Journal of Urban Intelligence and Smart Systems

JUISS

www.juiss.org
eISSN: 0000 -000

Governing the Ethical AI City: Mitigating Algorithmic Vulnerabilities in Smart Urban Infrastructure

Reazul Ahsan[1,*], Sung Jin Park[2], Yejin Kim[3]

1 Department of City and Metropolitan Planning, University of Utah (Asia Campus), Incheon, South Korea
2 CEO SPARK CITY, Seoul, South Korea
3 Department of City and Metropolitan Planning, University of Utah (Asia Campus), Incheon, South Korea

ARTICLE INFO

Article history:
Received 28 June 2026
Received in revised form 1 August 2026
Accepted 4 August 2026
Available online 7 August 2026

Keywords:

Algorithmic governance; Smart city ethics; Urban artificial intelligence deployment; Cyber-physical infrastructure; Spatial justice

ABSTRACT

Algorithmic governance must take a crucial urban turn in light of the swift shift from data-informed smart cities to autonomously adaptive artificial intelligence (AI) cities. Although there are international frameworks for AI ethics, they do not address the systemic, infrastructural, and spatial issues that arise when deploying AI at the city level. By presenting urban settings as dynamic ethical subjects rather than passive technology nodes, this article expands the conceptual paradigm of the Ethical AI City. We identify key vulnerabilities in existing smart-city cyber-physical systems by analyzing growing systemic occurrences ranging from algorithmic bias and autonomous car failures to mass surveillance. We contend that integrating ethical protections directly into intelligent public infrastructure is necessary to ensure democratic resilience, equality, and spatial justice. Lastly, we suggest a socio-technical roadmap that leverages cross-sectoral governance models, transparent algorithmic audits, and participatory design. This framework provides actionable insights for urban planners, system engineers, and policymakers designing the next generation of resilient, equitable, and sustainable cities.

1. Introduction: The Strategic Shift from Smart to AI Cities

The urban landscape around the world is undergoing a profound change in how people and technology interact. It is moving from the "Smart City" model, where Information and Communication Technologies (ICT) are gradually used to enhance government efficiency, to the "AI City," where Artificial Intelligence (AI) drives governance, predictive maintenance, and spatial organization [12,33,37,70]. This shift toward AI-driven urbanism extends beyond mere data collection to include highly customized services and algorithms that can forecast crises. This necessitates a complete reorganization of ethical considerations [13,33,94].

This transition is particularly significant in the South Korean context, where the government has articulated a national ambition to become one of the world's top three global AI leaders [77]. In pursuit of this goal, the government is advancing a new strategic direction toward AI cities, building

* Corresponding author.
E-mail address: reazul.ahsan@utah.edu

on its longstanding leadership in smart city development. Key initiatives, including the establishment of a central government-led "AI City" Task Force [95], the launch of the AI City R&D Program [76], and the announcement of AI City pilot projects [96], reflect ongoing efforts to develop and institutionalise a world-leading model for AI Cities [96]. While these initiatives demonstrate strong institutional commitment, parallel discussions on ethical governance and societal implications at the city scale remain relatively underdeveloped, thereby revealing a significant gap in ethical oversight in the development of AI cities.

Korea's National AI Ethics Guidelines [97] and the recently enacted Framework Act on AI [124], together with the Organization for Economic Co-operation and Development (OECD) AI Principles [113] and the European Union (EU) AI Act [51], are examples of current AI ethical benchmarks that establish important principles but do not fully address the challenges of AI-enabled urban service, infrastructure, and city-scale governance [15]. The "black-box" nature of AI reduces cities' resilience, increasing risks from minor disruptions to major calamities [116]. This article presents the "Ethical AI City" as a governance paradigm that expands Henri Lefebvre's "Right to the City" to encompass algorithmic co-creation and spatial equity [82,132].

1.1 The Urban Turn in AI Ethics

The "Urban Turn" distinguishes traditional AI ethics, which prioritizes individual rights and societal implications, from urban informatics, which focuses on communal spatial justice [151]. Widely recognized AI ethics frameworks, such as Fairness, Accountability, and Transparency (FAT), primarily address individual-level concerns, including personal biases, accountability, and transparency. However, urban ethics must also encompass the protection of city infrastructure, public platforms, and interrelated data ecosystems [94]. This shift makes the city an ethical entity, challenging the placelessness of digital technologies and grounding them in local geographies and power systems [72].

The "Urban Turn" concept advanced in this paper is distinct from prior contributions to smart city ethics in three key respects. First, whereas foundational smart city ethics scholars such as Kitchin [72] and Mattern [94] critique smart city ideology from a social science perspective exposing surveillance capitalism, corporate solutionism, and data extractivism they do not systematically theorize the city as an ethical subject with its own governance architecture and spatial agency. The "Urban Turn" takes this further by positioning the city as an active moral actor whose infrastructural systems, public platforms, and spatial configurations instantiate ethical commitments (or fail to). Second, prior AI ethics frameworks, including the widely cited FAT literature (e.g., [10,46,152,153]) and AI governance frameworks such as the OECD AI Principles and the EU AI Act, primarily operate at the level of individual rights, dataset properties, and algorithmic outputs. The "Urban Turn" argues that these frameworks are categorically insufficient for city-scale AI deployment, where the unit of ethical harm is not the individual algorithmic decision but the systemic, cumulative spatial effect of thousands of interconnected AI decisions on urban form, mobility, housing, and public life. Third, by grounding its ethical analysis in Lefebvre's [82] concept of the Right to the City, this paper insists that algorithmic co-creation of urban space is a political right, not merely a technical or regulatory problem. This distinguishes the "Urban Turn" from technocratic smart city governance approaches and aligns it with participatory urban design and spatial justice traditions [52,85,132].

