

Review Article

Received: 2026/05/26
Revised: 2026/09/13
Accepted: 2026/09/14



COPYRIGHTS

©2027 The author(s). This is an open access article distributed under the terms of the Creative Commons Attribution (CC BY 4.0), which permits unrestricted use, distribution, and reproduction in any medium, as long as the original authors and source are cited. No permission is required from the authors or the publishers.



HOW TO CITE THIS ARTICLE

Asgari A. Mirzajani M. Zalibeigi A. Examining thematic evolution and epistemological shifts in the literature on artificial intelligence and machine learning in urban planning and smart city studies (2020–2025): A scientometric and interpretive analysis. *Urban Economics and Planning* 8(2):4-28.

DOI: [10.22034/uep.2026.583459.1901](https://doi.org/10.22034/uep.2026.583459.1901)

Examining thematic evolution and epistemological shifts in the literature on artificial intelligence and machine learning in urban planning and smart city studies (2020–2025): A scientometric and interpretive analysis

Atiyeh Asgari^{1*}, Mahdi Mirzajani², Anahita Zalibeigi²

1. Assistant Professor, Department of Urban Planning, ST.C., Islamic Azad University, Tehran, Iran
2. PhD Candidate, Department of Urban Planning, ST.C., Islamic Azad University, Tehran, Iran

Abstract

The expansion of artificial intelligence and machine learning in urban planning and smart city studies, beyond analytical tool development, has been accompanied by changes in knowledge production, urban problem definition, and decision-making support. Despite rapid research growth, the relationship between thematic evolution and changes in the logic of urban knowledge production has not been systematically examined. This study investigates the thematic evolution and epistemological shifts in this literature during 2020–2025. Its novelty lies in integrating scientometric analysis with epistemological interpretation to explain the reconfiguration of the field's knowledge structure. Data were extracted from Web of Science and Scopus, and following PRISMA 2020 screening, 3,578 articles were selected. Using a scientometric–interpretive approach, we analyzed supervised thematic classification, temporal trend analysis, the smoothed relative growth index (RGI), epistemological paradigms, scientific collaboration, and the geographical distribution of knowledge production. The results showed that urban sustainability, accounting for 31.9%, is the largest thematic domain; the data-driven and predictive paradigm, accounting for 77.7%, is the dominant approach to knowledge production, and generative AI approaches are gradually expanding. Knowledge production is also interdisciplinary but remains geographically concentrated. Overall, the findings indicate an intra-disciplinary reconfiguration within the literature on AI in urban planning and a gradual shift of AI from analytical tools toward becoming part of the infrastructure for knowledge production and urban decision-making support.

Keywords

Epistemological shift
machine learning
scientometric analysis
smart city
supervised thematic classification
urban artificial intelligence
urban planning

* Corresponding Author: a_asgari@iau.ac.ir

1. Introduction

In recent decades, urban planning has evolved from a field grounded in static and hierarchical logics and technical rationality into an increasingly complex, data-driven, and networked domain, in which urban decision-making processes are more profoundly shaped than ever by digital technologies, big-data flows, and computational algorithms. Within the classical planning tradition, the city was largely conceived as a predictable and controllable system that could be managed through rational-comprehensive models, linear analyses, and centralized planning (Faludi, 1973; Hall, 2014). However, growing recognition of these approaches' limitations in addressing the nonlinear complexities, multilayered uncertainties, and socio-spatial dynamics of contemporary cities paved the way for the emergence of communicative, participatory, and critical paradigms in urban planning. These paradigms emphasized the roles of discourse, participatory knowledge, power, and lived experience in the production of space (Healey, 1997; Innes & Booher, 2010).

With the advent of the twenty-first century, the proliferation of the Internet of Things (IoT), urban sensors, large-scale geospatial data, and computational infrastructures facilitated the emergence of the "smart city" and "data-driven planning" paradigms, within which data have become one of the primary resources for producing urban knowledge (Batty, 2013; Kitchin, 2014). In this context, the growing penetration of artificial intelligence (AI) and machine learning (ML) signifies more than the development of advanced tools for urban analysis; it points to an emerging epistemological transformation in how cities are understood, explained, and governed. This transformation can be interpreted as a shift from classical planning logics toward a form of "algorithmic rationality," whereby the logic of urban decision-making is gradually moving away from theoretical models grounded in causal assumptions and normative reasoning toward data-driven inference, the discovery of latent patterns, and computational prediction (Kitchin, 2017; Batty, 2024).

Batty (2026) characterizes this transformation as a shift from deductive, theory-driven models toward inductive and data-driven approaches in urban planning, in which the production and validation of urban knowledge increasingly rely on machine learning, neural networks, and the analysis of complex patterns within big data. Within this emerging logic,

urban knowledge is generated less through predefined theoretical frameworks and more through processing large-scale datasets, identifying latent correlations, and algorithmic inference. Machine-learning algorithms, predictive models, and deep neural networks now play a central role across a wide range of urban planning domains, from intelligent transportation and traffic management to the analysis of citizen behavior, urban development modeling, resilience assessment, and spatial governance (Yigitcanlar et al., 2020; Bibri, 2021). At the same time, theoretical interpretations within the critical literature suggest that this transformation has also reconfigured the structure of epistemic power in cities, as urban decision-making has become increasingly dependent on computational infrastructures, data platforms, and opaque algorithms, while the boundary between "human expertise" and "algorithmic agency" has become progressively blurred (Leszczynski, 2020).

In recent years, particularly during 2024–2025, this trajectory has entered a more complex phase alongside the emergence of a new generation of frontier technologies. The rise of generative artificial intelligence (GenAI), large language models (LLMs), multimodal models, and urban digital twins has not merely expanded the capacity to analyze urban data; it has also enabled the generation of urban development scenarios, the simulation of alternative urban futures, and the formulation of planning policy recommendations (Cugurullo et al., 2024; Yigitcanlar et al., 2024). In this context, AI has shifted from an analytical and decision-support tool to a "generative agent" in processes of urban planning, design, and governance. This transformation raises fundamental questions concerning the legitimacy of algorithmic decisions, spatial justice, data bias, algorithmic transparency, and data governance in contemporary smart cities (Acuto et al., 2023; OECD, 2024).

Alongside the growing prominence of the concept of the "algorithmic city," these developments have contributed to a critical turn in the urban AI literature. From this perspective, urban AI can be understood as a stage extending beyond the classical smart-city paradigm, in which algorithms not only analyze urban data but increasingly operate as active agents in the production of knowledge, policymaking, and urban governance. Consequently, the emerging literature is concerned not only with the technical efficiency of algorithms but also with their social, ethical, and spatial implications (Kitchin, 2016; Leszczynski, 2020). Concepts such as explainable artificial intelligence

(XAI), responsible AI, algorithmic governance, data justice, and spatial algorithmic bias have therefore gradually moved toward the center of scholarly debates in urban studies and smart-city research.

Despite these developments, much of the existing scientometric research on urban AI remains largely confined to descriptive and quantitative analyses. Such studies commonly focus on co-authorship patterns, keyword co-occurrence, or citation networks, while paying comparatively limited attention to the epistemological and critical implications of these transformations (Chen, 2017; Donthu et al., 2021). In other words, although the knowledge structures of urban AI are evolving rapidly, an analytical framework that can explain precisely what these bibliometric changes reveal about transformations in the logic of urban knowledge production remains insufficiently developed. Moreover, many existing studies remain dominated by technocratic approaches and devote relatively limited attention to the relationships among algorithms, power, spatial inequality, and urban policymaking.

Accordingly, the present study seeks not merely to map the knowledge structures or thematic clusters of the literature on artificial intelligence and machine learning in urban planning and smart-city studies, but also to operationalize and explain the epistemological shifts within this field through an integration of scientometric analysis and conceptual-critical interpretation. This study does not claim to demonstrate a historical transition across the entire discipline of urban planning. Rather, it explicitly focuses on the “structural transformations and paradigmatic shifts within the urban AI literature during the 2020–2025 period (intra-disciplinary shift).” Evidence from the literature indicates that this field has undergone multidimensional transformations in recent years, including a gradual shift from predominantly technical and engineering-oriented research toward interdisciplinary approaches, changes in dominant concepts and vocabularies, the reconfiguration of disciplinary boundaries, and shifts in patterns of scientific collaboration. Nevertheless, the dimensions and mechanisms underlying these

transformations have yet to be systematically examined. The present study therefore analyzes these changes across the knowledge structure, thematic evolution, emerging trends, scientific collaboration networks, and the geographical distribution of knowledge production.

From this perspective, the present article conceptualizes artificial intelligence not merely as an urban technology, but as an emerging epistemic infrastructure within urban planning that reshapes knowledge production and validation, decision-making logic, and power relations in contemporary smart cities.

2. Theoretical and empirical background

With the rapid expansion of artificial intelligence (AI) and machine learning (ML) in urban planning and smart city studies, the literature in this field has grown substantially in recent years. Existing research encompasses a broad range of topics, including intelligent transportation, spatial analysis, urban digital twins, urban governance, environmental analysis, urban resilience, and decision-support systems. At the same time, a growing body of research has sought to move beyond technical applications by examining the social, ethical, and governance implications of AI in cities. However, the existing literature has followed diverse research trajectories, with individual studies addressing only specific dimensions of these broader transformations.

At the regional level, studies such as Behzadfar et al. (2026) have adopted a comparable approach to examine the evolution of research on sustainable transportation through bibliometric analyses, including co-word, co-authorship, and co-citation analyses. Nevertheless, such studies have largely remained confined to specific subfields of the smart city literature and have devoted relatively limited attention to epistemological interpretation or cross-domain comparison. To provide a clearer understanding of the trajectory of the literature and to identify existing research gaps, the most relevant previous studies are summarized in Table 1.

Table 1. Overview of previous studies on artificial intelligence in urban planning and smart city studies

Author/ institution	Year	Research focus	Method	Key findings	Limitation/research gap
Yigitcanlar et al.	2020	Role of AI in urban planning and smart cities	Literature review	AI plays an increasingly important role in areas such as transportation, energy, governance, and urban services.	Predominant focus on technical applications, with limited attention to epistemological implications
Janowicz et al.	2020	Emergence of GeoAI in spatial studies	Analytical review	AI has transformed spatial analysis and geospatial modeling.	Focus on technical and spatial dimensions, without examining their epistemological connections to urban planning
Bibri	2021	Convergence of big data, AI, and smart cities	Conceptual review	Big data and AI have become central components of data-driven planning.	Limited to a conceptual approach, without empirical analysis of knowledge structures
Donthu et al.	2021	Frameworks and methods for bibliometric analysis	Methodological review	Bibliometric analysis provides an appropriate approach for examining knowledge structures and research trends.	No direct application to the epistemological analysis of urban planning
UN-Habitat	2022	Opportunities, risks, and governance of AI in cities	Analytical report	Data justice, algorithmic transparency, and responsible governance have gained increasing importance.	No analysis of knowledge networks or patterns of epistemological change
Steenmans et al.	2023	Governance implications of AI in cities	Theoretical analysis	The need for responsible, transparent, and ethically oriented approaches to urban AI has become increasingly prominent.	Lacks empirical and network-based analysis
Cugurullo et al.	2024	Conceptualization of the algorithmic city	Conceptual–critical analysis	AI is becoming an active agent in urban knowledge production and governance.	Lack of scientometric evidence and analysis of knowledge trends
Si et al.	2025	Applications of AI in urban planning	Bibliometric and co-word analysis	Deep learning, machine learning, generative AI, and explainable AI (XAI) were identified as emerging research trends.	Primarily descriptive analysis of trends, with limited epistemological interpretation
Batty	2026	Transformation of AI in urban planning	Theoretical–interpretive article	A shift from deductive, theory-driven models toward inductive, data-driven approaches, alongside the emergence of generative AI in urban design, is identified.	Lacks empirical analysis of knowledge structures and thematic trends
Behzadfar et al.	2026	Emerging trends in sustainable transportation	Bibliometric analysis (co-word, co-authorship, and co-citation)	Electric vehicles and shared mobility were identified as emerging trends, while social equity in transportation remains comparatively underexplored.	Focused exclusively on a single subfield (transportation), without epistemological analysis or comparison with other smart city domains

A review of the existing literature indicates that previous research has primarily developed along two main trajectories. The first comprises applied and conceptual studies focusing on the technical and theoretical capacities of artificial intelligence in areas such as intelligent transportation, spatial analysis, urban governance, digital twins, and smart cities. The second encompasses review-based and scientometric studies that examine the knowledge structures,

citation networks, and the field's thematic trends. Despite these contributions, much of the existing research remains largely descriptive, focusing on patterns of knowledge production and research trends while devoting comparatively limited attention to the epistemological and critical implications of these transformations.

