City Development: Issues and Best Practices 2 (2026)

AI for Cities: Patterns, Gaps, and Future Directions

Seyed Navid Mashhadi Moghaddam[a], *, Huhua Cao[a], Benjamin Gianni[b]

[a] *Department of Geography, University of Ottawa, Ottawa, Canada*
[b] *Azrieli School of Architecture & Urbanism, Carleton University, Ottawa, Canada*

ARTICLE INFO

Keywords:
artificial intelligence
urban studies
systematic review
smart cities
sustainability

ABSTRACT

Artificial intelligence is increasingly regarded as crucial for tackling urban challenges; however, the alignment of the research enterprise with sustainability and governance requirements has yet to be scrutinized. In this study, we look at 7,660 peer-reviewed articles from 2020 to 2025 that look at AI's use in urban studies. We do this to describe the field's thematic structure, methodological landscape, and sustainability orientation. We observe that prediction and classification tasks predominate, constituting over 40%, whereas decision support is infrequent, and architectures designed for urban complexity are underutilized. Correspondence analysis shows that there is an interpretability axis that organizes research. Digital and Smart Cities have the weakest sustainability integration (37% compared to 48% in Resilience and Safety). Cross-sectional designs are the most common (74.3%), and equity issues are still not being looked at closely enough. We suggest a research agenda focused on architectural innovation that incorporates urban theory into models, governance integration that links AI to decision support, and participatory frameworks that promote democratic development and implementation. Urban governance is a high-stakes area where black-box prediction should give way to clear explanations, and fairness limits are needed to improve efficiency.

Main

Cities are becoming the most important places for people to live. Since 1950, the number of people living in cities has more than doubled, and about 4.5 billion people now live in cities. By 2050, two-thirds of the world's population is expected to live in cities (United Nations Department of Economic and Social Affairs, 2019) (United Nations Department of Economic and Social Affairs, 2025). This unprecedented urbanization brings both opportunities and risks. For example, cities make more than 80% of the world's GDP, use 75% of its natural resources, and release 70% of its carbon emissions(U.N.-Habitat & Mila, 2022). The problems that urban governance faces, such as climate change, affordable housing, public health, and social cohesion, require analytical skills that are as complex as the problems themselves. Artificial intelligence has surfaced as a potential technology to address this demand, offering the capability to derive actionable insights from the vast data streams produced by modern cities and to enhance urban systems at scales surpassing human cognitive abilities.(Batty et al., 2012; Yigitcanlar, Desouza, Butler, & Roozkhosh, 2020). However, the swift expansion of AI applications in urban settings has surpassed the necessary contemplation regarding their efficacy in achieving the fundamental objectives they purport to address.

The smart city paradigm that has dominated urban technology discourse for over a decade positioned AI as the primary vehicle for achieving efficiency, sustainability, and quality of life (Bibri & Krogstie, 2017) (Kitchin, 2014). This framing has drawn a lot of research funding, and AI applications range from improving transportation, managing energy, monitoring the environment, planning land use, and delivering public services(Allam & Dhunny, 2019; Sanchez, Shumway, Gordner, & Lim, 2023). Critics, on the other hand, have raised doubts about whether smart city projects live up to their promises or whether they are just a way for technology vendors to find places to deploy their products, city governments to push modernization stories, and researchers to focus on benchmark performance instead of governance needs(Robert G. Hollands, 2008) (Morozov & Bria, 2018b). The disparity between AI's theoretical potential and its actual incorporation into governance frameworks has become increasingly evident, with obstacles such as data quality, algorithmic transparency, interpretability, and ethical considerations hindering the transition from research prototypes to operational systems (Cugurullo, 2020)

* Corresponding author.
E-mail address: Navid.mm@uottawa.ca (Seyed Navid Mashhadi Moghaddam).

https://doi.org/10.18192/cdibp.v1i2.7961
Received 18 March 2026; Received in revised form 2 May 2026; Accepted 5 May 2026
Available online 5 May 2026

(Kitchin, 2017). These criticisms imply that the field necessitates a systematic evaluation: not only enumerating existing AI applications but also analyzing the evolution of the research domain, the prevailing methodological preferences, and the alignment of the resultant knowledge base with urban sustainability objectives.

This systematic review meets that need by looking at 7,660 peer-reviewed articles that looked at AI applications in urban studies that were published between 2020 and 2025. In contrast to earlier reviews that concentrate on particular application domains or technology families, we employ a holistic approach encompassing the entirety of urban AI research to delineate the field's thematic framework, methodological diversity, and sustainability focus. Our analysis reveals a rapidly growing field with persistent patterns that require critical attention: the predominance of prediction and classification tasks over causal explanation and decision support; the underutilization of architectures designed for urban complexity, such as graph neural networks and physics-informed approaches; an inverse correlation between technological sophistication and the depth of sustainability integration; and the systematic underrepresentation of equity considerations across various application domains. These results indicate that urban AI research has predominantly adopted methodologies from computer vision and natural language processing, rather than creating architectures specifically designed for the intrinsic multi-scale, temporally dynamic, and socio-technical complexity of urban environments.

We combine these patterns into a research agenda based on three main ideas: architectural innovation that incorporates urban theory into model design, governance integration that links AI capabilities to decision support needs, and participatory frameworks that make the development and use of urban AI systems more democratic. The analysis adds to ongoing discussions about responsible AI in high-stakes areas by showing that urban governance is the perfect place for black-box prediction to give way to clear explanation, for efficiency optimization to need equity constraints, and for technical sophistication to need institutional accountability (Cynthia Rudin, 2019) (Floridi et al., 2018). Cities all over the world are dealing with climate change, changing populations, and limited resources. The question is not if AI will change the future of cities, but if research will give us the information we need to do so in a fair and long-lasting way.

Results

Overview of the research landscape

Our systematic review found 7,660 articles about AI uses in urban studies (see Data S1) that were published between 2020 and 2025 in 109 countries and 1,547 different places (Fig. 1a,b). The field has grown quickly, with the number of articles published each year going from 574 in 2020 to 2,395 in 2025, more than four times as many. This growth trajectory shows how quickly artificial intelligence and urban science are coming together. SDG 11 (Sustainable Cities and Communities) is the main focus of 71.6% of all publications.

