POLITECNICO DI TORINO
Repository ISTITUZIONALE

City Decoder: An interpretablemultimodal framework forurban pattern recognition

Original
City Decoder: An interpretablemultimodal framework forurban pattern recognition / Arda, A.. - In: TRANSACTIONS IN PLANNING AND URBAN RESEARCH. - ISSN 2754-1223. - ELETTRONICO. - (2026). [10.1177/27541223261443485]

Availability:
This version is available at: 11583/3010187 since: 2026-04-23T04:54:09Z

Publisher:
SAGE Publications

Published
DOI:10.1177/27541223261443485

Terms of use:

This article is made available under terms and conditions as specified in the corresponding bibliographic description in the repository

Publisher copyright
Sage postprint/Author's Accepted Manuscript

(Article begins on next page)

25 November 2026

Title.

City Decoder: *An Interpretable Multimodal Framework for Urban Pattern Recognition*

Abstract.

Contemporary urban environments are increasingly shaped by intertwined layers of infrastructure, data, and regulation that produce spatial patterns difficult to interpret through single-modality analyses or purely predictive models. While recent AI-driven urban studies have advanced large-scale measurement and forecasting, they often struggle to (i) integrate heterogeneous modalities such as imagery, geospatial networks, archives, and planning documents, (ii) relate spatial change to regulatory and institutional timelines, and (iii) generate traceable outputs that support architectural and planning reasoning. This study addresses these limitations by proposing an interpretable, multimodal framework for urban pattern interpretation.

The paper introduces City Decoder, a spatiotemporal pattern-recognition framework designed to decode urban environments through evidence-linked interpretation rather than prediction or optimization. Methodologically, the research conducts a structured comparative review of representative trajectories from major global city laboratories—including Beijing City Lab, MIT Senseable City Lab, ETH Future Cities Laboratory, and Stanford Urban Informatics Lab—coding their dominant data modalities, analytical tasks, and output types. Based on this synthesis, City Decoder is specified as an operational pipeline that aligns time-series urban imagery, vector geospatial data, and regulatory texts within a shared spatial–temporal reference. The framework performs recurrence detection, cross-layer coupling analysis, and discontinuity identification between policy intent and material outcomes.

The framework is designed to produce a Pattern Atlas, relational mappings, and diagnostic reports intended to support architectural and urban analysis, policy evaluation, and scenario-based design inquiry. The primary contribution is a transferable methodological specification for interpretive, traceable pattern recognition in urban research—one that repositions multimodal AI from a predictive tool toward an instrument of spatial understanding.

keywords. *city decoder, urban AI framework, urban computing, spatial-temporal recognition model, city lab.*

Introduction.

Recent advances in artificial intelligence have significantly transformed the ways urban environments can be measured, classified, and modeled at scale. In contemporary urban research, machine learning techniques are increasingly employed to extract patterns from satellite and street imagery, detect land-use change, analyze mobility networks, and support large-scale environmental monitoring. While these approaches have produced valuable insights, they are

largely oriented toward prediction, optimization, or automated classification. As a result, they often provide limited support for interpretive urban reasoning—that is, the ability to relate observed spatial patterns to the infrastructural, regulatory, and institutional conditions through which urban space is produced and transformed.

This limitation becomes particularly evident in complex urban territories shaped by multiple overlapping systems. Architectural and urban design discourse has traditionally approached the city through distinct analytical lenses—such as morphology, infrastructure, governance, and environment—each examined as a separate layer. Yet many of the most consequential urban transformations occur precisely where these layers intersect: logistical corridors, peri-urban thresholds, infrastructural landscapes, and residual urban fragments. These in-between territories are often defined by unstable relations between regulatory frameworks, ecological processes, and everyday practices, producing spatial conditions that resist single-form or single-logic explanations. Engaging with such environments requires methodological approaches capable of working across multiple spatial scales, temporalities, and heterogeneous data modalities.

This paper addresses that need by proposing City Decoder, an interpretable hybrid AI–geospatial framework for multimodal spatiotemporal pattern recognition. Rather than treating the city as a system to be optimized or forecast, the City Decoder is conceived as a framework for evidence-linked decoding. Its aim is to make spatial patterns readable by relating them to the infrastructural relations, regulatory sequences, and spatial behaviors that co-produce them. In this sense, the framework is intended to support architectural and planning reasoning by generating traceable and interpretable outputs, rather than opaque predictions.

The proposal builds on the premise that contemporary urban environments increasingly operate through forms of "code." Historically, code referred to codified laws and regulations that inscribed authority into spatial form. In contemporary urban conditions, code also includes software-mediated infrastructures and data-driven governance systems that shape circulation, access, visibility, and control. Across urban theory and science-and-technology studies, scholars have argued that code is not an external layer applied to space but a material force in spatial production. Within architectural discourse, related traditions have approached the city as a legible system, a text, or a patterned language—suggesting that urban coherence emerges through recurring relationships rather than isolated forms. City Decoder draws from these conceptual positions but shifts the emphasis from metaphor to method: it treats "code" as a relational condition to be decoded through multiple forms of evidence distributed across space and time.

Methodologically, the paper proceeds from the observation that major global city laboratories have developed powerful techniques for extracting urban patterns, yet these approaches frequently remain modality-specific or output-oriented, focusing on prediction, simulation, or visualization without consistently linking spatial patterns to regulatory and institutional sequences. To clarify this landscape and position the proposed framework, the study conducts a structured comparative review of representative research trajectories from major global city laboratories. Each case is coded according to its dominant data modalities, analytical tasks, and output types. This synthesis reveals a methodological gap: while many approaches excel at measurement and prediction, fewer provide tools for cross-modal alignment, temporal coupling between spatial change and governance, and traceable interpretive outputs.