More specifically, it remains unclear to what extent the bibliometric changes observed in the urban AI

literature reflect deeper transformations in the logics through which urban knowledge is produced, validated, and distributed. This gap constitutes the point of departure for the present study. By integrating scientometric analysis with conceptual–critical interpretation, this study seeks to provide a multidimensional account of the knowledge structure, thematic evolution, and epistemological shifts characterizing the literature on artificial intelligence and machine learning in urban planning and smart city studies. To address the research gaps identified above, the following research questions were formulated:

RQ1. How are the knowledge structure and intellectual foundations of the literature on artificial intelligence and machine learning in urban planning organized during the 2020–2025 period?

RQ2. How have the major research themes and topics in this field evolved and shifted over the period under study?

RQ3. What are the most significant emerging trends, emerging concepts, and leading research fronts in this field?

RQ4. How has the expansion of artificial intelligence and machine learning contributed to epistemological and structural shifts in the logic of urban knowledge production?

RQ5. How have patterns of scientific collaboration and the degree of interdisciplinarity in this field evolved in recent years?

RQ6. What geographical and institutional inequalities characterize the production of urban AI-related knowledge across different regions of the world?

3. Research methodology

3.1. Research design

This structured review study adopts a scientometric–interpretive approach and employs an integrated framework comprising scientometric analysis, lexical text mining, and conceptual–critical interpretation. The unit of analysis consists of 3,578 English-language research and review articles indexed in the Web of Science and Scopus databases, selected after a screening and refinement process. In this study, we used a Thesaurus-based supervised lexical mapping method to organize the literature and distinguish it from conventional unsupervised clustering methods that rely solely on the data’s statistical structure.

Unlike conventional scientometric studies, which mainly focus on descriptive indicators, co-authorship networks, or keyword co-occurrence, the present

study considers scientometric patterns as indicators for interpreting potential changes in the knowledge structure and the logic of urban knowledge production. From this perspective, the evolution of keywords, thematic shifts, patterns of scientific collaboration, and the geographical distribution of knowledge production, alongside an overview of the literature structure, provide a basis for interpreting the field’s conceptual and epistemological transformations. All stages of data collection, cleaning, standardization, analysis, and visualization were carried out using a customized web-based analytical system, the specifications and capabilities of which are described in Section 2.4.

3.2. Data sources, search strategy, and PRISMA 2020 screening

To obtain a comprehensive set of bibliographic data and standardized metadata, two international scientific databases, Web of Science Core Collection and Scopus, were used. These two databases were selected because of their broad coverage of scientific journals, the high quality of their citation metadata, and their widespread use in scientometric studies and knowledge structure analysis. The research data were extracted on April 28, 2026, and all analyses were conducted based on this version of the databases.

The search strategy was designed iteratively and revised through several stages to achieve a balance between retrieval comprehensiveness and thematic precision. After testing the initial versions, the final query focused on two technological concepts, artificial intelligence and machine learning, and three urban concepts, urban planning, city planning, and smart city, to retrieve studies related to urban planning and smart cities with the greatest possible precision.

The final query in the Web of Science database was executed in the topic field as follows:

TS=((“artificial intelligence” OR “machine learning”) AND (“urban planning” OR “smart city” OR “city planning”))
AND PY=(2020-2025)

AND DT=(Article OR Review)

This search resulted in the retrieval of 1,909 records.

The corresponding query in the Scopus database was defined and executed as follows:

TITLE-ABS-KEY(
 (“artificial intelligence” OR “machine learning”) AND
 (“urban planning” OR “smart city” OR “city planning”))

)
AND PUBYEAR > 2019
AND PUBYEAR < 2026
AND (LIMIT-TO(DOCTYPE,"ar") OR LIMIT-
TO(DOCTYPE,"re"))
AND LIMIT-TO(LANGUAGE,"English")

This search retrieved 3,181 records.
The data refinement and screening process was conducted in accordance with the PRISMA 2020 guidelines and based on predefined inclusion and exclusion criteria. Table 2 shows the stages of this process.

Table 2. Data screening and refinement process based on the PRISMA 2020 framework (N = 3,578)

Prisma stage	Number of records	Methodological description
Identification	5,090	Initial extraction from Web of Science = 1,909 and Scopus = 3,181
Deduplication	1,010	Removal of duplicate records based on DOI, title, and metadata matching (remaining: 4,080)
Screening	460	Removal of studies unrelated to urban planning, including hardware, telecommunications, and non-urban sensors (233), agriculture and farming (135), and molecular biology and genetics (92) (remaining: 3,620)
Eligibility	42	Removal of records with structural deficiencies in the abstract or publication-year fields (remaining: 3,578)
Included	3,578	Final eligible articles included in the quantitative and qualitative synthesis

In total, 3,578 eligible articles out of the 5,090 records initially retrieved were included in the final analysis after duplicate removal, thematic screening, and metadata quality assessment. This dataset forms the basis for all scientometric analyses, text mining, thematic classification, and interpretations presented in the present study.

3.3. Data preprocessing and layered differentiation of general terms (stop-words)

The raw data extracted from the Web of Science and Scopus databases were imported into the preprocessing stage in RIS file format. At this stage, after identifying and removing duplicate records based on DOI, article title, and author metadata matching, terms were standardized with respect to spelling variations, abbreviations, synonyms, and different written forms. For example, different forms of concepts such as artificial intelligence/AI and machine learning/ML were unified to prevent artificial dispersion in term frequencies.

Since some highly frequent umbrella terms, such as

artificial intelligence, machine learning, smart city, and urban planning, appeared in a large proportion of articles, retaining them in the co-occurrence network would have resulted in the formation of very large nodes and reduced the discriminative power of the network. Therefore, the scope of applying stop-words was defined in a layered manner. These terms were excluded from the keyword co-occurrence network and keyword-based emerging trend analysis in order to reveal relationships among more specialized and contextual concepts, but were retained in the search query, thematic classification of articles, and coding of epistemological paradigms.

This differentiation ensured that the removal of stop-words was applied only to analyses in which their very high frequency could distort the network structure or the ranking of trends, without disregarding the role of these concepts in article retrieval and text-based analyses. Table 3 shows the scope of application of the selected terms, and the methodological justification for each case.

Table 3. Stop-words scope matrix

Methodological and technical rationale	Paradigm text-mining and filtering layer	Interactive graph network layer (vis-network)	Term/concept
Preventing an ineffective umbrella node in the graph / retained for filtering and paradigm coding	Retained	Removed	Artificial intelligence / AI
Removed from the lexical graph to allow contextual concepts to emerge/retained for the data-driven paradigm	Retained	Removed	Machine learning / ML

Methodological and technical rationale	Paradigm text-mining and filtering layer	Interactive graph network layer (vis-network)	Term/concept
Removed from the thematic network graph / retained in the initial query and for assessing the comprehensiveness of the articles	Retained	Removed	Smart city / smart cities
Removed from the network graph layer to neutralize non-informative connections / retained in text mining	Retained	Removed	Urban planning
Retained as a distinctive specialized keyword in the environment and sustainability cluster	Retained	Retained	Sustainable development

3.4. Customized analytical dashboard and implementation environment

To integrate the processing, analysis, and visualization stages of scientometric data, a web-based analytical dashboard was developed. This environment supported data standardization, the construction of keyword co-occurrence networks, thematic classification, temporal trend analysis, identification of emerging concepts, analysis of scientific collaboration, and examination of the geographical distribution of knowledge production within an integrated system. The dashboard was used as the environment for conducting the analyses and was not intended as a substitute for scientometric methods; rather, it enabled the consistent implementation of the various stages of analysis, data management, review of results, and generation of the outputs required for conceptual-critical interpretation. The system was developed using JavaScript. The Vis-Network library was used to display knowledge networks and interactive graphs, while Chart.js was used to generate statistical charts and temporal trends. Following data cleaning, duplicate removal, and term standardization, the data were imported into the analytical environment, and all analyses reported in this study were conducted within this environment.

To enhance the transparency and reproducibility of the analyses, the parameters used in the Barnes–Hut Force-Directed Layout algorithm for network layout were set to fixed values. The values used were as follows:

- gravitationalConstant = -2000
- centralGravity = 0.1
- springLength = 95
- springConstant = 0.04
- damping = 0.09

Furthermore, to reduce network noise, only edges with a weight of $W \geq 2$ were displayed in the analysis and visualization of the co-occurrence networks.

3.5. Analytical and validation techniques and their linkage to the research questions

To address the six research questions, a set of complementary techniques, including supervised thematic classification, network analysis, temporal trend analysis, text mining, co-authorship analysis, and geographical distribution analysis, was employed. Each technique examines one dimension of the field's knowledge structure and, together, they enable the interpretation of thematic evolution and changes in knowledge production patterns. Furthermore, to ensure the validity and robustness of the results, we applied appropriate evaluation indicators and quality-control measures at each stage of the analysis; the details are described in the subsections below.

Knowledge structure analysis and thematic classification

To address the first research question, we analyzed the knowledge structure of the literature on artificial intelligence and machine learning in urban planning using thesaurus-based supervised lexical mapping. Unlike conventional unsupervised clustering approaches, this study assigned articles to thematic domains based on predefined lexical rules to ensure greater alignment between the knowledge domains and specialized urban planning concepts.

The specialized thesaurus used in this study was developed through a systematic literature review, analysis of frequently occurring keywords, standardization of synonyms, and conceptual review by the researchers. Accordingly, five main thematic domains were defined: urban environment and sustainability; smart mobility and transportation; management, safety, and governance; spatial analysis and land use; and urban design and infrastructure. For each domain, a set of specialized keywords was specified as the rules for assigning articles.

During the classification process, the title, abstract, and author keywords of each article were examined

simultaneously, and each article was assigned to a single dominant thematic domain based on the highest degree of conceptual correspondence with the thesaurus. Therefore, thematic classification in this study was conducted as a single-label classification.

To assess the validity of the classification, two evaluators independently coded a sample of 400 articles. Cohen's kappa coefficient ($\kappa = 0.884$) indicated a very high level of agreement between the evaluators. Furthermore, to examine the robustness of the classification structure, the keyword co-occurrence network was analyzed using the Louvain community detection algorithm, and its convergence with the supervised classification was assessed using the normalized mutual information (NMI = 0.795) and adjusted Rand index (ARI = 0.821). The obtained values indicate an appropriate level of convergence between the keyword network structure and the proposed thematic classification.

