Received: 1 November 2025 | Revised: 22 July 2026 | Accepted: 28 August 2026 | Published online: 18 September 2026

RESEARCH ARTICLE

Journal of Data Science and Intelligent Systems
2026, Vol. 00(00) 1–22
DOI: 10.47852/bonviewJDSIS62028057

BON VIEW PUBLISHING

Mining Twitter Data to Inform Balcony Design in Residential Architecture: A Data-Driven Approach to Human Experience in Semi-open Spaces

Amir Shakibamanesh[1,*] and Seyedziya Hosseinzadeh[2]

[1] *Department of Urban Design, Iran University of Art, Iran*

[2] *Department of Architecture, University of Tehran, Iran*

Abstract: This study explores public perceptions and lived experiences of balconies in Iranian residential architecture through the lens of social media. Drawing on a curated dataset of over 10,000 Persian-language tweets posted by Iranian users between 2024 and 2025, this research adopts a data-mining approach to uncover how people emotionally and functionally relate to semi-open domestic spaces—particularly balconies and terraces. Tweets were collected through keyword-based filtering and then systematically preprocessed and classified into one or more of 16 architectural and environmental design categories (e.g., spatial dimensions, materiality, greenery, privacy, safety, and spatial perception) through a multi-stage thematic coding procedure. Sentiment analysis was then applied to determine the emotional tone—positive, negative, neutral, or mixed—of each thematic instance. Findings reveal a substantial gap between architectural intent and users' lived experience of their balconies. Categories related to functional use, privacy, and spatial adequacy showed high negative polarity (77.4% negative), largely reflecting complaints about small size, lack of privacy, and impracticality. In contrast, discourse on greenery and esthetics was markedly more favorable (29.7% positive). These results demonstrate the value of user-generated content as a source of data for human-centered design: architects, landscape designers, and urban planners can draw on lived spatial experiences captured in vernacular digital archives to gain empirical insight into occupants' spatial needs. Accordingly, the study proposes revised evaluation criteria for balcony design grounded in climatic factors and social behaviors.

Keywords: balcony design, human-centered architecture, Twitter data mining, sentiment analysis, user experience

1. Introduction

Balconies represent a critical yet often overlooked component of residential architectural design [1]. As an architectural element, the balcony is defined as a projecting platform attached to the exterior wall of a building at an upper-floor level, typically enclosed by a parapet, railing, or balustrade, and directly accessible from the interior through a door or French window [2, 3]. In formal typological terms, balconies are classified according to their structural configuration—including cantilevered slabs, bracket-supported projections, and recessed loggia variants—as well as their positional relationship to the floor plan (e.g., corner, frontal, or wrap-around) and their degree of enclosure [4, 5]. From a tectonic standpoint, the balcony integrates several interdependent architectural sub-systems: the structural support element (slab or brackets), the enclosure system (railing, glass screen, or solid parapet), the surface finish and drainage layer, the threshold element (door or glazed opening), and optional overhead cover (canopy or projecting slab from above) [5]. These components collectively determine the balcony's functional capacity, its spatial relationship to the interior, and its contribution to the building's façade composition and thermal envelope.

Beyond its technical constitution, the balcony occupies a distinct position within the spatial hierarchy of residential architecture: it is neither fully interior nor fully exterior, but rather a transitional semi-open threshold [4, 6]. This liminal condition—simultaneously part of the private dwelling and exposed to the public realm—gives the balcony its characteristic spatial ambiguity and makes it a site where architectural decisions intersect with environmental performance, social behavior, and cultural norms. Its depth, orientation, enclosure level, and materiality each carry direct consequences for how residents use, perceive, and emotionally experience the space [7, 8].

Although balconies are common in urban design and associated literature, there is a notable lack of empirical studies on the everyday experiences of the people who occupy them. A significant portion of current studies on semi-open spaces has concentrated on elements such as official design features, climate efficiency, or the appearance of balconies on façades—which is

***Corresponding author:** Amir Shakibamanesh, Department of Urban Design, Iran University of Art, Iran. Email: a.shakibamanesh@art.ac.ir

helpful yet overlooks a major aspect of the scenario. Balconies serve as a social and spatial interface, positioned at the boundary between private and public existence, and this nuance often goes unnoticed. Conventional methods such as post-occupancy surveys are not very effective, as they fail to accurately reflect the immediate, emotional, and socio-psychological reactions that occur in real time. This study follows a different path. It examines large amounts of Twitter/X data to explore those less visible sides of how people interact with balconies. The goal is to move past a purely functional view and reach a better understanding of balconies as architectural spaces that people really live in.

Balconies are often designed based on prevailing market/-typological norms rather than feedback from the in-use phase; industry actors show limited understanding of how balconies are actually used, and mainstream provision commonly results in narrow balconies (around 1.2–1.5 m depth) with design emphasis on safety and façade appearance [8]. This gap is particularly pronounced in contexts such as Iran, where balconies simultaneously embody cultural symbolism, climatic necessity, and contested social visibility [9]. Recent studies also show that balcony design has direct implications for environmental sustainability, including daylight access, thermal regulation, adaptive façade performance, and energy efficiency, reinforcing the need to treat balconies as active contributors to sustainable building performance rather than as purely formal appendages [10, 11].

Consequently, the supply side (developers and designers/-architects) often lacks systematic, user-informed insight from the occupancy (in-use) phase on how balconies support everyday life, which sustains reliance on supply-side assumptions rather than evidence-based understanding [8]. Traditional post-occupancy evaluation (POE) can generate valuable qualitative, user-perspective insights, but conventional data collection (e.g., surveys, interviews, and on-site observations) is often time- and resource-intensive and challenging to scale; moreover, findings are frequently context-specific, which can limit broader generalization [12, 13]. At the same time, computational social science has shown that geotagged Twitter data can be aggregated to capture place-based affect and subjective well-being at population scale [14, 15] and can also reflect everyday environmental concerns (e.g., air pollution and wildfire smoke) in ways that track measured conditions [16, 17]. However, social media analytics in the built-environment literature has been used predominantly at urban/public-space scales [18, 19] and has been less frequently mobilized to study specific architectural typologies—particularly semi-private domestic thresholds such as balconies. This dual gap—limited scalable user evidence in architectural research and limited typology-focused adoption of social media analytics—motivates the present study and frames its contribution as a bridge between experience-led architectural inquiry and digital urban analytics. The core problem addressed by this research is threefold and can be stated precisely as follows. **First**, balcony design in Iranian residential construction is predominantly driven by supply-side assumptions—typological conventions, regulatory minimums, and developer preferences—rather than by systematic evidence of how residents actually use, evaluate, and emotionally experience these spaces [8, 9]; the result is a persistent mismatch between designed intent and lived reality. **Second**, existing architectural methods for capturing user experience—particularly POE—are resource-intensive, geographically bounded, and difficult to scale, leaving no established pathway for generating large-sample, affect-tagged evidence about a single architectural element such as the balcony [12, 20]. **Third**, while computational methods for mining social media discourse are well-established in urban studies and computational social science [14, 21], they have not been systematically applied to specific residential architectural typologies, creating a methodological gap precisely where cross-disciplinary integration would be most productive [22, 23]. Together, these three gaps produce a situation in which architects and developers design balconies with limited empirical understanding of resident needs, frustrations, and aspirations—and no scalable tool exists to remedy this. This study is designed to address that situation directly.

This study addresses this gap by treating social media discourse as a form of post-occupancy evidence: using unsolicited, user-generated Persian-language tweets to capture experiential and affective evaluations of balconies at scale, while recognizing known ethical and representativeness challenges of such data. We are not claiming that Twitter users represent all residents perfectly. The data are just used as extra post-occupancy signals that can show common themes, problems, and design issues worth looking into more deeply. By using thematic coding together with sentiment analysis and design interpretation, this paper tries to give a clear data-driven picture of people's needs, difficulties, and perceptions about balcony design in Iranian residential buildings.

The study looks at 10,000 Persian tweets from Iranian users between July 2024 and June 2025. These tweets talk about balconies in real-time everyday language. An advantage of the data is that people wrote them naturally, without any questions or prompts from researchers. Because of that, we can see real thoughts about how people inhabit, perceive, and sometimes struggle with their balconies. This kind of natural talk shows emotions, social issues, and body experiences that are usually missing in formal architectural studies.

This study combines thematic content analysis and computational sentiment analysis in a mixed-method design. The purpose is to sort tweets into 16 architectural and environmental categories. They cover factors such as spatial adequacy, materiality, greenery, privacy, safety, and esthetics. The tweets are then labeled as positive, negative, neutral, or mixed in terms of sentiment polarity. Specifying such polarity across all the design themes enables us to see which features tend to make people satisfied and which ones lead to their discomfort. An important point is that the researchers took all this user-generated discourse from social media and actually connected it to architectural theory and real design analysis. That approach feeds into the bigger movement toward evidence-based and human-centered design. It basically shows that data on what people are saying online on a large scale can help plan semi-open spaces, particularly balconies, which are responsive, climate-aware, and culturally sensitive.

It is important to be clear about what this study adds to what is already known. The prior work at the intersection of social media and the built environment has mostly been limited to urban-scale phenomena—like public parks, streets, neighborhood affect, and city-wide well-being indicators [18, 19, 24, 25]—and has relied almost exclusively on English-language corpora from North American or Western European cities. Studies that have examined specific architectural elements have done so through small-sample surveys or controlled interviews, not through corpus-scale computational analysis [8, 26]. No published study, to our knowledge, has applied large-scale thematic coding and sentiment analysis to a single residential architectural typology—the balcony—using non-Western, non-English social media data. Furthermore, existing Persian-language social media studies in architecture have remained at the level of general housing satisfaction or urban perception, without

disaggregating findings by specific spatial element or applying a design-theory-grounded thematic schema [22, 23].

Against this backdrop, this paper makes three original and precisely bounded contributions. (1) It operationalizes Persian-language Twitter discourse as a form of scalable, complementary post-occupancy evidence for a specific and architecturally significant residential element—the balcony—extending both the linguistic scope and the spatial granularity of social media built-environment research beyond what has previously been attempted. (2) It develops and applies a novel 16-category, design-theory-grounded thematic framework for classifying balcony-related discourse, associates each category with a quantified Sentiment Polarity Index (SPI) derived from 10,000 tweets, and validates the framework through inter-coder reliability testing ($\kappa = 0.81$)—providing a replicable methodological schema for future discourse-based POE studies of other architectural typologies. (3) It translates these empirical sentiment patterns into context-sensitive design implications for Iranian high-density residential architecture, grounding design recommendations in resident-expressed needs and frustrations rather than typological convention—while explicitly acknowledging the representativeness, ethical, and methodological limitations inherent in social media data, including the absence of geographic disaggregation and the constraints of lexicon-based sentiment classification.

2. Literature Review

2.1. Balconies as semi-open interfaces in residential architecture

Balconies occupy a spatially and symbolically liminal position within residential architecture, functioning as semi-open thresholds that mediate between interior domestic life and the exterior public realm [27]. They serve a multiplicity of environmental and psychosocial roles—ranging from natural ventilation and daylight penetration to enabling social observation, ritual behaviors, solitude, and even micro-scale cultivation [28]. In climatic and urban settings such as the Middle East and Mediterranean, where high-density housing coexists with cultural preferences for outdoor living, balconies assume heightened ecological and symbolic relevance [29]. Yet despite their multifunctionality and embeddedness in everyday life, balconies are often marginalized in architectural practice and research—treated as residual extensions, façade ornaments, or token gestures rather than purposefully designed, inhabitable spaces [30]. The prevailing literature has focused predominantly on formal typologies, thermal optimization strategies, and regulatory constraints, with comparatively limited attention paid to the experiential, emotional, and socio-cultural dimensions of balcony use [31]. In the Iranian context, these dimensions are particularly salient. Balconies are deeply entangled with cultural tensions—particularly around gendered visibility, modesty norms, and the blurred boundary between private domesticity and public exposure [32, 33]. As Aronis [27] argues, occupying a balcony is a form of social navigation—inviting interaction while also exposing inhabitants to surveillance. In this sense, the balcony becomes not merely a spatial element but a site of ideological negotiation, where architecture intersects with social behavior, identity politics, and cultural expectations.

While approaches grounded in environmental psychology and POE have begun to acknowledge the importance of lived experience in the design and evaluation of built environments [34, 35], empirical studies of how users interact with and emotionally respond to balconies remain scarce [8]. This gap is especially critical given the growing recognition of semi-open spaces as central to urban mental health, domestic well-being, and social resilience [36]. This research aims to address that gap by analyzing unsolicited, user-generated discourse from social media platforms—capturing how individuals perceive, critique, and emotionally engage with balcony spaces in everyday life. In doing so, it shifts the architectural lens from top-down typological assumptions toward a bottom-up, culturally grounded understanding of the balcony as a lived and negotiated spatial interface.

2.2. Toward human-centered and experience-based architectural inquiry

A growing body of scholarship in architectural theory, environmental psychology, and critical spatial studies highlights the persistent disjunction between the intentions of designers and the experiences of users [37, 38]. While architectural practice often emphasizes formal coherence, regulatory compliance, and typological legibility, occupants engage with space through complex blends of functionality, emotion, social interaction, and symbolic meaning [39]. This divergence is especially pronounced in semi-open spatial types—such as balconies, terraces, and courtyards—which serve as both extensions of private life and points of visual and acoustic exchange with the external environment [40]. Semi-private zones expose residents to contradictory demands: they must provide a sense of retreat and security while supporting visibility, openness, and social interface [41]. In high-density or culturally conservative settings, these demands often become sources of spatial conflict or adaptive behavior, such as retrofitting balconies with screens, plants, or enclosures to reclaim privacy [8, 42]. Despite their everyday ubiquity, these nuanced tensions are rarely explored in design studios or policy documents, which tend to relegate balconies to formal or passive roles within façade systems [43].

To address this disjunction between what is built and what people experience in those spaces, researchers have turned to user-centered methodologies (POE [26], qualitative interviews, ethnographic studies, and behavioral mapping [44]). They have proven to be really useful in tackling issues such as comfort, usability, satisfaction, and behavior. They are central to what now motivates evidence-based and participatory design thinking. However, conventional POE methods face significant limitations: they are often resource-intensive, context-specific, and limited in scope, making it difficult to extrapolate generalizable patterns across different cultural or urban settings [20]. Additionally, many traditional methods struggle to capture the spontaneous, affective, and informal ways in which users describe, critique, and emotionally engage with space in everyday life [45]. This gap has led scholars to call for more adaptive, scalable, and user-generated forms of spatial knowledge [46]. In this context, social media platforms—particularly Twitter—present a novel opportunity to access real-time, unsolicited, and emotionally charged reflections on the built [47]. Tweets often contain rich clues about how users feel, adapt to, or reject architectural decisions, making them a valuable—if underutilized—resource for informing more responsive and nuanced design practice [48]. By embracing social media as a form of post-occupancy discourse, this study aims to capture the subjective, lived, and emotional dimensions of balcony use at scale—offering not only empirical data but also insights into how

architectural space becomes meaningful (or problematic) through everyday interaction.