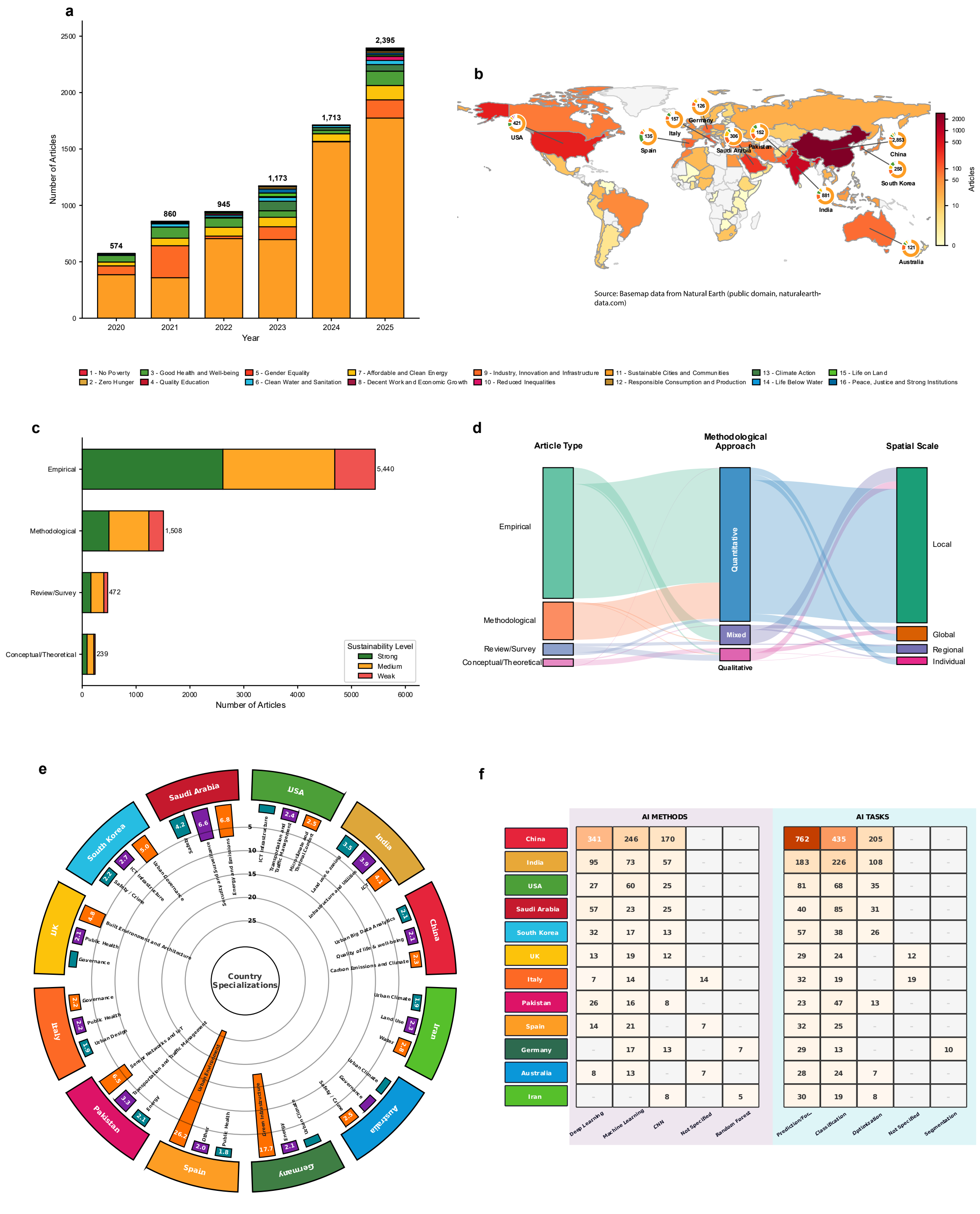

Figure. 1 | Overview of the research landscape. a, Temporal distribution of publications by year (2020–2025) with SDG alignment indicated by color. b, Geographic distribution of publications by first-author country affiliation. c, Distribution of articles by type (empirical, methodological, review/survey, conceptual/theoretical) with sustainability integration levels. d, Sankey diagram showing relationships between article type, methodological approach, and spatial scale. e, Radial visualization of country research specializations across thematic domains. f, Cross-tabulation of top contributing countries by AI methods (left) and AI tasks (right). Panel B Source: Basemap data from Natural Earth (public domain, naturalearthdata.com)

Geographic analysis shows that most publications come from the Global South and Asia, with China (37.2%), India (11.5%), and the United States (5.5%) making up more than half of all publications (Fig. 1b,e). Research priorities differ based on the country: China and India focus on deep learning methods for transportation and environmental uses, while European countries are more involved in governance, public health, and urban design. This geographic diversity shows that different regions have different levels of AI research capacity and face different problems with urbanization.

The corpus is predominantly comprised of empirical studies (71.0%) utilizing quantitative methods (82.5%) at local spatial scales (80.2%), with deep learning and machine learning serving as the principal methodological approaches (Fig. 1c,d,f). Prediction and classification tasks constitute more than 40% of AI applications, primarily focused on transportation, environmental monitoring, and land use issues (see Supplementary Section 2 for analyses of the 2,219 non-empirical studies, which include methodological contributions, reviews, and conceptual frameworks). However, only 43.9% of articles show strong integration of sustainability, and cross-sectional designs (74.3%) are much more common than longitudinal analyses. This shows that there are big gaps in our understanding of how AI-driven urban changes happen over time.

Thematic structure of empirical research

We looked at the 5,441 empirical studies and used a two-stage classification method that combined coding with a large language model and expert validation to find six main research areas through hierarchical thematic analysis (Fig. 2a). The largest cluster, with 2,052 items, is Digital and Smart Cities. It includes smart mobility, urban intelligence systems, and urban planning apps. Climate and Environment research (n=1,446) looks into things like air quality, urban heat islands, green infrastructure, and how to manage energy. Mobility and Transportation (n=539), Urban Development (n=506), Social Equity and Quality of Life (n=470), and Resilience and Safety (n=228) round out the thematic landscape. Each has its own unique methodological signatures and sustainability orientations (see Supplementary Figures S2-S7 and Tables S16-S27 for the full hierarchical taxonomy and non-empirical thematic distributions).

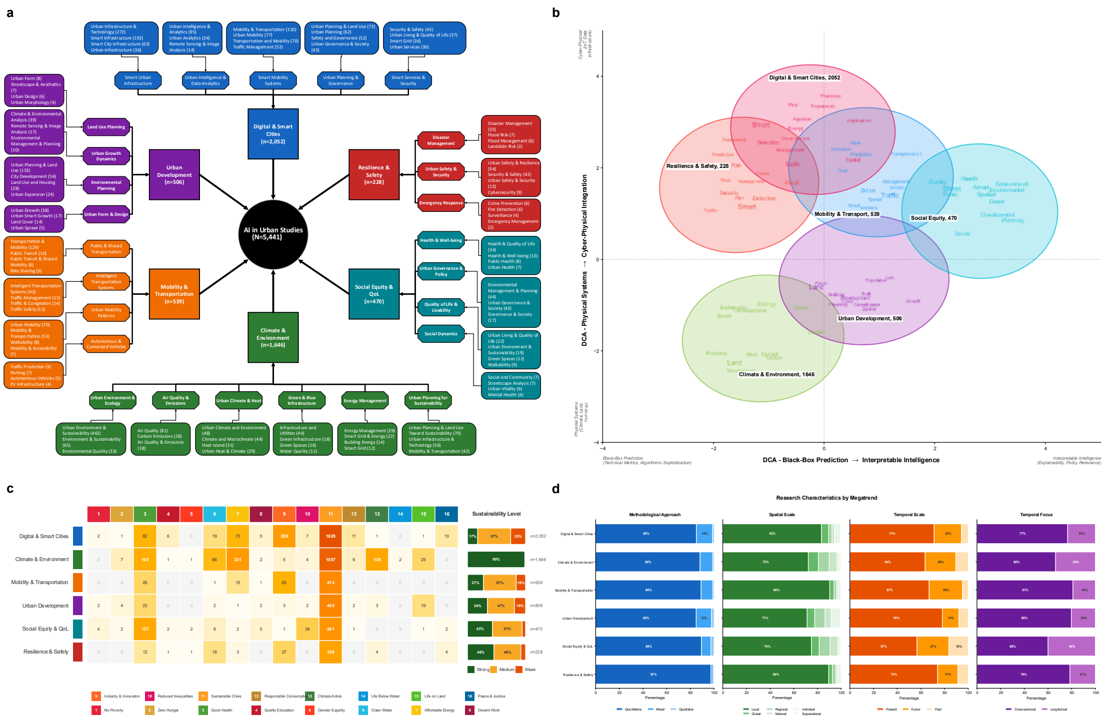

Figure. 2 | Thematic structure of empirical research. a, Hierarchical taxonomy of 5,441 empirical studies showing six megatrends, constituent research domains, and subject clusters with article counts. b, Detrended correspondence analysis positioning megatrends along two axes: interpretability (horizontal, from black-box prediction to interpretable intelligence) and system focus (vertical, from physical systems to cyber-physical integration). Word clouds indicate dominant terminology within each cluster. c, SDG alignment across megatrends with sustainability integration levels (strong, medium, weak). d, Research characteristics by megatrend showing methodological approach, spatial scale, temporal scale, and temporal focus.