Finally, the keyword co-occurrence network was visualized using the Barnes–Hut algorithm in Vis-Network to display relationships among concepts. This algorithm was used solely for node layout and to improve the visualization of the network structure and played no role in the thematic classification process or its validation.

Analysis of thematic evolution and shifts in research concepts

To address the second research question, the thematic evolution of the literature on artificial intelligence and machine learning in urban planning was examined through a time-series analysis of the distribution of articles across the five thematic domains during the 2020–2025 period. At this stage, the number and relative share of articles in each thematic domain were calculated for each year to identify gradual changes in thematic weight and the growth, decline, or stabilization of the different domains.

To avoid misleading interpretations from overall growth in publication volume, in addition to the absolute number of articles, the relative share of each domain in the total annual scientific output was also calculated. This approach enabled us to distinguish a domain's actual growth from changes resulting from the general increase in publication volume. Accordingly, the temporal trend analysis revealed which thematic domains accounted for a larger share of the literature during the study period, which domains experienced a decline in their relative share, and, consequently, how research priorities in the field of urban AI were reconfigured over time. The results of this analysis

form the basis for interpreting the thematic evolution and shifts in research priorities presented in the Results section.

In addition, to compare the rate of expansion of scientific publications across the different domains, the simple growth rate of articles in each domain relative to the base year (2020) was calculated using the following formula:

$$\text{Growth Rate} = ((N_t - N_{2020}) / N_{2020}) \times 100$$

This rate was reported for two time intervals: 2020–2025 (the entire study period) and 2020–2024 (to control for and adjust the effect of the indexing lag of articles in the final year), allowing the quantitative changes in each domain to be compared without being confused with the compound annual growth rate (CAGR).

Analysis of emerging trends and research fronts

To address the third research question, emerging keyword trends were identified and used to interpret emerging research fronts in the literature on artificial intelligence and machine learning in urban planning. This analysis assumed that concepts experiencing greater growth in recent periods than in earlier periods may indicate shifts in research priorities and the formation of emerging research areas. Therefore, emerging trends were identified based on the intensity of keyword growth rather than merely their absolute frequency.

To compare the growth intensity of concepts, a relative growth index (RGI) was defined and calculated. For this purpose, the study period was divided into two time intervals: the baseline period (2020–2022) and the recent period (2023–2025). Subsequently, for each keyword, its total frequency in each period was calculated, and the RGI was obtained according to the following equation:

$$RGI(kw) = \frac{F_{recent}(kw) - F_{baseline}(kw)}{F_{baseline}(kw) + 5}$$

where:

$$F_{recent}(kw) = \sum_{year=2023}^{2025} Freq(kw, year)$$

and

$$F_{baseline}(kw) = \sum_{year=2020}^{2022} Freq(kw, year)$$

represent the total frequency of the keyword in the recent period and the baseline period, respectively.

To prevent artificial inflation of growth rates for keywords with very low frequencies during the baseline period, a constant value of five was added to

the denominator as a smoothing constant. Furthermore, to reduce noise and increase the reliability of the results, only keywords with at least 10 occurrences in the recent period were included in the analysis. In addition, general terms lacking discriminative power that were included in the Stop-words list were removed before calculating the index. After calculating the RGI, eligible keywords were ranked in descending order according to their index values, and concepts with the highest RGI values were reported as the most prominent emerging trends. Subsequently, to assess the stability of these trends, annual changes in the frequency and relative share of the selected keywords were examined in the analytical dashboard. Thus, the identified emerging trends provided the basis for interpreting shifts in research agendas and emerging research fronts in the urban AI literature.

In this study, the RGI was used as an operational indicator to compare the relative growth intensity of concepts, with the purpose of ranking emerging trends, rather than statistically estimating growth rates or testing the significance of temporal changes.

Analysis of epistemological shifts and classification of knowledge-production paradigms

To address the fourth research question, epistemological shifts in the literature on artificial intelligence and machine learning in urban planning were analyzed through text mining of article titles, abstracts, and keywords. This analysis aimed to examine changes in the logic of knowledge production and validation and to identify shifts in the dominant epistemological approaches in this field.

The analytical framework of the study was based on four epistemological paradigms: 1) the classical statistical and analytical paradigm (deductive/statistical empiricism), 2) the simulation- and rule-based paradigm (complexity/process-driven), 3) the data-driven and predictive paradigm (inductive/algorithmic epistemology), and 4) the generative paradigm based on modern artificial intelligence (synthetic/generative epistemology). These paradigms were identified based on a set of specialized multi-word expressions (MWEs) defined as lexical indicators for each paradigm. Using these specialized expressions, rather than individual terms, reduced the likelihood of incorrect assignments resulting from ambiguous or general terms and increased classification accuracy.

The classical statistical and analytical paradigm included studies that relied on hypothesis testing, statistical analyses, and regression models, identified

through expressions such as spatial regression, linear regression, and geographically weighted regression (GWR). The simulation- and rule-based paradigm included studies that used models such as agent-based modeling, cellular automata, and system dynamics to represent and simulate urban processes. The data-driven and predictive paradigm included studies based on machine learning, deep learning, and predictive algorithms, identified through indicators such as deep learning, neural networks, random forests, and XGBoost. Finally, the generative paradigm, based on modern artificial intelligence, included studies that employed generative models, large language models, and generative design in urban analysis, design, or decision-making and was identified through expressions such as generative adversarial network (GAN), large language model (LLM), ChatGPT, and generative design.

During the coding process, the title, abstract, and keywords of each article were examined simultaneously. When an article contained lexical indicators from more than one paradigm, it was assigned to the paradigm with the highest frequency and the strongest conceptual correspondence within the article's text. To prevent statistical overlap among the paradigms and maintain the comparability of the results, classification was conducted as a single-label classification; therefore, each article was assigned to only one dominant paradigm.

After completing the coding process, the frequency and relative share of each paradigm were calculated for the entire study period and separately by year to examine the evolution of knowledge production logics from 2020 to 2025. Furthermore, to analyze the relationship between thematic domains and epistemological approaches, a cross-tabulation matrix between the four epistemological paradigms and the five thematic clusters was constructed to analyze the distribution of knowledge-production paradigms across different domains of the urban AI literature.

To evaluate the proposed framework's performance, the coding process was validated using a reference dataset. An F1-Score of 0.907 indicated the framework's high accuracy in identifying epistemological paradigms and reducing errors from assigning individual terms.

Analysis of scientific collaboration and the interdisciplinarity of the field

To address the fifth research question, patterns of scientific collaboration and the degree of interdisciplinarity in the literature on artificial

intelligence and machine learning in urban planning were analyzed. At this stage, the collaboration index was calculated based on the average number of authors per article, and co-authorship networks, prolific authors, and their institutional affiliations were examined. In addition, the diversity of institutional affiliations and the subject categories of the articles were analyzed as an indicator of the interdisciplinarity of the field. This analysis aimed to assess the extent to which knowledge production shifts from individual and single-disciplinary patterns toward team-based, multi-institutional, and interdisciplinary research networks and to explain the role of scientific collaboration in the expansion of the urban AI literature.

Analysis of geographical and institutional inequalities in knowledge production

To address the sixth research question, information on the authors' institutional affiliations and the countries in which they were located was extracted from the article metadata, integrated, and standardized. Subsequently, the geographical distribution of scientific publications, the shares of different countries, patterns of international collaboration, and the degree of concentration of knowledge production at the global scale were analyzed. This analysis was conducted to identify the main centers of knowledge production, examine the degree of concentration or dispersion of research activities, and assess the distribution of scientific capacities in artificial intelligence and machine learning in urban planning. In this study, patterns of geographical distribution were not regarded merely as indicators for counting scientific publications or ranking countries; they were rather interpreted as evidence of institutional differences and existing inequalities within the global knowledge-production ecosystem. Accordingly, the concentration of scientific publications in certain countries may reflect differences in access to data infrastructures, computational capacity, research investment, international collaboration networks, and institutional capacities for the development and application of artificial intelligence.

From this perspective, the analysis of the geography of knowledge production is not limited to describing where research is published; rather, it makes it possible to examine how the unequal distribution of scientific and technological resources may affect the orientation of research agendas, the formation of knowledge frontiers, the development of artificial intelligence methods, and even the generalizability and

transferability of research findings in the context of urban planning. For this reason, this section's results provide a basis for interpreting the institutional and geographical dimensions of the evolution of the artificial intelligence literature in urban planning.

3.6. Validation, robustness, and reviewability of the analysis

To strengthen the validity of the findings and ensure their robustness, methodological controls were applied at four levels: data quality, validity of thematic classification, conceptual coherence of the results, and reviewability of the analytical process. Because the present study combines scientometric analysis, supervised lexical classification, and conceptual interpretation, the validity of the results was not limited solely to data processing accuracy; rather, the conceptual consistency of the classifications and the transparency of the analytical process were also evaluated.

In the first step, data quality was controlled by removing duplicate records, standardizing keywords, integrating spelling variants and synonyms, and applying the layered differentiation policy for general terms in accordance with Section 2.3. This process prevented the removal of conceptual information required for text-mining analyses and thematic classification, while reducing noise in the co-occurrence networks.

In the second step, two evaluators independently coded 400 articles to examine the validity of the thematic classification. The Cohen's kappa coefficient ($\kappa = 0.884$) indicated a very high level of agreement between the evaluators. To assess the robustness of the extracted structure, the results of the supervised classification were compared with the output of the unsupervised Louvain Community Detection clustering algorithm. The values of $NMI = 0.795$ and $ARI = 0.821$ also indicated acceptable convergence between the structure of the keyword network and the study's conceptual classification.

In the third step, the conceptual coherence of the thematic domains was evaluated based on central keywords, co-occurrence relationships, and their correspondence with the specialized literature. In addition, we manually reviewed a portion of the classification results to identify potential cases of incorrect assignment, conceptual ambiguity, or inconsistencies in domains' naming. This review complemented the statistical indicators and increased the interpretability of the results.

Finally, to strengthen the study's transparency and reviewability, we documented all stages of data collection, cleaning, standardization, and analysis, and made a viewable version of the analytical dashboard, along with explanations of the processing procedure, available on the research website (futureresearcher.ir). However, fully reproducing the results requires access to the original data, the specialized thesaurus, the

standardization rules, and identical analytical settings; therefore, this study emphasizes methodological transparency, reviewability, and conditional reproducibility.

Figure 1 presents the study's overall methodological framework for examining thematic evolution and changes in approaches to knowledge production.