2.3. Social media as an emerging tool in architectural and urban research

The proliferation of digital communication platforms has introduced novel opportunities for spatial inquiry, particularly in contexts where traditional forms of data collection are constrained by scale, access, or responsiveness [49]. Among these platforms, Twitter has emerged as a uniquely valuable resource, offering large volumes of unsolicited, timestamped, and geolocated narratives that reflect how individuals perceive, critique, and emotionally engage with their environments in real time [50]. These narratives often contain spontaneous expressions of satisfaction, frustration, or desire—making them fertile ground for affective urban analytics and experience-centered spatial research [51]. In recent years, researchers across urban studies, geoinformatics, and computational social science have employed techniques such as text mining, sentiment analysis, natural language processing (NLP), and topic modeling to reveal spatial patterns embedded in social media discourse [52].

Beyond general urban analytics, recent work has demonstrated the relevance of text mining and NLP for sustainability-oriented built-environment research, including analyses of green housing discourse and social media–based decision support in smart-city contexts [53, 54].

These methods have been used to identify emotional geographies [24], monitor public perception of urban policy and infrastructure [55], and map inequalities in access, exposure, or amenity [25]. Such work demonstrates the capacity of user-generated content to complement and challenge top-down datasets like census records, land use maps, or planning [56]. Despite this progress, the architectural field—particularly in its design-oriented branches—has been slow to integrate these digital methodologies into its analytic repertoire [22]. Most existing applications of social media in architecture are limited to marketing analytics, brand sentiment, or post-occupancy surveys tied to specific developments [12]. Rarely are platforms like Twitter or Instagram treated as vernacular archives of spatial experience, capable of revealing how architectural elements—such as balconies, thresholds, or façades—are emotionally and behaviorally inhabited [57]. This gap is especially salient in Global South contexts, where planning regimes are often less participatory, and residents' voices are underrepresented in official evaluation frameworks [58]. In such settings, social media may serve as a form of grassroots spatial testimony, articulating needs, discomforts, and micro-political struggles that do not surface in professional discourse or regulatory policy [59]. While ethical and linguistic challenges remain—such as data privacy, representativeness, and local idioms—there is growing recognition that digitally harvested, affect-rich, and culturally embedded data can provide invaluable insight into everyday urbanism [60]. By applying social media mining to architectural elements such as balconies, this study contributes to the nascent but promising intersection of computational urbanism, emotive spatial analysis, and experience-led design research.

2.4. Research gap and contribution

While balconies have been explored through multiple lenses—ranging from technical performance and environmental behavior [43] to formal esthetics [61] and policy compliance [31]—there remains a marked absence of scholarship focused on their affective, perceptual, and everyday functional dimensions [62]. Existing research tends to privilege designer-centric narratives, often overlooking how inhabitants emotionally inhabit and evaluate balcony spaces in practice [63]. This is especially true within Iranian urban contexts, where balconies are shaped by religious modesty codes, climatic constraints, high-density living patterns, and evolving cultural norms around public/private spatial boundaries [32, 33]. Moreover, no prior architectural or urban study has leveraged large-scale, user-generated social media content to systematically investigate balconies as emotionally charged, socially contested, and functionally complex spaces. While computational social science has demonstrated the potential of platforms like Twitter to reveal spatial affect, behavioral insight, and emergent urban concerns [21], such methods have rarely been applied to architectural typologies—let alone semi-private, domestic thresholds like balconies [23]. This study tries to fill some of these gaps. It works with 10,000 Persian tweets from July 2024 to June 2025 and combines thematic coding, sentiment analysis, and statistical charts in a mixed-methods design. The goal is to achieve a bottom-up account of how people really see, judge, and adapt to their balconies. This means that lived experience and emotional resonance are to be recognized as legitimate dimensions of spatial evaluation, not as afterthoughts. The research is situated at the intersection of architectural theory, environmental psychology, and digital urban studies and presents a novel conceptual and methodological approach to the study of balconies. The argument is that these spaces, so easily overlooked, are in fact key nodes of environmental, cultural, and emotional negotiation. So, architects and planners need empirical tools that reflect this reality if they want to design housing that is really responsive, adaptive, and user-centered.

2.5. Theoretical framework

Three main theories support this study and help conceptualize and understand what users say about balconies on the internet.

The first is Attention Restoration Theory by Kaplan and Kaplan (1989). This theory explains that contact with nature, even in small amounts, helps people recover from mental tiredness and stress.

The second is Ulrich's Theory of Supportive Design (1997) [29]. It suggests that buildings and spaces can reduce stress and improve well-being when they include natural elements, offer pleasant distractions, and give people some control over their surroundings. In practice, this theory guides our interpretation of user complaints about balconies, such as cramped, exposed, or noisy, as indicators that the space was not supportive or comfortable.

The third theory relates to threshold and liminal spaces in architecture theory [4, 6]. The idea is of balconies as liminal spaces—spaces on the threshold of private domesticity and public exposure. That spatial liminality generates real tensions about privacy, social visibility, and cultural norms, and those tensions are evident in the discourse patterns we observed in Category D (View and Privacy) and Category K (Social Interaction). Together, these three frameworks allow the study's empirical findings to be situated within established theoretical traditions, enabling a more theoretically grounded interpretation of sentiment patterns and design implications.

3. Research Methodology

3.1. Research design

This research employs a mixed-methods, data-driven approach combining computational text analytics with thematic content analysis. The objective is to uncover how Iranian social media users perceive, evaluate, and emotionally respond to architectural features of balconies in residential buildings. Drawing on public, user-generated data, the study bridges the gap between designed intent and lived experience, contributing to human-centered and evidence-based design research in architecture and urbanism.

The methodological framework consists of four main phases:

- Data Collection
- Data Preprocessing and Cleaning
- Thematic Categorization
- Sentiment Analysis and Statistical Interpretation

3.2. Data collection

The data source was Twitter. Persian-language tweets from Iranian users were pulled specifically to capture organic, unsolicited reflections on balconies in residential settings—the kind of commentary people post without being asked, which is exactly what makes it useful. Keyword-based filtering with temporal constraints kept the collection both relevant and contextually coherent. One thing worth flagging upfront: in Persian, "balcony" and "terrace" get used interchangeably, so both terms were used as the sole filters for extracting tweets. That yielded 10,843 original tweets posted between July 1, 2024, and June 30, 2025. Retweets, replies, promotional posts, and unrelated tweets were removed to keep only original user-generated content. Tweets that were clearly promotional, made by bots, or not related to the topic were removed during preprocessing. When location and time information was available, it was saved for possible future use, for example, spatiotemporal analysis of the discourse on balconies. The result was a pruned collection that shows public feelings about balconies in Iran, which is suitable for big content and sentiment analysis.

3.3. Data cleaning and preprocessing

Before the tweets could be used for thematic and sentiment analysis, they had to go through a careful cleaning process. This step was necessary to make sure the data were reliable and clear from a linguistic and analytical point of view. Several steps were followed to make the collection of tweets more accurate, as detailed below:

1) **Language Filtering:** A dedicated language detection model was applied to verify that all tweets were in Persian, eliminating non-Persian content and multilingual noise.
2) **Stopword Removal:** Common Persian stopwords (e.g., "که", "برای", "از")—words that carry minimal semantic weight—were removed using curated Persian NLP stopword lists.
3) **Emoji, Hashtag, URL, and Punctuation Cleaning:** Non-textual elements such as emojis, hashtags, mentions, web links, and excess punctuation were stripped to reduce processing bias and enhance clarity.
4) **Tokenization and Lemmatization:** Tweets were tokenized into constituent word units, and lemmatization was applied using Persian NLP libraries such as Hazm and Parsivar, standardizing words to their base grammatical forms.
5) **Deduplication:** Identical and near-identical tweets—often originating from spam accounts or coordinated campaigns—were removed to prevent distortion of sentiment distributions.
6) **Relevance Filtering:** A special classifier (a computer tool built for this purpose) was used to remove tweets that were not actually about balconies as physical or architectural spaces, as in metaphorical or poetic uses of the word.

The final dataset was cleaned, normalized, and made linguistically consistent. This dataset then became the basis for the next steps: text mining, thematic categorization, and sentiment scoring. This cleaning stage (called preprocessing) was necessary because it reduced noise in the data, kept the interpretation accurate, and protected the overall reliability of the study.

To ensure analytical reliability, the raw dataset underwent a rigorous, systematic preprocessing and review pipeline (summarized in Table 1). The data cleaning, filtering, and coding processes were conducted utilizing spreadsheet software (Microsoft Excel) to facilitate systematic search, sorting, and manual review.

Instead of using automated scripts, a careful human-in-the-loop method was used. Stopword removal and noise reduction were handled by systematically filtering out common Persian functional words alongside domain-specific promotional terms (e.g., real-estate jargon) identified through repeated passes over the corpus. Tweets containing non-architectural, metaphorical, or poetic uses of "balcony" and "terrace" were flagged using an exclusion list of contextual keywords and then reviewed by the researchers directly. That screening process removed 843 irrelevant or noisy posts from the initial pool of 10,843, leaving a clean, standardized dataset of exactly 10,000 tweets ready for thematic coding and sentiment analysis.

Table 1
Data processing pipeline—tweet counts at each stage

Stage	Description	Tweets remaining
Initial collection	Keyword-based extraction ("تراس" / "بالکن") and excluding retweets/replies	10,843
Data cleaning and noise reduction	Systematic removal of duplicates, malformed, and non-Persian text	10,000
Relevance filtering	Search-based and manual exclusion of non-architectural metaphors/promotions	10,000
Final dataset	Final curated corpus retained for thematic and sentiment analysis	10,000

Note: After cleaning and relevance filtering, 843 tweets were excluded. No stable intermediate subtotal was retained.

3.4. Thematic categorization

The thematic categorization worked in two stages, combining a deductive framework with inductive refinement. In the first stage, an initial set of 12 thematic categories was built deductively, drawing on established frameworks from architectural theory, environmental psychology, and POE literature. That initial framework was then tested on a pilot sample of 500 tweets—and in the process, four additional categories emerged inductively from recurring discourse patterns that the original schema hadn't captured. That brought the total to the final 16-category framework shown in Table 2.

One deliberate methodological choice worth flagging was that classification was done through expert manual coding rather than computational topic modeling like Latent Dirichlet Allocation (LDA) or BERTopic. The reasoning is straightforward. The vocabulary here is highly domain-specific, individual tweets often carry multiple labels at once—describing spatial dimensions and emotional experience in the same breath—and the categories needed to align with architectural theory rather than statistical co-occurrence patterns. Unsupervised topic modeling just wasn't the right tool for that job.

Each of the 10,000 preprocessed tweets was coded by the primary researcher using a structured codebook and then assigned to 16 thematic clusters grounded in principles from architecture, landscape, and environmental design. A single tweet could be assigned to more than one category where the content warranted it:

To check whether the thematic coding was actually consistent, an inter-coder reliability procedure was run. A stratified random sample of 543 tweets—5% of the full dataset—was independently coded by a second researcher with expertise in architectural theory and qualitative content analysis, someone who hadn't been involved in the original classification at all. Both coders worked independently, assigning each sampled tweet to one or more of the 16 thematic categories using only the definitions laid out in Table 2. Agreement was then calculated using Cohen's kappa (κ) [64], the standard measure of inter-rater reliability that adjusts for the agreement you'd expect by chance alone. The result ($\kappa = 0.81$) indicates strong inter-coder agreement, clearing the commonly accepted threshold of $\kappa \geq 0.80$ for reliable qualitative coding schemes. Where disagreements did come up, they mostly involved tweets with overlapping spatial cues—simultaneous references to spatial dimensions and functional use, for instance. These tweets were worked out through structured discussion until both coders agreed on the same label. Thematic assignments were also verified by domain experts in architecture and landscape studies. In total, the 10,000 tweets produced 17,305 category-level thematic instances, a figure that reflects the multiple category assignments applied to individual tweets. The complete set of coded and sentiment-labeled tweets, together with their thematic categories, is provided as Supplementary Dataset S1 (bilingual, Persian–English).

3.5. Sentiment analysis

To understand how users felt about different balcony features, each tweet was analyzed for its sentiment (whether it expressed a positive, negative, neutral, or mixed feeling). This was done by looking at the tone and language style of each tweet. A hybrid method was used for this sentiment analysis: it combined rule-based Persian sentiment lexicons (lists of words with assigned emotional value) with contextual scoring heuristics (rules that help interpret meaning based on context). This combination helped improve accuracy, especially with informal, idiomatic, and emotional language, which is very common on social media.

The choice of a lexicon-based approach—rather than fine-tuned transformer models such as ParsBERT [65] or XLM-RoBERTa [66]—was motivated by three domain-specific constraints. First, no publicly available labeled corpus exists for Persian-language architectural or residential discourse, making supervised training of neural models infeasible without prohibitive annotation cost. Second, the sentiment vocabulary in this domain is highly specialized and spatially grounded (e.g., terms denoting cramped conditions, visual exposure, or thermal discomfort),

Table 2
Thematic categories for balcony-related tweets

Code	Thematic category	Example keywords
A	Dimensions & spatial scale	small, cramped, spacious
B	Form, geometry, & spatial configuration	rectangular, narrow, curved
C	Position in floor plan	corner, front-facing, central
D	View & visual privacy	view, overlooking, privacy
E	Greenery & vegetation	plants, flowerpots, garden
F	Materiality & surface quality	wood, ceramic, aluminum
G	Cover & shading elements	shade, canopy, awning
H	Light, ventilation, & climatic response	breeze, sunlight, hot, suffocating
I	Safety & physical security	railing, children, falling, unsafe
J	Daily use & functional purpose	coffee, smoking, drying clothes
K	Social interaction & neighbor noise	neighbors, shouting, eavesdropping
L	Indoor–outdoor connection	door, transition, threshold
M	Esthetic quality & decoration	beautiful, ugly, well-designed
N	Spatial perception & psychological impact	calm, disturbing, peace
O	Architectural style & design value	modern, brutalist, fake-classic
P	Urban role of balconies	cityscape, skyline, façade

Note: Each tweet may fall under one or more categories, based on architectural and spatial features.

which are systematically underrepresented in general-purpose Persian sentiment benchmarks trained on news or product reviews. Third, the lexicon-based approach preserves full interpretability: each sentiment assignment can be traced to specific lexical triggers, which is essential for validating design-relevant inferences and communicating findings to architectural practitioners unfamiliar with machine learning outputs.