Against this background, the paper introduces City Decoder as an operational framework. It is specified as a pipeline that aligns heterogeneous datasets—such as time-series urban imagery, vector geospatial networks, and regulatory texts—within a shared spatial–temporal reference. Through this alignment, the framework performs three core analytical operations: (i) recurrence detection across temporal layers, (ii) cross-layer coupling analysis between spatial, infrastructural, and regulatory domains, and (iii) discontinuity identification between policy intent and material outcomes. The framework is designed to produce a Pattern Atlas, relational mappings, and diagnostic reports intended to support architectural analysis, policy evaluation, and scenario-based design inquiry.

This leads to the paper's central research question:

How can hybrid AI–geospatial models assist in decoding coded urban environments by revealing relational patterns across heterogeneous spatiotemporal data in a form that remains interpretable and evidence-linked for architectural and planning reasoning?

The paper makes three contributions. First, it proposes a conceptual reorientation of urban AI from prediction toward interpretive, traceable pattern recognition, grounded in a comparative synthesis of global city-laboratory methodologies. Second, it specifies City Decoder as a transferable framework—defining its input modalities, three analytical operations, and diagnostic output types. Third, it demonstrates the framework's analytical logic through a structured illustrative application to Kartal district, Istanbul, showing how cross-layer coupling and discontinuity detection surface relational patterns that no single-modality analysis could produce.

The paper is organized into four sections. Section 2 reviews key theoretical debates on code, urban pattern reasoning, and interpretive approaches to spatial analysis. Section 3 presents the comparative analysis of global city laboratories and extracts a structured typology of their methodological trajectories. Section 4 specifies the City Decoder framework and its analytical pipeline, and includes an illustrative application of the framework's three analytical operations to Kartal district, Istanbul, demonstrating the kind of relational patterns and diagnostic outputs the framework produces when applied to real data. Section 5 concludes by summarizing the contributions and discussing implications for urban research and practice. Figure 1 presents an overview of the City Decoder framework.

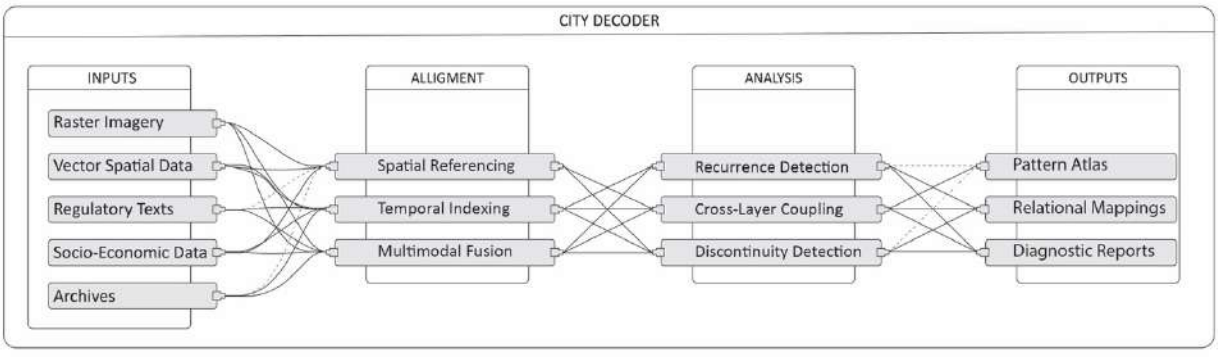


Figure 1. City Decoder Framework

Literature Review.

Research on the city as a dynamic and relational system has deep conceptual roots at the intersection of architecture, urban studies, and information technologies. While the recent rise of artificial intelligence and large-scale urban data has intensified this research direction, its theoretical foundations predate contemporary computational approaches. This literature review situates the present study within three interconnected trajectories: (1) the conceptualization of urban space as a coded system, (2) the development of computational and data-driven urban analysis, and (3) pattern-oriented approaches in architectural and urban theory. Together, these strands form the epistemological and methodological basis for a decoding-oriented framework.

1. Code as a Spatial and Relational System

The conceptual foundation of this study rests on the notion of code as both metaphor and governing mechanism, an idea that has evolved across cryptography, law, linguistics, and spatial theory. In its technical and historical sense, code refers to a cryptosystem operating on linguistic units, transforming meaning through processes of encryption and translation. Etymologically derived from the Latin *codex*, meaning a bound collection of laws, the term historically referred to the inscription of authority into material form.

This legal–material genealogy is extended by Mitchell (1995), who famously asserted that "code is law," arguing that computational systems reorganize social and spatial orders through executable scripts. Hillier and Hanson (1984) similarly describe code as an underlying system of rules that structures spatiotemporal events, suggesting that such rules precede and organize spatial interactions. Kitchin and Dodge (2011) further demonstrate that software, though often invisible, exerts material agency over spatial production by shaping infrastructures, services, and everyday routines. Collectively, these interpretations position code as an operative structure—an invisible architecture that regulates how both information and space are produced and experienced.

Bernstein (1981) extends this logic by defining code as a regulative principle that selects and integrates meanings and their realizations. In this view, code mediates between contexts and the relationships within them, governing both what meanings can coexist and how they are expressed. This relational interpretation situates code as a system of translation linking the social, the semantic, and the spatial.

Within spatial theory, Lefebvre advanced this position by describing the city as a coded yet unfinished text. Rejecting structuralist conceptions of the city as a closed signifying system, he argued that "the city cannot therefore be conceived as a signifying system, determined and closed as a system" (Lefebvre, 1968 [2003]). Instead, urban space is continuously rewritten through social practice: "the city writes itself on its walls and in its streets, but that writing is never completed" (Lefebvre, 1970, 1974 [2003]). In this sense, the city becomes a code in constant revision—a living script that records, erases, and rewrites itself over time.