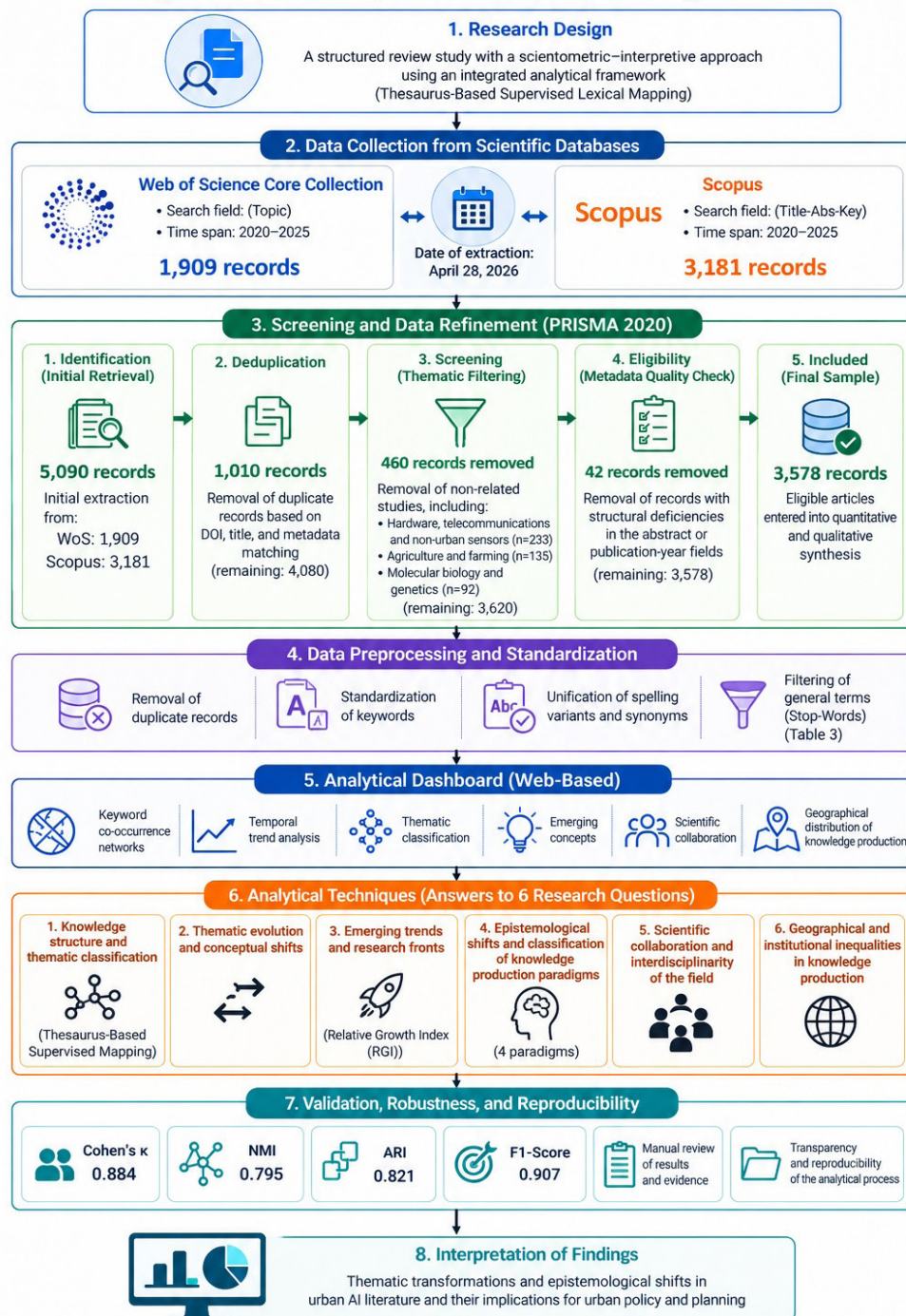


Figure 1. Thematic design of the research methodological process

4. Results

The study results are organized to address the six research questions and are based on the analytical framework presented in the Methodology section. This section first examines the field's knowledge structure and then analyzes thematic evolution, emerging trends, epistemological shifts, patterns of scientific collaboration, and the geographical distribution of knowledge production. This sequence provides a coherent picture of how the knowledge structure, research orientations, and logic of knowledge production have evolved in the fields of artificial intelligence and machine learning in urban planning and smart city studies. All reported results were extracted and analyzed based on the final dataset of 3,578 articles.

4.1. Knowledge structure and intellectual cores of the field

To address the first research question, we analyzed the knowledge structure of the literature on artificial intelligence and machine learning in urban planning during the 2020–2025 period. Of the 5,090 records initially extracted from the Web of Science and Scopus databases, 3,578 articles were included in the final sample after duplicate removal, thematic screening, and metadata assessment. Based on the thesaurus-based supervised lexical mapping method described in Section 2.5.1, each article was assigned to only one of the five main thematic domains according to the highest degree of conceptual correspondence between its title, abstract, and keywords and the specialized thesaurus.

As shown in Table 4, urban sustainability, with 1,143 articles (31.9%), constitutes the largest thematic domain in the literature. It is followed by smart mobility (968 articles, 27.1%), and governance &

management (910 articles, 25.4%), ranking second and third, respectively. The two domains of spatial analytics & land use (305 articles, 8.5%) and urban design & infrastructure (252 articles, 7.0%) account for smaller shares of the article set. Given the single-label nature of the classification system, each article was classified into only one of these five domains; therefore, the sum of the domain frequencies equals the total number of articles in the final sample (3,578 articles). The slight difference between the sum of the percentages and 100% is due to the rounding of decimal values.

The thematic distribution pattern indicates that the field's knowledge structure is primarily organized around three axes: environmental sustainability, smart mobility, and urban governance, which together account for 84.4% of all articles. The dominance of the urban sustainability domain indicates that artificial intelligence applications in urban planning have shifted toward sustainability, climate change, pollutant emissions, and urban environmental management. Alongside this, the high share of the smart mobility domain indicates the continued prominence of topics such as traffic management, autonomous vehicles, and mobility optimization, while the substantial presence of the governance & management domain indicates the growing importance of decision-support systems, data governance, cybersecurity, and smart management in the literature of this field. In contrast, although spatial analytics & land use and urban design & infrastructure account for smaller shares of scientific publications, they are still recognized as distinct and complementary components of the knowledge structure of urban AI. Table 4 reports the distribution of these five thematic domains, along with each domain's frequency and share.

Table 4. Distribution of thematic domains in the literature on artificial intelligence and machine learning in urban planning based on supervised thematic classification (2020–2025)

Thematic cluster	Frequency	Percentage	Most frequent keywords
Urban sustainability	1,143	31.9%	sustainability, carbon emissions, ...
Smart mobility	968	27.1%	congestion, autonomous vehicles, ...
Governance & management	910	25.4%	governance, decision support system, ...
Spatial analytics & land use	305	8.5%	geographic information systems, land use, ...
Urban design & infrastructure (3d modeling)	252	7.0%	information modeling, BIM, digital twin, ...
Total articles	3,578	100%	—

To complete the representation of the field's conceptual structure, Figure 2 presents a word cloud of the frequently occurring concepts extracted from the article keywords. In this visualization, each term's size is proportional to its frequency in the dataset; therefore, larger terms represent concepts that appeared in more articles. Concepts such as Internet of Things, remote sensing, learning systems, big data, sustainable development, decision-making, digital twin, GIS, blockchain, and cybersecurity are among

the most frequently observed concepts. This pattern indicates that the urban AI literature is based on a diverse set of topics including smart technologies, environmental sustainability, spatial analysis, digital infrastructures, and urban governance. While the word cloud provides an overall representation of the distribution and relative importance of concepts, the thematic classification presented in Table 4 systematically organizes these concepts into five main knowledge domains.

Word Cloud of Key Concepts



Figure 2. Word cloud of frequently occurring concepts in the literature on artificial intelligence and machine learning in urban planning (2020–2025)

4.2. Thematic evolution and shifts in research focus

To address the second research question, we examined the evolution of thematic domains during the 2020–2025 period from two perspectives: relative share and absolute number of articles. Figure 3 shows changes in the relative shares of the domains; Figure 4 shows trends in the number of articles, and Table 5 shows the quantitative data for both indicators.

The data indicate that the relative shares of the thematic domains were not constant throughout the study period and underwent considerable changes. The share of the urban sustainability domain increased from 17.5% in 2020 to 48.8% in 2025 and, particularly in the final years, became the unequivocally leading domain in the literature. In contrast, the governance & management domain changed from 31.1% to 16.4%, while the smart mobility domain changed from 29.1% to 20.3%. During the same period, the spatial analytics & land use domain declined from 13.6% to 10.2%, and that of urban design & infrastructure decreased from 8.7% to 4.3%.

These changes indicate a gradual reconfiguration of research priorities in the literature on artificial intelligence and machine learning in urban planning. The continuous increase in the share of the urban sustainability domain indicates that topics such as sustainability, climate change, urban resilience, and

environmental management have gradually become the main focus of research in this field. In contrast, the decline in the relative shares of governance & management and smart mobility does not necessarily indicate a decline in research activity in these domains; rather, it reflects the more rapid growth of other thematic domains, particularly urban sustainability. Examining the absolute number of articles provides a complementary picture of this trend. As shown in Table 5, the number of articles in all thematic domains increased consistently from 2020 to 2024. For example, the number of articles in the urban sustainability domain increased from 18 articles in 2020 to 376 articles in 2024, representing a growth of 1,988.9%. Over the same period, the number of articles in governance & management increased from 32 to 246, smart mobility from 30 to 274, spatial analytics & land use from 14 to 89, and urban design & infrastructure from 9 to 74.

A simultaneous comparison of the relative shares and absolute numbers of articles indicates that the decline in some domains is not due to decreased scientific output, but rather to uneven growth across the different domains. For example, although the number of articles in smart mobility increased from 30 to 274, its relative share of the overall literature declined because the urban sustainability domain expanded at a much higher growth rate. The same pattern is

observed in governance & management, where, despite a substantial increase in the number of articles, its relative share of the total scientific output declined. Caution is also required when interpreting the final year's trends. The decrease in the number of articles in

2025 does not necessarily indicate an actual decline in research activity. But it may result from an indexing lag in the Web of Science and Scopus databases. Therefore, the trends observed in 2025 should be interpreted with this limitation in mind.

Changes in the Relative Share of Thematic Domains (2020–2025)

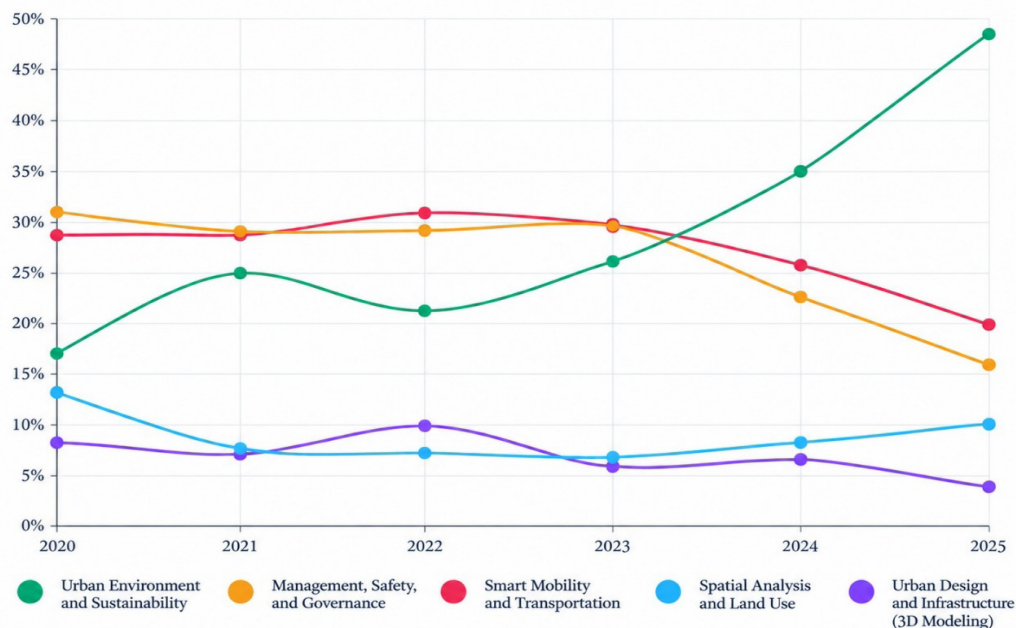


Figure 3. Changes in the relative shares of thematic domains in the urban AI literature during 2020–2025

Trends in the Number of Articles in Five Thematic Domains (2020–2025)