Detrended correspondence analysis shows that the field is built around two main axes (Fig. 2b). The first axis separates research based on how easy it is to understand, from black-box prediction methods that focus on technical performance metrics to interpretable intelligence frameworks that focus on explainability and policy relevance. The second axis shows the shift from physical-technical systems to socio-spatial integration. It separates studies that look at how to improve infrastructure from those that look at how people and the environment interact. The technical-predictive quadrant is where Digital and Smart Cities research fits, while the Social Equity and Climate and Environment clusters show a stronger focus on interpretable, socially-embedded approaches. There is a lot of difference in how SDGs fit into different areas of research (Fig. 2c). SDG 11 is the most important goal in all clusters. Climate and Environment research focuses more on SDG 13 (Climate Action) and SDG 7 (Clean Energy), while Social Equity studies focus more on SDG 3 (Health and Well-being) and SDG 10 (Reduced Inequalities). Integration levels for sustainability vary greatly. Resilience and Safety has the highest percentage of strong sustainability engagement (48%), while Digital and Smart Cities has a lower percentage (37%). This suggests that being technologically advanced does not mean having a deep understanding of sustainability.Methodological characteristics indicate domain-specific research cultures (Fig. 2d). Quantitative methods are the most common in all clusters, but Social Equity research uses mixed methods much more often. Local-scale analysis is the most common type of analysis, but Climate and Environment studies often include regional and global points of view. Longitudinal designs are still not used enough in all areas, but Mobility and Transportation have slightly more temporal engagement. These patterns reveal a pervasive constraint within the field: notwithstanding AI's ability for dynamic modeling, the majority of empirical applications are restricted to static, cross-sectional representations of urban phenomena.

AI methods and tasks in urban applications

The methodological landscape shows that the field is based on tried-and-true methods but is slower to adopt architectures that work well in complex urban settings (Fig. 3a). Machine learning and deep learning are the most popular methods. Convolutional neural networks are the main way to do image-based urban sensing, from classifying satellite images to understanding scenes at street level. But this reliance on CNNs shows a bigger trend: urban AI research has mostly used techniques from computer vision and natural language processing instead of creating architectures that are better suited to the inherently relational, multi-scale, and changing nature of urban systems.

Figure. 3 | AI methods and tasks in urban applications. a, Distribution of AI methods across megatrends. b, Proportional area chart showing AI task distribution by SDG alignment. c, Distribution of AI tasks across megatrends. d, Funding source analysis showing AI methods distribution (left) and SDG alignment (right) for major funding agencies.

Even though graph neural networks are theoretically good for modeling urban networks, transportation systems, infrastructure interdependencies, and social connectivity, they are still surprisingly underused (n=111). Their most widespread use in Mobility and Transportation research shows that people are aware of network-centric approaches, but even here, simpler methods are more common. Likewise, transformers (n=70) and large language models (n=23) embody significant, yet largely unexplored, potential for the incorporation of qualitative urban knowledge, planning documents, and citizen feedback into analytical frameworks. This conservative approach to methods may be because publishers prefer small improvements to benchmark datasets over big changes to architecture for specific problems. Domain-specific methodological preferences show how different epistemic cultures exist within urban AI (Fig. 3a). Digital and Smart Cities research prioritizes predictive accuracy, utilizing deep learning for real-time monitoring and automated decision support, a methodology conducive to operational efficiency yet frequently compromising interpretability. Climate and Environment studies, on the other hand, show that random forest and ensemble methods are used more often. This is because environmental science traditions value feature importance and model transparency for policy communication. Research on social equity shows that people are more likely to use simulation and analysis tasks. This suggests that they are more interested in understanding how things work than in making predictions, which is an important difference for programs that aim to fix structural inequalities.Prediction and forecasting represent the predominant paradigm, comprising more than one-third of all AI tasks (Fig. 3b,c). While useful for operational uses like predicting traffic flow, energy demand, and air quality, this focus shows that the field is more interested in predicting the future of cities than in explaining how cities work or judging the results of interventions. Classification tasks, while abundant, primarily facilitate pattern recognition rather than diagnostic or evaluative functions. The limited number of decision support and monitoring applications (n=47) indicates that AI in urban studies is predominantly restricted to research settings, with minimal integration into practical urban governance systems. The proportional distribution of tasks across SDGs reveals discrepancies between AI capabilities and sustainability priorities (Fig. 3b). Prediction and optimization tasks focus a lot on SDG 11, which shows that smart cities talk is mostly about technology. On the other hand, SDGs that deal with structural problems, like poverty (SDG 1), inequality (SDG 10), and gender (SDG 5), show very little AI involvement in any of the task categories. This pattern suggests that current AI applications do a good job of optimizing for measurable urban performance metrics but not so well for goals that need a qualitative understanding of social processes and distributional outcomes. Funding geography affects both the methods used and the focus on long-term sustainability (Fig. 3d). Chinese funding agencies support the most research, with a strong focus on deep learning and a strong alignment with SDG 11. This shows that smart city development and modernizing urban infrastructure are important to the country. EU Horizon 2020 and North American funding show a wider range of methods and more widespread participation in the SDGs. This could be because different research governance structures encourage interdisciplinary and socially-oriented approaches. These correlations between funding, methodology, and SDGs suggest that making AI better for sustainable urban development requires not only new technology but also strategic research funding that clearly values interpretability, equity, and methodological pluralism. Supplementary Tables S40-S41 show detailed cross-tabulations of AI methods, tasks, and SDG alignment across all megatrends.

Discussion

Before turning to the substantive interpretation, we draw an explicit line between the descriptive findings of this review and the interpretive claims that follow. The frequency distributions, cross-tabulations, and correspondence-analysis geometry presented in the Results section are direct features of the corpus. The arguments advanced below regarding interpretability deficits, institutional incentives, publisher preferences, funding effects, and the causal limitations of the field are abductive readings that draw on prior literature to make sense of those patterns; they are not themselves tested in this paper, and we phrase them throughout as interpretive readings rather than as demonstrated findings. Empirical evaluation of these interpretations would require primary survey or interview data with researchers, funders, and editors and is out of scope for the present synthesis.

