-818).

[13] Chen, L., & Ng, E. (2012). Outdoor thermal comfort and outdoor activities: A review of research in the past decade. Cities, 29(2), 118-125.

[14] Cresswell, T. (2004). *Place: A Short Introduction*. Blackwell.

[15] Devlin, J., Chang, M. W., Lee, K., & Toutanova, K. (2019, June). Bert: Pre-training of deep bidirectional transformers for language understanding. In Proceedings of the 2019 conference of the North American chapter of the association for computational linguistics: human language technologies, volume 1 (long and short papers) (pp. 4171-4186).

[16] d'Ignazio, C., & Klein, L. F. (2016, October). Feminist data visualization. In Workshop on Visualization for the Digital Humanities (VIS4DH), Baltimore. IEEE (Vol. 2, No. 10).

[17] Erell, E., Pearlmutter, D., & Williamson, T. (2012). Urban microclimate: designing the spaces between buildings. Routledge.

[18] Ewing, R., & Handy, S. (2009). Measuring the unmeasurable: Urban design qualities related to walkability. Journal of Urban design, 14(1), 65-84.

[19] Friedman, J. H. (2001). Greedy function approximation: a gradient boosting machine. Annals of statistics, 1189-1232.

[20] Gehl, J. (2012). Life between buildings: using public space. Princeton University Press.

[21] Goodchild, M. F. (2007). Citizens as sensors: the world of volunteered geography. GeoJournal, 69(4), 211-221.

[22] Ibrahim, M. R., Haworth, J., & Cheng, T. (2020). Understanding cities with machine eyes: A review of deep computer vision in urban analytics. Cities, 96, 102481.

[23] Jacobs, J. (1961). The Death and Life of Great American Cities. vintage book, a division of random house. New York, 178.

[24] Kaplan, R., & Kaplan, S. (1989). The experience of nature: A psychological perspective. Cambridge university press.

[25] Kitchin, R. (2014). The real-time city? Big data and smart urbanism. GeoJournal, 79(1), 1-14.

[26] Koskela, H. (1997). 'Bold walk and breakings': Women's spatial confidence versus fear of violence. Gender, Place and Culture: A Journal of Feminist Geography, 4(3), 301-320.

[27] Kuo, F. E. (2001). Coping with poverty: Impacts of environment and attention in the inner city. Environment and behavior, 33(1), 5-34.

[28] Lin, T. P. (2009). Thermal perception, adaptation and attendance in a public square in hot and humid regions. Building and environment, 44(10), 2017-2026.

[29] Long, J., Shelhamer, E., & Darrell, T. (2015). Fully convolutional networks for semantic segmentation. In Proceedings of the IEEE conference on computer vision and pattern recognition (pp. 3431-3440).

[30] Loukaitou-Sideris, A. (2006). Is it safe to walk? 1 neighborhood safety and security considerations and their effects on walking. Journal of Planning Literature, 20(3), 219-232.

[31] Flinton, P. (2009). Sidewalks: Conflict and Negotiation over Public Space. Electronic Green Journal, (29), 1.

[32] Low, S. M. (1997). Urban public spaces as representations of culture: The plaza in Costa Rica. Environment and behavior, 29(1), 3-33.

[33] McQuire, S. (2006). The politics of public space in the media city. First Monday.

[34] Lundberg, S. M., & Lee, S. I. (2017). A unified approach to interpreting model predictions. Advances in neural information processing systems, 30.

[35] Lynch, K. (1960). The Image of the. City, 19, 50.

[36] Madanipour, A. (2003). Public and private spaces of the city. Routledge.

[37] Mehta, V. (2014). Evaluating public space. *Journal of Urban Design*, *19*(1), 53–88.

[38] Mikolov, T., Chen, K., Corrado, G., & Dean, J. (2013). Efficient estimation of word representations in vector space. *arXiv preprint arXiv:1301.3781*.

[39] Mitchell, D. (1995). The end of public space? People's Park, definitions of the public, and democracy. *Annals of the Association of American Geographers*, *85*(1), 108–133.

[40] Mitchell, D. (2003). *The Right to the City: Social Justice and the Fight for Public Space*. Guilford Press.

[41] Nasar, J. L. (1998). *The Evaluative Image of the City*. SAGE Publications.

[42] Nikolopoulou, M., & Steemers, K. (2003). Thermal comfort and psychological adaptation as a guide for designing urban spaces. *Energy and Buildings*, *35*(1), 95–101.

[43] Oldenburg, R. (1999). *The Great Good Place*. Marlowe & Company.

[44] Pain, R. (2000). Place, social relations and the fear of crime: A review. *Progress in Human Geography*, *24*(3), 365–387.

[45] Redmon, J., Divvala, S., Girshick, R., & Farhadi, A. (2016). You only look once: Unified, real-time object detection. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition* (pp. 779–788).

[46] Relph, E. (1976). *Place and Placelessness*. Pion.

[47] Ren, S., He, K., Girshick, R., & Sun, J. (2015). Faster R-CNN: Towards real-time object detection with region proposal networks. In *Advances in Neural Information Processing Systems* (pp. 91–99).

[48] Salesses, P., Schechtner, K., & Hidalgo, C. A. (2013). The collaborative image of the city: Mapping the inequality of urban perception. *PLOS ONE*, *8*(7), e68400.

[49] Tuan, Y.-F. (1977). *Space and Place: The Perspective of Experience*. University of Minnesota Press.

[50] Varna, G. (2014). *Measuring Public Space: The Star Model*. Ashgate.

[51] Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., Kaiser, L., & Polosukhin, I. (2017). Attention is all you need. In *Advances in Neural Information Processing Systems* (pp. 5998–6008).

[52] Wang, J., & Biljecki, F. (2022). Unsupervised machine learning in urban studies: A systematic review of applications. *Cities*, *129*, 103925.

[53] Whyte, W. H. (1980). *The Social Life of Small Urban Spaces*. Conservation Foundation.