Critically, the "Urban Turn" is not merely a conceptual reframing: it carries direct implications for urban design practice and the physical form of cities. Ethical AI governance shapes how cities look, feel, and function spatially. Algorithmic Transparency affects how intelligible urban infrastructure sensor arrays, automated systems, data collection nodes is to residents as they navigate everyday street environments. Accountable Governance determines who controls decisions about land use, zoning, and public-space allocation in AI-enabled districts, including Songdo International Business District (Songdo IBD) and future AI pilot cities envisioned under MOLIT's 2026 programme. Structural equity directly conditions whether AI-optimized mobility systems serve or bypass low-income neighborhoods, and whether the design of smart districts produces inclusive or exclusionary spatial and economic outcomes [134,158]. The Quality of Life-Privacy Balance governs the extent to which ubiquitous CCTV, facial recognition, and Internet of Things (IoT) sensor networks embedded in building facades, streetscapes, and public plazas enhance or degrade the sense of place and civic comfort. Fair AI Economy principles determine whether the economic value generated by AI data centers and platform economies is reinvested in urban infrastructure or extracted by private actors. Urban designers engaging with AI cities must therefore treat ethical governance not as an external regulatory constraint, but as an intrinsic dimension of the design brief one that shapes street sections, building programs, public-space typologies, and the governance structures of smart districts from the outset. The following table 1 distinguishes between individual- and city-scale ethics.

Table 1
Differing Individual vs. City-Scale Ethics

Criterion	Individual & Society Lens	City Lens (The Urban Turn)
Ethical Subjects	Individuals, users, vulnerable groups	City systems, public platforms, urban data
Key Questions	How does AI affect personal rights and society?	How does AI reshape urban systems, governance and spatial conditions?
Data Structure	Personal data, behavioral patterns, consumer profile	Large-scale urban data (e.g., CCTV, IoT networks, sensor infrastructure)
Accountability Actors	Users, private companies, tech/platform firms	Municipalities, agencies, and infrastructure operators
Volume of Impact	Personal injustice, bias, social exclusion	Systemic collapse, urban inequality, distrust

Note. Author's adaptation and synthesis based on Mattern [94] and Batty [12]

1.2 Urbanism Beyond Frankenstein

Federico Cugurullo's "Frankenstein Urbanism" compares experimental cities to Shelley's monster: ambitious but flawed creations that escape control and prioritize corporate interests over human needs [38]. In eco-cities like Masdar City (Abu Dhabi, UAE), green technologies promise sustainability but often result in elitist enclaves with rebound effects that exacerbate inequality [27,38]. Smart city solutionism worsens this by deploying siloed technologies like IoT sensors that emphasize efficiency metrics and data monetization by companies like IBM, at the expense of public welfare and enabling surveillance [63,139]. This leads to systems that invade privacy and widen social and digital divides, such as biased algorithmic policing in disadvantaged areas [72].

The AI City amplifies risks by embedding opaque neural networks into decision-making, threatening democracy without participatory governance [26,159]. Black-box AI could produce unfair outcomes, such as prioritizing corporate logistics over community access, unless citizens collaborate in workshops to ensure inclusivity and transparency [28,90]. Nordic examples, such as citizen-centric data governance initiatives in Helsinki and other Finnish cities, demonstrate that ethical AI can build trust by allowing public input into data policies [90].

AI cities possess autonomous agency (the ability to make independent decisions), swarm intelligence (the ability to coordinate AI agents for tasks like disaster response), and generative capacities (the ability to create adaptive designs). This enables cities to evolve in novel ways but risks excluding people [11,43]. Hybrid models that combine technology and participation can foster inclusive cities [11].

1.3 The Paradigm Shift: Automation vs. Autonomy

Smart systems automate rule-based tasks with human oversight. For example, programmed traffic lights in cities like Pittsburgh reduce delays by 25% [135]. However, these systems struggle with unpredictable scenarios [11,12]. AI shifts to autonomy, enabling probabilistic and adaptive actions in unstructured environments such as self-driving vehicles interpreting social cues or algorithms predicting floods from real-time data [11,70,140]. In urban planning, automation handles routine tasks, while autonomy learns and innovates. For instance, AI-optimized urban systems in Singapore predict needs [37,74]. This requires ethical constraints for manageability, defining objectives like “minimize latency at the lowest cost” to support human collaboration within resilient ecosystems [11,94]. Table 2 illustrates the paradigm shift from Smart City to AI City.

Table 2
Smart City vs. AI City Paradigms

Feature	Smart City Paradigm	AI City Paradigm
System	Integrated Monitoring Platform	AI enabled / personalized AI agent
Goal	Efficiency / Productivity	Adaptability / Flexibility
Means	Automation	Autonomy
Core Function	Data Collection & Monitoring	Prediction, Adaptation, & Action
Human Role	Decision-Maker (Data-Informed)	Supervisor / Human-on-the-Loop
Decision Logic	Rule-Based / Deterministic	Probabilistic / Neural / Generative
Temporal Scale	Retrospective / Real-Time	Predictive / Pre-emptive
Key Risk	Surveillance / Privacy Breach	Loss of Agency / Systemic Fragility

Note. Author’s adaptation and synthesis based on Batty [12], Cugurullo [37], Barns [11], and Jang [70].

2. Chronology of Failure: Analyzing Systemic Risks in Urban AI Systems

The systemic failures of AI systems worldwide stem from recurring structural weaknesses that become especially challenging at the urban scale. These include accountability gaps (where responsibility for harm diffuses across developers, operators, and regulators), algorithmic bias (which reproduces societal inequalities), algorithmic transparency deficits (black-box decisions that obscure outcomes and erode public trust), infrastructural fragility (single points of failure in critical services), surveillance-privacy risks, and power concentration (over-reliance on a few tech providers).

The table 3 below focuses on AI deployments and governance controversies at the urban scale. The cases were primarily identified through the AI Incident Database[1] and the OECD AI Incidents Monitor (AIM)[2].Two publicly accessible repositories document real-world AI failures and governance-related harms across sectors and jurisdictions. Because they encompass AI incidents across diverse

[1] *An open-source, community-led database that collects and classifies reports of AI-related failures, harms, and safety issues to promote transparency and inform responsible AI development. Website: https://incidentdatabase.ai/*

[2] *A monitoring system developed by the OECD to track global trends in AI incidents, providing data-driven insights on safety, ethical risks, and impacts for policymakers and researchers. Website: https://oecd.ai/en/incident-monitor*

domains, cases were screened for relevance to urban systems and for demonstrated impacts on urban governance, public services, mobility, public safety, and city residents.