2. From Spatial Code to Computational Urban Analysis

The idea of spatial lawfulness and systemic order was later formalized through analytical and computational approaches. Alexander et al. (1987) argued that cities evolve through interacting layers of spatial rules, emphasizing the need to understand the laws that produce urban wholeness. Hillier and Hanson (1984) translated similar principles into a formal methodology through space syntax, which treats urban space as a language whose grammar can be analyzed through patterns of connection and integration. By relating spatial configurations to social variables, space syntax introduced one of the earliest computer-aided analytical approaches to urban form.

Building on these developments, urban computing emerged as a data-driven paradigm for analyzing cities. Zheng et al. (2014) define urban computing as the acquisition, integration, and analysis of large, heterogeneous datasets generated by multiple urban sources and agents. These approaches successfully merge computation and urban analysis but often frame the city primarily as a measurable and optimizable system, rather than as an interpretive or relational field (Batty, 2013).

The notion of the dynamic city gained further traction through the work of Batty (1994), who used fractal geometry to describe urban form as a real, measurable structure rather than an abstract ideal. Later, Batty (2013) articulated the concept of the computable city, arguing that urban understanding emerges from the relations between places rather than from the intrinsic attributes of individual objects. This relational perspective forms a key foundation of the emerging science of cities. However, much of this work remains oriented toward simulation, prediction, and formal modeling, leaving open the question of how computational methods might support interpretive reasoning about urban transformation.

3. Pattern-Based Reasoning in Architectural Theory

Parallel to these computational developments, architectural theory has long approached the city as a coded and patterned environment. Lynch (1960) introduced the concept of urban legibility as a cognitive code through which inhabitants read spatial order. Van Eyck, in the *Team 10 Primer* (Smithson, 1968), described code as a relational syntax mediating time, movement, and experience. Rossi (1982) conceptualized the urban locus as a coded memory of collective identity. Tschumi (1996) further identified new codes of assemblage emerging from sequences of events, programs, and movements within the contemporary metropolis.

Among these contributions, Christopher Alexander's work remains central to the idea of pattern-based spatial reasoning. His concept of patterns, understood as recurring social–spatial relationships forming a language of form and structure , provided a systemic approach to design. In *A Pattern Language* (Alexander, 1977), he defined each pattern as a recurring problem accompanied by a solution adaptable to different contexts. However, Alexander later argued that patterns alone could not generate living environments; instead, he proposed a morphogenetic understanding of urban formation that links design processes to evolving spatial structures (Alexander, 1987). His question—what kinds of laws, and at what levels, are required to produce urban wholeness—remains central to contemporary systemic approaches to design.

Watanabe (2002) extended these ideas by advocating bottom-up decision-making processes that define relationships between parts while maintaining overall coherence. This perspective aligns

with complexity science, which reveals the simple rules underlying apparently chaotic systems. Within this framework, the theories of Alexander's pattern language, Hillier's space syntax, and Batty's computable city can be understood as related attempts to interpret the city as a self-organizing and evolving system.

4. Toward a Decoding-Oriented Framework

Across these trajectories, a shared conceptual thread emerges: the city is repeatedly described as a system structured by rules, relations, and recurring patterns. However, existing computational approaches often emphasize measurement, prediction, or simulation, while architectural theory has traditionally focused on conceptual or formal interpretations of spatial codes. What remains underdeveloped is a methodological framework capable of decoding urban patterns across heterogeneous datasets, linking spatial configurations to regulatory, infrastructural, and institutional processes over time.

Artificial intelligence research defines its central task as recognizing patterns within large volumes of data and making them visible or actionable (Batty, 2024). Yet, within urban research, this pattern-recognition capacity is frequently directed toward forecasting or optimization rather than interpretive understanding. The present study builds on this gap by proposing a decoding-oriented approach: a framework that uses multimodal pattern recognition not to predict future states, but to reveal the relational structures and temporal dynamics underlying urban transformation.

The following section examines contemporary global city laboratories and their methodological trajectories, providing a comparative basis for the specification of the City Decoder framework.

Comparative Analysis of Global City Laboratories.

Despite decades of technological progress, artificial intelligence research has largely evolved within linguistic and symbolic domains rather than spatial ones. As Zhao et al. (2025) note, language models have advanced from simple text generators to complex systems capable of reasoning and problem-solving, yet they remain grounded in textual data and generative processes. This dominance of language-based systems has shaped analytical cultures that privilege semantic patterning over spatial interpretation.

Batty (2024) observes that most neural architectures are trained to reproduce patterns within data rather than to explain or construct them, generating outputs that mimic observed conditions without exposing their underlying logic. He later distinguishes between discriminative models, which read existing conditions, and generative models, which create new possibilities under uncertainty (Batty, 2025). Yet both paradigms often operate within isolated modalities, focusing on singular data streams rather than the overlapping layers that characterize urban environments. Cities, however, are constituted through interdependent infrastructural, regulatory, ecological, and behavioral systems that cannot be reduced to a single analytical logic.

To address this gap, several research laboratories have developed data-driven approaches that integrate computation with spatial reasoning. The field of urban computing (Zheng et al., 2014) emerged precisely to bridge computer science and urban studies. More recent initiatives—such as the MIT Senseable City Lab, Beijing City Lab, ETH Future Cities Laboratory, and the Stanford Urban Informatics Lab—have extended this agenda through multimodal urban analytics. As Zou et al. (2025) emphasize, no single dataset can capture the full complexity of territorial systems; meaningful insights require the integration of spatial, environmental, textual, and behavioral data. Yet as Long and Zhang (2024) observe, many of these approaches remain oriented toward prediction, optimization, or simulation, with limited emphasis on interpretive reasoning.

Selection logic and comparative framework

The following laboratories were selected as representative case studies based on three criteria:

1. Institutional influence within computational urban research,
2. Methodological diversity across data modalities and analytical tasks, and
3. Demonstrated application of AI or data-driven methods to spatial and urban questions.