The interpretability imperative: from prediction to explanation

Rudin's important criticism said that black-box machine learning models shouldn't be used for important decisions. Instead, he said that models that are naturally interpretable and don't need post-hoc explanation should be used (C. Rudin, 2019). Urban governance exemplifies a high-stakes arena, where choices regarding infrastructure investment, land use regulation, and service distribution impact millions of inhabitants and influence urban development for decades. But our study shows that urban AI research has mostly ignored this warning. The prevalence of prediction and classification tasks, which together make up more than 40% of all applications, shows that people are more interested in predicting the future of cities than in learning about how cities work. While predicting traffic jams, air quality, or energy demand is useful for operations, these models usually work as advanced correlation engines that find patterns without explaining how they happen. The difference is very important for policy: to make effective changes, you need to know why things happen, not just when they will happen.

The correspondence analysis that shows an interpretability axis structuring the field implies that researchers are aware of this tension. Climate and Environment studies exhibit a pronounced inclination towards interpretable methodologies, such as random forest and ensemble techniques that elucidate feature importance, mirroring the traditions of environmental science where policy communication necessitates transparency (Breiman, 2001). Social equity research also focuses on simulation and analysis more than just making predictions. This is because it is important to understand how urban configurations create disadvantage, not just where disadvantage is concentrated. On the other hand, Digital and Smart Cities research is in the technical-predictive quadrant. It focuses on performance metrics that meet engineering standards but often hide mechanisms that are important for making decisions. This pattern reflects apprehensions articulated in the extensive machine learning literature regarding the exaggerated or entirely fictitious accuracy-interpretability tradeoff (Hastie, Tibshirani, Friedman, & Friedman, 2009; C. Rudin, 2019). Nevertheless, urban AI persists in functioning under the assumption that black-box complexity is essential for predictive performance.

The insufficient use of architectures that can encode urban theory is a significant missed chance. Graph neural networks are still surprisingly rare (n=111), even though they fit well with urban phenomena. For example, transportation networks, infrastructure interdependencies, and social connectivity patterns all have relational structures that GNNs can easily represent (Jiang & Luo, 2022; Li et al., 2024). Recent surveys show that GNNs are the best at predicting traffic because they can capture the graph structures that are common in transportation systems (Jiang & Luo, 2022). However, our corpus shows that these methods are not spreading beyond a few specific uses. Physics-informed neural networks, which incorporate domain knowledge as constraints, could amalgamate established urban models (gravity models, spatial interaction theory, urban scaling laws) with data-driven learning (Raissi, Perdikaris, & Karniadakis, 2019; Shao, Liu, Zhang, Zhao, & Hu, 2023), yet they seem to be nearly absent from the empirical literature. This methodological conservatism implies that urban AI has appropriated tools from computer vision and natural language processing instead of creating architectures specifically designed for the intrinsic multi-scale, temporally dynamic, and socio-technical complexity of urban environments.

The shift from prediction to causal inference signifies a pivotal methodological frontier. Pearl's hierarchy of causal reasoning differentiates associations, interventions, and counterfactuals, positing that the majority of machine learning systems function at the associational level, irrespective of data volume (Pearl & Mackenzie, 2018). Urban policy inquiries are inherently causal: does transit-oriented development enhance ridership, or does it simply correlate with unquantified residential preferences? Does green infrastructure enhance health outcomes, or do healthier populations gravitate towards greener neighborhoods? Recent developments in causal machine learning, particularly in heterogeneous treatment effects and counterfactual reasoning (Athey, 2015; Kreif & DiazOrdaz, 2019), have not been extensively applied in urban contexts, despite their significant implications for policy evaluation. The lack of causal approaches in our corpus (we found very few clear examples of causal inference) suggests that urban AI is still more focused on recognizing patterns than on the counterfactual reasoning that intervention design needs. To move the field forward, we need to do more than just use existing methods on urban data. We also need to create AI architectures that are native to cities and encode theoretical priors about how cities work. These architectures should also help answer the causal questions that urban governance needs to answer.

Sustainability depth versus technological sophistication

A notable paradox arises from the sustainability analysis: technological advancement does not ensure the integration of sustainability. Digital and Smart Cities research, the largest and most technically advanced cluster using deep learning, federated learning, and new large language models, shows the least interest in sustainability. Only 37% of articles in this area had strong integration, compared to 48% in Resilience and Safety research and almost all articles in Climate and Environment studies. This inverse relationship necessitates thorough scrutiny, especially considering that smart city discourse has traditionally regarded technology as the principal means for attaining urban sustainability objectives (Batty et al., 2012; Bibri & Krogstie, 2019).

We highlight one mechanism that we believe is central to this pattern: in much of the Digital and Smart Cities subfield, technical optimization is implicitly treated as a proxy for sustainability. Lower energy intensity per task, higher throughput, and reduced computational latency are reported as sustainability gains even when distributional outcomes, rebound effects, and the lifecycle costs of the underlying digital infrastructure are not assessed. This conflation helps explain why a cluster that is both technically advanced and rhetorically aligned with SDG 11 simultaneously scores lowest on our integration scale: the framing under which sustainability is being measured is itself narrower in this cluster than in fields such as Climate and Environment, where sustainability framings have a longer disciplinary history and require explicit ecological or distributional accounting. We present this mechanism as an interpretive hypothesis suggested by our cross-tabulations rather than as a tested causal claim.

There are a number of ways that this pattern could happen. First, smart city frameworks have mostly focused on making things more efficient and have seen sustainability as a technical problem that can be solved with algorithms (Robert G Hollands, 2020). This framing favors measurable performance metrics, such as lowering energy use, improving traffic flow, and automating services, while pushing aside questions about who benefits from these changes. optimized based on whose choices? The focus of AI applications on SDG 11 (Sustainable Cities) with limited involvement in SDGs related to poverty, inequality, and gender indicates that contemporary research enhances overall urban performance while neglecting objectives that necessitate consideration of social processes and equitable outcomes. This pattern shows Goodhart's Law: when metrics become goals, they stop being good measures (Strathern, 1997). Urban AI systems that optimize for measurable efficiency may reach their goals without making progress toward the real sustainability goals that those metrics were meant to show.

Second, critics say that the smart city model is "solutions in search of problems," and that it is driven by tech companies with new products instead of a systematic study of urban problems (Morozov & Bria, 2018a). Our discovery that the research cluster with the most advanced technology has the least depth of sustainability supports this criticism. When AI methods are chosen based on how new or good they are at solving problems, rather than how well they fit with sustainability challenges, the applications that come out of them may be very technically advanced but not very helpful for sustainable development. Digital and Smart Cities research has a lot of prediction and classification tasks but very few decision support or monitoring tasks. This suggests that the field is more focused on showing what AI can do than on solving governance problems.

Third, the correlations between funding, methodology, and SDGs show how research governance structures affect sustainability orientation. Chinese funding agencies support the most research with a focus on SDG 11 alignment and deep learning. This shows that the government wants to modernize smart city infrastructure. Funding from the EU Horizon 2020 and North America shows a wider range of SDG engagement and a greater variety of methods. This could be because different institutions value interdisciplinary and socially-oriented approaches. These patterns indicate that sustainability depth arises not solely from researcher decisions but from institutional contexts that variably prioritize efficiency over equity, prediction over explanation, and technical performance over social accountability (Chib, Alvarez, & Todorovic, 2022).