Table 3
Chronology of AI-Related Urban Incidents and Governance Controversies

Incident / urban system	Core ethical risk	Impact on public trust
2016–2020: Los Angeles PredPol / predictive policing programs	Algorithmic bias; feedback loops; weak transparency	Critiques that historical policing data reproduced over-policing in already surveilled neighbourhoods contributed to program abandonment and broader scepticism toward predictive policing.
2016 and 2018: autonomous-vehicle fatalities in public roadways (Tesla Autopilot; Uber ATG)	Accountability gap; safety-case failure; human-machine responsibility	Fatal crashes showed that automated mobility failures are urban public-safety issues, not only product-liability events.
2020–2024: Detroit police facial-recognition false arrests	Racial bias; automation bias; due-process failure	Wrongful arrests of Black residents damaged trust in police AI and led to litigation, audits, and stricter local governance requirements.
2021–2024: Chicago ShotSpotter acoustic gunshot detection	Surveillance, reliability, and racialized deployment	High false or unverified alert burdens and concentrated deployment in minority neighborhoods undermined community trust and intensified calls to cancel or restrict the system.
2021–2023: Rotterdam municipal welfare-fraud risk scoring	Algorithmic discrimination; opacity in public services	Risk scoring disproportionately flagged migrants, single mothers, and vulnerable residents, weakening trust in automated welfare administration.
2023: San Francisco Cruise robotaxi pedestrian-dragging incident	Safety failure; incomplete incident reporting; private-platform accountability	Regulatory suspension and national service pause showed that city governments need enforceable incident disclosure, audit access, and emergency protocols for autonomous fleets.
2023–2024: robotaxis blocking emergency response and urban traffic in San Francisco	Infrastructural fragility; coordination failure	Stalled vehicles and emergency-service interference raised doubts about whether private autonomous fleets were adequately integrated with city operations.
2025: New Orleans live facial-recognition camera network controversy	Surveillance creep; procurement opacity; local-law compliance	Reports that police used a privately operated real-time facial-recognition network despite local restrictions intensified concerns about democratic oversight of urban AI surveillance.
2025: Baltimore County school AI gun-detection false alarm	False positives; escalation risk; child and civil-rights harms	A student was handcuffed after an AI camera reportedly misread a snack bag as a weapon, demonstrating how automated safety systems can rapidly convert model error into embodied harm.
2025: expanding live facial-recognition deployments in city centers (e.g., Liverpool/London)	Privacy, proportionality, and bias concerns	Large-scale public-space identification systems normalized continuous biometric surveillance and renewed debate over consent, redress, and equal treatment in smart cities.

Note. Author's adaptation and synthesis based on the AI Incident Database and OECD AI Incidents Monitor

These cases demonstrate that urban AI failures are rarely isolated technical glitches; they systematically amplify existing societal fault lines at the city scale. Rather, they reveal systemic governance challenges that emerge when AI systems are integrated into complex urban environments. Accountability remains elusive when legal frameworks and procurement contracts cannot keep pace with privately operated autonomous systems. Bias persists because training data and operational data reflect historical inequities. Fragility increases when AI is tied to mobility, emergency response, schools, policing, or public-benefit systems. Surveillance erodes civil liberties, and power concentrates in vendors whose incentives may conflict with long-term public accountability.

3. Methodology: A Decolonial Qualitative Systematic Literature Review

This study employs a systematic qualitative literature review (SLR) to interrogate the ethical challenges of AI-driven urbanism, using Songdo International City (hereafter, Songdo), South Korea, as an illustrative case, with particular attention to the smart city governance of Songdo IBD and IFEZ. Rather than a conventional aggregative review that seeks to summarize “what works,” this is an interpretive, configurative synthesis [57]: its purpose is theory-building, generating conceptual insight that equips urban governments to adopt proactive, human-centered AI governance rather than reacting to technological momentum after harms occur.

The review is explicitly framed through a decolonial research lens. Global smart city discourse and the more recent AI City narrative emerging in Korean policy since 2025 have been disproportionately shaped by technology firms, consultancies, and research institutions that embed assumptions about efficiency, datafication, and technological solutionism as universally desirable and beneficial outcomes [98,103]. A decolonial framework treats these assumptions as situated rather than neutral, and directs analytic attention to questions a mainstream review would leave implicit: whose knowledge counts as evidence, whose harms are rendered visible, who is positioned as the designer versus the governed, and how imported smart-city models reproduce extractive or hierarchical relations in a Korean context [9,121]. This reframing matters analytically because Songdo was conceived in part as a globally exportable, corporate-partnered model city; reading it through a decolonial lens exposes the power asymmetries that a purely technical or efficiency-oriented evaluation would naturalize.

3.1 Rationale for the Songdo Case

Songdo was selected as a case through purposive, critical-case sampling for four compounding reasons. First, it is paradigmatic: as South Korea’s most extensively documented purpose-built smart city, developed under the u-City and smart city frameworks and anchored by the Songdo IBD, it serves as both a national flagship and a reference point for subsequent AI City policy discussions [88,95]. Second, it is a critical case for anticipatory ethical governance: even if Songdo itself has relatively few documented AI-harm incidents, it concentrates the structural conditions this paper theorizes accountability gaps, privacy concerns, infrastructural fragility, public-private dependence, and platform governance making it an analytically productive site for testing the explanatory reach of the Ethical AI City Framework [127]. Third, it is institutionally proximate: the research team is based at the University of Utah Asia Campus in Songdo, affording situated access to planning documents and local discourse a positionality the decolonial frame requires us to declare rather than treat as a neutral vantage (see Section 3.5). Fourth, it is temporally relevant: as Korea has moved from smart city policy toward explicit AI City initiatives since 2025, Songdo’s trajectory offers actionable lessons for live design and governance decisions. The case is therefore diagnostic rather than illustrative selected not because Songdo is already an incident-rich AI City, but because its documented smart-city experience exposes the structural conditions that an Ethical AI City Framework must confront [133,134].