Each laboratory was examined through a structured review of key publications, coding their dominant data modalities, analytical techniques, and primary output types. This comparative approach provides a basis for identifying methodological patterns and gaps across contemporary urban AI research.

Beijing City Lab: Policy-Embedded Spatial Modeling

Research at the Beijing City Lab demonstrates a strong connection between spatial analysis and planning frameworks. Early work by Long et al. (2012) introduced constrained cellular automata and regional sensitivity analysis to translate planning parameters into spatial outcomes. This approach revealed how planning intentions can operate as testable spatial codes.

Lang et al. (2018) expanded this perspective by combining spatial entropy, dissimilarity indices, and structural equation modeling, framing the city as a dynamic equilibrium between order and variation. Subsequent research incorporated deep learning and street-view imagery to operationalize visual cognition (Chen et al., 2023; Ma et al., 2025), transforming perceptual qualities into computationally readable datasets.

Parallel studies used segmentation and classification to map vacant and abandoned structures, revealing cycles of urban growth and decay (Mao et al., 2022; Li et al., 2023). More recent work integrates environmental sensing with behavioral data to analyze adaptive urban dynamics, including soundscapes and post-pandemic public spaces (Zhang et al., 2023; Li et al., 2024). Across these studies, the laboratory progressively moves from predictive modeling toward multimodal pattern interpretation.

ETH Future Cities Laboratory: Integrated Urban Intelligence

The ETH Future Cities Laboratory follows a parallel but distinct trajectory, moving from computational visualization toward integrated urban intelligence. Zeng et al. (2016) employed visual analytics to translate complex transportation data into interpretable visual structures. Sun and Axhausen (2016) advanced this work through probabilistic tensor factorization, uncovering hidden mobility patterns across large datasets.

Later research integrated generative and evolutionary algorithms into architectural design processes, allowing AI to participate in design decision-making (Koenig et al., 2020). Chadzynski et al. (2023) synthesized these directions through a dynamic geospatial knowledge graph, combining semantic and spatial reasoning within a continuously updated urban model. Collectively, ETH's work demonstrates how analytical modeling, visualization, and design-oriented computation can operate within multi-scalar urban systems.

MIT Senseable City Lab: Multimodal Perceptual Urbanism

The MIT Senseable City Lab has long pioneered perceptual and multimodal approaches to computational urbanism. Kan et al. (2020) coupled space-time GIS with vehicular modeling to transform emissions and traffic data into interpretable spatiotemporal structures.

More recent research explores human–AI collaboration in planning contexts. Zheng et al. (2025) proposed frameworks in which large language models interact with visual and behavioral data, enabling shared decision-making processes between human designers and computational systems. Guo et al. (2025) further extended this line of inquiry by applying AI-driven image clustering to measure the visual identity of urban spaces, effectively translating perceptual qualities into quantifiable datasets. Together, these works conceptualize the city as a layered informational environment, where environmental, behavioral, and perceptual data require different forms of analysis and interpretation.

Stanford Urban Informatics Lab: Transparent Energy Modeling

At the Stanford Urban Informatics Lab, research led by Roth et al. (2019, 2020) demonstrates how open data, machine learning, and physics-based simulation can converge to analyze urban energy systems. Their Urban Building Energy Model (UBEM) and its extended version (A-UBEM) generate synthetic hourly energy profiles for over a million buildings in New York City, combining random forest algorithms, convex optimization, and Monte Carlo validation.

Beyond technical performance, this work emphasizes transparency and accessibility. By building replicable and open models, the research highlights the importance of data openness and methodological clarity in urban AI. It suggests that the intelligence of cities depends not only on the sophistication of analytical models but also on the accessibility of the data infrastructures supporting them.

Comparative Observations

Across these laboratories, a common trajectory emerges: a shift from isolated data analysis toward more integrated and multimodal forms of urban computation. However, their primary outputs

remain oriented toward prediction, simulation, classification, or visualization. While these approaches provide powerful analytical tools, they rarely offer a unified framework for interpreting spatial patterns across heterogeneous modalities and institutional timelines.

In other words, existing laboratories produce valuable fragments of urban intelligence—visual perception, mobility structures, energy systems, or planning simulations—but these insights often remain methodologically separated. What is still lacking is a framework capable of aligning these diverse modalities into a coherent interpretive system.

Positioning the City Decoder

The City Decoder is proposed in response to this methodological gap. It builds on the interpretive tendencies observed across these laboratories—policy-linked modeling at Beijing, integrated reasoning at ETH, multimodal perception at MIT, and open-data transparency at Stanford—but seeks to articulate them within a single, transferable framework.

Rather than focusing on prediction or generative design, the City Decoder emphasizes multimodal alignment, spatiotemporal pattern recognition, and evidence-linked interpretation. Its goal is not to forecast urban futures, but to reveal the layered relations already embedded within urban environments by connecting spatial data, infrastructural systems, and regulatory sequences into a unified analytical field.

In this sense, the City Decoder extends the trajectory of contemporary urban computing by repositioning artificial intelligence as a tool for decoding relational urban patterns, supporting architectural reasoning and policy interpretation rather than purely predictive or generative tasks.

Synthesis.

The City Decoder stands at the center of this research as both a conceptual and methodological proposal: a multimodal spatiotemporal pattern-recognition framework designed to decode the relational logics embedded within complex urban environments. Its objective is to make patterns apparent patterns that emerge from the interaction of spatial, legal, ecological, infrastructural, and socio-economic systems across time. These patterns often remain difficult to interpret through conventional architectural or urban analysis, as they are produced through overlapping processes that extend beyond the limits of single datasets or disciplinary lenses. The Decoder therefore does not aim to predict or generate future states; instead, it seeks to recognize, align, and interpret the relational structures that shape territorial transformation.