This methodological choice entails several acknowledged limitations. Lexicon-based classifiers are inherently more sensitive to sarcasm, irony, and context-dependent polarity shifts than supervised neural models, particularly in informal social media discourse where negation structures and colloquial idioms are prevalent. Additionally, out-of-vocabulary terms—including neologisms, slang, and domain-specific architectural jargon not captured in general Persian sentiment lexicons—may be misclassified or assigned neutral sentiment by default. These limitations are strongest in tweets with hidden emotional hints, figurative language, or culture-specific expressions that need pragmatic inference beyond simple word matching. To push back against such weaknesses, the lexicon was manually expanded with 127 domain-relevant terms pulled from a preliminary qualitative review of a 500-tweet subsample. Next, contextual heuristics were applied to handle common negation patterns—for example, negation markers preceding positive terms. Even so, the sentiment classifications are best understood as approximations of emotional valence rather than ground-truth labels. Future work should look at domain-adapted transformer models as annotated Persian architectural corpora become available, and consider hybrid approaches that pair lexicon-based baselines with active learning or few-shot fine-tuning strategies.

Tweets were assigned to one of four sentiment categories:

1) Positive—Expressions of satisfaction, comfort, admiration, or delight
2) Example: "I love sitting out here with my coffee. The view is perfect."
3) Negative—Expressions of dissatisfaction, annoyance, critique, sarcasm, or distress
4) Example: "This balcony is a smoking hell—I can't even breathe in my own house."
5) Neutral—Descriptive or observational statements with little to no emotional valence
6) Example: "The balcony is two meters wide and faces the street."
7) Mixed—Simultaneous expressions of both positive and negative affect
8) Example: "Nice view but absolutely no privacy from the neighbors."

The sentiment detection was applied at the tweet level, but results were subsequently linked to each of the tweet's thematic codes, allowing the identification of emotional patterns across architectural categories (e.g., how often users reacted negatively to daily functional issues vs. positively to greenery or decoration). To quantify the emotional valence of each thematic category in a comparable and reproducible manner, an SPI was computed for each of the 16 categories. The SPI is defined as:

$$\mathrm{SPI}_c = \frac{N_c^+ - N_c^-}{N_{\mathrm{total},c}} \tag{1}$$

where N_c^+ denotes the number of positively coded tweets in category c, N_c^- denotes the number of negatively coded tweets in category c, and $N_{\mathrm{total},c}$ is the total number of tweets assigned to category c. The SPI ranges from −1 (all tweets negative) to +1 (all tweets positive), with values near 0 indicating either predominantly neutral discourse or an equal balance of positive and negative expressions. This index allows direct comparison of emotional valence across categories of varying size, normalizing for differences in tweet frequency. Categories with an SPI below −0.30 were flagged as high-dissatisfaction domains and moved to the front of the queue in the design implications analysis.

That cross-sectional sentiment-thematic mapping became the backbone for the statistical comparisons and visualizations reported in the Findings section. It also made it possible to spot polarized or emotionally charged design features—which in turn allowed for a richer, evidence-informed reading of what users actually prefer, what frustrates them, and what they expect.

3.5.1. Validation of the sentiment classification approach

To check how well the lexicon-based sentiment classifier performed, a random sample of 150 tweets was selected from the final dataset. The sample was stratified, meaning it kept the same proportion of the four sentiment types (positive, negative, neutral, mixed) as found in the whole dataset. A second researcher labeled these tweets independently. This researcher has expertise in Persian computational linguistics and did not see the automated classification output before. The human-annotated labels acted as the ground-truth reference set to evaluate the lexicon-based predictions. Standard metrics (precision, recall, and F1-score for each sentiment class, plus macro-averaged F1 across all four classes) were applied to measure classification performance. The results showed acceptable performance for an interpretable lexicon-based approach. The overall accuracy was 80.0% (120/150 correct), and the macro-averaged F1-score was 0.778 (Table 3; see also Supplementary Dataset S2 for full validation results). As expected, performance was strongest for the Negative class and comparatively lower for Mixed tweets, which are inherently more difficult to classify using rule-based sentiment methods.

3.6. Statistical and visual analysis

After tagging the tweets by theme and sentiment, the dataset was analyzed using descriptive and comparative statistics. This

Table 3
Validation of lexicon-based sentiment classification against human-annotated ground truth (n = 150)

Sentiment class	Precision	Recall	F1-score	Support (n)
Positive	0.775	0.795	0.785	39
Negative	0.883	0.898	0.891	59
Neutral	0.725	0.674	0.699	43
Mixed	0.700	0.778	0.737	9
Macro average	0.771	0.786	0.778	150

Note: Overall accuracy: 80.0% (120/150 correct).

way, trends, differences, and patterns in how users perceive balconies were identified. The objectives of this analytical phase were to:

1) Quantify the frequency distribution of tweets across the 16 thematic categories (A–P);
2) Measure the sentiment composition (positive, negative, neutral, and mixed) within each category;
3) Identify category-level divergence (i.e., determine which themes generated the most polarized or emotionally complex responses); and
4) Detect salient overlaps between categories, indicating compound spatial concerns or correlated experiences.

To operationalize these goals, the study employed Python 3.10 with the following libraries: pandas (v2.1) for data structuring and tabular transformation; matplotlib (v3.8) and seaborn (v0.13) for graphical output; scikit-learn (v1.3) for co-occurrence matrix computation; and custom scripts for SPI calculation and category-level statistical aggregation.

Thematic co-occurrence was quantified using the Jaccard Similarity Coefficient (J), defined as:

$$J(A, B) = \frac{|A \cap B|}{|A \cup B|} \tag{2}$$

where $|A \cap B|$ is the number of tweets assigned to both categories A and B and $|A \cup B|$ is the number of tweets assigned to at least one of the two categories. Jaccard values range from 0 (no co-occurrence) to 1 (complete overlap), with higher values indicating stronger thematic coupling in the discourse. This coefficient was computed for all 120 possible category pairs ($16 \times 15/2$) and visualized as a symmetric heatmap.

For presentation clarity, only the most analytically salient subset of category pairs is shown in the heatmap below, rather than the full 16×16 matrix.

The following outputs were generated and incorporated into the analysis:

1) Horizontal bar charts allowed easy visual side-by-side scanning of emotional responses by comparing positive, negative, neutral, and mixed sentiment across all categories.
2) Summary tables laid out both category frequency and sentiment breakdown for each thematic code.
3) A co-occurrence matrix was used to dig into multi-thematic tweets and map out the relational structure of balcony features as they appeared in user discourse.
4) The analysis also flagged outlier categories—cases where emotional intensity diverged sharply from tweet volume, like themes that generated strong reactions despite relatively low representation in the dataset.

Together, those visualizations and statistical summaries gave the interpretation of results a solid empirical foundation. In particular, they allowed evidence-based evaluation of which architectural elements users found most dissatisfying, emotionally meaningful, or socially contested. This directly informed the design implications discussed in subsequent sections.

4. Findings

This section presents the analytical results of the thematic and sentiment analysis conducted on 10,000 curated Persian-language tweets. Findings are structured around the 16 thematic categories (A–P) introduced in Table 2 and supported by corresponding sentiment metrics and visualizations.

4.1. Overview of thematic distribution

Among the full dataset, users frequently mentioned design-related attributes such as daily functionality (J), spatial dimensions (A), privacy and view (D), and greenery (E). Table 4 summarizes the distribution of tweets across thematic categories.

Table 4
Frequency of tweets per thematic category. Higher frequencies indicate features users frequently referenced—positively or negatively

Code	Category name	Tweet count	% of total tweets
J	Daily functional use	3227	29.7
A	Dimensions & spatial scale	2310	21.3
D	View & visual privacy	1874	17.3
E	Greenery & vegetation	1320	12.2
M	Esthetics & decoration	984	9.1
H	Light, ventilation, & climate	932	8.6
K	Social interaction & noise	817	7.5
F	Materiality & surface quality	778	7.2
I	Safety & physical security	733	6.8
G	Cover & shading elements	702	6.5
N	Spatial perception & mental experience	693	6.4
C	Position in floor plan	665	6.1
O	Architectural style & design value	622	5.7
L	Indoor–outdoor connection	598	5.5
B	Form, geometry, & spatial configuration	561	5.2
P	Urban role of balconies	489	4.5

Note: Categories are not mutually exclusive; some tweets appear in multiple categories.

Table 5
Sentiment breakdown by thematic category

Code	Category name	Positive %	Negative %	Neutral %	Mixed %
J	Daily functional use	9.3	77.4	8.8	4.5
A	Dimensions & spatial scale	12.2	36.3	42.6	8.9
D	View & visual privacy	21.7	31.2	38.5	8.6
E	Greenery & vegetation	29.7	17.4	42.1	10.8
M	Esthetics & decoration	27.8	20.5	40.6	11.1
H	Light, ventilation, & climate	18.6	33.1	39.4	8.9
K	Social interaction & noise	10.4	41.5	39.3	8.8
F	Materiality & surface quality	22.0	24.6	40.8	12.6
I	Safety & physical security	13.1	28.3	47.7	10.9
G	Cover & shading elements	19.2	27.4	45.0	8.4
N	Spatial perception & mental experience	25.9	30.5	32.4	11.2
C	Position in floor plan	12.8	22.7	53.4	11.1
O	Architectural style & design value	23.5	19.9	44.1	12.5
L	Indoor–outdoor connection	19.4	27.1	45.2	8.3
B	Form, geometry, & spatial configuration	16.1	26.4	49.7	7.8
P	Urban role of balconies	21.1	23.7	47.0	8.2

Note: Top negative and positive values are highlighted.

4.2. Sentiment distribution across categories

Sentiment analysis reveals meaningful distinctions across thematic areas. These category-level sentiment patterns should be interpreted in light of the validated performance of the sentiment classifier (overall accuracy = 80.0%, macro-F1 = 0.778), which supports the reliability of the broad polarity trends while still warranting caution in borderline or mixed cases. As seen in Table 5 and Figures 1, 2, 3, and 4, users express particularly negative sentiment toward functional and spatial inadequacies, while showing more appreciation for vegetation, design esthetics, and experiential elements.

4.3. Sentiment visualizations

All figures are shown with clear captions and standardized formatting to make comparison easy across sentiment classes and thematic categories. Category codes A–P correspond to the thematic framework defined in Table 2.

As Figure 1 shows, Category J (Daily Use) had the highest negative sentiment of any category—77.4% of related tweets expressed frustration, annoyance, or discomfort. That negativity was driven largely by recurring complaints about smoking neighbors, the visibility of drying laundry, and a general lack of privacy during balcony use. Social Interaction (K) and Dimensions & Scale (A) were two other categories with significantly high negative sentiment, indicating spatial constraints and neighbor behavior as two of the larger sources of dissatisfaction in the dataset.

Conversely, Figure 2 shows that Category E (Greenery) elicited the strongest positive sentiment, with 29.7% of tweets expressing joy, relaxation, or admiration. Users frequently praised balconies that featured plants, flowerpots, climbing vines, or miniature gardens, describing them as calming, refreshing, and esthetically enriching. Esthetics (M) and Spatial Perception (N) also ranked high in positivity, reflecting a strong appreciation for balconies that provided visual delight or psychological retreat.

As illustrated in Figure 3, Position in Plan (C) and Form & Geometry (B) had the highest proportions of neutral sentiment, with over 50% of tweets in these categories being purely descriptive. Users in these tweets often referred to the layout, location, or configuration of balconies in architectural plans without conveying a clear emotional tone. These findings suggest that certain design attributes are more frequently discussed in functional or observational terms, rather than affective ones.

Figure 5 consolidates these patterns into a single heatmap, cross-tabulating sentiment class against thematic category to allow direct visual comparison of polarity concentration across all 16 categories.

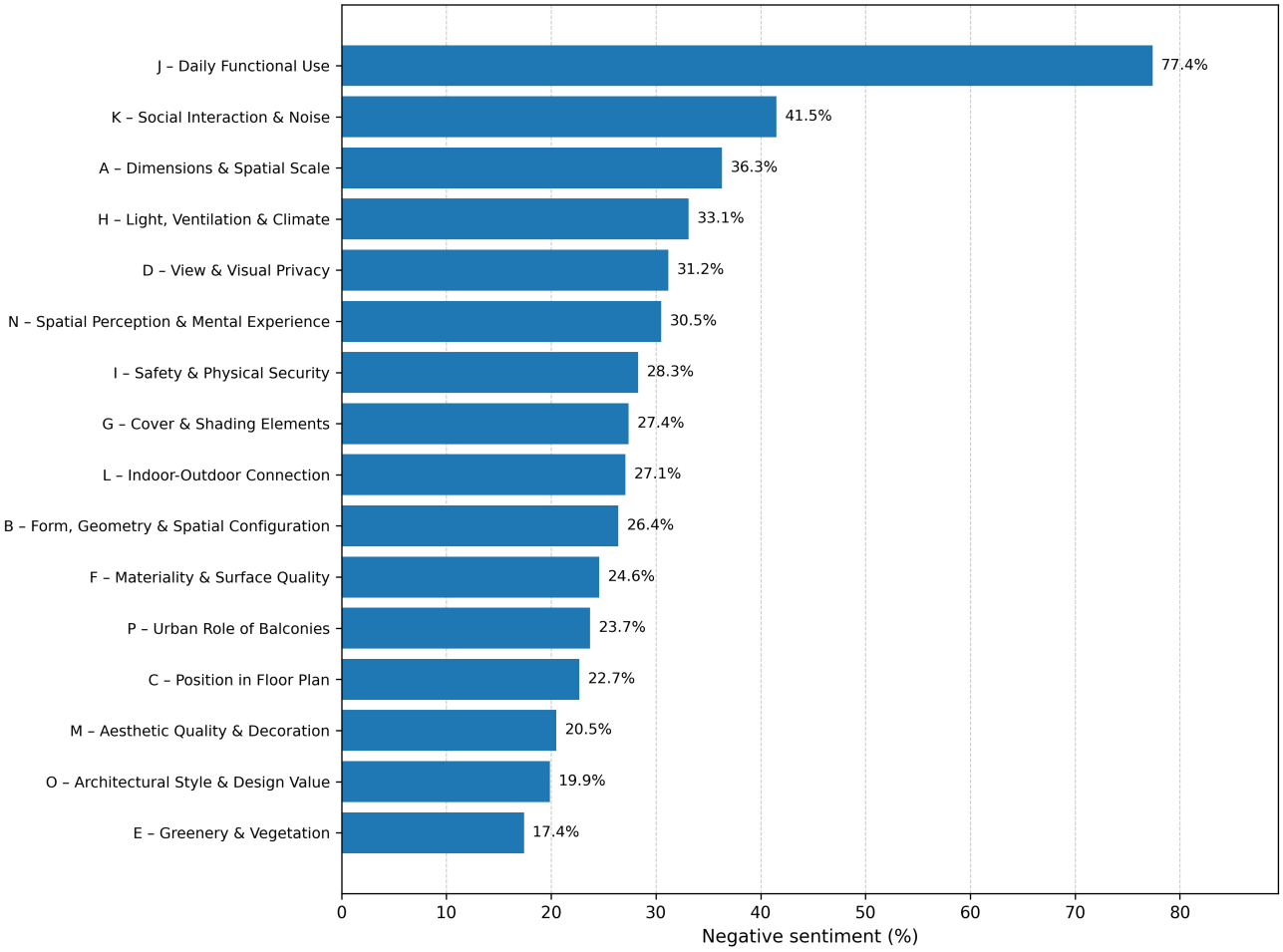
Figure 6 presents these same sentiment–category relationships as a Sankey diagram, tracing how thematic instances flow from each sentiment class into the 16 design categories.