3.2 Review Protocol and Search Strategy

Following established SLR conventions adapted for qualitative synthesis [21,119], a review protocol was defined in advance specifying research questions, search sources, eligibility criteria, and the analytic approach. Searches were conducted across Web of Science, Scopus, and Google Scholar, supplemented by purposive grey-literature and primary-source searches in the AI Incident Database and OECD AIM, major Korean urban planning journals, the MOLIT Smart City portal, and reputable news archives (e.g., Reuters, The Guardian, BBC).

Search terms combined: "smart city ethics," "urban AI governance," "algorithmic bias cities," "Songdo smart city," "responsible AI urban," "urban AI failure," "surveillance capitalism cities," and "digital twin urban ethics." The search window was 2008–2025 for smart city and u-City literature, with 2008 chosen to capture the emergence of Korea's formal u-City legal framework. The keyword "AI City Korea" was applied only to 2025 materials, because explicit AI City policy narratives emerged later than the broader smart city literature.

The corpus spanned three intersecting evidence streams: (i) Songdo-specific materials official planning reports, feasibility studies, and evaluation documents from IFEZ and New Songdo International City LLC (NSIC), materials related to the IFEZ Smart City Operation Center, and corporate documentation on Songdo IBD managed services (e.g., Cisco), and academic and media case studies; (ii) global AI failure chronologies technical reports, the AI Incident Database, investigative journalism, government audits, and academic analyses of urban AI incidents; and (iii) ethical AI and smart city governance literature policy frameworks, white papers, Korean-language journal articles, and scholarly critiques of accountability, bias, transparency, privacy, infrastructural fragility, and power concentration in AI urban systems.

3.3 Screening and Eligibility

The initial corpus comprised approximately 120 records. Documents were included if they satisfied all three criteria: (1) direct relevance to AI or ICT governance in urban contexts; (2) engagement with at least one core ethical dimension (accountability, bias, transparency, privacy, fragility, or power concentration); and (3) availability in full text via open academic databases or official institutional sources. Documents were excluded if they (a) addressed AI ethics solely at the individual or national level without urban application; (b) were superseded by later policy editions; (c) were duplicates or near-duplicates; or (d) could not be independently verified for provenance. Screening reduced the corpus from 120 to 68 documents; this flow is reported in a PRISMA-style diagram (Figure 1), which presents counts of identification, screening, eligibility, and inclusion.

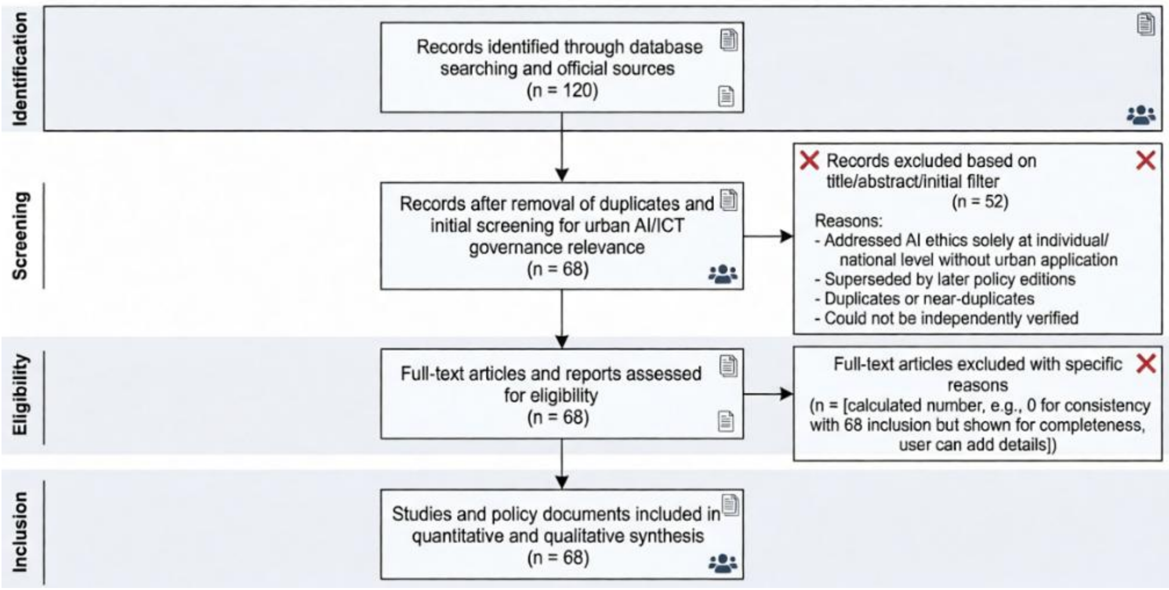

Fig. 1 PRISMA-style flow diagram of document selection

3.4 Analytical Procedure: Critical Thematic Synthesis

Included documents were analyzed through thematic synthesis [143], conducted in four interpretive passes adapted from Bowen [23]: (1) skimming for relevance, (2) close reading and analytic memo-writing, (3) line-by-line and document-level coding, and (4) synthesis of descriptive codes into higher-order analytic themes. Coding was inductive but read against a critical-decolonial sensitizing framework, so that each emergent theme accountability gaps, algorithmic bias, infrastructural fragility, surveillance and privacy erosion, and power concentration was interrogated not only for its presence but for its distributive logic: who benefits, who bears the risk, and whose knowledge the source treats as authoritative.

Constant comparison across the three evidence streams continued until theoretical saturation, operationalized as the point at which additional documents yielded no new analytic codes [84]. In practice, saturation was reached after roughly 55 documents; the remaining 13 were reviewed to confirm consistency. A reflexive audit trail of coding decisions and analytic shifts was maintained throughout. A manual approach was chosen over CAQDAS to enable deeper interpretive engagement with the texts.