Fourth, the temporal framework of the research foundation significantly constrains sustainability evaluation. Cross-sectional designs are the most common (74.3%). They only give us snapshots, so we can't see how AI interventions change over time, whether the benefits last, or how the initial benefits grow or fade across groups of people. Longitudinal analysis is still uncommon in all fields, even though AI can do dynamic modeling. Without temporal depth, sustainability assertions are based on anticipated rather than substantiated results. This methodological gap is especially important because sustainability is a concept that changes over time. To meet the needs of the present without hurting future generations, we need to understand trajectories, not just states (United Nations Development Programme, 2015).

Lastly, the equity aspects of urban AI are still not being looked at closely enough. Our corpus indicates that informal settlements and vulnerable populations often remain excluded from training datasets; or, more detrimentally, are primarily represented as subjects of surveillance rather than recipients of services (Datta, 2018; Luque-Ayala & Marvin, 2016). Predictive policing algorithms that use past crime data to train police officers tend to send more police to neighborhoods that are already heavily policed. This creates feedback loops that make spatial inequality worse instead of better (Lum & Isaac, 2016). Facial recognition systems make more mistakes with people with darker skin. Studies have shown that error rates can be as high as 35% for women with darker skin and less than 1% for men with lighter skin (Buolamwini & Gebru, 2018). These patterns indicate that contemporary AI applications might encode and exacerbate existing urban disparities while purporting to maintain technical objectivity. To achieve real sustainability, we need to do more than just improve urban systems; we also need to make sure that the benefits of those improvements are shared fairly. This is a problem that most of the current research doesn't address.

Toward urban-native artificial intelligence

The previous analysis shows that the field is at a turning point: it has advanced methods but not enough theoretical development, and it has new technologies but is not connected to the needs of governance. We suggest a research agenda that goes beyond criticism and is based on three important ideas: architectural innovation that incorporates urban theory into model design, governance integration that links AI capabilities to decision support needs, and participatory frameworks that make the development and use of urban AI systems more democratic.

The architectural imperative requires transcending borrowed methodologies in favor of urban-native AI designs. Most of the time, architects who work on urban applications use architectures that were made for computer vision, natural language processing, or general time series forecasting. They see urban applications as places to put their work, not as design limits. This method gives up theoretical grounding in favor of benchmark performance. Graph neural networks, although underutilized in our corpus (n=111), present a promising alternative due to the inherently networked nature of urban systems: transportation infrastructure, social relationships, economic flows, and ecological processes all demonstrate graph structures conducive to relational learning (Battaglia et al., 2018). More ambitiously, physics-informed neural networks could directly incorporate established urban theory, such as gravity models, spatial interaction frameworks, and scaling laws, into model architectures. This would limit learned representations to be physically plausible while still allowing for data-driven flexibility (Bettencourt, 2013; Raissi et al., 2019). These kinds of mixed methods would solve the problem of interpretability that was brought up before. Models that encode theoretical mechanisms are easier to explain than black-box models because their parameters correspond to meaningful urban processes instead of random feature weights.

The rise of urban digital twins brings both chances and dangers for this architectural plan. Digital twins, which are virtual copies of city systems that let you watch them in real time and run scenarios, have grown quickly. By 2025, there will be more than 500 city-scale deployments around the world (A. B. I. Research, 2020). These platforms could be places where physics-informed models, graph-based learning, and causal inference come together to help test out policies before they are put into practice (Batty, 2018). For example, the Herrenberg digital twin in Germany uses virtual reality to combine 3D models of cities with mobility simulation, wind flow analysis, and data from citizens to help with participatory planning (Dembski, Wössner, Letzgus, Ruddat, & Yamu, 2020). However, digital twins may also replicate the efficiency-over-equity bias identified in smart city studies if they prioritize overall system performance without regard for distributional impacts. The difficulty is in creating digital twin architectures that put equity metrics first, along with efficiency indicators. This will let planners see not only how changes affect system-level outcomes, but also how benefits and burdens are shared among different groups and neighborhoods.

The governance imperative necessitates closing the significant disparity between AI capabilities and the decision support requirements identified in our task analysis. Even though decision support applications are the most relevant to policy, there are still very few of them (n=47). Instead, prediction and classification are the most common types. This imbalance shows a basic misalignment: researchers focus on tasks that can be evaluated with benchmarks, while practitioners need tools that help them make decisions instead of automating them (Veale & Brass, 2019). To make good decisions, you need more than just accurate predictions. You also need to be able to measure uncertainty, compare scenarios, and think about what would have happened if things had gone differently. Most current architectures don't have these features. The previously identified causal inference frontier is especially significant in this context: urban policy inquiries primarily concern interventions (e.g., what are the effects of implementing transit-oriented development?) rather than predictions (e.g., where will population growth occur?). However, the majority of applications function at the associational level of Pearl's causal hierarchy (Pearl & Mackenzie, 2018).

Algorithmic accountability frameworks provide institutional mechanisms for linking AI development to governance requirements. Public algorithm registers, which started in Amsterdam and Helsinki, require that automated systems used in city decision-making be open and honest (Floridi, 2020). Algorithmic impact assessments, analogous to environmental impact assessments, require systematic evaluation of potential harms before deployment (Reisman, Schultz, Crawford, & Whittaker, 2018). Yet evidence from early implementations, including New York City's Automated Decision Systems Task Force, suggests that accountability mechanisms often narrow to definitional debates and inventory exercises rather than substantive evaluation of algorithmic effects on communities (Baykurt, 2022). Effective accountability requires not merely transparency about which algorithms are used but meaningful oversight of how algorithmic outputs shape decisions affecting residents, particularly those in communities historically subject to discriminatory treatment. The finding that predictive policing algorithms and facial recognition systems disproportionately affect communities of color underscores the stakes: algorithmic accountability is not merely a procedural nicety but a civil rights imperative (Benjamin, 2019).

The participatory imperative addresses the democratic deficit in current urban AI development. Our corpus reveals that affected communities rarely participate in defining problems, selecting methods, or evaluating outcomes. This exclusion is both ethically problematic and epistemically costly: local knowledge about neighborhood dynamics, community needs, and historical context could substantially improve model specification and interpretation (Sloane, Moss, Awomolo, & Forlano, 2022). Emerging participatory AI frameworks demonstrate feasibility: Helsinki's deployment of generative AI tools for urban co-design enabled citizens to visualize and iterate on public space proposals in real time, while intergenerational workshops in Panama used similar tools to bridge perspectives between elderly residents and university students (Interoperable Europe, 2024; Labs, 2024). These applications suggest that AI can democratize rather than technocratize urban planning when designed with participation as a first-order objective rather than an afterthought.

To make this vision of participation a reality, we need to deal with the power imbalances that are built into AI governance today. Most city AI systems come from private vendors who say they have trade secret protection over the details of their algorithms. This means that public decision-making processes are effectively privatized (Brauneis & Goodman, 2018). Community groups and civil society advocates have fought for procurement rules that require participatory design, public consultation, and ongoing community oversight as part of municipal AI contracts (Ada Lovelace Institute & Partnership, 2021). These kinds of changes in institutions could change the way AI is developed from a technical task done by experts to a sociotechnical process in which communities that are affected have real power over the systems that affect their lives. The other option, keeping marginalized voices out of algorithmic governance, could encode current inequalities into urban infrastructure for many years to come.