4.4. Interpretation of key patterns

The combination of thematic classification and sentiment analytics reveals a layered and complex public perception of residential balconies. Several key patterns emerge that are instructive for both architectural theory and design practice:

1) **Functional constraints dominate negative sentiment**. Category J (Daily Use) shows the highest concentration of negative sentiment (77.4% negative tweets). Many tweets describe balconies as too narrow, cluttered, or non-functional and refer to secondary uses such as laundry drying, smoking, or storage. Together, these posts indicate that a substantial share of balcony-related discourse in our dataset frames balconies primarily through everyday functional constraints rather than leisure-oriented use. Because the dataset is not geographically disaggregated, this pattern should be interpreted as a

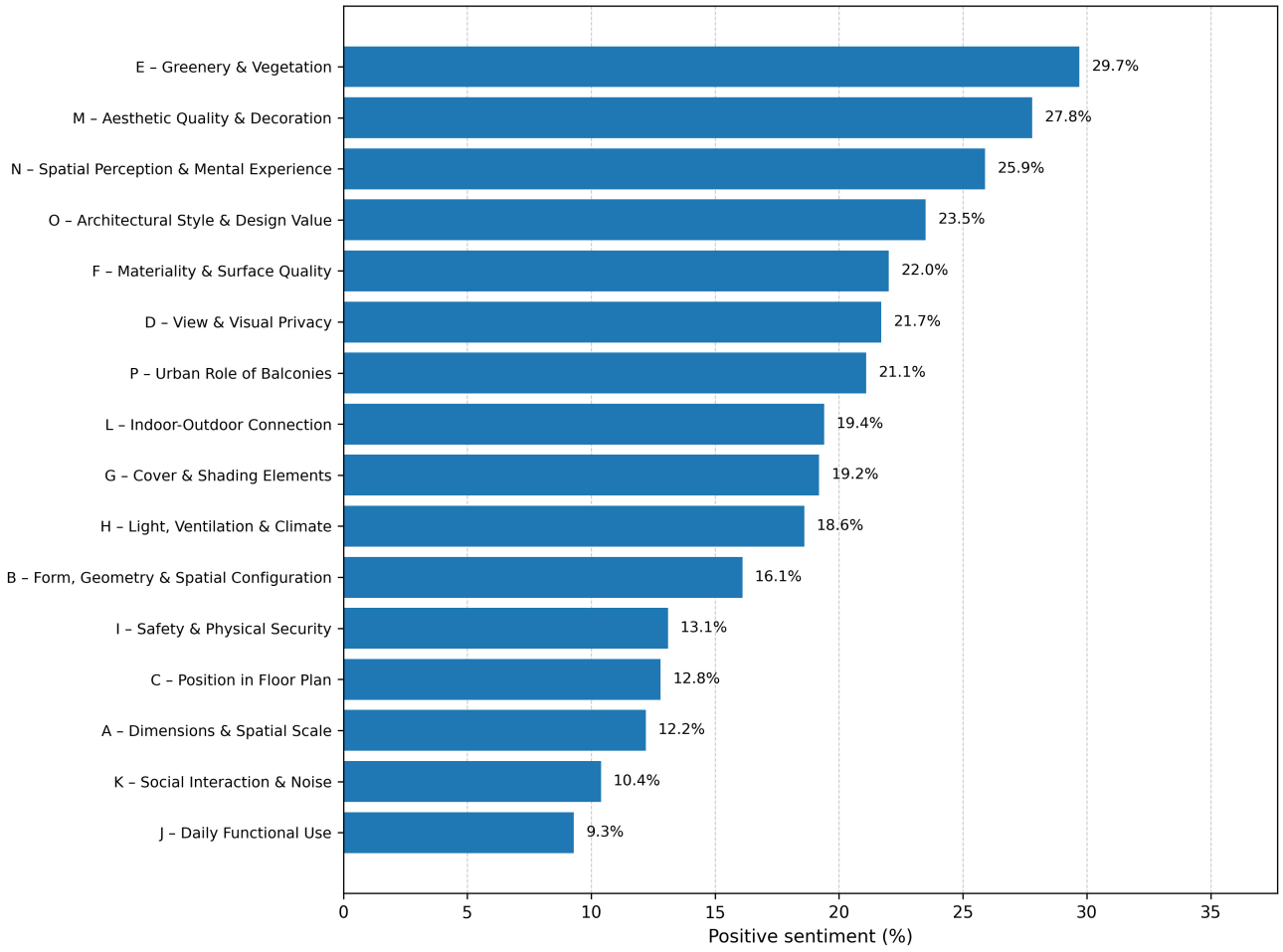
Figure 1
Distribution of negative sentiment across the 16 balcony design categories



corpus-level tendency rather than as a uniform condition across all Iranian cities, housing typologies, or climatic regions.

2) **Design sensitivity is appreciated.** In contrast, categories such as Greenery (E) and Esthetics (M) received strong positive reactions. Tweets praising balconies with vegetation, decorative detailing, or well-composed materials show that some users value balconies not only as functional spaces but also as areas for visual and sensory enjoyment in this dataset. This shows the significance of biophilic and affective design principles, even in small urban settings. However, positive comments about greenery and esthetics should not be read as the same in all places, since the dataset does not allow distinction of whether these preferences are more strongly expressed in specific urban or climatic contexts.
3) **Privacy remains contested.** Sentiment results for Category D (View & Privacy) were mixed. Many users appreciated open views but, at the same time, expressed concern about being seen by others, lack of screening, and too much exposure to neighbors. This mixed reaction reveals a common design conflict between openness and privacy. This prompts the exploration of context-sensitive interventions (e.g., shading, planting-based screening, or adjustable enclosure strategies) as possible design responses. This ambivalence about privacy can be understood as a recurring motif in the entire dataset, but the strength of the ambivalence might be different in parts of the dataset characterized by different cultural norms, densities of urbanization, and types of housing that cannot be untangled in the present dataset.
4) **Urban identity is underdiscussed.** Despite growing discourse on the balcony's role in urban form and façade articulation, Category P (Urban Role) had low tweet frequency and modest sentiment variation within this dataset. This pattern may indicate that balconies, in the minds of users, are understood primarily as personal or domestic spaces, rather than elements with public or civic identity. This detachment calls attention to the underutilized potential of balconies as contributors to urban vitality, ecological performance, and architectural storytelling—roles that remain unrecognized in public discourse. Because no regional or city-level breakdown is available in the present dataset, this limited attention to the urban role of balconies should be interpreted as an aggregate discourse pattern rather than a geographically invariant feature of public perception across Iran.
5) **Contradictions exist within single tweets.** The presence of mixed sentiment in multiple categories (e.g., F, O, D) indicates that users sometimes express positive and negative evaluations in the same post (e.g., "Nice view but absolutely no privacy from the neighbors."). This pattern suggests that balcony experience is discussed through multiple, sometimes competing lenses (practical, emotional, social, and environmental), and that design implications should be framed as trade-offs rather than one-dimensional fixes.

Figure 2
Distribution of positive sentiment across the 16 balcony design categories



Taken together, these findings suggest that when people talk about balconies on Twitter, they frame them as emotionally salient and socially negotiated domestic thresholds with recurring environmental themes running through the discourse. That's a meaningful signal for housing research and practice—though it comes with the caveat that social media evidence has real limits. These patterns should be read as aggregated tendencies within the observed Twitter corpus, not as geographically uniform statements about balcony experience across all Iranian residential contexts.

5. Discussion

This section explains the findings using architectural theory, environmental psychology, and human-centered design.

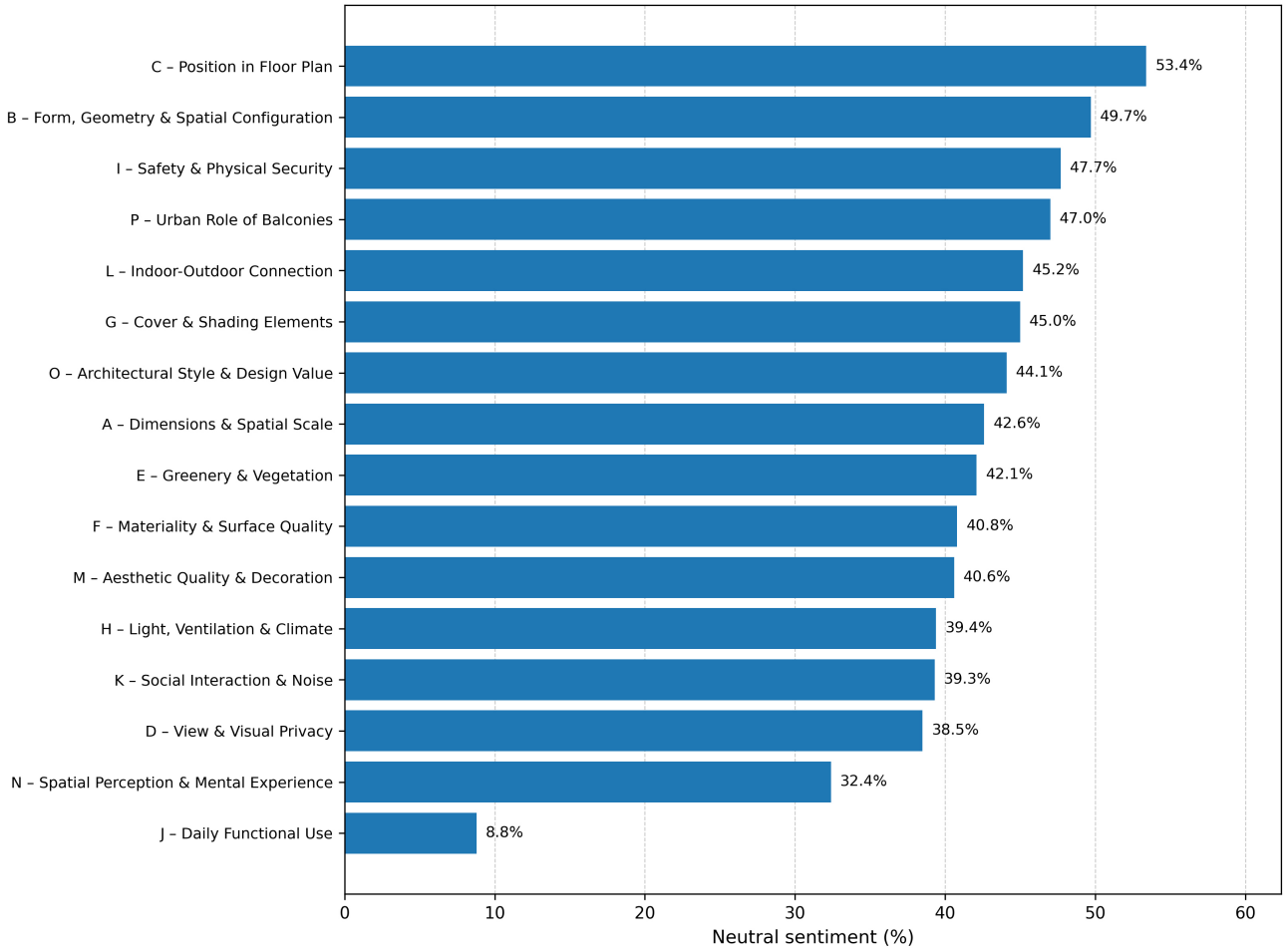
Based on the analysis of 10,000 selected Persian-language tweets, the study offers an unprompted view of how users describe their balcony experience in emotional, spatial, and social terms—within the limits of this dataset. By comparing these results with common design approaches, several important themes appear. To help explain these themes, Figure 7 shows a combined sentiment profile for all 16 design categories. The radar chart shows that functional categories like Daily Use (Category J) and Dimensions & Scale (Category A) often come with negative sentiment. In contrast, sensory and emotional categories like Greenery (Category E) and Esthetics (Category M) are more often linked to positive sentiment. Neutral and mixed responses fall in between, showing uncertain or situation-based feelings.

5.1. Reframing balcony design through user-centered evidence

The findings suggest a potential mismatch between some emphases in prevailing balcony design discourse (e.g., typological integration, façade articulation, climatic responsiveness) and the concerns that users most frequently articulate in everyday talk about balconies [8, 67]. Within this dataset, residents often evaluate balconies in terms of day-to-day usability, emotional comfort, and social frictions, rather than formal or environmental performance alone. This user-oriented emphasis is consistent with prior studies showing that residents tend to evaluate balconies primarily through everyday use, social norms, and adaptability in practice rather than through formal design intentions alone [8, 30, 31].

This is reflected in the sentiment patterns: Category J (Daily Use)—covering routine practices such as smoking, drying clothes, or sitting outdoors—shows the highest share of negative expressions (77.4%). These complaints show dislike not with the idea of the balcony itself but with felt limits on usability, tight space conditions, and frictions with nearby units. Examples include discomfort said to come from secondhand smoke affecting close apartments, worries about visual intrusion due to low screening, and balconies described as utility zones instead of thoughtfully

Figure 3
Distribution of neutral sentiment across the 16 balcony design categories



designed spaces. Taken as user-reported views, these patterns may show shortfalls in spatial programming and socioenvironmental sensitivity, rather than small mistakes in detailing or looks. Similar tensions between how balconies are meant to be used and how people actually use them show up in post-occupancy and context-based studies of apartment housing too. People often end up dissatisfied because their balconies are too narrow, because they have to juggle different needs in a small space, or just because of issues with the neighbors [13, 29, 43].