3.5 AI Support and Researcher Reflexivity

Given the paper's focus on AI ethics, the use of AI-assisted tools during the research is disclosed for transparency. The primary AI tool during the initial review stage was Grok 3 (xAI, via Grok.com, February–March 2026); ChatGPT (GPT-4o) was used to cross-check AI-generated summaries for consistency. AI assistance was confined to: (1) generating initial keyword combinations; (2) producing preliminary summaries of lengthy policy documents, subsequently read in full by the researcher; (3) first-pass translation support for Korean-language policy excerpts, later verified by a Korean-proficient co-author. All AI outputs were treated as draft material subject to mandatory human verification, with every AI summary cross-referenced against its source; any discrepancy resulted in the output being discarded and the passage rewritten from the primary source. The following were performed exclusively by human researchers: inclusion/exclusion decisions, thematic coding, interpretive analysis, theoretical framing, synthesis of the Ethical AI City Framework, source selection, and all substantive scholarly judgements.

4. The Ethical AI City Framework: A Multidimensional Model

The Ethical AI City Framework integrates responsible AI principles into city governance, moving beyond mere compliance checklists. It transforms key ideas about fairness, accountability, and transparency into practical steps that cities can implement [117,164]. This framework recognizes cities' critical role in addressing broader ethical issues and risks, especially where unchecked AI systems could lead to privacy violations or disputes over decisions.

At its core, the framework includes six interconnected pillars that tackle specific risks associated with AI-enabled urban systems. These pillars function as a system to ensure adaptability across different city types, whether newly built smart cities or legacy ones adapting to new technologies [26,102]. Implementing this framework involves engaging stakeholders including city officials, residents, and businesses to align AI with values of fairness and democracy. Each pillar addresses key risks, providing examples and policy suggestions for effective urban AI governance.

4.1 Algorithmic Transparency

Algorithmic transparency aims to ensure that the collection, use, and sharing of urban data, together with the AI algorithms that shape decision-making processes, are visible and understandable to citizens and other stakeholders. As cities increasingly rely on AI to support decisions in areas such as land-use planning, participatory urban planning, and urban environment monitoring, it becomes essential that citizens understand what data are being collected, how AI algorithms use those data, and which institutions have access to them [14,19,33]. Without such transparency, AI-enabled systems risk operating as opaque infrastructures beyond democratic scrutiny [4,6,25].

The importance of transparency became widely recognized following the 2018 Facebook-Cambridge Analytica data privacy scandal, which revealed how large-scale data harvesting and algorithmic profiling could influence political processes. More broadly, the scandal exposed the risks associated with the secondary use of personal data beyond its original purpose and without meaningful public awareness or consent, as evidenced by practices such as facial recognition in public spaces, AI-based workplace monitoring, and the algorithmic evaluation of employee performance [3,65]. These cases highlight the risks of data misuse and underscore the need for greater transparency in data and algorithm governance, providing citizens with clearer visibility into how data and algorithms are collected, processed, shared, and deployed in systems that shape public outcomes and urban decision-making.

Initiatives such as Helsinki's AI Register and Amsterdam's Algorithm Register provide greater transparency regarding how algorithms are used in public services and urban decision-making [34,35].

In Songdo, for example, extensive AI CCTV deployment and sensor-based urban management systems rely on large volumes of data to operate effectively. Greater transparency would allow residents to understand how these systems function, what data they rely on, and how that information is used in decision-making processes.

4.2 Accountable Governance

This principle establishes clear governance guidelines for the responsible use of AI in cities. AI is increasingly being employed in public administration for tasks such as policy prediction, fraud detection, and resource allocation [8,150]. Yet in many cases, public authorities are not required to disclose how, where, or for what purposes AI algorithms are used. As AI deployment expands particularly with the growing use of generative AI opaque black-box algorithms may undermine due process by obscuring decision-making logic and procedural mechanisms from affected individuals and oversight bodies [5,25,128].

The Netherlands' SyRI (System Risk Indication) welfare monitoring system provides a notable example of the risks associated with opaque algorithmic governance. Municipalities used algorithmic profiling to identify potential welfare fraud, yet some residents were labeled "high risk" without explanation, triggering protests and a court ruling that struck down the system for violating human-rights protections [110,148]. Similar concerns have been observed in the Arkansas Medicaid allocation algorithm, which reportedly reduced care services for vulnerable beneficiaries, and the UK welfare fraud detection algorithm, which faced criticism for potentially discriminating against people with disabilities [22,80]. Such opacity arising from the deployment of AI systems without legitimate due process can weaken public trust and contestability in urban governance. In Songdo, the automated waste system prioritizes engineering efficiency over procedural fairness. The lack of administrative appeal mechanisms denies residents the due process to challenge automated operational rules, effectively sidelining democratic participation [161].

To operationalize accountable governance, cities can introduce enforceable mechanisms such as algorithmic impact assessments, AI ethics boards, and public AI registries throughout the AI system development and deployment process, as demonstrated by Helsinki's AI Register [35]. Institutionalizing these governance practices from the outset can help build public trust and ensure accountability in AI-driven urban systems.

4.3 Accountability Chain

This pillar addresses the growing challenge of establishing clear lines of liability and responsibility when AI system failures, errors, or faulty actions cause harm to people, property, or economic interests in urban environments. Beyond the widely discussed autonomous vehicle incidents involving Tesla, Uber, and Waymo, comparable challenges are also becoming evident in service robots and AI-enabled public systems. AI-related incidents have included robots causing injuries or property damage, AI surveillance and decision-support systems contributing to wrongful enforcement actions, such as false penalties or arrests, and AI-generated misinformation triggering unnecessary emergency responses, such as school lockdowns and emergency dispatches [2,20,154].

As AI becomes more deeply embedded in urban systems, cities are likely to encounter a growing number of AI incidents in which responsibility is shared among developers, operators, public agencies, and regulators, yet remains poorly defined, creating new governance and liability challenges. When responsibility remains unclear, public trust can erode, turning technical failures into broader urban social conflicts, as seen in public backlash and vandalism targeting robotaxi services in San Francisco, while leaving affected individuals with limited avenues for redress [71,126]. The December 2024 collision in Songdo, involving a remotely operated delivery robot, illustrates the breakdown of the accountability chain, as the blurred line between autonomous logic and human-in-the-loop control obscures liability and complicates the due process for affected citizens [67].