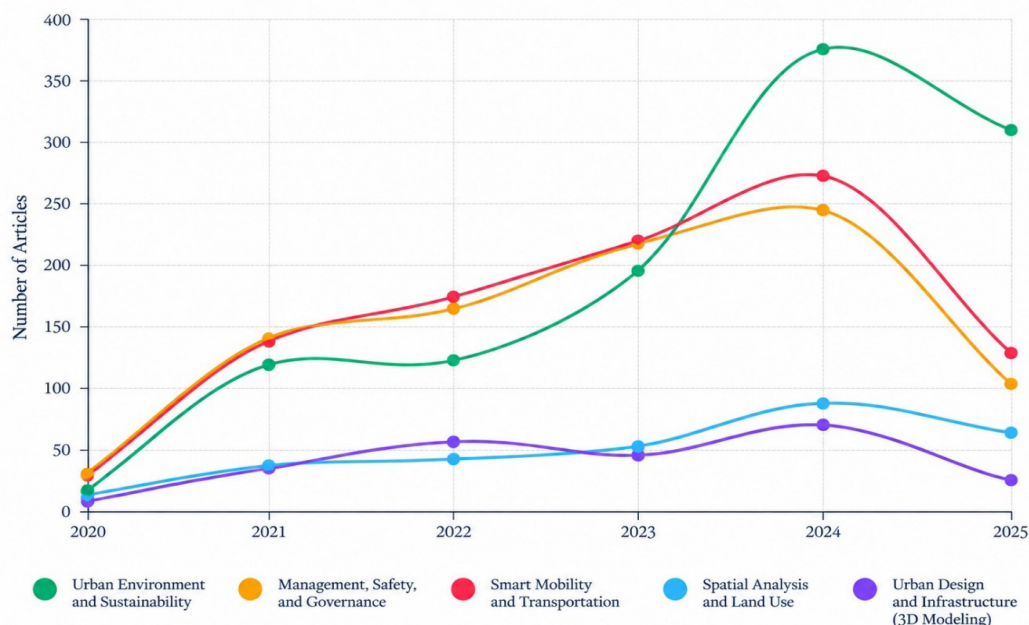


Figure 4. Annual trend in the number of articles across the five thematic domains of the urban AI literature during 2020–2025

Table 5. Annual distribution of scientific publications, relative shares, and growth rates of the five thematic domains in the urban AI literature (2020–2025)

Domain/keyword	2020	2021	2022	2023	2024	2025	Growth rate (2020–2025)	Growth rate (2020–2024)
Urban sustainability	18	120	123	196	376	310	1,622.2%	1,988.9%
Governance & management	32	140	168	220	246	104	225.0%	668.8%
Smart mobility	30	138	175	222	274	129	330.0%	813.3%
Spatial analytics & land use	14	39	44	54	89	65	364.3%	535.7%
Urban design & infrastructure (3D modeling)	9	37	57	48	74	27	200.0%	722.2%
Total annual publications (fully consistent with N = 3,578)	103	474	567	740	1,059	635	516.5%	928.2%

4.3. Emerging trends and research fronts

Analysis of the relative growth index (RGI) indicates that emerging trends in the literature on artificial intelligence and machine learning in urban planning during the 2020–2025 period have primarily formed around five thematic areas. As shown in Table 6, these concepts, in addition to having considerable frequencies within the dataset, also experienced the greatest growth intensity between the baseline period (2020–2022) and the recent period (2023–2025).

Among the concepts examined, urban resilience recorded the highest growth rate, with RGI = 7.76. It was followed by carbon emissions, with RGI = 6.00; digital twin, with RGI = 4.70; cybersecurity, with RGI = 3.37; and forestry (monitoring of forests and urban green spaces), with RGI = 3.00 (Table 6). In contrast, in terms of articles' total frequency, urban resilience, carbon emissions, and digital twin accounted for the largest volumes of scientific publications among the emerging topics (Table 6).

The annual changes in these concepts also reveal different patterns (Figure 5). Although frequencies of all topics increased during the recent period, their growth trajectories were not identical. Urban resilience

and carbon emissions show substantial increases from 2023 onward, whereas digital twin and cybersecurity followed more continuous trends but with lower growth intensity. Despite its lower frequency, forestry also experienced noticeable growth during the final years of the study period.

A comparison of the RGI values (Figure 6) indicates that the growth intensity of emerging concepts does not necessarily correspond to their total frequency. For example, despite having fewer articles, forestry ranks among the top five topics in terms of growth rate; in contrast, some concepts with higher frequencies experienced more moderate relative growth. This difference indicates that the RGI can identify emerging topics independently of the absolute volume of scientific publications.

Overall, the results of this section indicate that the most important emerging trends in the urban AI literature are concentrated around three broad axes: resilience and environmental sustainability, digital representation of the city, and the security of smart infrastructures. These axes exhibited the greatest growth intensity during the period under study.

Table 6. Periodic frequencies and relative growth index (RGI) of selected emerging concepts in the literature on artificial intelligence and machine learning in urban planning and smart city studies (2020–2025)

Keyword/specialized concept	Baseline-period frequency (2020–2022), Fbase	Recent-period frequency (2023–2025), Frecent	Relative growth index (RGI)	Trend interpretation
Urban resilience	12	144	7.76	Real and very strong growth
Carbon emissions	15	135	6.00	High real growth
Digital twin	18	126	4.70	Genuine emerging trend
Cybersecurity	14	78	3.37	Emerging trend in urban governance
Forestry	4	31	3.00	Emerging environmental trend

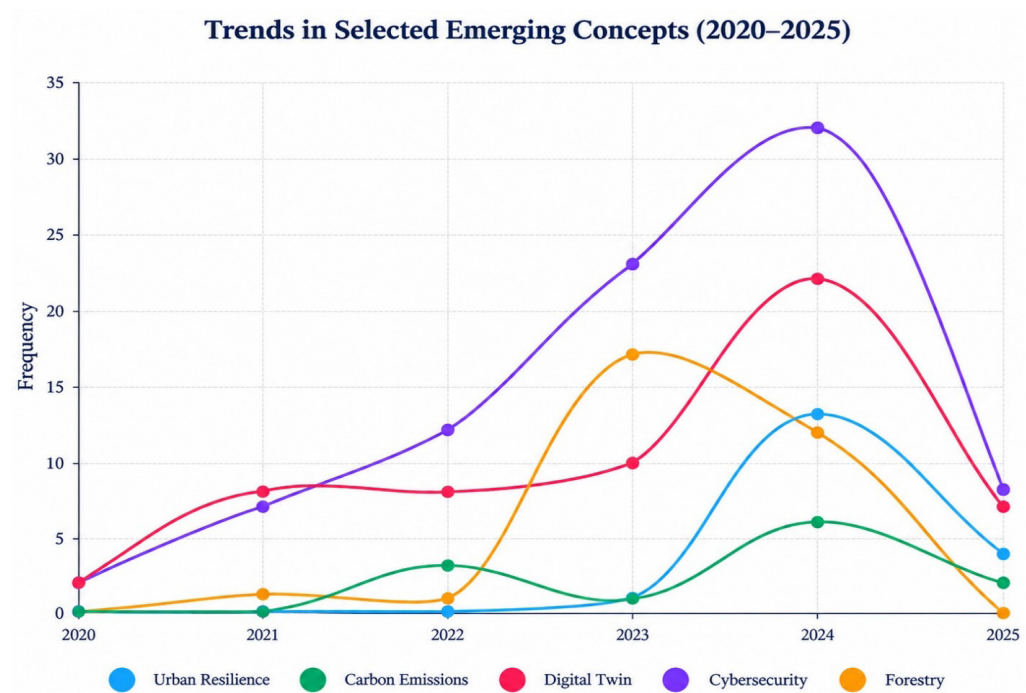


Figure 5. Temporal trends of selected emerging concepts (2020–2025)

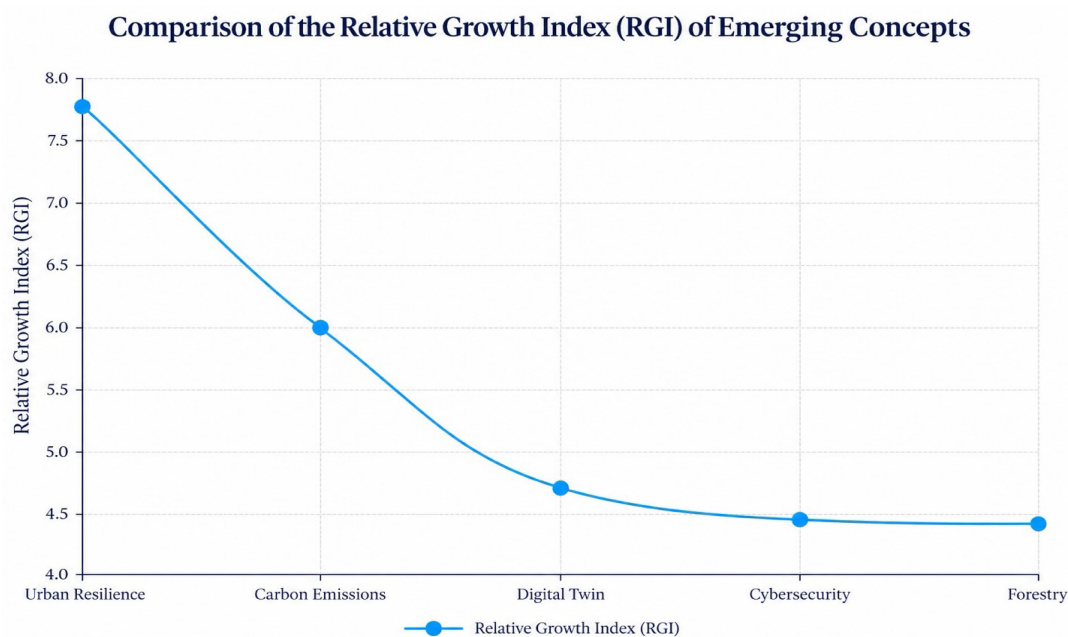


Figure 6. Comparison of the relative growth index (RGI) of selected emerging concepts (2020–2025)

4.4. Epistemological shifts and the logic of knowledge production

The analysis of epistemological paradigms indicates that research on artificial intelligence and machine learning in urban planning during the 2020–2025 period was predominantly data-driven and predictive. As shown in Table 7, of the 3,578 articles examined,

2,781 articles (77.7%) were classified within the data-driven and predictive paradigm. This was followed by the generative paradigm based on modern artificial intelligence (296 articles, 8.3%), the classical statistical and analytical paradigm (99 articles, 2.8%), and the simulation- and rule-based paradigm (73 articles, 2.0%). In addition, 329 articles (9.2%) were not

assigned to any of the four epistemological paradigms because of their conceptual or review nature and the absence of a clearly specified analytical method.

An examination of the temporal trends of the paradigms indicates that the data-driven and predictive paradigm maintained its dominant position throughout the study period (Table 7). Although this paradigm's relative share declined slightly in the final years (Figure 7), its absolute number of articles continued to increase, indicating the continued expansion of the use of machine learning, deep learning, and predictive models in urban studies (Figure 7).

In contrast, the generative paradigm based on modern artificial intelligence experienced an upward trend from 2023 onward (Figures 9 and 10). The gradual increase in the frequency and relative share of this paradigm indicates that generative models, large language models (LLMs), digital twins, and generative design are gradually entering the mainstream of urban AI research. During the same period, the classical statistical and analytical paradigm and the simulation- and rule-based paradigm accounted for smaller shares

of scientific publications compared with the other two paradigms.

The distribution of epistemological paradigms across the five thematic domains also shows a similar pattern. As shown in Figure 8, the data-driven and predictive paradigm has the highest frequency across all thematic domains, although its dominance varies among the different clusters. In contrast, the generative paradigm is more prominent in domains such as urban design & infrastructure and governance & management, while the statistical and simulation paradigms account for limited shares across all domains (Figure 8).