The results support a design framing in which balconies are not treated only as decorative or secondary appendages. Rather, they should be treated as functionally and socially sensitive semi-private zones. Using user experience evidence may help create more adaptive, context-aware, and meaningful design responses. An example of this evidence is unsolicited social media narratives. At the same time, the present corpus does not allow city-level or climate-zone-level disaggregation. Thus, these design implications should be seen as indicative and context-sensitive rather than universally representative of residential conditions across Iran.

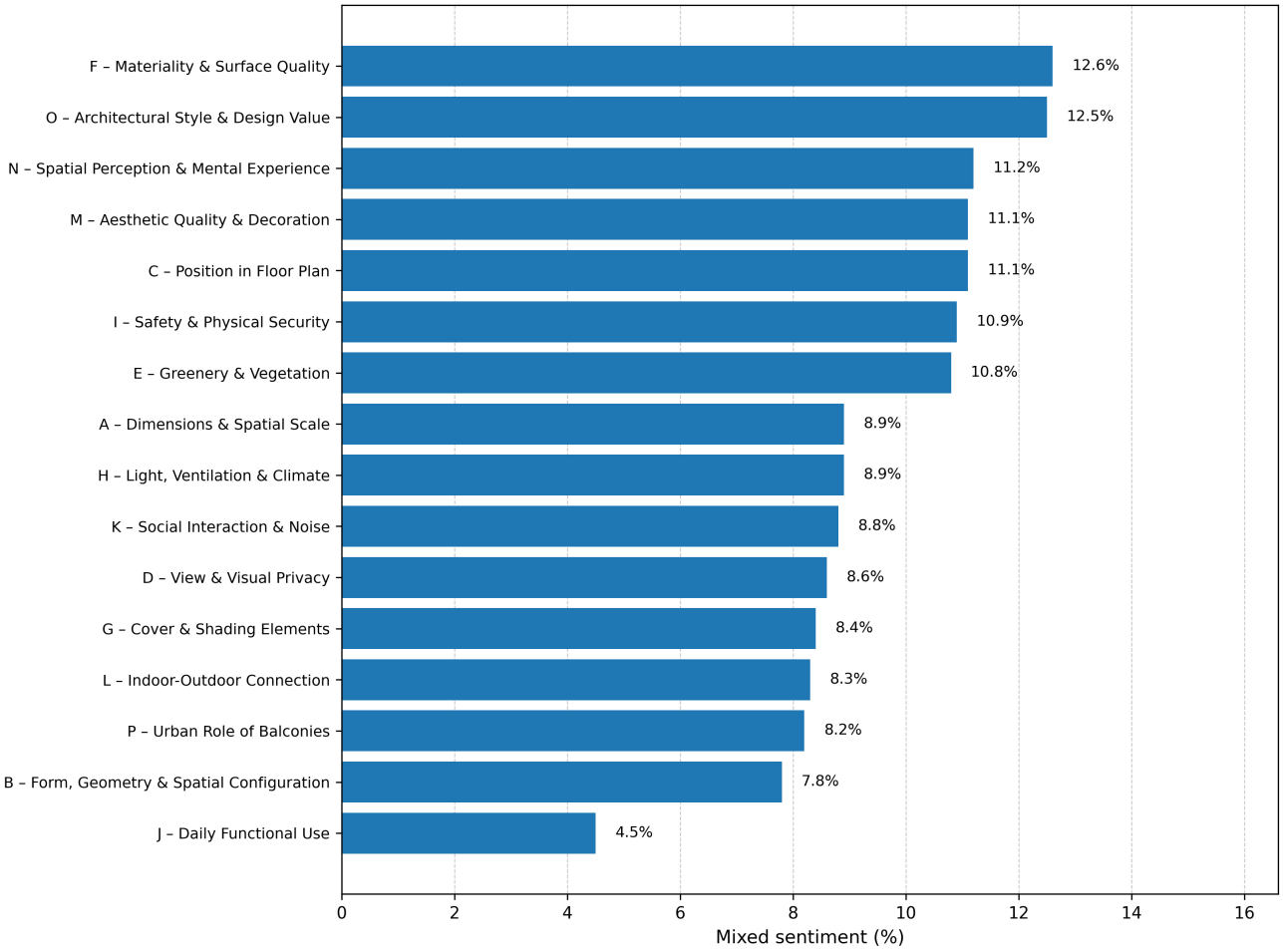
5.2. The role of emotional and sensory design

The categories with comparatively higher positive expressions—particularly Greenery (Category E) and Esthetics (Category M)—highlight how users' balcony talk frequently foregrounds sensory appeal, affective comfort, and biophilic elements. Tweets often praise balconies with plants, decorative features, natural textures, or pleasing views, suggesting that—even in small doses—such elements are associated with more favorable user descriptions of balcony experience. These patterns resonate with Kaplan and Kaplan's argument that even modest incorporations of greenery and environmental variety can enhance environmental preference, attachment, and restorative benefits [68]. Similarly, Ulrich's Theory of Supportive Design (1997) argues that spaces offering natural elements, esthetic softness, and opportunities for retreat are more likely to reduce stress and promote psychological well-being [69]. While not a causal test, the discourse patterns observed here are broadly consistent with these theoretical perspectives.

This interpretation is also consistent with recent sustainability-oriented balcony research showing that balcony geometry, shading capacity, and adaptive configuration can influence daylighting, thermal comfort, and building energy performance across different façade conditions [10, 11].

Within this dataset, Category E (29.7% positive) and Category M (27.8% positive) show relatively higher positive sentiment compared to most other categories, indicating that greenery and esthetic qualities are recurrently discussed in favorable terms. This may signal that users often value balconies not only for technical adequacy but also for emotional and experiential qualities. From the perspective of landscape architecture, even minimal

Figure 4
Distribution of mixed sentiment across the 16 balcony design categories



green interventions—such as potted plants, vertical trellises, or hanging baskets—have been shown to enhance perceptions of privacy, serenity, and domesticity [69, 70]. Taken together, the discourse patterns here suggest that "soft" experiential attributes may be underweighted in performance-driven balcony design routines, despite their salience in everyday user talk.

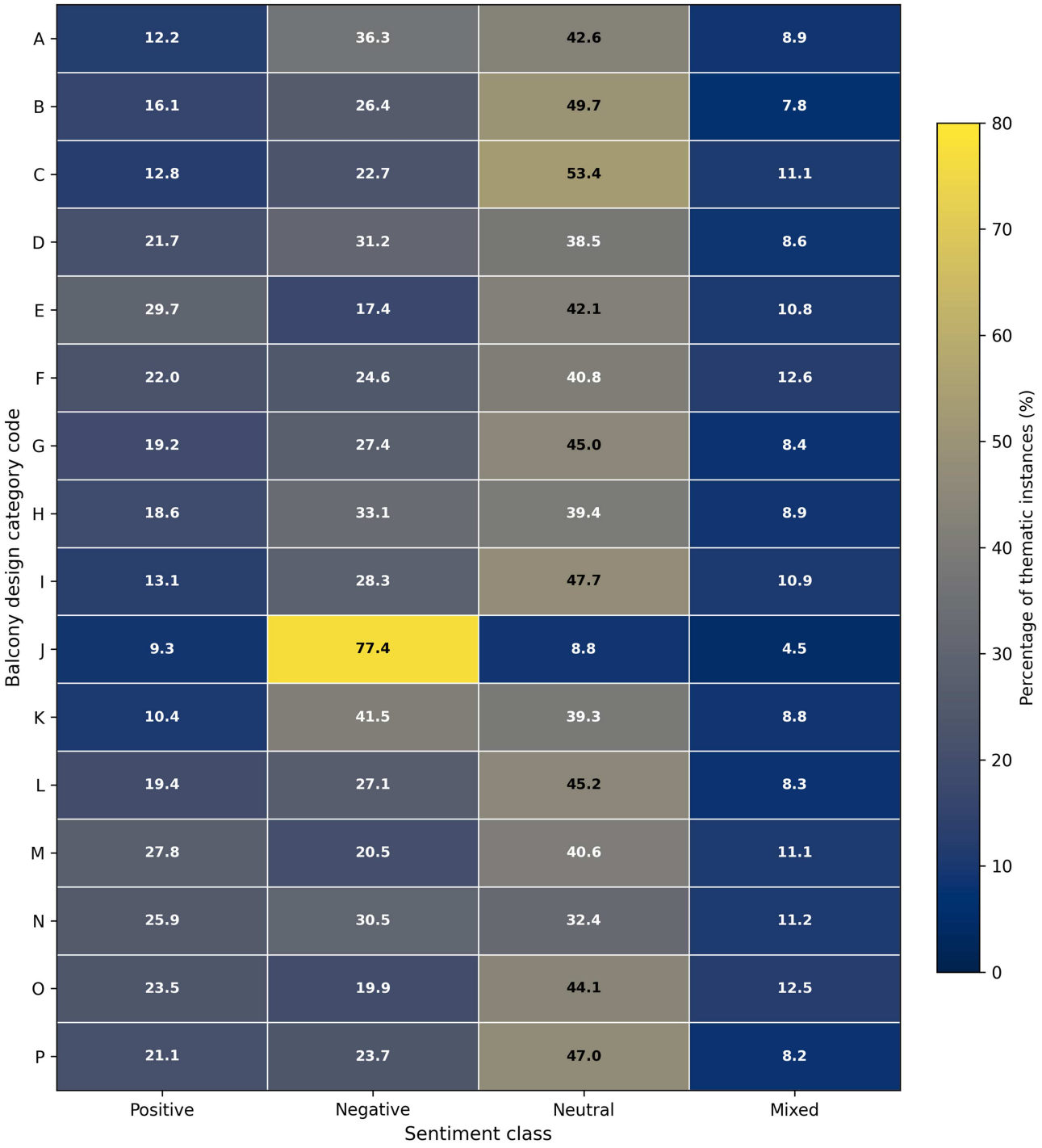
5.3. Cultural context and the semi-private dilemma

In the Iranian context, balconies occupy a uniquely ambivalent spatial and social position. As architectural extensions of private domestic life, they physically belong to the household; yet, by virtue of their exposure to the exterior, they exist on the threshold between public and private domains [71]. This semi-private status introduces complex cultural and behavioral dynamics that deeply influence how balconies are perceived, inhabited, and emotionally experienced [72]. In this discourse, concerns around privacy kept coming up in the tweets—people talked about being watched, about neighbors being too close for comfort, and how uncomfortable it is to feel exposed in crowded apartment buildings. These anxieties tie back to higher-level questions around modesty, privacy, and working out boundaries in daily life, all of which run deep in Iranian social norms [71]. Sometimes, standard building designs—what might be called typological defaults—like fully open balconies or ones with glass railings, just do not match what makes people feel socially comfortable. When this happens, residents might avoid using their balconies altogether or try fixes like putting up makeshift screens or using retrofitted concealment practices. Instead of assuming that a completely open balcony is always the perfect design for everyone everywhere, the discourse shows the need to frame balcony design as a culturally responsive negotiation of visibility and control. In simple terms, it is all about finding a healthy balance between how much people want to see outside and how much privacy they want from their neighbors, based on their local culture. Design moves such as adjustable shading, layered planting for visual filtering, perforated screens, or modular enclosure systems can give residents more control over being seen, letting balcony use change with time of day, activity, and felt social comfort. Importantly, the return of privacy-related expressions (including both bad and mixed sentiment) shows that privacy works not only as a geometric parameter but also as an affective condition that shapes whether a balcony feels welcoming or stressful from the start. Taken together, these patterns make a quite strong case for moving away from fixed, form-driven answers and toward more flexible setups that balance outside connection with optional enclosure.

5.4. Implications for future design practice

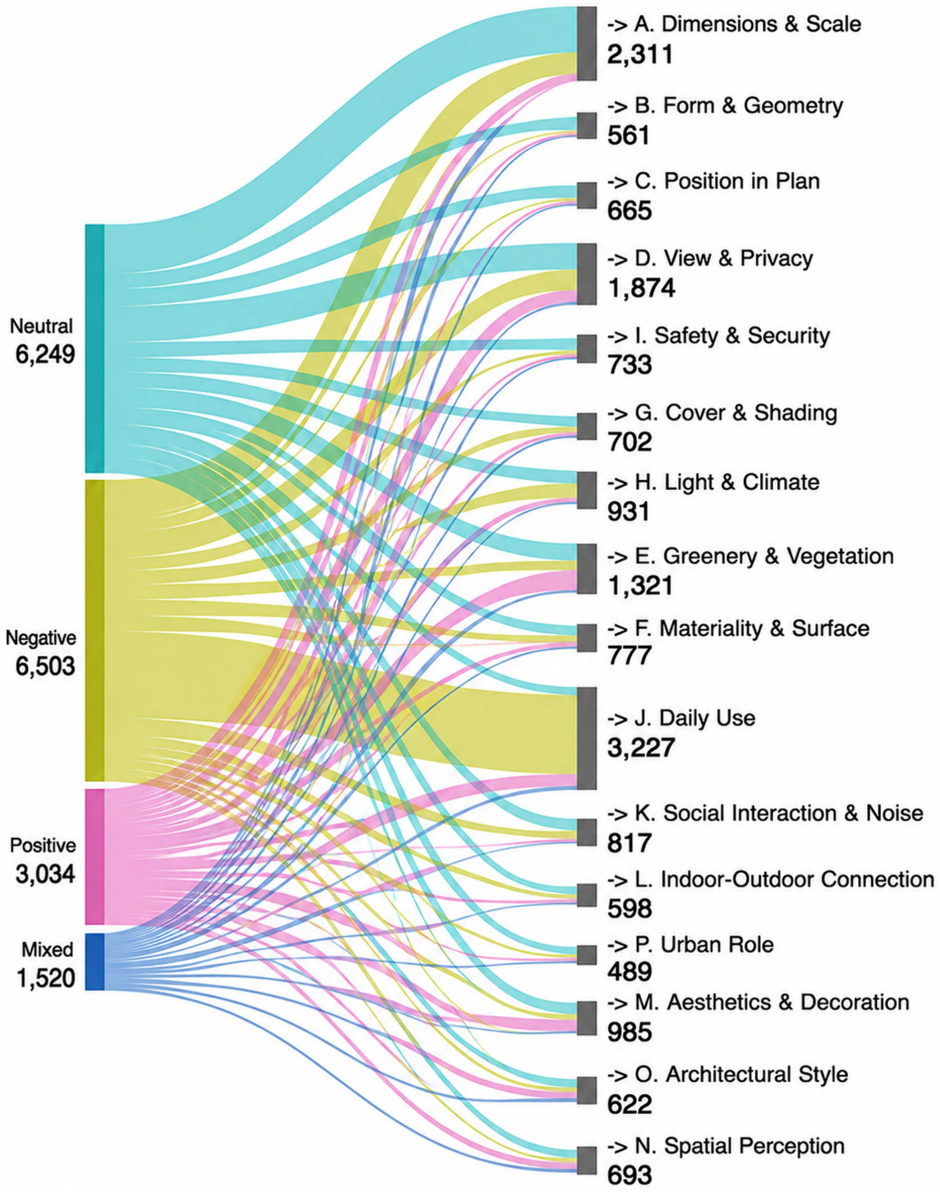
Figure 8 presents a Jaccard-based co-occurrence heatmap of balcony design categories, showing how often two features are mentioned together relative to their total appearances. The

Figure 5
Heatmap of sentiment distribution across the 16 balcony design categories



strongest association is between H (Lighting & Ventilation) and K (Social Interaction/Noise) ($J = 0.26$), indicating that roughly one quarter of mentions involving either category also invoke the other. This co-mention pattern suggests that discussions of environmental openness are frequently articulated alongside concerns about social exposure and noise, rather than in isolation. A second, smaller association appears between M (Esthetics) and P (Urban Role) ($J \approx 0.13$), indicating that tweets about appearance sometimes also reference façade or neighborhood image. A third link, H–N (Lighting & Ventilation ↔ Spatial Perception, $J \approx 0.09$), suggests that references to daylight and airflow are occasionally discussed alongside perceptions of spaciousness. Several moderate ties cluster around $J \approx 0.05$–0.06—including D–P (0.064), L–N (0.063), A–L (0.060), D–E (0.056), H–J (0.055), and A–D (0.054)—while most other pairs remain weak. Notably, F–P is low ($J \approx 0.03$), indicating limited co-mention between materiality and the urban/collective role of balconies in this dataset. Overall, these patterns indicate that residents do not perceive physical qualities in isolation: concerns about airflow and sunlight often intertwine with social dynamics (privacy/noise), and esthetic

Figure 6
Sankey diagram illustrating flows between sentiment classes and balcony design categories



appeal sometimes aligns with perceptions of the balcony's role in the urban façade.