Lastly, the funding and institutional frameworks that influence urban AI research require rigorous scrutiny. Our analysis showed that the geography of funding is related to the depth of sustainability and the orientation of methods. For example, Chinese funding focuses on deep learning and SDG 11, while EU and North American funding shows a wider range of methods and a more distributed approach to SDG engagement. These patterns indicate that fulfilling the research agenda delineated herein necessitates not only individual researcher decisions but also institutional incentives that prioritize interpretability alongside accuracy, equity alongside efficiency, and governance relevance alongside benchmark performance. Funders, journal editors, and tenure committees all have a say in what is considered valid urban AI research. To move the field toward urban-native, governance-integrated, and participatory approaches, the criteria for evaluation and the way resources are distributed must also change.

This synthesis suggests that urban AI is at a crossroads. One path keeps things going the way they are: borrowed architectures used on urban data, prediction over explanation, efficiency without equity limits, and communities left out of decisions that affect them. To get to urban-native AI, we need new ways of building based on urban theory, new ways of integrating AI capabilities with decision support needs, and new ways of involving everyone in both development and deployment. Urban problems like climate change, affordable housing, and public health are so big that we need the best tools we can find. But cities that are fair, just, and long-lasting won't come from having the ability to do things without being held accountable, being able to predict things without explaining them, or being able to do things quickly without being fair. The field needs to make a choice.

Methods

Search strategy and study selection

We performed a systematic literature review in accordance with PRISMA 2020 guidelines (Page et al., 2021) to investigate artificial intelligence applications in urban research published from January 2020 to December 2025 (see Supplementary Figures S1). Scopus was the main database because it has a wide range of engineering, computer science, environmental science, and urban studies journals, it has a strict Content Selection and Advisory Board that chooses the content, and it is already known as a reliable source of bibliometric data for large-scale research assessments (Baas, Schotten, Plume, Côté, & Karimi, 2020) (Mongeon & Paul-Hus, 2016). The search strategy used both AI-related words (like "artificial intelligence," "machine learning," "deep learning," specific architectures like CNN, LSTM, GNN, and transformer, and algorithm families like random forest, support vector machine, and reinforcement learning) and urban-related words (like "urban," "city," "smart city," "metropolitan," and domain-specific phrases like "urban planning," "urban mobility," "urban climate," and "land use"). The search was limited to English-language peer-reviewed journal articles, and it only looked at titles, abstracts, and keywords. This strategy produced 8,430 initial records over the study period: 618 in 2020, 921 in 2021, 1,188 in 2022, 1,363 in 2023, 1,801 in 2024, and 2,539 in 2025 (see Supplementary Section 1.1-1.3 and Table S1-S15 for the full search string and PRISMA flow diagram).

Screening continued by checking for relevance against set inclusion criteria: peer-reviewed journal articles that present original research, methodology, or systematic reviews; a clear focus on AI or machine learning applications in urban settings; and abstracts that provide enough methodological detail. The exclusion criteria left out conference papers, book chapters, and editorials (which had already been filtered in Scopus); articles that talked about AI applications but didn't have anything to do with cities (about 400); articles that only mentioned AI or cities in passing and didn't really engage with them (about 320); and duplicate entries (about 50). This process left out 770 articles (9.1%), leaving a total of 7,660 articles. The corpus was then divided into two groups: empirical studies that included original data collection and analysis (n=5,441, 71.0%) and non-empirical studies that included methodological contributions, reviews, and conceptual frameworks (n=2,219, 29.0%). In-depth examinations of thematic structure, AI methodologies, and task allocations concentrated on the empirical subset to guarantee that results represent practical research applications rather than idealistic or theoretical assertions.

Data extraction and classification

Data extraction integrated automated metadata collection with systematic content analysis. The metadata fields in Scopus gave information about the publication, the author's affiliations, the abstract, the number of citations, the funding acknowledgments, and the open access status. Country attribution came after the first author's institutional affiliation, which made it possible to do geographic analysis across 109 countries. Funding data was obtained from acknowledgment sections and categorized by source type (national research councils, international organizations, private foundations) to analyze the correlations between research governance and methodological orientation.

Content classification utilized a hybrid methodology that combined automated keyword extraction, large language model-assisted coding, and expert validation. This approach embodies evolving best practices in systematic review methodology, wherein LLMs have proven effective in text classification tasks such as abstract screening, thematic coding, and content categorization (Delgado-Chaves et al., 2025) (Gilardi, Alizadeh, & Kubli, 2023). We used a predefined taxonomy with 16 method categories (such as deep learning, convolutional neural networks, random forest, graph neural networks, reinforcement learning, and natural language processing) to find AI methods by systematically matching keywords. We then manually verified any cases that were unclear. AI tasks were also split into 14 groups, including prediction and forecasting, classification, optimization, object detection, segmentation, simulation, and decision support. Thematic classification occurred via a two-stage process: an initial LLM-assisted allocation to 163 subject categories and 428 detailed clusters, succeeded by hierarchical consolidation into six megatrends (Digital Transformation and Smart Cities, Climate Change and Environmental Sustainability, Urban Mobility and Transportation, Urban Development and Land Use, Social Equity and Quality of Life, and Urban Resilience and Safety), which were validated through expert review (the comprehensive coding scheme, including all 15 variables and code definitions, is presented in Supplementary Tables S6-S14).

To support reproducibility, the LLM-assisted classification workflow is summarized as follows. Titles, abstracts, and author keywords for the 5,441 empirical studies were processed through a frontier-class large language model accessed via API under deterministic decoding (temperature = 0). Each prompt enumerated the candidate categories with one-sentence operational definitions, required the model to flag low-confidence cases for human review, and elicited a brief textual rationale for every assignment to enable downstream audit. Each record was processed in duplicate with prompt-order randomization, and disagreements between the two passes were routed to expert adjudication. The complete prompt templates, the JSON output schema, and an audit log of low-confidence flags are provided in Supplementary Section 1.5.

The assessment of sustainability included many different aspects. We coded SDG alignment against all 17 Sustainable Development Goals based on clear mentions, implicit thematic relevance, and contributions to sustainability outcomes. Articles could be aligned with more than one goal. We used a three-level scale to measure how deeply sustainability was integrated. Strong integration meant that there were clear references to sustainability frameworks, environmental or social impact assessments, and long-term effects. Medium integration meant that there were improvements in efficiency or resource use without clear references to sustainability. Weak integration meant that the contributions were purely technical and did not take sustainability into account. We coded the methodological characteristics of the articles by type (empirical, methodological, review, conceptual), research approach (quantitative, qualitative, mixed), spatial scale (local, regional, national, global), and temporal design (cross-sectional versus longitudinal).

Quality assurance adhered to established systematic review protocols(Higgins, 2011). A random 10% sample was double coded by independent reviewers to check for inter-coder reliability. Any differences were worked out through discussion and codebook improvement. Automated checks for consistency found logical problems, like when temporal designs didn't match up with the stated spatial scales. Domain experts then looked over the whole coding scheme to make sure it was face valid. Edge cases and outliers were manually reviewed to make sure they were classified correctly. The goal of these procedures was to find a middle ground between the size needed for a full corpus analysis and the strictness needed for a meaningful thematic synthesis (see Supplementary Section 1.7 and Table S15 for more information on quality assurance protocols and validation metrics).