Overall, the results of this section indicate that the methodological structure of the literature on artificial intelligence and machine learning in urban planning in recent years has been predominantly based on data analysis, algorithmic learning, and predictive models. At the same time, the gradual growth of the generative paradigm from 2023 onward indicates that approaches based on content generation, scenario development, and intelligent design are also gradually consolidating their position in this field.

Table 7. Frequency and relative share of epistemological paradigms in the literature on artificial intelligence and machine learning in urban planning and smart city studies (2020–2025, N = 3,578)

Methodological and epistemological logic	Percentage of total (n = 3,578)	Number of articles	Epistemological paradigm
Dominance of deep learning, big data, and predictive algorithms	77.7%	2,781	Data-driven and predictive
Generative models (GANs), digital twins, and LLMs	8.3%	296	Generative and based on modern artificial intelligence
Spatial regression, GWR, and classical statistical analyses	2.8%	99	Classical statistical and analytical
Agent-based modeling (ABM) and cellular automata (CA)	2.0%	73	Simulation- and rule-based
Articles with explicit methodological coding	90.8%	3,249	Total articles with methodological labels
Purely review and conceptual articles without methodological coding	9.2%	329	Conceptual/qualitative articles without methodological labels

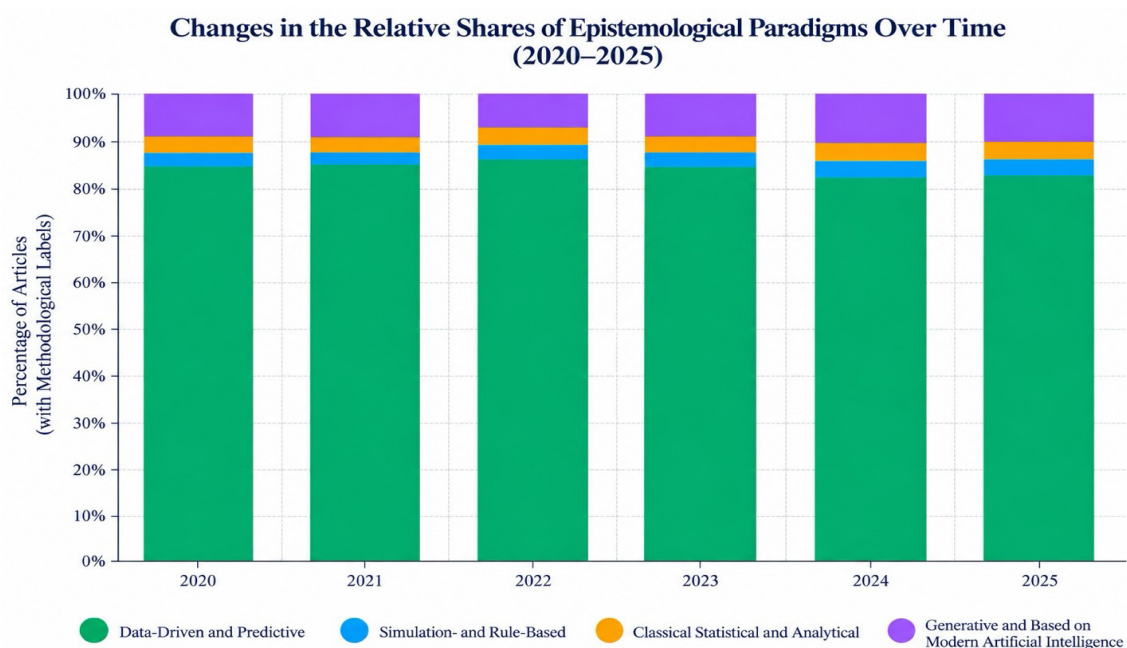


Figure 7. Changes in the relative shares of epistemological paradigms in the literature on artificial intelligence and machine learning in urban planning and smart city studies (2020–2025)

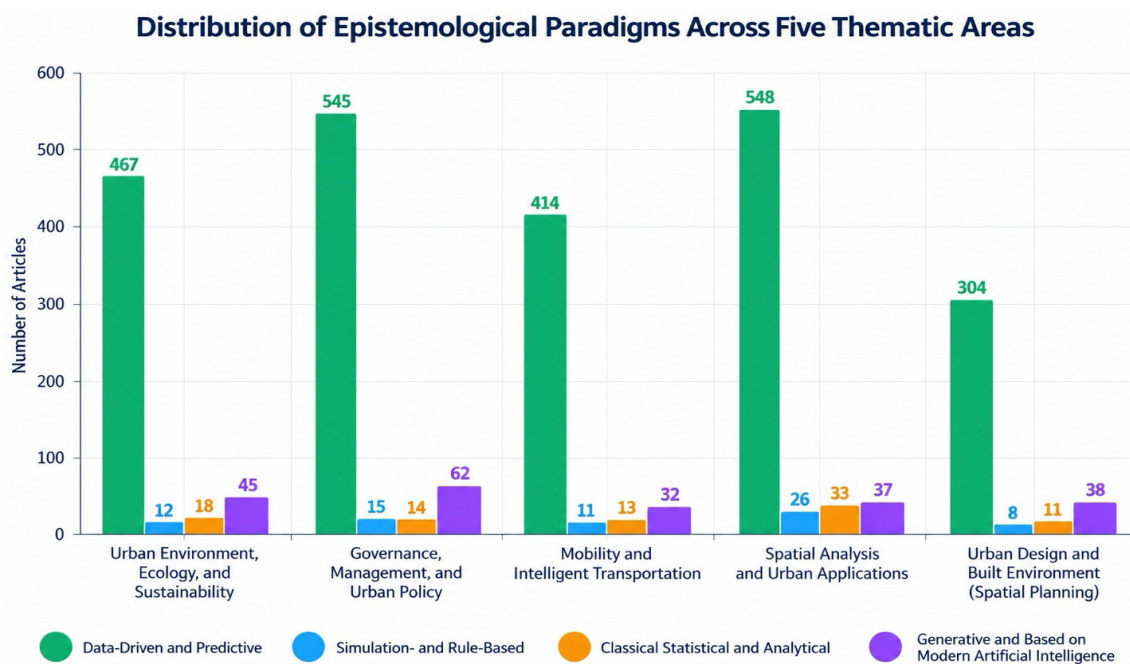


Figure 8. Distribution of Epistemological paradigms across the five thematic domains of the literature on artificial intelligence and machine learning in urban planning and smart city studies

4.5. Evolution of scientific collaboration and the interdisciplinarity of the field

The results concerning patterns of scientific collaboration indicate that knowledge production in artificial intelligence and machine learning in urban planning has predominantly been conducted through

multi-authored research. The average number of authors per article ranged from 4.2 to 4.5 during 2020–2025, indicating the predominance of team-based research in this field (Figure 9). Furthermore, a Stirling–Rao diversity index value of 0.74 and the distribution of Web of Science subject categories,

including urban studies (28%), computer science (24%), and environmental sciences (22%), indicate a high level of interaction among different disciplines in the production of knowledge in this field (Table 8). The temporal trend of the collaboration index indicates that, following a relative decline in 2021, the average number of authors increased through 2023 and 2024 and declined slightly in 2025 (Figure 9). Despite these fluctuations, the index remained relatively high throughout all years of the study period, indicating the continued predominance of collaborative research in knowledge production in this field.

An examination of the distribution of prolific authors also indicates that scientific publications are distributed among a large number of researchers, with little concentration among particular individuals. As shown in Table 8 and Figure 10, the most prolific researcher (Li, X.), with 26 articles, accounted for only 0.7% of all articles. This was followed by Wang, Y., with

23 articles; Liu, Y., with 22 articles; and Li, Y. and Zhang, J., each with 20 articles. The low shares of even the most active researchers indicate that knowledge production in this field is based on a broad network of researchers and scientific collaborations, with little dependence on a limited number of individuals or groups.

Overall, the results of this section indicate that the literature on artificial intelligence and machine learning in urban planning has developed within an interdisciplinary context and through scientific collaboration. The relatively high average number of authors per article, the value of the Stirling–Rao diversity index, and the broad distribution of participation among researchers all indicate that knowledge production in this field is based on interaction among different areas of expertise and extensive research collaboration.

Trend in the Average Number of Co-authors per Article (2020–2025)

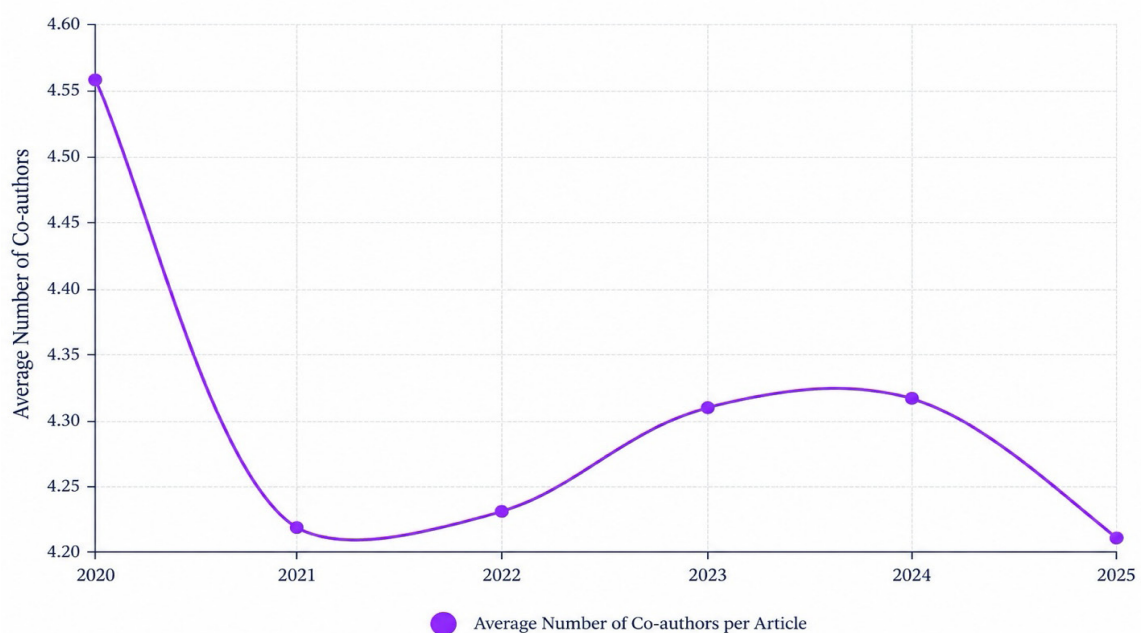


Figure 9. Trend in the average number of authors per article (collaboration index) in the literature on artificial intelligence and machine learning in urban planning and smart city studies (2020–2025)

Table 8. Leading authors based on the number of articles and their share of participation in knowledge production in the field of artificial intelligence and machine learning in urban planning and smart city studies (2020–2025)

Rank	Author	Number of articles	Participation (%)
1	Li, X.	26	0.7%
2	Wang, Y.	23	0.6%
3	Liu, Y.	22	0.6%
4	Li, Y.	20	0.6%

Rank	Author	Number of articles	Participation (%)
5	Zhang, J.	20	0.6%
6	Zhang, Y.	19	0.5%
7	Khan, M.A.	18	0.5%
8	Li, J.	18	0.5%
9	Wang, X.	18	0.5%
10	Zhang, Z.	17	0.5%