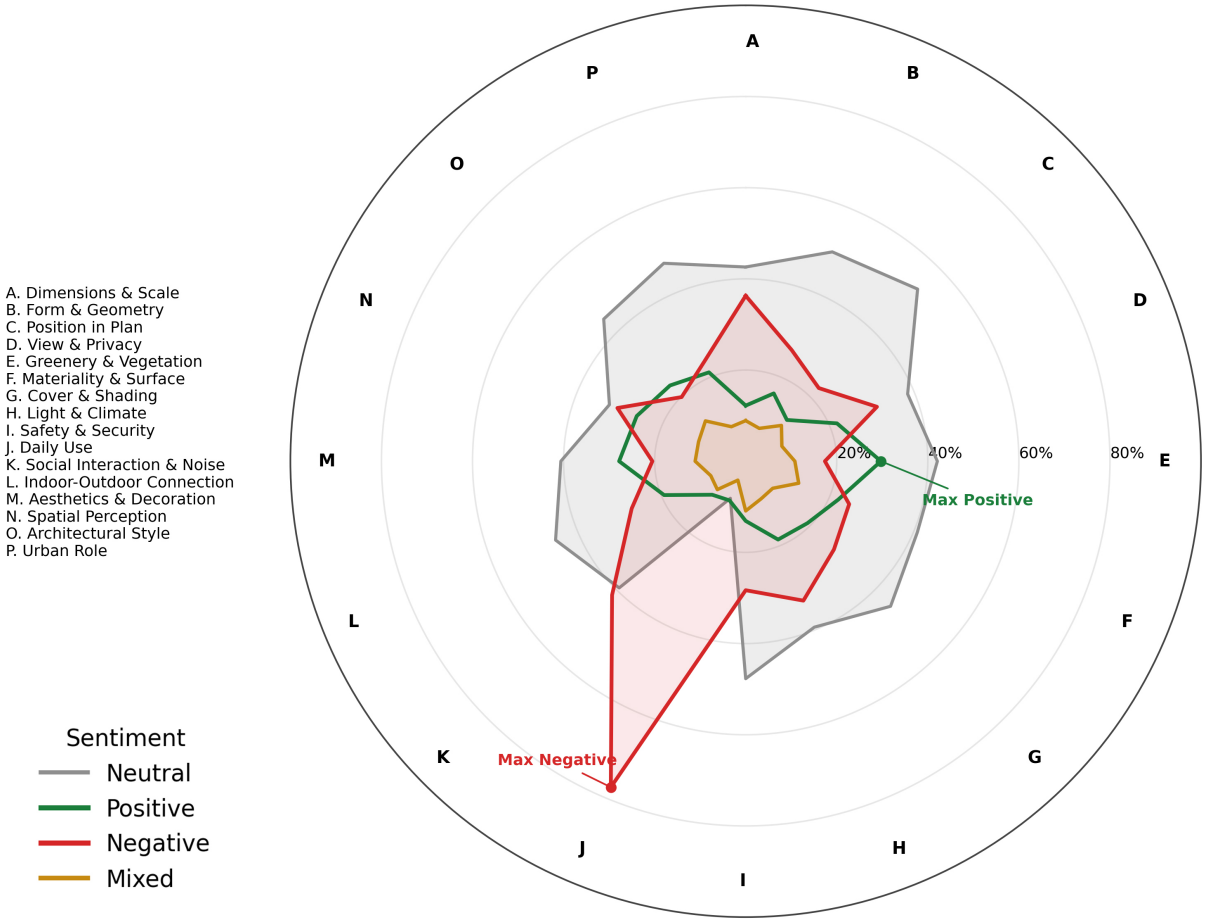
5.5. Design implications from social media evidence

By leveraging unsolicited, user-generated content, this study illustrates the potential of social media discourse as a complementary source of POE—one that can surface emotive, unprompted, and time-sensitive descriptions of lived spatial experience within the scope and limits of the dataset. This methodological positioning is further supported by recent studies applying social media text mining and NLP to sustainability-related building discourse and smart-city decision contexts, suggesting that digitally generated textual data can serve as a meaningful supplementary evidence base for built-environment evaluation [53, 54].

Unlike structured interviews or survey instruments, which are often mediated by predefined questions and academic assumptions, social media commentary may capture spontaneous, context-rich reflections on how people describe using, adapting, or struggling with domestic architectural spaces. At the same time, Twitter is best positioned as a complement rather than a replacement for conventional POE. Social media discourse shows what users choose to express. The expression is often spontaneous. But it reflects a self-selecting population. It does not give the demographic, spatial, or behavioral controls that field-based methods usually give. Therefore, the patterns mentioned here should be taken as showing signs that can extend the proof base, bring up less-talked-about worries, and help create design-relevant ideas to test. These signs can then be made stronger through follow-up real-world work (e.g., targeted POE, interviews, or in situ observation) to test how common they are, how they work, and context-specific limits.

The findings motivate a shift of balcony design goals toward adaptive, context-aware, and evidence-informed ways (Figure 9). These recurrent themes and co-mentions in the discourse suggest a number of design-relevant directions:

Figure 7
Radar chart showing the integrated sentiment profile across the 16 balcony design categories

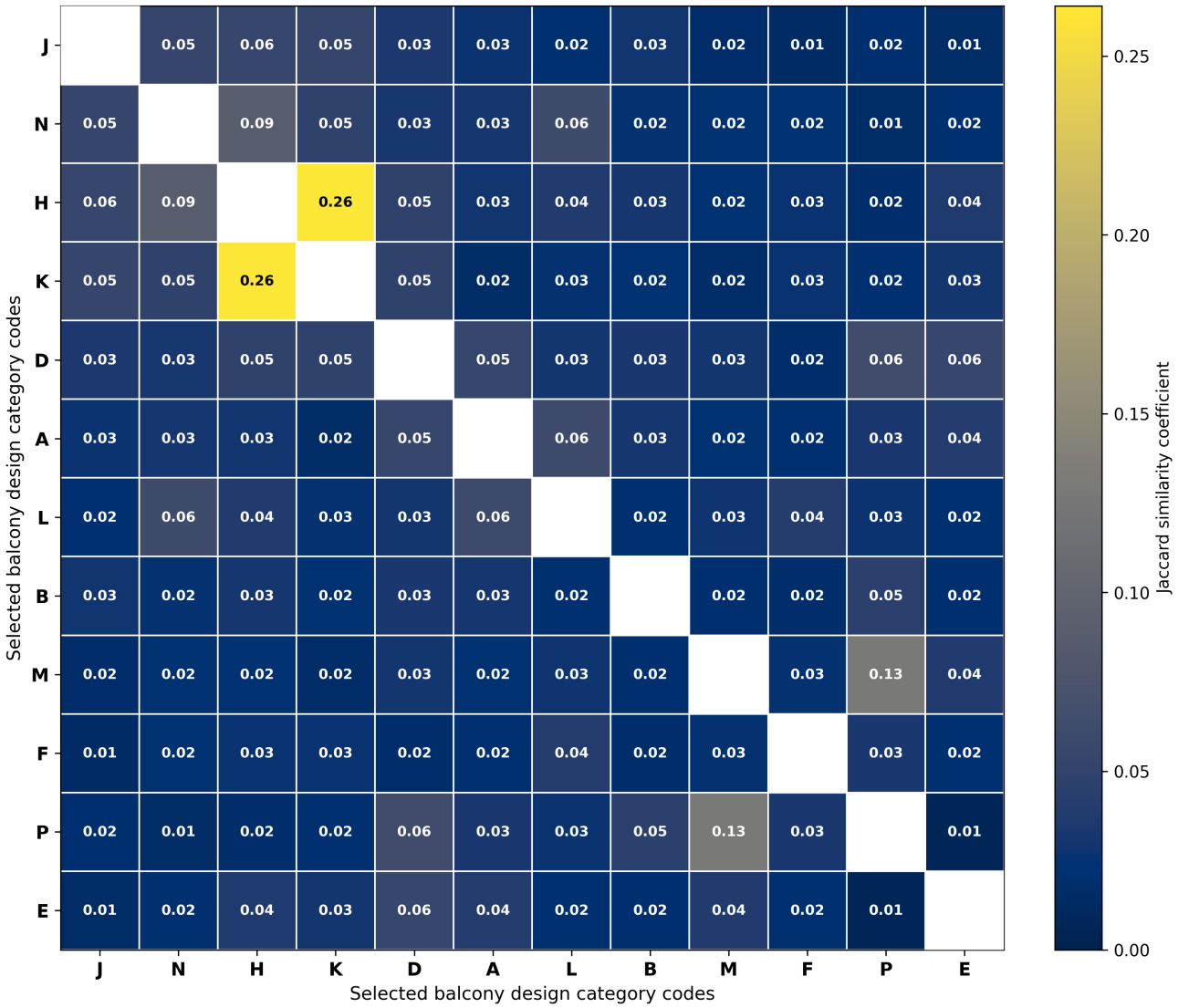


1) **User agency in spatial modulation**: Balconies should permit resident control over exposure, granting people the ability to modulate privacy, ventilation, and shading, as environmental and social conditions evolve. Fixed, one-size-fits-all solutions do not do well in culturally diverse and spatially dense urban contexts.
2) **Green infrastructure integration**: Whether it's a few potted plants or a full vertical garden, vegetation needs to enter the design conversation early and not get bolted on at the end as decoration. It pulls real weight as a core spatial feature with thermal comfort, emotional well-being, and visual buffering. That's too much to leave as an afterthought.
3) **Recognition of everyday rituals and affective use**: The tweets also make clear that balconies carry a lot of everyday emotional weight—rest, reflection, the occasional small gathering. That matters for design. Experiential affordances need to sit alongside technical adequacy: seating niches, micro-layout flexibility, lighting that actually responds to how people use the space. These details, in other words, signal the space was designed for living in, not just built to code.
4) **Social tension mitigation**: In dense housing, the talk again and again links balconies to frictions around smoke, noise, and visual intrusion. Design answers may include buffering strategies (setbacks, partitions, planting screens) and clearer boundary-making to lower the chance of conflict in the first place.

Together, these implications suggest treating the balcony not as a side leftover but as a key semi-private interface where privacy, comfort, and everyday social life are handled. Importantly, because this study uses Twitter as the only main data source, these implications should be understood as tied to the context, exploratory, and complementary to current literature. They are useful for forming design hypotheses and guiding stakeholder discussions, but they also call for further validation through additional data sources and on-site evaluation.

Methodological and Contextual Limitations. It should be noted that this study has some limitations. Most importantly, the dataset does not allow us to consider the geographic distribution of tweets across Iran's different urban and climatic regions. Iran has a very wide range of climatic zones—from the humid subtropical Caspian littoral around Rasht to the semi-arid central plateau of Isfahan and Yazd, the cold mountainous northwest around Tabriz, and the hot-humid Persian Gulf littoral near Ahvaz and Bandar Abbas—as well as great variation in urban density, housing typology, and socioeconomic conditions. The design of balconies, patterns of use, and expectations of residents are very likely to be very different in these contexts: a balcony in a high-rise apartment block in Tehran serves different functional and cultural roles than a balcony in a low-density courtyard house in Shiraz. By treating Iran as a uniform context without looking into various areas separately, the analysis risks obscuring regionally specific patterns. This can result in mixing up observations

Figure 8
Jaccard-based co-occurrence heatmap of selected balcony design categories

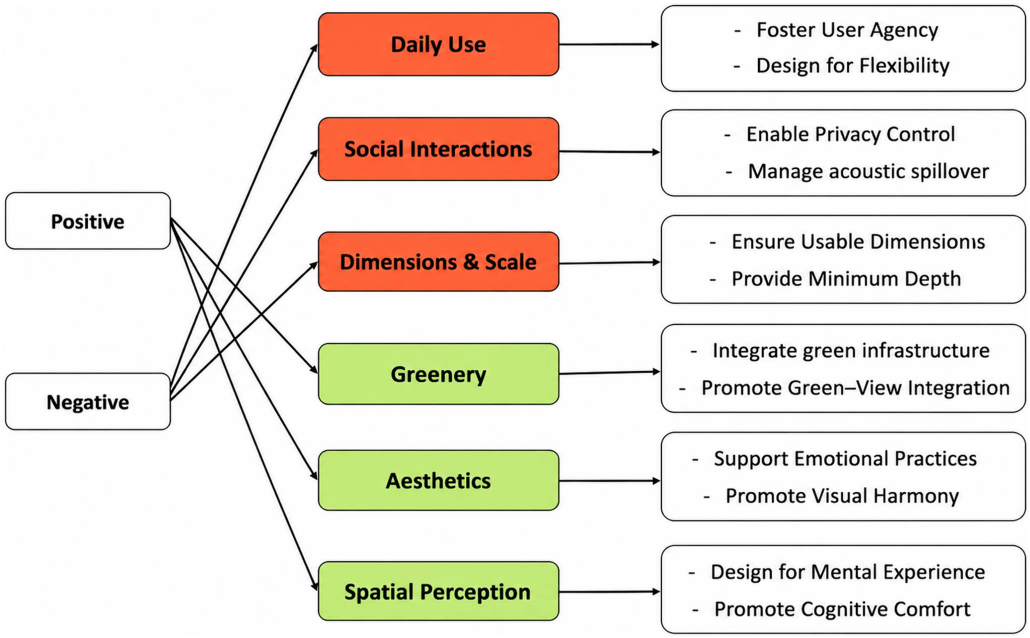


from places with totally different climates, and it might end up making general statements that don't really fit every part of the country. For any future work, it would be a good idea to add geolocation filtering. That way, the analysis could be done at a city or regional level, and it could also take into account factors like climate zone, housing type, and local cultural norms. Second, the dataset includes a group of people who chose to be in it themselves—Persian-speaking Twitter users who live in cities, are younger, and use digital technology a lot—and this group does not represent all the people who use balconies in Iran. Third, the lexicon-based sentiment method, even though it's easy to understand and works well for this specific task, doesn't perform as accurately in classification as models like ParsBERT or XLM-RoBERTa that have been fine-tuned. This is a real methodological limitation, and future work should look for model-based alternatives as annotated Persian architectural corpora become available. Fourth, although the thematic coding was done by human coders with an appropriate evaluation of inter-coder reliability ($\kappa = 0.81$), the category framework itself reflects the theoretical assumptions of the research team. It may not contain all the balcony experience elements that users really appreciate.