In both architecture and computational research, a model functions as a structured abstraction that organizes complexity into interpretable form. In computational logic, models mediate between data and meaning, transforming information into recognizable structures. In architecture, models act as relational constructs, linking spatial systems, actors, and temporal processes. The City Decoder adopts this dual understanding. It operates as a decoding interface—a methodological environment in which heterogeneous datasets, regulatory frameworks, and spatial traces are

aligned and interpreted together. Its value lies not in simulating reality, but in clarifying it: exposing how infrastructural, institutional, and spatial systems intersect across time.

At its core, the Decoder treats the city as a layered and coded environment, composed of multiple interacting systems. Urban space is understood not as a stable object but as the product of continuous negotiation among overlapping temporalities: legal time, when regulations are enacted; infrastructural time, when systems are built or modified; ecological time, when environmental processes adapt; and social time, when patterns of occupation and use emerge. These temporal layers are read together rather than in isolation, allowing the framework to connect decisions, actions, and spatial outcomes. For example, a regulatory change can be traced through subsequent infrastructural developments and ecological effects, revealing how policy decisions translate into material transformations. Figure 2 illustrates the City Decoder's layered structure and the relationships between these temporal domains.

-

Figure 2. City Decoder Layered Structure

Operational Structure

Methodologically, the City Decoder is specified as a multimodal spatiotemporal pattern-recognition pipeline composed of three main stages: data alignment, relational analysis, and interpretive output generation.

The first stage involves the collection and alignment of heterogeneous datasets within a shared spatial–temporal reference. These inputs are deliberately diverse, reflecting the layered nature of urban environments. Spatial datasets provide the physical substrate, including parcels, infrastructural networks, ecological zones, and administrative boundaries. Raster imagery, such as satellite or aerial data, captures temporal changes in land use, vegetation, and construction. Textual and regulatory sources—planning documents, legal codes, environmental assessments—introduce institutional and normative dimensions. Historical archives reconstruct longer genealogies of territorial transformation, while socio-economic indicators provide additional behavioral and demographic correlations. Through spatial referencing and temporal indexing, these datasets are aligned into a unified analytical frame.

The second stage performs relational pattern analysis across these aligned layers. Rather than analyzing each dataset independently, the framework searches for cross-layer relationships and temporal correspondences. Three primary analytical operations structure this process. First, recurrence detection identifies spatial configurations or transformations that repeat across time or across different contexts. Second, cross-layer coupling analysis examines how changes in one domain—such as regulatory decisions or infrastructural interventions—correlate with spatial or ecological transformations. Third, discontinuity detection identifies moments where expected relationships break down, revealing gaps between policy intention and material outcomes, or between infrastructural logic and social use.

Through these operations, the Decoder identifies relational patterns that would remain invisible within isolated datasets.

The third stage translates these analytical results into interpretive outputs designed for architectural and planning reasoning. Rather than producing deterministic predictions, the system generates a set of diagnostic representations. These include pattern atlases that document recurring territorial configurations, relational mappings that illustrate interactions between spatial, legal, and infrastructural systems, and diagnostic reports that highlight inconsistencies or tensions within territorial processes. Such discontinuities—where regulations fail to produce intended spatial effects, or where infrastructural systems produce unintended consequences—are treated not as errors, but as critical entry points for interpretation.

Illustrative Application: Kartal District, Istanbul

The present paper proposes City Decoder as a methodological framework specification rather than a fully implemented computational system. To demonstrate that the framework's analytical logic is both operable and productive when applied to real data, this section presents a structured illustrative application conducted manually across three heterogeneous data layers for Kartal district on Istanbul's Asian shore as presented in Figure 3. The application does not claim to replicate what an implemented City Decoder system would produce at scale; it demonstrates, using traceable and publicly available data, the kind of relational patterns and interpretive outputs the framework is designed to surface. The site was selected because it concentrates, within a single compact district, the very conditions the framework is designed to decode: overlapping regulatory regimes, a building stock of sharply divergent ages and construction types, a documented seismic risk profile, and a persistent gap between governance intention and material outcome.

-

Figure 3. Kartal District in Istanbul.

Data Alignment. The first stage of the City Decoder pipeline requires the collection and alignment of heterogeneous datasets within a shared spatial–temporal reference. For Kartal, three primary layers were assembled and co-registered. The spatial layer draws on the Istanbul Metropolitan Municipality's open building inventory (IBB, 2023), comprising approximately 30,000 building records classified by construction era, structural type, floor count, and neighbourhood. The regulatory layer draws on the municipal zoning register and the 1947 industrial zoning designation that formally structured Kartal's western neighbourhoods, subsequently overlaid by Gecekondu Law No. 775 (1966), the 1980 prevention zone classification, and Urban Transformation Law No. 6306 (2012). The risk layer draws on the seismic scenario developed by Erdik et al. (2003) for the Istanbul Metropolitan Area, adopted as the operative legal basis for transformation decisions, and the building-level damage estimates published by the Istanbul Metropolitan Municipality Earthquake Scenario Analysis (IBB, 2023). These three layers—spatial, regulatory, and risk—were indexed against a shared temporal reference spanning 1945 to 2023, with population census records from TÜİK (1945–2022) providing a fourth socio-demographic layer confirming the pace of settlement relative to regulatory action. Figure 4 presents the resulting land use and urban fabric

structure of Kartal district, showing the spatial co-presence of residential, industrial, and transitional zones alongside neighbourhood boundaries and infrastructure networks.

-

Figure 4. Kartal District: Urban Fabric and Land Use Structure. Sources: IBB Open Data (2023); TÜİK building inventory; municipal zoning register.