Top 10 Authors by Number of Articles in the Dataset

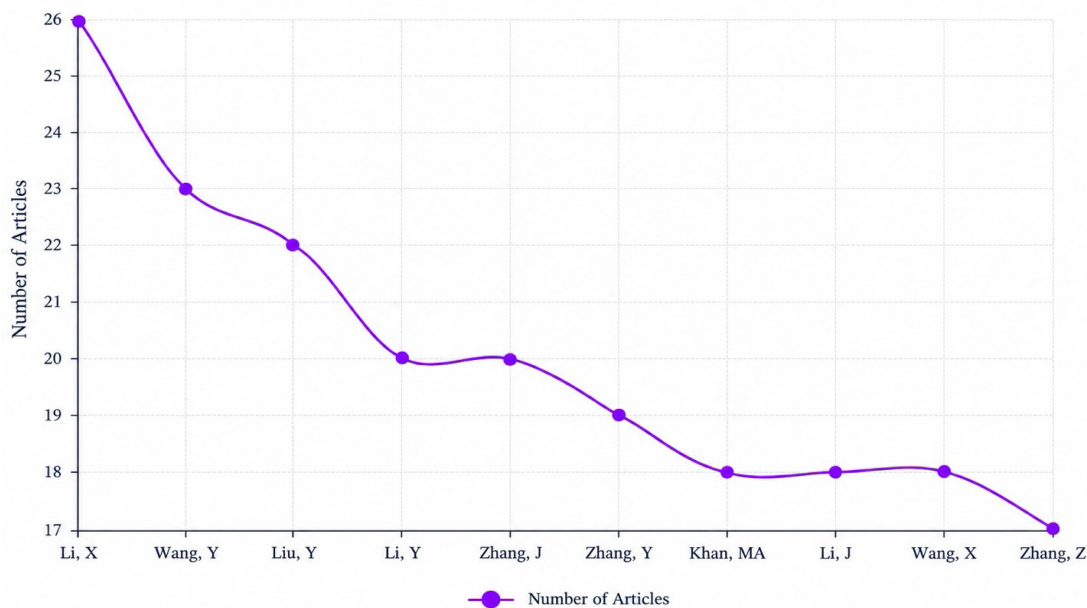


Figure 10. Comparison of the number of articles published by the ten leading authors in the field of artificial intelligence and machine learning in urban planning and smart city studies (2020–2025)

4.6. Geographical pattern and concentration of knowledge production

Analysis of the geographical distribution of knowledge production, based on authors' institutional affiliations, indicates that research on artificial intelligence and machine learning in urban planning is noticeably concentrated in a limited number of countries. In this method, each country was counted based on the institutional affiliations of the article authors; therefore, in articles resulting from international collaborations, a single article may be attributed to more than one country, and consequently, the total number of country contributions does not necessarily equal the total number of articles.

As shown in Table 9, China ranks first with 1,017 contributions (28.4%). The United States, with 778 contributions (21.7%), and India, with 558

contributions (15.6%), rank second and third, respectively. Saudi Arabia (362 contributions, 10.1%), South Korea (200 contributions, 5.6%), and Australia (186 contributions, 5.2%) are also among the leading countries in this field. Canada, Italy, the United Kingdom, and Spain, although accounting for smaller but still considerable shares, are also among the ten leading countries in knowledge production (Table 9). The distribution of contributions indicates a noticeable gap between the top three countries and the remaining countries. China alone accounts for more than one-quarter of all scientific contributions and, together with the United States and India, constitutes a substantial proportion of the scientific output in this field (Figure 15). In contrast, the shares of the other leading countries are considerably smaller and declining, indicating a substantial concentration of

knowledge-production capacity in a limited number of countries. At the same time, the presence of countries from East Asia, North America, Europe, Oceania, and the Middle East among the top ten indicates that research on artificial intelligence and machine learning in urban planning has expanded globally; however, the intensity of participation remains uneven across

countries (Table 9).

Overall, the results of this section indicate that although knowledge production in this field has spread geographically across different regions of the world, the pattern of scientific participation is still significantly concentrated, and the majority of scientific production comes from a limited number of countries.

Table 9. Leading countries based on the number of institutional contributions and their share of total scientific contributions in research on artificial intelligence and machine learning in urban planning and smart city studies (2020–2025)

Rank	Country	Number of institutional contributions	Percentage share
1	China	1,017	28.4%
2	USA	778	21.7%
3	India	558	15.6%
4	Saudi Arabia	362	10.1%
5	South Korea	200	5.6%
6	Australia	186	5.2%
7	Canada	159	4.4%
8	Italy	154	4.3%
9	UK	146	4.1%
10	Spain	121	3.4%

5. Discussion

The findings of the present study can be interpreted in relation to the theoretical issue raised in the Introduction: whether the expansion of artificial intelligence and machine learning in urban planning and smart city studies has merely led to the development of analytical tools or has also been accompanied by changes in the logic of knowledge production. The results indicate that the observed developments cannot be reduced solely to an increase in the volume of scientific publications or the expansion of technological applications. Between 2020 and 2025, a form of intra-disciplinary reconfiguration within the urban AI literature is observed, in which thematic changes, shifts in research priorities, and transformations in inferential logics have occurred simultaneously. However, these findings do not constitute evidence of a historical transition across the entire discipline of urban planning; rather, they indicate that, within the urban AI literature, the structure and orientation of knowledge production are changing.

The first indication of this reconfiguration is observed in the gradual shift in the literature's thematic center of gravity. The dominance of urban sustainability and the rapid growth of topics such as urban resilience,

carbon emissions, and digital technologies indicate that the value of AI in the new literature is increasingly defined in relation to addressing complex urban problems rather than merely enhancing the efficiency of computational tools. In the first generation of smart city studies, technology was often evaluated in terms of connectivity, the development of digital infrastructure, and the optimization of system performance; however, the present findings indicate that a substantial proportion of current research examines AI in relation to climate adaptation, energy management, carbon-emission reduction, and urban resilience. From this perspective, a gradual shift from "technology-centered" to "problem-centered" approaches can be identified. This interpretation is consistent with studies emphasizing the convergence of AI, big data, and sustainable urban planning (Yigitcanlar et al., 2020; Bibri, 2021; Son et al., 2023). However, the findings of the present study clarify two points. First, growth across the domains is uneven, and some topics, such as urban resilience, cybersecurity, and digital twins, have grown more rapidly than established domains. Second, a decline in a domain's relative share does not necessarily indicate a decline in its scientific output. For example, smart mobility continued to grow in absolute number of

articles but expanded at a slower rate than domains such as sustainability. Therefore, the observed transformation is better understood as an “uneven expansion of domains” rather than the complete replacement of one research agenda by another.

Within this framework, the continued prominence of smart mobility indicates that established AI applications, ranging from traffic prediction and congestion management to route optimization and the development of intelligent transportation systems, remain one of the field’s main pillars. This finding is also consistent with Behzadfar et al. (2026), who, in a separate bibliometric analysis of the sustainable transportation literature, showed that electric vehicles and shared mobility systems are among the most important emerging trends, while topics such as social equity in transportation remain relatively underexplored. The latter point is also consistent with the present study’s finding regarding the growing importance of justice and inequality in the urban AI literature. The difference is that these issues are no longer framed independently and solely as engineering problems but are increasingly linked to energy consumption, pollutant emissions, land-use patterns, equity of access, and data governance. In other words, previous research agendas have not disappeared; rather, they have been redefined within a broader network of urban issues.

The prominent presence of governance & management also indicates that the urban AI literature has increasingly attended to the institutional and governance implications of technology. Cities’ growing dependence on sensors, data platforms, decision-support systems, and predictive models has increasingly linked urban management to data security, privacy, trust, transparency, and algorithmic accountability. This interpretation aligns with Kitchin’s discussions in 2016 and 2017 on algorithmic rationality and with Leszczynski’s 2020 analysis of urban life dependence on digital infrastructures. Similarly, Kitchin and Moore-Cherry (2021) also show that data ecosystem governance in smart city districts is often fragmented and multi-institutional, highlighting the need to consider institutional coordination alongside the technical development of smart systems. In this sense, the research question is shifting from “How can we build a more accurate model?” toward “How should this model be evaluated, governed, and held accountable?”

This transformation can also be understood through the lens of AI urbanism. Cugurullo et al. (2024) argue

that AI in contemporary cities is no longer merely a supplementary tool for urban management but can play a more active role in the organization, planning, and experience of the city. The findings of the present study align with this perspective, as they indicate that topics such as cybersecurity, data governance, and decision-support systems are not peripheral to the literature but have become part of its core. Nevertheless, this consistency should be interpreted cautiously, because the present study did not directly measure power, the legitimacy of decisions, or the actual consequences of algorithmic governance. Another dimension of this reconfiguration is the strengthening position of spatial analysis and technologies based on geospatial data. The development of GeoAI has demonstrated that geographical data are no longer used merely to describe and map space but have become a basis for pattern discovery, predicting change, and modeling spatial futures. This transformation aligns with Janowicz et al. (2020) and with emerging directions in GeoAI, including geospatial foundation models, interpretability, and data justice (Mai et al., 2025). From this perspective, the evolution of GeoAI can be regarded as both technical and epistemological, because data quality, scale of analysis, and the manner in which space is represented affect what can be known about the city.

At the level of emerging concepts, the growth of urban resilience, digital twins, and cybersecurity highlights three important orientations: managing uncertainty, dynamic and scenario-based representation of the city, and governance of digital infrastructures. The growth of urban resilience indicates that the literature is shifting from managing normal conditions toward prediction, adaptation, and recovery under crisis conditions. The growth of digital twins also indicates a shift from static representations toward real-time, interactive, and scenario-based models, a trend that has coincided with the expansion of generative AI and multimodal models (Xu et al., 2024; Zhang et al., 2024). On the other hand, the growth of cybersecurity indicates that as the capacity for digital understanding and control of the city increases, risks related to privacy, trust, and the continuity of public services also become increasingly important (UN-Habitat, 2022).

At the epistemological level, the dominance of the data-driven and predictive paradigm indicates that a substantial proportion of the literature examined is based on pattern extraction, learning from data, and prediction. This finding is consistent with Batty’s (2026)

discussion of the shift from deductive and theory-driven models toward inductive and data-driven models. But it does not imply eliminating theory, statistical analysis, or simulation. The single-label classification recorded only each article's dominant logic, and many studies may have employed combinations of several methods. Therefore, what has changed is not the elimination of other approaches but a shift in the relative weight of knowledge-production logics.

In contrast, the gradual growth of generative approaches indicates that urban AI is moving from "analysis and prediction" toward "scenario and option generation." This transformation also changes the evaluation criteria. Whereas predictive accuracy and classification performance were central concerns in the previous generation, questions are now emerging regarding the validity of generated scenarios, the values embedded in outputs, responsibility for algorithmic recommendations, and their explainability. From this perspective, the planner's role is not eliminated but redefined; the planner is gradually becoming a mediator among data, models, institutional rules, and public values.

The findings on scientific collaboration also indicate that knowledge production in this field has acquired an interdisciplinary character. Research team size alone is not sufficient to demonstrate interdisciplinarity; however, when considered alongside the diversity of subject categories and the Stirling–Rao index, it indicates that urban AI has developed through interactions among urban planning, computer science, environmental sciences, and other related fields. In contrast, knowledge production remains geographically concentrated. This concentration may be related to differences among countries in data infrastructure, computational capacity, and research investment, although these factors were not directly measured in the present study.

The importance of this finding lies in the fact that the geography of knowledge production can influence the types of problems defined, the data used, and the generalizability of models. Therefore, justice in urban AI cannot be examined solely at the level of model outputs. Access to data, the capacity to develop models, computational resources, and opportunities to participate in problem definition are also part of the issue of justice. This is particularly important for cities in the Global South, as they may become consumers of models and standards developed in different social

and institutional contexts.