5.6. Toward transferable insights

The dataset is firmly within the Iranian urban context, but some of the themes that emerged—tensions around privacy, emotional and ritualized use, appreciation of greenery, discomfort in high-density housing—translate fairly clearly to other architectural and cultural environments. These are not to be taken as universal statements, but they do describe recurring problems that seem to be related to semi-private threshold spaces wherever they arise. The methodological framework itself—combining social media mining, thematic classification, and sentiment analysis—could be transferred to other spatial typologies, such as courtyards, rooftop terraces, and shared circulation spaces, and to other cultural contexts where locally specific concerns about privacy norms,

Figure 9
Conceptual mapping linking dominant sentiment patterns to balcony design implications



climate adaptation, and social interaction need to be surfaced. What is particularly useful about this type of approach is that unsolicited, emotionally expressive user discourse captures affective and experiential dimensions that structured instruments tend to underrepresent, which makes it a genuinely complementary perspective to conventional POE rather than a replacement for it.

Taken together, the study contributes to an ongoing disciplinary shift toward incorporating experience-oriented evidence into architectural thinking—while recognizing that broader generalization requires triangulation with additional empirical methods and contexts.

6. Conclusion

This study offers a data-informed rethinking of balcony design by mining over 10,000 Persian-language tweets and examining user sentiment across spatial, material, and experiential dimensions. The researchers combined thematic classification and sentiment analysis to discover how residents actually describe their functional and emotional engagement with balconies—often in terms that rarely make it into conventional feedback surveys. The findings stress the need to avoid looking at balconies as residual or purely esthetic appendages. What emerges instead is a picture of these spaces as emotionally charged, socially negotiated, and spatially consequential micro-environments in the fabric of everyday domestic life. They are important spots where daily life, emotions, and social needs happen.

6.1. Design insights informed by user discourse

The study extracted a set of design-relevant insights from unsolicited, user-generated material that reflects the lived conditions, frustrations, and aspirations balcony users actually express. These are not rules but patterns that emerge in the way balconies are experienced and talked about, and they provide a human counterpoint to formal typological or regulatory methods.

1) **User-centered functionality:** The social media discourse shows that people mostly talk about using their balconies for everyday tasks—like drying laundry, smoking, or finding room for basic seating and storage. This shows that residents do not just value balconies as pretty, scenic extensions. Instead, especially in crowded apartments, balconies are contested functional spaces where different everyday needs compete for the same small area. To realize these pressures requires an honest recognition of competing uses rather than idealizing balconies as passive leisure zones.
2) **Emotional and Sensory Integration**: People consistently show a much more favorable sentiment when they talk about plants, decorations, and climate-sensitive detailing. These features should never be dismissed as just optional embellishments. Instead, they are essential parts of the physical space that create real comfort, build a strong emotional attachment, and greatly improve perceived domestic quality.
3) **Privacy modulation:** Balconies live on the edge between public and private. Lots of people mention such a tension, especially where privacy connects with dignity and cultural norms. Flexible strategies like screens, planting, or adjustable enclosures are a more useful response as they give residents actual agency over their own level of visibility, rather than fixed degrees of openness.
4) **Social and Environmental Resilience**: As a threshold space, balconies mediate between the individual and the urban collective, and between interior microclimates and exterior environmental conditions. Their design must seriously consider the intrusion of noise, air quality, and thermal comfort and treat balconies as resilient, livable interfaces of the broader housing ecology, not residual appendages that happen to be attached to apartments.

All of the above implies that balcony design could benefit a lot from closer attention to lived routines, emotional comfort, and social negotiation—not as an alternative to formal and environmental concerns, but complementary to them.

6.2. Toward a responsive balcony design framework

In response to these patterns, the study sketches out a flexible, user-informed framework for balcony design. Such a framework tries to balance practical space needs with how people feel and experience the space, while making room for both shared design criteria and cultural expectations. The framework is structured around four interrelated domains:

1) **Spatial Dimensions and Scale.** User discourse suggests that excessively narrow or symbolic balconies often fail to support basic activities such as seating, planting, or movement. Treating the balcony as a functional extension of domestic space—rather than a façade gesture—appears central to its everyday viability.
2) **Materiality and atmosphere**. Beyond durability and performance, material choices influence sensory perception and emotional tone. Users often respond to material continuity between interior and exterior spaces, as well as to textures and finishes that support comfort and visual coherence.
3) **Connection to Green and View**. Visual and ecological connections are frequently discussed in positive terms. Balconies should either offer direct access to planting elements—such as railing planters, vertical gardens, or integrated soil beds—or be oriented toward natural vistas when site conditions permit. These features promote biophilic engagement, mental restoration, and environmental awareness, even in compact urban settings.
4) **Psychological Flexibility and Use.** Balconies are often places for personal routines like having coffee, reading, thinking, or just taking a break from inside [30]. People often talk about them as spots for being alone, reflecting, or taking short pauses in the day. So, the design needs to offer adaptable space conditions to respond to different emotional states, times of day, and user moods. Movable shading, adjustable seating, and surfaces that can flex between functions aren't luxuries. They're what makes a balcony genuinely usable across different moments in a day. Flexibility enhances psychological ownership and personalization, empowering users to configure the space in line with evolving needs [73].

This framework is intended as a heuristic tool rather than a fixed template, supporting design exploration rather than replacing context-specific judgment.

6.3. Reconsidering small spaces through social media evidence

The analysis shows that large-scale social media data, which users post without being asked, can give a helpful new view for architectural study. This is especially true for capturing emotional, everyday descriptions of how people use space. Those kinds of details are hard to get from structured evaluations alone. In this study, balconies—which are often seen as unimportant in design talks—come across as places where privacy, routine, identity, and environmental conditions all meet in real ways. This kind of data cannot replace standard POE. But it can help bring up hidden concerns, show repeated conflicts, and suggest ideas for more empirical research later.

6.4. Future research directions

This research sets up a methodological and conceptual base for more interdisciplinary work at the intersection of architecture, data analysis, and user-centered design. Some potential avenues include:

1) **Multimodal Analysis**: Integrating visual content from image-based platforms such as Instagram can complement textual sentiment data with information on spatial esthetics, material expression, and personalization behaviors. This would provide a more comprehensive perspective on how balconies are experienced, decorated, and disseminated within digital culture.
2) **Geospatial Mapping**: Linking sentiment-labeled tweets to urban morphology, housing typologies, or environmental variables could open up a much richer picture of where balcony satisfaction and dissatisfaction actually cluster spatially. Geolocated analysis might, for instance, help correlate specific design attributes—orientation, density, and greenery coverage—with emotional outcomes in ways that aggregate datasets simply can't.
3) **Simulated User Behavior**: Combining the preferences derived from sentiment analysis with agent-based models or immersive simulations (i.e., virtual reality environments) is recommended. This might enable researchers to investigate the impact of various balcony design changes on the occupancy of the space, privacy negotiation, and emotional engagement, especially across varying cultural or climatic settings.
4) **Cross-Cultural Comparisons**: Running the same text mining and sentiment analysis framework on datasets in other languages and cultural contexts could do two things at once: reveal which design drivers seem to show up regardless of cultural context and expose where expectations around semi-private outdoor space are genuinely culturally contingent rather than universal. Such comparisons may support the development of more globally attuned balcony design guidelines.

Together, these directions highlight how social media–based analysis can function as a supplementary lens—broadening the evidentiary base for architectural decision-making while remaining grounded in real-world perceptions and lived experience.

Ethical Statement

The study was reviewed by the Ethics Committee for Human Research, University of Art, Tehran, and granted an exemption from full ethical review (Ref. IR.AU.REC.1404.568) on November 10, 2025. The study analyzed publicly accessible Twitter/X posts without recruitment, direct interaction, or intervention with human participants. All procedures complied with relevant ethical guidelines, privacy requirements, and applicable platform policies.

Conflicts of Interest

The authors declare that they have no conflicts of interest to this work.

Data Availability Statement

The supplementary datasets supporting the findings of this study are provided with the submission as Supplementary Dataset S1 (Tweets with Sentiment, bilingual) and Supplementary Dataset S2 (Validation Results, English). Additional tweet-level metadata can be made available by the corresponding author upon reasonable request, subject to ethical and platform constraints.

Author Contribution Statement

Amir Shakibamanesh: Conceptualization, Methodology, Formal analysis, Investigation, Data curation, Writing – original draft, Writing – review & editing, Visualization, Supervision, Project administration. **Seyedziya Hosseinzadeh:** Methodology, Validation, Formal analysis, Data curation, Writing – review & editing.

AI Disclosure

ChatGPT (OpenAI) was used only for language refinement and editorial polishing of selected parts of the manuscript. The authors take full responsibility for the study design, analysis, results, and final content.

Supplementary Materials

The supplementary materials are available from the article webpage at: https://doi.org/10.47852/bonviewJDSIS62028057.

References

[1] Bayazit Solak, E., & Kisakurek, S. (2023). A study on the importance of home and balcony during the COVID-19 pandemic. *Environment, Development and Sustainability*, *26*(10), 26345–26363. https://doi.org/10.1007/s10668-023-03732-w

[2] Dictionary.com. (n.d). *Balcony*. Retrieved December 24, 2025, from https://www.dictionary.com/browse/balcony

[3] Collins English Dictionary. (n.d). *Balcony*. Retrieved December 24, 2025, from https://www.collinsdictionary.com/dictionary/english/balcony

[4] Emekci, Ş. (2021). Balcony: A remembered architectural element amid pandemic: Evidences from digital media. *İDEALKENT, COVID-19 Sonrası Kentsel Kamusal Mekânların Dönüşümü*, *12*, 609–630. https://doi.org/10.31198/idealkent.870385

[5] Hasan, H. A., Ismail, K. J. A., & Shekho, I. O. (2020). Impact of design requirements for facade formation on the functional performance of the balcony space in residential apartments. *IOP Conference Series: Materials Science and Engineering*, *928*(2), 022114. https://doi.org/10.1088/1757-899X/928/2/022114

[6] Ek, F. İ., Can Traunmüller, I., & Yaylalı, B. (2020). Balcony as an intermediate space: A prosthetic threshold in the domestic-urban life of izmir. *Meltem İzmir Akdeniz Akademisi Dergisi*, *7*, 6–27. https://doi.org/10.32325/iaad.2020.1

[7] Philokyprou, M., Michael, A., & Malaktou, E. (2021). A typological, environmental and socio-cultural study of semi-open spaces in the Eastern Mediterranean vernacular architecture: The case of Cyprus. *Frontiers of Architectural Research*, *10*(3), 483–501. https://doi.org/10.1016/j.foar.2021.03.001

[8] Smektała, M., & Baborska-Narożny, M. (2022). The use of apartment balconies: Context, design and social norms. *Buildings and Cities*, *3*(1), 134. https://doi.org/10.5334/bc.193

[9] Akbari, P., Yazdanfar, S.-A., Norouzian-Maleki, S., & Hosseini, S.-B. (2025). Balconies' visual preferences: The case of residential apartments in Tehran, Iran. *PLOS One*, *20*(7), e0324874. https://doi.org/10.1371/journal.pone.0324874

[10] Shamseldin, A. (2023). Adaptation opportunities for balconies to achieve continuity of their environmental functions. *Alexandria Engineering Journal*, *67*, 287–299. https://doi.org/10.1016/j.aej.2022.12.037

[11] Loche, I., Loonen, R., & Neves, L. O. (2024). Balcony design recommendations to enhance daylight, thermal and energy performance of mixed-mode office buildings. *Energy and Buildings*, *321*, 114702. https://doi.org/10.1016/j.enbuild.2024.114702

[12] Zhang, B., Song, Y., Liu, D., Zeng, Z., Guo, S., Yang, Q., . . . , & Shen, X. (2023). Descriptive and network post-occupancy evaluation of the urban public space through social media: A case study of Bryant Park, NY. *Land*, *12*(7), 1403. https://doi.org/10.3390/land12071403

[13] Elsayed, M., Pelsmakers, S., Pistore, L., Castaño-Rosa, R., & Romagnoni, P. (2023). Post-occupancy evaluation in residential buildings: A systematic literature review of current practices in the EU. *Building and Environment*, *236*, 110307. https://doi.org/10.1016/j.buildenv.2023.110307

[14] Park, J., Tsou, M.-H., Nara, A., Cassels, S., & Dodge, S. (2024). Developing a social sensing index for monitoring place-oriented mental health issues using social media (Twitter) data. *Urban Informatics*, *3*(1), 2. https://doi.org/10.1007/s44212-023-00033-5

[15] Lin, Y., Wood, S. A., & Lawler, J. J. (2022). The relationship between natural environments and subjective well-being as measured by sentiment expressed on Twitter. *Landscape and Urban Planning*, *227*, 104539. https://doi.org/10.1016/j.landurbplan.2022.104539