Addressing AI-related risks requires a clear liability and accountability framework that allocates responsibilities across the AI lifecycle and ensures oversight, transparency, and redress mechanisms for affected residents [24]. Cities should establish protocols for documenting, auditing, and accessing operational data to identify responsible actors and investigate system failures. Tools such as AI decision logs, audit trails, smart contracts, and blockchain-based tracking can enhance traceability and accountability. Canada’s Directive on Automated Decision-Making offers a useful model by requiring documentation and logging of AI decisions for accountability [144]. Such measures can strengthen urban resilience by improving cities’ capacity to monitor, investigate, and respond to AI-related failures.

4.4 Structural Equity and Inclusion

This pillar addresses the risks of algorithmic bias, data deserts, and representation gaps, which reinforce structural inequalities and undermine spatial justice.

AI applications such as résumé screening, facial recognition, and welfare automation often reproduce historical biases embedded in training data, reinforcing existing social, economic, and spatial inequalities. Such outcomes contribute to what scholars describe as a “data divide” that compounds the traditional digital divide [10,17,41,50]. Structural inequities may arise not only from biased data but also from the absence of data itself. These “data deserts” emerge when certain populations or places generate insufficient data to be meaningfully represented in AI-driven decision-making processes [17,78,125]. Examples may include temporary residents, informal economies, marginalized communities, or complex ecological systems. As cities rely on data-intensive governance, those with data scarcity risk receiving less policy attention and reduced visibility in urban decision-making.

Even with data, structural inequities can persist if certain actors, interests, or impacts remain underrepresented in decision-making processes. Representation gaps occur when available data are systematically underweighted, overlooked, or excluded from AI-driven governance systems. AI systems prioritize measurable datasets while marginalizing hard-to-quantify values, including ecological systems, future generations, international residents, and other non-human stakeholders [16,79,157]. Consequently, AI cities may reinforce inequalities not only through biased data or data scarcity but also through the systematic exclusion of diverse perspectives.

An ethical AI city must therefore move beyond algorithmic fairness to recognize representation as a fundamental governance challenge. If appropriately designed and governed, AI can amplify the voices of those who have traditionally been fragmented, underrepresented, or excluded from urban governance [7]. This includes not only diverse human groups such as international residents in Songdo but also environmental systems and emerging urban actors such as robots and AI-enabled infrastructures [156,162]. Ultimately, AI-enabled urban governance should move beyond a technocratic reliance on data, acknowledging that all data is partial and value-laden. By critically compensating for these inherent flaws, cities can evolve from mere data-driven efficiency toward an inclusive architecture that fosters more equitable and adaptive urban futures.

4.5 Quality of Life-Privacy Balance

This pillar addresses the growing tension between citizen welfare and the protection of personal privacy in AI-driven smart city systems. The same data fusion that enables citizen welfare through predictive and personalized services can foster pervasive surveillance and “function creep,” in which urban data becomes a tool for eroding civil liberties [39,48,53,72].

This concern is already evident in AI-enabled social care services. Examples include Seoul’s IoT-based care services for older adults living alone, Singapore’s Smart Elderly Monitoring System (SEMS), and the EU’s Ambient Assisted Living (AAL) programme. These initiatives leverage residential sensors and behavioral data to detect signs of distress and deliver proactive welfare interventions [18,45,83]. Similar approaches are being implemented in Songdo through Incheon’s IoT Safety and Care Service (in-home activity sensors for older adults), Care Plug Service (utility consumption monitoring), and AI Safety Check Service (utility consumption monitoring) to support vulnerable populations and enable preventive care [68,69].

Toronto’s Sidewalk Labs smart district project ultimately collapsed despite its promise to enhance urban quality of life through data-driven services, sustainable infrastructure, and personalized urban experiences. Public opposition emerged over the extensive collection of personal and behavioral data, unresolved questions regarding data ownership and control, and the lack of transparent governance mechanisms to ensure meaningful public oversight [61]. The case illustrates that improvements in urban livability and service personalization cannot be pursued at the expense of privacy and individual autonomy.

To mitigate these risks, this pillar advocates governance mechanisms that safeguard personal data while balancing public safety and AI-enabled personalized urban services. Approaches such as anonymized data flows, consent-based data governance, and differential privacy techniques can reduce risks of mass surveillance [40]. Embedding such privacy protections into urban AI governance is essential to ensure that smart city technologies enhance urban life without undermining civil liberties.

4.6 Fair AI Economy

This pillar addresses the risk of economic power concentration and uneven value distribution within AI-driven urban economies [31,145].

Historical experiences with technology-led urban development suggest that concentrated investments often generate uneven economic outcomes across and within cities. New towns, innovation districts, and smart city initiatives have frequently enhanced economic opportunities and quality of life in selected locations while contributing to disparities between core and peripheral areas [62,89,101,127]. Songdo, for example, illustrates how strategic investments in advanced infrastructure can create significant economic advantages for specific districts while raising concerns about uneven regional development [160]. Similar dynamics may emerge in the AI era. Investment, innovation activities, and high-skilled employment are likely to concentrate in locations with AI infrastructure and strong digital connectivity, while less-connected areas may decline relatively [108]. As a result, disparities in access to AI infrastructure and AI-enabled industrial development may reinforce existing inequalities and create new economic divides among cities and regions.

At the same time, cities must contend with the growing influence of global digital platforms on local economies. AI algorithms embedded within dominant platforms may also distort local market dynamics and strengthen platform market power [104,131]. Algorithmic pricing systems employed by platforms such as Zillow, RealPage, and Airbnb have been linked to rent inflation, speculative behavior, and increased volatility in local housing markets [1,55,66]. These effects can extend beyond housing by influencing investment patterns, affordability, and broader urban economic conditions.

Together, these developments highlight the need for AI city governance frameworks that address not only technological innovation but also the equitable distribution of economic value [54]. In response, this pillar proposes governance mechanisms aimed at ensuring that AI-driven productivity gains are shared equitably across places and sectors. These include conceptualizing AI data centers as digital commons collectively governed computational infrastructures that support shared access to AI capabilities and fairer AI value distribution as well as strengthening institutional capacities to monitor and regulate the economic impacts of AI-driven platform economies. Through these measures, cities can foster more inclusive and spatially balanced AI-enabled economic development.