Overall, the main contribution of the present study is to demonstrate that scientometric analysis, when combined with conceptual and epistemological interpretation, can move beyond describing frequencies and networks and provide indications of changes in the dominant logics of knowledge production. On this basis, AI in the urban planning literature is gradually evolving from a set of analytical tools into part of the infrastructure of knowledge production and decision support. This interpretation should be understood within the limits of the study's evidence and as an interpretive conclusion, rather than as direct evidence of changes in power relations or governance.

6. Conclusion

This study aimed to analyze thematic evolution and epistemological shifts in the literature on artificial intelligence and machine learning in urban planning and smart city studies during the 2020–2025 period. Unlike the majority of scientometric studies, which mainly described citation structures, collaboration networks, or publication trends, this study sought to interpret bibliometric patterns in light of theoretical discussions in urban planning, data governance, and smart cities. The study's main contribution is demonstrating that the growth of the urban AI literature can be regarded as an indication of an internal reconfiguration in this field's logic of knowledge production. In this sense, AI in the literature examined has gradually moved beyond its position as an analytical tool and can be interpreted as part of the epistemic infrastructure of knowledge production and support for urban decision-making. However, this interpretation is theoretical, based on scientometric evidence, and does not claim an empirical or definitive transformation across the entire discipline of urban planning or the urban governance system.

The findings showed that urban sustainability, with a share of 31.9%, is the largest thematic domain in the literature, while topics such as urban resilience, digital twins, and cybersecurity are emerging with the highest relative growth rates. At the epistemological level, the data-driven and predictive paradigm, with a share of 77.7%, remains the dominant framework of knowledge production, while approaches based on generative AI are gradually expanding. These findings indicate that the evolution of the urban AI literature is not limited merely to changes in research topics but also reflects

changes in the ways problems are formulated, evidence is produced, and urban decision-making is supported.

From a theoretical perspective, the present study argues that the concept of an “epistemological shift” in this field should not be interpreted as a transformation of the entire discipline of urban planning, but rather as an internal change within the urban AI literature – a change in which data-driven and predictive logics, alongside the emerging capabilities of generative AI, are gradually redefining how knowledge is produced, validated, and organized. Furthermore, findings on interdisciplinary collaboration and the geographical concentration of knowledge production indicate that the development of urban AI depends not only on technological advances but also on access to data, computational infrastructures, research capacities, and scientific collaboration networks. From this perspective, issues such as data governance, algorithmic justice, and transparency should be considered not only at the level of smart-system applications but also at the level of knowledge production and the formation of research agendas.

From a methodological perspective, the present study demonstrated that integrating supervised thematic classification, text mining, and scientometric analysis can provide an appropriate framework for examining knowledge transformations in interdisciplinary fields. In this approach, bibliometric indicators are not used merely to describe publication patterns but are employed as evidence to interpret changes in knowledge structures, research priorities, and the logic of knowledge production, a feature that distinguishes the present study from many previous scientometric studies.

Nevertheless, the findings should be interpreted within the study’s limitations. The present analysis is based solely on English-language articles indexed in the Web of Science and Scopus databases and does not include non-English-language literature or grey literature. In addition, the thematic and epistemological classifications used a single-label approach; the analyses were based on article metadata and short texts, and the 2025 values may have been underestimated because of indexing lag. Furthermore, the present study did not directly measure power, the legitimacy of decisions, or the actual consequences of algorithmic governance. Therefore, the links between the scientometric findings and concepts such as algorithmic justice or algorithmic agency should be

regarded as theoretical interpretations based on the existing literature rather than as direct findings derived from the research data.

Accordingly, future research could provide a more precise understanding of the epistemological transformations of urban AI by integrating scientometric analyses with qualitative content analysis of influential articles, examining the role of large language models and generative AI in urban planning processes, and conducting comparative studies of geographical differences in access to data and computational infrastructures. Overall, the evidence from this study indicates that the literature on artificial intelligence and machine learning in urban planning and smart city studies is undergoing an internal reconfiguration in which thematic changes, transformations in dominant paradigms, and changes in the logic of knowledge production are occurring simultaneously. From this perspective, the future significance of urban AI will not be limited to developing more advanced algorithms but will also depend on how these capacities are guided within frameworks of data governance, transparency, accountability, and public values.

Authors’ Contributions

The first author contributed 35%, the second author 35%, and the third author 30% to the preparation of this article.

Acknowledgments

This article is the result of the authors’ research efforts and scholarly work. The research involved no commercial interests, and the authors received no financial compensation for presenting this work. Furthermore, the study received no financial or non-financial support, and the authors covered all expenses associated with the research. OpenAI’s ChatGPT Plus (GPT-5.6) was used solely for the English translation and linguistic refinement of this article. Its use was limited to translation and linguistic improvement and did not involve data analysis, interpretation of the results, or generation of the scientific content of the study. The authors reviewed, revised, and approved all translated sections to ensure translation accuracy and full consistency with the original Persian version.

Conflict of Interest

The authors declare no conflict of interest.

References

- Acuto, M., et al. (2023). Urban governance in the digital age. *Nature Cities*, 1(1), 12–19. <https://doi.org/10.1038/s44284-023-00002-3>

- Batty, M. (2013). *The new science of cities*. MIT Press. <https://doi.org/10.7551/mitpress/9399.001.0001>
- Batty, M. (2024). AI and design. *Environment and Planning B: Urban Analytics and City Science*, 51(4), 799–802. <https://doi.org/10.1177/23998083241236619>
- Batty, M. (2026). The evolution of AI in city planning. *Discover Cities*, 3, 47. <https://doi.org/10.1007/s44327-026-00224-0>
- Behzadfar, M., Rezaei, Sh., Torkaman, M., & Razaghi, F. (2026). Analyzing emerging trends in sustainable transportation: Designing a framework for transforming urban transportation policy and planning models. *Journal of Urban Economics and Planning*, 7(9), 18–40. <https://doi.org/10.22034/uep.2026.577735.1853>
- Bibri, S. E. (2021). *Data-driven smart sustainable cities of the future: Urban computing and intelligence for strategic, short-term, and joined-up planning*. Springer. <https://doi.org/10.1007/978-3-030-68059-1>
- Chen, C. (2017). Science mapping: A systematic review of the literature. *Journal of Data and Information Science*, 2(2), 1–40. <https://doi.org/10.1515/jdis-2017-0006>
- Cugurullo, F., Caprotti, F., Cook, M., Karvonen, A., McGuirk, P., & Marvin, S. (2024). The rise of AI urbanism in post-smart cities: A critical commentary on urban artificial intelligence. *Urban Studies*, 61(6), 1168–1182. <https://doi.org/10.1177/00420980231203386>
- Donthu, N., Kumar, S., Mukherjee, D., Pandey, N., & Lim, W. M. (2021). How to conduct a bibliometric analysis: An overview and guidelines. *Journal of Business Research*, 133, 285–296. <https://doi.org/10.1016/j.jbusres.2021.04.070>
- Faludi, A. (1973). *Planning theory*. Pergamon Press.
- Hall, P. (2014). *Cities of tomorrow: An intellectual history of urban planning and design since 1880* (4th Ed.). Wiley-Blackwell.
- Healey, P. (1997). *Collaborative planning: Shaping places in fragmented societies*. Macmillan.
- Innes, J. E., & Booher, D. E. (2010). *Planning with complexity: An introduction to collaborative rationality for public policy*. Routledge.
- Janowicz, K., Gao, S., McKenzie, G., Hu, Y., & Bhaduri, B. (2020). GeoAI: Spatially explicit artificial intelligence techniques for geographic knowledge discovery and beyond. *International Journal of Geographical Information Science*, 34(4), 625–636. <https://doi.org/10.1080/13658816.2019.1684500>
- Kitchin, R. (2014). *The data revolution: Big data, open data, data infrastructures and their consequences*. SAGE. <https://doi.org/10.4135/9781473909472>
- Kitchin, R. (2016). The ethics of smart cities and urban big data. *Philosophical Transactions of the Royal Society A*, 374(2083), 20160115. <https://doi.org/10.1098/rsta.2016.0115>
- Kitchin, R. (2017). Thinking critically about and researching algorithms. *Information, Communication & Society*, 20(1), 14–29. <https://doi.org/10.1080/1369118X.2016.1154087>
- Kitchin, R., & Moore-Cherry, N. (2021). Fragmented governance, the urban data ecosystem and smart city-regions: The case of Metropolitan Boston. *Regional Studies*, 55(12), 1913–1923. <https://doi.org/10.1080/00343404.2020.1735627>
- Leszczynski, A. (2020). Glitchy vignettes of platform urbanism. *Environment and Planning D: Society and Space*, 38(2), 189–208. <https://doi.org/10.1177/0263775819878721>
- Mai, G., Xie, Y., Jia, X., Lao, N., Rao, J., Zhu, Q., Liu, Z., Chiang, Y.-Y., & Jiao, J. (2025). Towards the next generation of geospatial artificial intelligence. *International Journal of Applied Earth Observation and Geoinformation*, 136, 104368. <https://doi.org/10.1016/j.jag.2025.104368>
- OECD. (2024). *Governing smart cities and artificial intelligence*. OECD Publishing. <https://doi.org/10.1787/g2g55432-en>
- Si, S., Yao, Y., & Wu, J. (2025). Artificial intelligence in urban planning: A bibliometric analysis and hotspot prediction. *Land*, 14(11), 2100. <https://doi.org/10.3390/land14112100>
- Son, T. H., Weedon, Z., Yigitcanlar, T., Sanchez, T., Corchado, J. M., & Mehmood, R. (2023). Algorithmic urban planning for smart and sustainable development: Systematic review of the literature. *Sustainable Cities and Society*, 94, 104562. <https://doi.org/10.1016/j.scs.2023.104562>
- Steenmans, K., Robin, E., Acuto, M., Iwaszuk, E., & Ortega Garza, L. (2023). Governing the informed city: Examining local government strategies for information production, consumption, and knowledge sharing across ten cities. *Urban Governance*, 3(4), 243–251. <https://doi.org/10.1016/j.ugi.2023.09.001>
- United Nations Human Settlements Programme. (2022). *AI & cities: Risks, applications and governance*. UN-Habitat. <https://unhabitat.org/ai-cities-risks-applications-and-governance>
- Xu, H., Omitaomu, F., Sabri, S., Zlatanova, S., Li, X., & Song, Y. (2024). Leveraging generative AI for urban digital twins: A scoping review on the autonomous generation of urban data, scenarios, designs, and 3D city models for smart city advancement. *Urban Informatics*, 3, 29. <https://doi.org/10.1007/s44212-024-00060-w>
- Yigitcanlar, T., Agdas, D., Xia, B., Degirmenci, K., & Inkinen, T. (2024). Unlocking artificial intelligence adoption in local governments: Best practice lessons from real-world implementations. *Smart Cities*, 7(4), 1576–1593. <https://doi.org/10.3390/smartcities7040064>
- Yigitcanlar, T., Desouza, K. C., Butler, L., & Roozkhosh, F. (2020). Contributions and risks of artificial intelligence (AI) in building smarter cities: Insights from a systematic review of the literature. *Energies*, 13(6), 1473. <https://doi.org/10.3390/en13061473>
- Zhang, W., Han, J., Xu, Z., Ni, H., Lyu, T., Liu, H., & Xiong, H. (2026). Towards urban general intelligence: A review and outlook of urban foundation models. *ACM Transactions on Intelligent Systems and Technology*. <https://doi.org/10.1145/3821528>