Quantitatively, inter-coder reliability for the random 10% double-coded sample (n=545) reached Cohen's kappa = 0.83 for thematic megatrend assignment, 0.87 for AI method category, 0.81 for AI task type, and 0.78 for sustainability integration depth, all of which fall within the conventional "substantial" to "almost perfect" agreement bands. Disagreements were resolved through structured discussion and led to two iterative refinements of the codebook before the final pass.

We do not report a separate study-level quality appraisal because the unit of analysis here is the published article as a bibliometric and methodological signal rather than as a body of evidence to be synthesized for an effect estimate. Cochrane- and Joanna Briggs-style risk-of-bias instruments are designed for intervention-effect synthesis and are not informative for a thematic and methodological mapping at this scale. Methodological characteristics that would be inputs to such an appraisal (study design, spatial scale, temporal design, validation procedures) are instead encoded directly as variables in our coding scheme so that readers can stratify the descriptive findings by these markers.

Analytical framework

Descriptive analyses delineated temporal trends, geographic distributions, and methodological profiles throughout the corpus. Cross-tabulations analyzed the correlations among country of origin, AI methodologies, task orientations, thematic domains, and sustainability indicators. Funding analysis connected research traits to funding sources by looking at the difference in citation impact between funded and unfunded research and mapping funder priorities across methods, topics, and SDGs. Correspondence analysis is a multivariate method for examining relationships between categorical variables, yielding graphical representations that disclose patterns imperceptible through pairwise analysis alone(Greenacre, 2017) (Sourial et al., 2010). It situates thematic domains along interpretive axes that reflect methodological orientation (technical-predictive versus socio-spatial integration) and epistemological stance (black-box prediction versus interpretable explanation), thereby elucidating the conceptual framework of the field beyond mere frequency counts.

Correspondence analysis was performed on a contingency table cross-classifying the six megatrends against AI method × AI task combinations restricted to combinations with cell counts above ten. The first two principal axes accounted for 41.7% and 23.4% of total inertia, respectively, and were retained for interpretation. The horizontal axis was labeled "technical-predictive versus socio-spatial integration" because its positive pole loaded heavily on prediction, classification, and deep neural architectures, while its negative pole loaded on simulation, decision support, qualitative-inflected methods, and the Social Equity and Resilience megatrends. The vertical axis was labeled "black-box versus interpretable" because its poles separated tree-based ensembles and statistical learners (interpretable side) from deep and transformer architectures (black-box side). These labels are an analyst-imposed gloss on the underlying mathematical solution and should be read as interpretive scaffolding rather than as latent constructs identified by the data themselves.

The empirical subset (n=5,441) was the main focus of the analysis of the results about AI methods, tasks, and domain-specific patterns. This made sure that the reported distributions were based on real applications and not just proposed frameworks. Geographic analyses used the first author's affiliation to avoid giving too much weight to contributions from big international collaborations. However, this method may not show all of the collaborative networks. Temporal analyses looked at how the use of methods, thematic focus, and sustainability integration changed from year to year to find new trends and gaps that have been around for a long time. We focused on substantive interpretation over purely statistical description, trying to explain patterns we saw in terms of disciplinary incentives, institutional contexts, and epistemological commitments instead of just recording frequencies. Some of the problems are that it only uses one database (which could mean that it misses relevant work that is indexed in other places), it is biased toward English-language research (which could mean that it doesn't represent non-English-speaking research traditions well), and there is always some subjectivity in thematic classification, even with quality assurance measures in place.

Two limitations bear directly on how our findings should be read. First, restricting the search to English-language venues has uneven consequences across regions. China contributes 37.2% of the corpus despite the language filter, because Chinese researchers routinely target international English-language journals to maximize global visibility; nevertheless, Chinese-language indexing systems (e.g., CNKI), Lusophone urban research from Brazil, and Spanish-language research from Latin America are systematically underrepresented, and conclusions about regional governance practices and locally calibrated AI implementations should be read against that boundary. Second, the inferences advanced in the Discussion regarding institutional incentives, publisher preferences, and funding effects are interpretive readings of correlational patterns observed in the corpus rather than tests of those mechanisms; they are framed throughout as hypotheses for further investigation.

Data availability

The dataset underlying this study was retrieved from Scopus and is available in the Supplementary Data S1. Data S1 contains the 7,660 articles analyzed, including thematic classifications, AI method codes, and associated metadata. The complete dataset including codes is accessible at https://github.com/navid-nsk/ai-for-cities.

Author Contributions

S.N.M.M. conceived and designed the study, developed the systematic review methodology, conducted the literature search and data collection of 7,660 articles, performed the LLM-assisted classification and thematic analysis, executed the correspondence analysis and statistical analyses, created all figures and tables, wrote the original manuscript draft, and approved the final version for submission. H.C. supervised the research, contributed to the conceptual framework and research design, participated in the validation of thematic classifications, provided critical revision of the manuscript, and approved the final version for submission. M.S. contributed to the methodological framework development, provided expertise on spatial analysis and quantitative methods, reviewed and validated the analytical approach, participated in manuscript revision, and approved the final version for submission. B.G. contributed expertise on urban planning and governance perspectives, participated in the interpretation of findings related to smart cities and urban design, provided critical feedback on the discussion and policy implications, and approved the final version for submission.

Competing Interests

The authors declare no competing interests. None of the authors participated in any capacity as guest editor or reviewer to the journal or any collections.

Acknowledgements

No funding was received for this research.

References

A. B. I. Research. (2020). More than 500 Cities Will Have Digital Twins by 2025. Retrieved from

Ada Lovelace Institute, A. N., & Partnership, O. G. (2021). Algorithmic accountability for the public sector: AI Now Institute and Open Government Partnership.

Allam, Z., & Dhunny, Z. A. (2019). On big data, artificial intelligence and smart cities. Cities, 89, 80-91.

Athey, S. (2015). Machine learning and causal inference for policy evaluation. Paper presented at the Proceedings of the 21th ACM SIGKDD international conference on knowledge discovery and data mining.

Baas, J., Schotten, M., Plume, A., Côté, G., & Karimi, R. (2020). Scopus as a curated, high-quality bibliometric data source for academic research in quantitative science studies. Quantitative science studies, 1(1), 377-386.

Battaglia, P. W., Hamrick, J. B., Bapst, V., Sanchez-Gonzalez, A., Zambaldi, V., Malinowski, M., . . . Faulkner, R. (2018). Relational inductive biases, deep learning, and graph networks. arXiv preprint arXiv:1806.01261.

Batty, M. (2018). Digital twins (Vol. 45, pp. 817-820): Sage Publications Sage UK: London, England.

Batty, M., Axhausen, K. W., Giannotti, F., Pozdnoukhov, A., Bazzani, A., Wachowicz, M., . . . Portugali, Y. (2012). Smart cities of the future. The European Physical Journal Special Topics, 214(1), 481-518.

Baykurt, B. (2022). Algorithmic accountability in US cities: Transparency, impact, and political economy. Big Data & Society, 9(2), 20539517221115426.

Benjamin, R. (2019). Race after technology: Abolitionist tools for the new Jim code: Polity.

Bettencourt, L. M. (2013). The origins of scaling in cities. Science, 340(6139), 1438–1441. doi:10.1126/science.1235823

Bibri, S. E., & Krogstie, J. (2017). Smart sustainable cities of the future: An extensive interdisciplinary literature review. Sustainable Cities and Society, 31, 183-212.