Cross-Layer Coupling Analysis. The second analytical operation examines how changes in one domain correlate with transformations across others. Aligning the regulatory and spatial layers reveals a pattern of coupling that conventional single-modality analysis cannot surface. Kartal's population expanded from approximately 21,000 in 1945 to over 572,000 by 1985—a 27-fold increase over four decades driven by industrial migration (TÜİK, 1945–2000). This growth occurred overwhelmingly on land that remained formally zoned as industrial or undesignated under the 1947 classification. The building inventory records that of approximately 30,000 structures in Kartal, 9,211 are coded to construction era 1 (pre-1980) and 14,491 to era 2 (1980–2000)—together representing approximately 79 per cent of the building stock, the majority of which is reinforced concrete (Betonarme, ~25,500 structures) but with a significant minority of unreinforced masonry (Yığma, ~3,887 structures) concentrated in the neighbourhoods that grew earliest and fastest. Cross-layer coupling analysis reveals that the spatial distribution of Yığma structures is not random: it is strongly correlated with the areas of earliest informal settlement, those where rapid population growth preceded both regulatory coverage and structural standards. The regulatory layer confirms that the 1947 industrial zoning designation remained formally unrepealed throughout this settlement process, producing a legal condition in which the buildings that came to house the majority of Kartal's population were constructed on land whose designated use had never been revised to reflect residential occupation. This is the coupling pattern the City Decoder's cross-layer analysis is designed to detect: the co-evolution of spatial form and regulatory status across time, and the structural tensions their divergence produces.

Discontinuity Detection. The third operation identifies moments where the expected relationship between policy intent and material outcome breaks down. In Kartal, the most analytically significant discontinuity concerns seismic governance. The operative risk evidence base for transformation decisions across Kartal is the 2003 site-dependent deterministic intensity distribution developed by Erdik et al. (2003) for the Istanbul Metropolitan Area (Erdik et al., 2003), which classifies the district's western and central neighbourhoods at seismic intensity 7–8 and portions of the coastal edge at 8–9 under an Mw 7.4 scenario. This assessment, now over two decades old, remains the legal basis on which Urban Transformation Law No. 6306 applications are assessed and prioritised. The gap between the 2003 calibration date of this risk instrument and the urban condition it currently governs represents a measurable discontinuity between regulatory intent and material reality: the instrument was designed for a building stock and settlement pattern that have continued to evolve since its calibration, while the governance decisions it authorises remain indexed to a fixed analytical baseline. The spatial distribution of seismic damage estimates across Kartal's neighbourhoods illustrates the material consequence of this discontinuity directly. Figure 5 maps the expected severely damaged buildings per neighbourhood under the Mw 7.5 scenario (IBB, 2023): Orhantepe records the highest concentration (13–22 severely damaged buildings per unit), followed by Hürriyet and Yunus. These are the neighbourhoods whose building

stock is most densely composed of pre-1980 construction, most heavily associated with informal settlement, and most exposed to seismic intensity 8–9. The discontinuity between what was regulated and what was built, sustained over eight decades, has produced a measurable risk gradient whose spatial logic is only legible through the simultaneous alignment of the regulatory, spatial, and risk layers.

-

Figure 5. Kartal District: Seismic Damage Distribution and Building Stock Vulnerability under Mw 7.5 Scenario. Sources: IBB Earthquake Scenario Analysis Results (2023); Erdik et al. (2003).

Pattern Atlas Entry. The interpretive output of this application takes the form of a prototype Pattern Atlas entry for Kartal—one of the diagnostic representations the City Decoder is designed to produce. The entry registers four relational findings. First, a recurrence pattern: the spatial concentration of vulnerable building stock in the western neighbourhoods is not a random artefact of urban growth but a recurring outcome of the structural coupling between early informal settlement, regulatory non-coverage, and construction-period constraints. Second, a cross-layer coupling pattern: population density, construction vulnerability, and seismic exposure are co-located precisely in the zones where regulatory intent diverged earliest and most persistently from settlement reality. Third, a discontinuity: the seismic governance instrument calibrated in 2003 continues to authorise transformation decisions for a building stock and risk profile that have evolved across the two decades since its adoption, producing a regulatory lag whose spatial consequences are measurable at neighbourhood scale. Fourth, a governance collision: the co-presence of the 1947 industrial designation, gecekondu cadastral geometry, and the 2012 transformation law has produced irresolvable property-rights conflicts in the zones of highest vulnerability, stalling precisely the interventions the regulatory framework was designed to enable. Together, these four findings constitute the kind of relational, evidence-linked diagnostic output the City Decoder framework is designed to generate—not a prediction of what Kartal will become, but an interpretive account of the structural logics that have produced its current condition and that constrain every governance decision taken within it.

This illustrative application demonstrates that the City Decoder's three analytical operations—data alignment, cross-layer coupling, and discontinuity detection—are not only conceptually specified but analytically productive when applied to traceable, publicly available data. The computational implementation of the framework would extend this logic to larger territorial scales, additional data modalities, and automated pattern detection across comparable urban contexts, producing the kind of systematic and scalable interpretive intelligence the present specification anticipates.

Epistemological Position

The Decoder's epistemological position is defined by recognition rather than prediction. It does not attempt to forecast urban futures or generate optimized forms. Instead, it operates as a diagnostic instrument that reveals the underlying structures through which territories evolve. Its function is to render legible the interactions between space, time, governance, and infrastructure, transforming fragmented datasets into coherent interpretive fields.

By aligning multiple systems within a single analytical framework, the Decoder constructs what can be described as a relational grammar of the territory—a set of recurring configurations through which spatial transformations occur. This grammar is neither fixed nor universal; it is context-dependent and continuously evolving, reflecting the adaptive nature of urban systems. The model therefore emphasizes situated interpretation rather than universal prediction.

Transferability and Scope

The City Decoder is conceived as a transferable framework rather than a site-specific tool. Its structure allows it to operate across different territorial scales, from metropolitan regions to infrastructural corridors, without losing analytical coherence. Its scalability lies not in computational magnitude alone, but in its conceptual capacity to align diverse temporalities and spatial systems within a unified interpretive logic. By working through recognition rather than generalization, the framework respects local specificity while enabling comparative readings across contexts.