To visualize the framework's interconnected nature, Figure 2 depicts a hexagonal model in which the central "Ethical AI City" hub connects to the six pillars, illustrating their mutual reinforcement. This structure emphasizes that the isolated adoption of a single pillar (e.g., transparency without accountability) is insufficient; full integration is required to address urban-scale risks holistically.

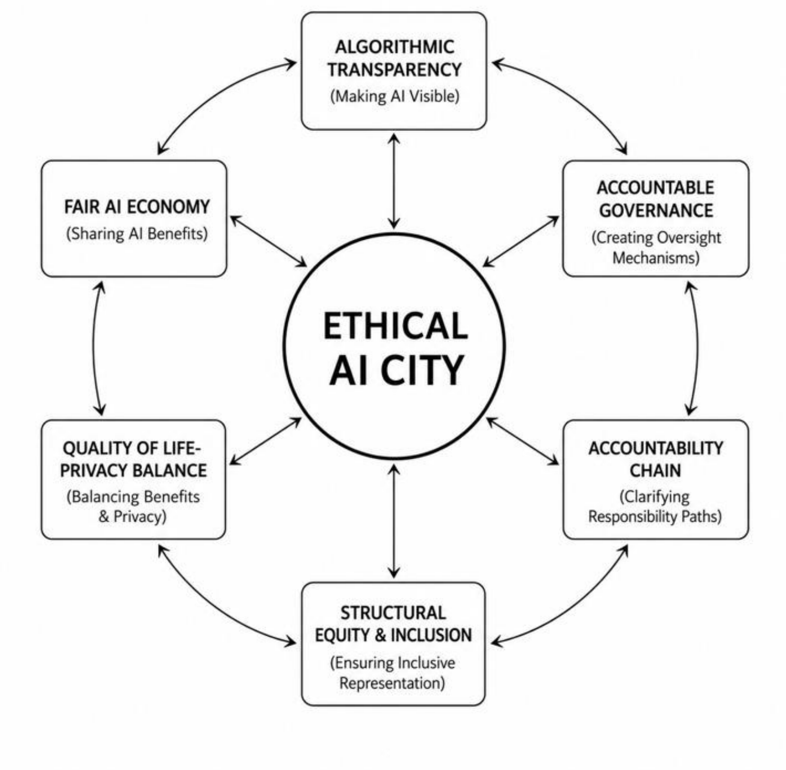

Fig. 2. The Ethical AI City Framework: six interconnected pillars

Note. The outer circular arrows represent a continuous governance feedback loop and mutual reinforcement among the six pillars rather than a fixed causal sequence.

Urban designers can apply this framework at multiple scales of practice. At the district scale as in Songdo or forthcoming Korean AI pilot cities the framework can serve as a design brief checklist: does the street environment embed transparency through accessible sensor data displays or community data dashboards? Does the proposed spatial layout enable accountable governance through legible infrastructure? Does the mobility network reflect structural equity, ensuring equal access across all socioeconomic groups? At the building and block scale, architects and designers can apply the Quality of Life-Privacy Balance pillar when specifying the placement and coverage of surveillance systems, ensuring that camera fields of view are documented and subject to public consent mechanisms. The Fair AI Economy pillar has direct implications for land-use programming: ethical AI Cities should resist single-use data-center zones that generate value without local redistribution, instead integrating productive uses that create local employment and public amenity. The Accountability Chain pillar translates into design documentation practices requiring that AI systems embedded in public infrastructure (traffic signals, waste systems, emergency response networks) are subject to auditable logs and clearly designated operator responsibility, mapped against the spatial assets they govern. In practice, urban designers working on AI City projects should insist on governance annexes to design briefs that specify: which AI systems will be deployed; who is responsible for each system; how residents can contest AI-driven decisions affecting their built environment; and how algorithmic impacts on spatial equity will be monitored over time.

5. Conclusion

The purpose of this paper is to draw attention to a growing yet underexamined reality: as cities transition from smart city models toward increasingly autonomous AI cities, the scale and complexity of ethical risks may be substantially greater than current policy discussions suggest. This study argues that integrating AI into urban systems requires an explicit urban turn in AI ethics. To address this gap, this paper proposes a six-pillar Ethical AI City Framework consisting of algorithmic transparency, accountable governance, accountability chains, structural equity and inclusion, quality-of-life–privacy balance, and a fair AI economy. The framework extends AI ethics beyond individual systems and users to the urban scale, where AI systems interact with complex socio-technical infrastructures and influence the distribution of opportunities, resources, risks, and public value.

The framework is not intended solely as a retrospective assessment tool, but as a proactive governance instrument to anticipate and mitigate emerging urban AI risks. As cities increasingly embed AI into public infrastructure, mobility systems, urban management, and service delivery, the absence of systematic ethical evaluation may expose urban systems to a progression of risks. In the short term, operational failures, service disruptions, and accountability ambiguities may emerge. Over time, these challenges may evolve into broader sociotechnical concerns, including algorithmic bias, exclusion, and declining public trust. In the longer term, increasing dependence on AI systems, concentration of decision-making power, and reduced democratic oversight may create structural vulnerabilities at the urban scale. The current stage of AI urbanization therefore represents a critical window for cities to diagnose vulnerabilities, build governance capacity, and embed ethical safeguards into the foundations of future AI-enabled urban development.

More broadly, the framework offers a governance lens through which cities can evaluate whether AI deployment aligns with democratic values, social equity, public accountability, and long-term urban resilience. Rather than treating ethics as a downstream compliance exercise, cities should integrate ethical evaluation into the planning, design, procurement, deployment, and governance of AI-enabled urban systems from the outset.

Future research should focus on operationalizing and validating the Ethical AI City Framework through multiple empirical pathways. One promising direction is to apply the framework to a single AI-enabled city and systematically evaluate how the six pillars manifest in practice across governance, public services, mobility, data management, and urban development. Such a case study would provide empirical validation and refinement of the framework while generating practical governance insights for AI-enabled cities.