Bibri, S. E., & Krogstie, J. (2019). Generating a vision for smart sustainable cities of the future: a scholarly backcasting approach. European Journal of Futures Research, 7(1), 1-20.

Brauneis, R., & Goodman, E. P. (2018). Algorithmic transparency for the smart city. Yale JL & Tech., 20, 103.

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

Buolamwini, J., & Gebru, T. (2018). Gender shades: Intersectional accuracy disparities in commercial gender classification. Paper presented at the Conference on fairness, accountability and transparency.

Chib, A., Alvarez, K., & Todorovic, T. (2022). Critical perspectives on the smart city: Efficiency objectives vs inclusion ideals. Journal of Urban Technology, 29(4), 83-99.

Cugurullo, F. (2020). Urban artificial intelligence: From automation to autonomy in the smart city. Frontiers in Sustainable Cities, 2, 38.

Datta, A. (2018). The digital turn in postcolonial urbanism: Smart citizenship in the making of India's 100 smart cities. Transactions of the Institute of British Geographers, 43(3), 405-419.

Delgado-Chaves, F. M., Jennings, M. J., Atalaia, A., Wolff, J., Horvath, R., Mamdouh, Z. M., . . . Baumbach, L. (2025). Transforming literature screening: The emerging role of large language models in systematic reviews. Proceedings of the National Academy of Sciences, 122(2), e2411962122.

Dembski, F., Wössner, U., Letzgus, M., Ruddat, M., & Yamu, C. (2020). Urban digital twins for smart cities and citizens: The case study of Herrenberg, Germany. Sustainability, 12(6), 2307.

Floridi, L. (2020). AI and its new winter: From myths to realities. Philosophy & Technology, 33(1), 1-3.

Floridi, L., Cowls, J., Beltrametti, M., Chatila, R., Chazerand, P., Dignum, V., . . . Vayena, E. (2018). AI4People—An ethical framework for a good AI society: Opportunities, risks, principles, and recommendations. Minds and Machines, 28, 689-707.

Gilardi, F., Alizadeh, M., & Kubli, M. (2023). ChatGPT outperforms crowd workers for text-annotation tasks. Proceedings of the National Academy of Sciences, 120(30), e2305016120.

Greenacre, M. (2017). Correspondence analysis in practice: chapman and hall/crc.

Hastie, T., Tibshirani, R., Friedman, J. H., & Friedman, J. H. (2009). The elements of statistical learning: data mining, inference, and prediction (Vol. 2): Springer.

Higgins, J. (2011). Cochrane handbook for systematic reviews of interventions. Version 5.1. 0 [updated March 2011]. The Cochrane Collaboration. www. cochrane-handbook. org.

Hollands, R. G. (2008). Will the real smart city please stand up? City, 12(3), 303-320.

Hollands, R. G. (2020). Will the real smart city please stand up?: Intelligent, progressive or entrepreneurial? The Routledge companion to smart cities (pp. 179-199): Routledge.

Interoperable Europe. (2024). Reimagining Helsinki: Participatory Urban Planning with Generative AI. Retrieved from

Jiang, W., & Luo, J. (2022). Graph neural network for traffic forecasting: A survey. Expert systems with applications, 207, 117921.

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

Kitchin, R. (2017). Thinking critically about and researching algorithms. Information, Communication & Society, 20(1), 14-29.

Kreif, N., & DiazOrdaz, K. (2019). Machine learning in policy evaluation: new tools for causal inference. arXiv preprint arXiv:1903.00402.

Labs, U. A. (2024). Bringing Communities Together Through AI-Driven Urban Planning. Retrieved from

Li, H., Zhao, Y., Mao, Z., Qin, Y., Xiao, Z., Feng, J., . . . Zhang, M. (2024). A survey on graph neural networks in intelligent transportation systems. arXiv preprint arXiv:2401.00713.

Lum, K., & Isaac, W. (2016). To predict and serve? Significance, 13(5), 14-19. doi:10.1111/j.1740-9713.2016.00960.x

Luque-Ayala, A., & Marvin, S. (2016). The maintenance of urban circulation: An operational logic of infrastructural control. Environment and Planning D: society and Space, 34(2), 191-208.

Mongeon, P., & Paul-Hus, A. (2016). The journal coverage of Web of Science and Scopus: a comparative analysis. Scientometrics, 106(1), 213-228.

Morozov, E., & Bria, F. (2018a). Rethinking the smart city. Democratizing Urban Technology, 2.

Morozov, E., & Bria, F. (2018b). Rethinking the Smart City: Democratizing Urban Technology. Retrieved from

Page, M. J., McKenzie, J. E., Bossuyt, P. M., Boutron, I., Hoffmann, T. C., Mulrow, C. D., . . . Brennan, S. E. (2021). The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. bmj, 372.

Pearl, J., & Mackenzie, D. (2018). The book of why: The new science of cause and effect: Basic Books.

Raissi, M., Perdikaris, P., & Karniadakis, G. E. (2019). Physics-informed neural networks: A deep learning framework for solving forward and inverse problems involving nonlinear partial differential equations. Journal of Computational Physics, 378, 686-707.

Reisman, D., Schultz, J., Crawford, K., & Whittaker, M. (2018). Algorithmic impact assessments: a practical Framework for Public Agency. AI Now, 9.

Rudin, C. (2019). Stop explaining black box machine learning models for high stakes decisions and use interpretable models instead. Nature Machine Intelligence, 1(5), 206–215. doi:10.1038/s42256-019-0048-x

Rudin, C. (2019). Stop explaining black box machine learning models for high stakes decisions and use interpretable models instead. Nature Machine Intelligence, 1, 206-215.

Sanchez, T. W., Shumway, H., Gordner, T., & Lim, T. (2023). The prospects of artificial intelligence in urban planning. International Journal of Urban Sciences, 27(sup1), 179-194.

Shao, X., Liu, Z., Zhang, S., Zhao, Z., & Hu, C. (2023). PIGNN-CFD: A physics-informed graph neural network for rapid predicting urban wind field defined on unstructured mesh. Building and Environment, 232, 110056.

Sloane, M., Moss, E., Awomolo, O., & Forlano, L. (2022). Participation is not a design fix for machine learning. Paper presented at the Proceedings of the 2nd ACM Conference on Equity and Access in Algorithms, Mechanisms, and Optimization.

Sourial, N., Wolfson, C., Zhu, B., Quail, J., Fletcher, J., Karunananthan, S., . . . Bergman, H. (2010). Correspondence analysis is a useful tool to uncover the relationships among categorical variables. Journal of clinical epidemiology, 63(6), 638-646.

Strathern, M. (1997). 'Improving ratings': audit in the British University system. European review, 5(3), 305-321.

U.N.-Habitat, & Mila. (2022). Quebec Artificial Intelligence Institute. UN-Habitat.

United Nations Department of Economic and Social Affairs. (2019). World Urbanization Prospects 2018: Highlights. Retrieved from

United Nations Department of Economic and Social Affairs. (2025). World Urbanization Prospects 2025: Summary of Results. Retrieved from

United Nations Development Programme. (2015). What are the Sustainable Development Goals. Retrieved from https://www.undp.org/content/undp/en/home/sustainable-development-goals.html/

Veale, M., & Brass, I. (2019). Administration by algorithm? Public management meets public sector machine learning. Public management meets public sector machine learning.

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.