Contribution of the Framework

Within the broader landscape of computational urban research, the City Decoder proposes a shift from predictive and generative paradigms toward interpretive, traceable pattern recognition. It transforms artificial intelligence from a tool primarily associated with forecasting or simulation into a method for understanding how spatial transformations occur across layered systems. In this sense, decoding becomes both an analytical operation and a methodological stance: an attempt to make visible the relational structures that organize contemporary urban environments.

Rather than producing singular solutions, the framework offers a structured way of reading complex territories. It aligns data, law, and spatial form into a shared analytical field, enabling architects, planners, and researchers to identify patterns, tensions, and discontinuities within urban systems. The City Decoder thus positions recognition as the foundation for a new mode of spatial reasoning—one that seeks not to predict the future of cities, but to render their existing logics legible.

Conclusion

This paper has proposed the City Decoder as a methodological framework for advancing multimodal pattern recognition within architectural and urban research. Conceived as a spatiotemporal decoding system, the framework responds to the limitations of predominantly predictive and generative AI approaches by repositioning artificial intelligence as an instrument of spatial interpretation rather than production. Through the alignment of heterogeneous datasets—including geospatial layers, regulatory texts, archival records, and socio-behavioral indicators—the City Decoder establishes an interpretive pipeline capable of revealing the relational processes through which urban environments are continuously structured and restructured.

The comparative analysis of global city laboratories demonstrates both the analytical potential and the methodological fragmentation of current computational approaches to urban analysis. While these laboratories have developed advanced techniques for measurement, simulation, and visualization, their outputs often remain modality-specific or oriented toward prediction. Against this landscape, the City Decoder proposes a unifying framework that foregrounds decoding as both a conceptual and operational act. By aligning spatial, regulatory, infrastructural, and ecological datasets within a shared spatiotemporal reference, the framework enables recurrence detection, cross-layer coupling analysis, and the identification of discontinuities between policy intentions and material outcomes.

The study makes three primary contributions. First, at a theoretical level, it articulates a shift from predictive and generative paradigms toward interpretive, traceable pattern recognition, reframing the role of artificial intelligence within architectural and urban inquiry. Second, at a methodological level, it introduces a transferable multimodal pipeline that specifies inputs, alignment procedures, analytical operations, and interpretive outputs for decoding complex urban systems. Third, at a practical level, it demonstrates the framework's analytical logic through an illustrative application to Kartal district, Istanbul, showing how data alignment, cross-layer coupling analysis, and discontinuity detection surface relational patterns—including a 21-year seismic governance lag and a structural coupling between informal settlement history and present-day vulnerability—that no single-modality analysis could produce.

Rather than forecasting urban futures or generating optimized forms, the City Decoder operates as a diagnostic instrument. Its primary function is to reveal the layered logics already embedded within urban territories, tracing how decisions, infrastructures, regulations, and everyday practices interact across time. In doing so, it positions recognition as a critical mode of design intelligence, one that complements existing predictive and generative tools by providing interpretable knowledge about the structures that shape spatial transformation.

Ultimately, the City Decoder reframes artificial intelligence as a medium of spatial understanding. By transforming fragmented datasets into coherent interpretive fields, it offers a methodological foundation for reading cities as coded, relational environments. This approach does not seek to replace existing computational methods, but to extend them toward a more interpretive and architecturally grounded mode of reasoning—one capable of making the hidden structures of urban space visible and actionable.

References.

Alexander C, Ishikawa S and Silverstein M (1977) A Pattern Language: Towns, Buildings, Construction. Oxford: Oxford University Press.

Alexander C (ed.) (1987) A New Theory of Urban Design. Oxford: Oxford University Press.

Batty M (2024) AI and design. Environment and Planning B: Urban Analytics and City Science 51(4): 799–802. DOI: 10.1177/23998083241236619.

Batty M (2012) Building a science of cities. Cities 29: S9–S16. DOI: 10.1016/j.cities.2011.11.008.

Batty M (2025) Generative AI. Environment and Planning B: Urban Analytics and City Science 52(5): 1031–1034. DOI: 10.1177/23998083251332093.

Batty M (2024) The Computable City: Histories, Technologies, Stories, Predictions. Cambridge, MA: MIT Press.
Batty M (2013) The New Science of Cities. Cambridge, MA: MIT Press.
Batty M and Longley P (1994) Fractal Cities: A Geometry of Form and Function. London: Academic Press.
Bernstein B (1981) Codes, modalities, and the process of cultural reproduction: a model. Language in Society 10(3): 327–363.
Chadzynski A, Li S, Grišiūtė A, et al. (2023) Semantic 3D city interfaces—intelligent interactions on dynamic geospatial knowledge graphs. Data-Centric Engineering 4: e20. DOI: 10.1017/dce.2023.14.
Chen J, Chen L, Li Y, et al. (2023) Measuring physical disorder in urban street spaces: a large-scale analysis using street view images and deep learning. Annals of the American Association of Geographers 113(2): 469–487. DOI: 10.1080/24694452.2022.2114417.
Erdik M, Aydinoglu N, Fahjan Y, et al. (2003) Earthquake risk assessment for Istanbul metropolitan area. Earthquake Engineering and Engineering Vibration 2(1): 1–23. DOI: 10.1007/BF02857534.
Istanbul Metropolitan Municipality (IBB) (2023) Earthquake Scenario Analysis Results. Istanbul: IBB.
Guo S, Jang KM, Duarte F, et al. (2025) Urban visual uniqueness: a landmark-free framework to quantify city's identity and distinctiveness from everyday scenes. Computers, Environment and Urban Systems 122: 102351. DOI: 10.1016/j.compenvurbsys.2025.102351.
Hillier B (1989) The architecture of the urban object. Ekistics 56(334/335): 5–21.
Hillier B and Hanson J (1984) The Social Logic of Space. Cambridge: Cambridge University Press.
Kan Z, Wong MS and Zhu R (2020) Understanding space-time patterns of vehicular emission flows in urban areas using geospatial technique. Computers, Environment and Urban Systems 79: 101399. DOI: 10.1016/j.compenvurbsys.2019.101399.
Kitchin R and Dodge M (2011) Code/Space: Software and Everyday Life. Cambridge, MA: MIT Press.
Koenig R, Bielik M, Dennemark M, et al. (n.d.) Levels of Automation in Urban Design Through Artificial Intelligence. Unpublished manuscript.
Koenig R, Miao Y, Aichinger A, et al. (2020) Integrating urban analysis, generative design, and evolutionary optimization for solving urban design problems. Environment and Planning B: Urban Analytics and City Science 47(6): 997–1013. DOI: 10.1177/2399808319894986.
Koselleck R (2002) The Practice of Conceptual History: Timing History, Spacing Concepts. Stanford: Stanford University Press.
Lang W, Long Y and Chen T (2018) Rediscovering Chinese cities through the lens of land-use patterns. Land Use Policy 79: 362–374. DOI: 10.1016/j.landusepol.2018.08.031.
Lefebvre H and Bononno R (2003) The Urban Revolution. Minneapolis: University of Minnesota Press.
Li W, Zhang E and Long Y (2024) Unveiling fine-scale urban third places for remote work using mobile phone big data. Sustainable Cities and Society 103: 105258. DOI: 10.1016/j.scs.2024.105258.
Li Y, Meng X, Zhao H, et al. (2023) Identifying abandoned buildings in shrinking cities with mobile sensing images. Urban Informatics 2(1): 3. DOI: 10.1007/s44212-023-00025-5.