[16] Hswen, Y., Qin, Q., Brownstein, J. S., & Hawkins, J. B. (2019). Feasibility of using social media to monitor outdoor air pollution in London, England. *Preventive Medicine*, *121*, 86–93. https://doi.org/10.1016/j.ypmed.2019.02.005

[17] Sachdeva, S., McCaffrey, S., & Locke, D. (2017). Social media approaches to modeling wildfire smoke dispersion: Spatiotemporal and social scientific investigations. *Information, Communication & Society*, *20*(8), 1146–1161. https://doi.org/10.1080/1369118X.2016.1218528

[18] Yang, C., & Liu, T. (2022). Social media data in urban design and landscape research: A comprehensive literature review. *Land*, *11*(10), 1796. https://doi.org/10.3390/land11101796

[19] Zabelskyte, G., Kabisch, N., & Stasiskiene, Z. (2022). Patterns of urban green space use applying social media data: A systematic literature review. *Land*, *11*(2), 238. https://doi.org/10.3390/land11020238

[20] Artan, D., Ergen, E., Kula, B., & Guven, G. (2022). Rateworkspace: BIM integrated post-occupancy evaluation system for office buildings. *Journal of Information Technology in Construction*, *27*, 441–485. https://doi.org/10.36680/j.itcon.2022.022

[21] Hu, M., & Kejriwal, M. (2022). Measuring spatio-textual affinities in Twitter between two urban metropolises. *Journal of Computational Social Science*, *5*(1), 227–252. https://doi.org/10.1007/s42001-021-00129-5

[22] Algassim, H., Sepasgozar, S. M. E., Ostwald, M., & Davis, S. (2023). A qualitative study on factors influencing technology adoption in the architecture industry. *Buildings*, *13*(4), 1100. https://doi.org/10.3390/buildings13041100

[23] Mesa-Pedrazas, Á., Nogueras-Zondag, R., & Duque-Calvache, R. (2023). The new town square: Twitter discourses about balconies during the 2020 lockdown in Spain. *Cities*, *143*, 104595. https://doi.org/10.1016/j.cities.2023.104595

[24] Liu, M., Shibuya, Y., & Sekimoto, Y. (2024). Emotions, behaviors and places: Mapping sentiments with behaviors in

Japanese tweets. *Cities*, *155*, 105449. https://doi.org/10.1016/j.cities.2024.105449

[25] Michelangeli, A., Östh, J., Toger, M., & Türk, U. (2025). Inequality in access to urban amenities. *Npj Urban Sustainability*, *5*(1), 54. https://doi.org/10.1038/s42949-025-00248-2

[26] Cranz, G., Morhayim, L., Lindsay, G., & Sagan, J. (2021). Post occupancy evaluation in architectural education and practice. *Technology|Architecture + Design*, *5*(1), 25–30. https://doi.org/10.1080/24751448.2021.1863663

[27] Aronis, C. (2022). Architectural liminality: The communicative ethics of balconies and other urban passages. *Cultural Studies*, *36*(3), 475–501. https://doi.org/10.1080/09502386.2020.1844254

[28] Giang Thi Ngoc, N., Tsaih, L. S.-J., Chen, J. C.-P., Tamariska, S. R., Coelho, A. M. F., & Kung, H.-Y. (2025). Balcony usage as a space to achieve human well-being during pandemic COVID-19. *Journal of Asian Architecture and Building Engineering*, *24*(4), 2339–2351. https://doi.org/10.1080/13467581.2024.2370408

[29] Al-Gebeily, R. M., Sherif, A., & Aly, R. (2025). The role of the residential balcony in local context; perception of Cairene women. *Archnet-IJAR: International Journal of Architectural Research*, *19*(4), 947–969. https://doi.org/10.1108/ARCH-03-2024-0082

[30] Peters, T., & Masoudinejad, S. (2022). Balconies as adaptable spaces in apartment housing. *Buildings and Cities*, *3*(1), 265–278. https://doi.org/10.5334/bc.191

[31] Smektała, M., & Baborska-Narożny, M. (2026). Balcony use: Value pluralism, house rules and social norms. *Housing Studies*, *41*(5), 1088–1121. https://doi.org/10.1080/02673037.2025.2527078

[32] Askarizad, R., He, J., & Dastoum, M. (2025). Gender disparity in public spaces of Iran: Design for more inclusive cities. *Cities*, *158*, 105651. https://doi.org/10.1016/j.cities.2024.105651

[33] Makino, K., & Natsume, Y. (2024). Secondary spaces of entrance, corridor, and balcony in Tehran housing—A comparative analysis with Nagoya residences. *Athens Journal of Architecture*, *10*(4), 329–360. https://doi.org/10.30958/aja.10-4-1

[34] Lolli, F., Marinello, S., Coruzzolo, A., & Butturi, M. (2022). Post-occupancy evaluation's (POE) applications for improving indoor environment quality (IEQ). *Toxics*, *10*(10), 626. https://doi.org/10.3390/toxics10100626

[35] Gifford, R. (2014). Environmental psychology matters. *Annual Review of Psychology*, *65*(1), 541–579. https://doi.org/10.1146/annurev-psych-010213-115048

[36] Grigoriadou, E. T. (2021). The urban balcony as the new public space for well-being in times of social distancing. *Cities & Health*, *5*(sup1), S208–S211. https://doi.org/10.1080/23748834.2020.1795405

[37] Mohammadi, M., & Koutamanis, A. (2025). Affordance-based design evaluation: Bridging architectural intention and adaptive user behavior. *Design Studies*, *101*, 101361. https://doi.org/10.1016/j.destud.2025.101361

[38] Stevens, Q., Daly, J., & Dovey, K. (2024). Designing for possibility in public space: Affordance, assemblage, and ANT. *URBAN DESIGN International*, *29*(2), 94–104. https://doi.org/10.1057/s41289-023-00235-y

[39] Qi, J., Mazumdar, S., & Vasconcelos, A. C. (2024). Understanding the relationship between urban public space and social cohesion: A systematic review. *International Journal of Community Well-Being*, *7*(2), 155–212. https://doi.org/10.1007/s42413-024-00204-5

[40] Kousidi, S. (2025). Space-containing façades: Mediating environmental, social, and urban dynamics in collective housing design. *Architecture*, *5*(1), 12. https://doi.org/10.3390/architecture5010012

[41] Amirjani, R., Tabrizi, N., & Kashi, S. M. H. (2025). Uncovering key components of semi-public spaces in multi-unit residences for enhancing social sustainability. *Sustainable Cities and Society*, *121*, 106155. https://doi.org/10.1016/j.scs.2025.106155

[42] Pelsmakers, S., & Warwick, E. (2022). Housing adaptability: New research, emerging practices and challenges. *Buildings and Cities*, *3*(1), 605–618. https://doi.org/10.5334/bc.266

[43] Jegard, L., Castaño-Rosa, R., Kilpeläinen, S., & Pelsmakers, S. (2024). Finnish glazed balconies: Residents' experience, wellbeing and use. *Buildings and Cities*, *5*(1), 540–561. https://doi.org/10.5334/bc.473

[44] Bishop, K., Marshall, N., Rahmat, H., Thompson, S., Steinmetz-Weiss, C., Corkery, L., . . . , & Tietz, C. (2024). Behavior mapping and its application in smart social spaces. *Encyclopedia*, *4*(1), 171–185. https://doi.org/10.3390/encyclopedia4010015

[45] Galvez-Pol, A., Nadal, M., & Kilner, J. M. (2021). Emotional representations of space vary as a function of peoples' affect and interoceptive sensibility. *Scientific Reports*, *11*(1), 16150. https://doi.org/10.1038/s41598-021-95081-9

[46] Gao, S., Liu, Y., Kang, Y., & Zhang, F. (2021). User-generated content: A promising data source for urban informatics. In W. Shi, M. F. Goodchild, M. Batty, M.-P. Kwan, & A. Zhang (Eds.), *Urban informatics* (pp. 503–522). Springer. https://doi.org/10.1007/978-981-15-8983-6_28

[47] Dorostkar, E., & Najarsadeghi, M. (2022). How to evaluate urban emotions using Twitter social media? *Cities*, *127*, 103713. https://doi.org/10.1016/j.cities.2022.103713

[48] Safa, R., Bayat, P., & Moghtader, L. (2022). Automatic detection of depression symptoms in Twitter using multimodal analysis. *The Journal of Supercomputing*, *78*(4), 4709–4744. https://doi.org/10.1007/s11227-021-04040-8

[49] Strauss, C., Harr, M. D., & Pieper, T. M. (2025). Analyzing digital communication: A comprehensive literature review. *Management Review Quarterly*, *75*(4), 3119–3157. https://doi.org/10.1007/s11301-024-00455-8

[50] Niu, H., Seraphim, A. P., Morgado, P., Miranda, B., & Silva, E. A. (2025). Mapping urban emotion from geotagged social media data: Age, gender and spatial heterogeneity. *Applied Geography*, *185*, 103768. https://doi.org/10.1016/j.apgeog.2025.103768

[51] Chapple, K., Poorthuis, A., Zook, M., & Phillips, E. (2022). Monitoring streets through tweets: Using user-generated geographic information to predict gentrification and displacement. *Environment and Planning B: Urban Analytics and City Science*, *49*(2), 704–721. https://doi.org/10.1177/23998083211025309

[52] Chai, Y., Kakkar, D., Palacios, J., & Zheng, S. (2023). Twitter sentiment geographical index dataset. *Scientific Data*, *10*(1), 684. https://doi.org/10.1038/s41597-023-02572-7

[53] Shen, C., & Li, P. (2023). Green housing on social media in China: A text mining analysis. *Building and Environment*, *237*, 110338. https://doi.org/10.1016/j.buildenv.2023.110338

[54] Hodorog, A., Petri, I., & Rezgui, Y. (2022). Machine learning and natural language processing of social media data for event

detection in smart cities. *Sustainable Cities and Society*, *85*, 104026. https://doi.org/10.1016/j.scs.2022.104026

[55] Ogunkan, D. V., & Ogunkan, S. K. (2025). Exploring big data applications in sustainable urban infrastructure: A review. *Urban Governance*, *5*(1), 54–68. https://doi.org/10.1016/j.ugj.2025.02.003

[56] Reveilhac, M., Steinmetz, S., & Morselli, D. (2022). A systematic literature review of how and whether social media data can complement traditional survey data to study public opinion. *Multimedia Tools and Applications*, *81*(7), 10107–10142. https://doi.org/10.1007/s11042-022-12101-0

[57] Wagiri, F., Wijaya, D. C., & Sitindjak, R. H. I. (2024). Embodied spaces in digital times: Exploring the role of Instagram in shaping temporal dimensions and perceptions of architecture. *Architecture*, *4*(4), 948–973. https://doi.org/10.3390/architecture4040050

[58] Mvuyana, B. Y. C. (2025). The role of social actors in building sustainable cities in the Global South. *Frontiers in Sustainable Cities*, *7*, 1536656. https://doi.org/10.3389/frsc.2025.1536656

[59] Chavanayarn, S. (2024). Epistemic injustice and ideal social media: Enhancing X for inclusive global engagement. *Topoi*, *43*(5), 1355–1368. https://doi.org/10.1007/s11245-024-10070-7

[60] Hemphill, L., Schöpke-Gonzalez, A., & Panda, A. (2022). Comparative sensitivity of social media data and their acceptable use in research. *Scientific Data*, *9*(1), 643. https://doi.org/10.1038/s41597-022-01773-w

[61] Lee, J. H., & Ostwald, M. J. (2023). The 'visual attractiveness' of architectural facades: Measuring visual complexity and attractive strength in architecture. *Architectural Science Review*, *66*(1), 42–52. https://doi.org/10.1080/00038628.2022.2137458

[62] Song, T., Xu, L., Zhao, F., & Du, Y. (2024). Healing properties of residential balcony: Assessment of the characteristics of balcony space in Shanghai's collective housing. *Journal of Building Engineering*, *87*, 108992. https://doi.org/10.1016/j.jobe.2024.108992

[63] Chowdhury, S., Noguchi, M., & Doloi, H. (2021). Conceptual parametric relationship for occupants' domestic environmental experience. *Sustainability*, *13*(5), 2982. https://doi.org/10.3390/su13052982

[64] Landis, J. R., & Koch, G. G. (1977). The measurement of observer agreement for categorical data. *Biometrics*, *33*(1), 159–174. https://doi.org/10.2307/2529310

[65] Farahani, M., Gharachorloo, M., Farahani, M., & Manthouri, M. (2021). ParsBERT: Transformer-based model for Persian language understanding. *Neural Processing Letters*, *53*(6), 3831–3847. https://doi.org/10.1007/s11063-021-10528-4

[66] Conneau, A., Khandelwal, K., Goyal, N., Chaudhary, V., Wenzek, G., Guzmán, F., . . . , & Stoyanov, V. (2020). Unsupervised cross-lingual representation learning at scale. In *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics*, 8440–8451. https://doi.org/10.18653/v1/2020.acl-main.747

[67] Hakim, L., Gul, Y., & Noaime, E. (2026). Balcony morphologies in contemporary Parisian housing (2007–2020): A qualitative typological framework for environmental mediation and socio-spatial interfaces. *Architecture*, *6*(1), 38. https://doi.org/10.3390/architecture6010038

[68] Kaplan, R., & Kaplan, S. (1989). *The experience of nature: A psychological perspective*. Cambridge University Press.

[69] Ulrich, R. S. (1997). A theory of supportive design for healthcare facilities. *Journal of Healthcare Design: Proceedings from the Symposium on Healthcare Design*, *9*, 3–7.

[70] Dominici, L., Comino, E., Torpy, F., & Irga, P. (2022). Vertical greening systems: A critical comparison of do-it-yourself designs. *Plants*, *11*(23), 3230. https://doi.org/10.3390/plants11233230

[71] Stender, M. (2022). The battle of the balconies—Scandinavian sameness and urban-domestic boundaries. *Home Cultures*, *19*(3), 171–192. https://doi.org/10.1080/17406315.2022.2152177

[72] Safarkhani, M. (2016). *Balconies consigned to oblivion in Iranian residential buildings: The case of Tehran, Iran*. Middle East Technical University, Master's Thesis

[73] Jagannath, S., Gatersleben, B., & Ratcliffe, E. (2024). Flexibility of the home and residents' psychological wellbeing. *Journal of Environmental Psychology*, *96*, 102333. https://doi.org/10.1016/j.jenvp.2024.102333

How to Cite: Shakibamanesh, A., & Hosseinzadeh, S. (2026). Mining Twitter Data to Inform Balcony Design in Residential Architecture: A Data-Driven Approach to Human Experience in Semi-open Spaces. *Journal of Data Science and Intelligent Systems*. https://doi.org/10.47852/bonviewJDSIS62028057