Long Y, Shen Z and Mao Q (2012) Retrieving spatial policy parameters from an alternative plan using constrained cellular automata and regionalized sensitivity analysis. Environment and Planning B: Planning and Design 39(3): 586–605. DOI: 10.1068/b37122.
Long Y and Zhang E (2024) City laboratory: embracing new data, new elements, and new pathways to invent new cities. Environment and Planning B: Urban Analytics and City Science 51(5): 1068–1072. DOI: 10.1177/23998083241246630.
Long Y and Zhang E (2023) Promoting urban studies and practice with emerging technologies: city laboratory, new city, and future city exploration. International Journal on Smart and Sustainable Cities 1(2): 2371004. DOI: 10.1142/S2972426023710046.
Lynch K (1960) The Image of the City. Cambridge, MA: MIT Press.
Ma X, Yang G and Yang Q (2019) Application of deep convolution neural network in automatic classification of land use. Journal of Physics: Conference Series 1187(4): 042104. DOI: 10.1088/1742-6596/1187/4/042104.
Ma Y, Li Y and Long Y (2025) Measuring temporal evolution of nationwide urban physical disorder: an approach combining time-series street view imagery with deep learning. Annals of the American Association of Geographers 115(4): 923–948. DOI: 10.1080/24694452.2025.2467330.
Mao L, Zheng Z, Meng X, et al. (2022) Large-scale automatic identification of urban vacant land using semantic segmentation of high-resolution remote sensing images. Landscape and Urban Planning 222: 104384. DOI: 10.1016/j.landurbplan.2022.104384.
Mitchell WJ (1995) City of Bits: Space, Place, and the Infobahn. Cambridge, MA: MIT Press.
Rossi A (1982) The Architecture of the City. Cambridge, MA: MIT Press.
Roth J, Bailey A, Choudhary S, et al. (2019) Spatial and temporal modeling of urban building energy consumption using machine learning and open data. In: Proceedings of Computing in Civil Engineering 2019, Atlanta, GA, June 2019, pp. 459–467. DOI: 10.1061/9780784482445.059.
Roth J, Martin A, Miller C, et al. (2020) SynCity: using open data to create a synthetic city of hourly building energy estimates by integrating data-driven and physics-based methods. Applied Energy 280: 115981. DOI: 10.1016/j.apenergy.2020.115981.
Smithson A (ed.) (1968) Team 10 Primer. London: Studio Vista.
Sun L and Axhausen KW (2016) Understanding urban mobility patterns with a probabilistic tensor factorization framework. Transportation Research Part B: Methodological 91: 511–524. DOI: 10.1016/j.trb.2016.06.011.
Tschumi B (1996) Architecture and Disjunction. Cambridge, MA: MIT Press.
Watanabe MS (2002) Induction Design: A Method for Evolutionary Design. Basel: Birkhäuser.
Zeng W, Fu C-W, Müller Arisona S, et al. (2016) Visualizing waypoints-constrained origin-destination patterns for massive transportation data. Computer Graphics Forum 35(8): 95–107. DOI: 10.1111/cgf.12778.
Zhang Y, Zhao H, Li Y, et al. (2023) Predicting highly dynamic traffic noise using rotating mobile monitoring and machine learning method. Environmental Research 229: 115896. DOI: 10.1016/j.envres.2023.115896.
Zhao WX, Zhou K, Li J, et al. (2025) A survey of large language models. arXiv preprint arXiv:2303.18223. Available at: https://arxiv.org/abs/2303.18223.
Zheng Y, Capra L, Wolfson O, et al. (2014) Urban computing: concepts, methodologies, and applications. ACM Transactions on Intelligent Systems and Technology 5(3): 1–55. DOI: 10.1145/2629592.

Zheng Y, Xu F, Lin Y, et al. (2025) Urban planning in the era of large language models. Nature Computational Science 5(9): 727–736. DOI: 10.1038/s43588-025-00846-1.
Zou X, Yan Y, Hao X, et al. (2025) Deep learning for cross-domain data fusion in urban computing: taxonomy, advances, and outlook. Information Fusion 113: 102606. DOI: 10.1016/j.inffus.2024.102606.