As AI becomes increasingly embedded in urban systems, planners, urban designers, policymakers, and technology providers must play a more active role in shaping how AI is deployed, governed, and spatially integrated into cities. Ultimately, responsible AI should be regarded not as a supplementary consideration but as a prerequisite for sustainable, democratic, and human-centered urban development.

Author Contributions

Conceptualization, R.A.; methodology, R.A. and S.J.P.; investigation and data curation, R.A. and Y.K.; formal analysis, R.A.; writing original draft preparation, R.A.; writing review and editing, R.A., S.J.P. and Y.K.; visualization, R.A.; supervision, R.A.; project administration, R.A. All authors have read and agreed to the published version of the manuscript.

Funding

This research received no external funding.

Data Availability Statement

This study is a qualitative systematic literature review; no new primary dataset was generated. The documents synthesized are publicly available through the sources cited in the text and reference list, including the AI Incident Database (https://incidentdatabase.ai/) and the OECD AI Incidents Monitor (https://oecd.ai/en/incident-monitor). The coding framework and analytic memos developed during the review are available from the corresponding author upon reasonable request.

Conflicts of Interest

Author Sung Jin Park is Chief Executive Officer of SPARKCITY, a smart-city technology company with commercial interests in AI-enabled urban infrastructure. To avoid any perceived influence on the study’s design, case selection, or conclusions, S.J.P.’s role was limited to methodology input and manuscript review; S.J.P. was not involved in document screening, thematic coding, or the interpretive synthesis of findings, which were conducted independently by the other authors. The remaining authors declare no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgement

This research was not funded by any grant.

AI Use Statement

The authors used Grok 3 (xAI) for keyword generation, preliminary document summarization, and first-pass translation of Korean-language sources. Claude (Anthropic) provided assistance in finding Korean literature on AI and the urban planning process, and in generating the edited reference list during manuscript revision, as well as support including expansion of the Songdo case analysis and methodology sections. All AI-generated content was reviewed, verified against primary sources, and rewritten or approved by the human authors of this article, who take full responsibility for the accuracy and integrity of the work.

References

[1] Acemoglu, D., Autor, D., Hazell, J., & Restrepo, P. (2022). Artificial intelligence and jobs: Evidence from online vacancies. Journal of Labor Economics, 40(S1), S293–S340. https://doi.org/10.1086/718327

[2] Adeolu, A. (2025). AI gun detection system triggers false alarm at Baltimore County high school. CBS Baltimore. https://www.cbsnews.com/baltimore/news/false-alarm-parkville-high-school-omnilert-ai-gun-detection-system/

[3] AI Incident Database. (n.d.). AI Incident Database. https://incidentdatabase.ai/

[4] AI Now Institute. (2018). Algorithmic accountability policy toolkit. AI Now Institute. https://ainowinstitute.org/publications/algorithmic-accountability-policy-toolkit

[5] Alon-Barkat, S. and Busuioc, M. (2022) 'Human–AI interactions in public sector decision making: “Automation bias” and “Selective adherence” to algorithmic advice,' Journal of Public Administration Research and Theory, 33(1), pp. 153–169. https://doi.org/10.1093/jopart/muac007.

[6] Ananny, M., & Crawford, K. (2018). Seeing without knowing: Limitations of the transparency ideal and its application to algorithmic accountability. New Media & Society, 20(3), 973–989.

[7] Arana-Catania, M., van Lier, F.-A., Procter, R., Tkachenko, N., He, Y., Zubiaga, A., & Liakata, M. (2021). Citizen participation and machine learning for a better democracy. Digital Government: Research and Practice, 2(3), Article 27, 1–22. https://doi.org/10.1145/3452118

[8] Babšek, D., Ravšelj, D., Umek, L., & Aristovnik, A. (2025). Artificial intelligence in public administration: Uses, benefits and governance challenges. Administrative Sciences, 6(3), 44.

[9] Bhambra, G. K. (2014). Connected sociologies. Bloomsbury Academic.

[10] Barocas, S., & Selbst, A. D. (2016). Big data’s disparate impact. Harvard Law Review, 129(3), 671–732.

[11] Barns, S. (2021). Platform urbanism: Negotiating platform ecosystems in connected cities. Palgrave Macmillan.

[12] Batty, M. (2018). Artificial intelligence and smart cities. Environment and Planning B: Urban Analytics and City Science, 45(1), 3–6.

[13] Beastrom, J. (2025). The ethics of smart cities. Viterbi Conversations in Ethics. https://vce.usc.edu/volume-8-issue-3/the-ethics-of-smart-cities/

[14] Belfer Center. (2023). Municipal data sharing mandate policy hub. Harvard Kennedy School.

[15] Belgian Data Protection Authority. (2024). Data protection in smart cities: Report on the study day. [Report]. https://www.dataprotectionauthority.be/citizen/news/2024/10/09/the-report-on-the-smart-cities-study-day-is-now-available

[16] Benedikt, O. (2015). The valuable citizens of smart cities: The case of Songdo city. Global Journal of Social Sciences, 12(2), 17-36.

[17] Benjamin, R. (2019). Race after technology: Abolitionist tools for the new Jim Code. Polity.

[18] Berridge, C. (2016). Breathing room in monitored space: The impact of passive monitoring technology on privacy in independent living. The Gerontologist, 56(5), 807–816. https://doi.org/10.1093/geront/gnv067

[19] Bibri, S. E., & Krogstie, J. (2021). Data-driven smart sustainable cities of the future: An evidence synthesis approach to a comprehensive state-of-the-art literature review. Sustainable Futures, 3, 100047. https://doi.org/10.1016/j.sftr.2021.100047

[20] Bonnefon, J.-F., Shariff, A., & Rahwan, I. (2016). The social dilemma of autonomous vehicles. Science, 352(6293), 1573–1576. https://doi.org/10.1126/science.aaf2654

[21] Booth, A., Sutton, A., & Papaioannou, D. (2016). Systematic approaches to a successful literature review (2nd ed.). SAGE.

[22] Booth, R. (2024). Revealed: Bias found in AI system used to detect UK benefits fraud. The Guardian. https://www.theguardian.com/society/2024/dec/06/revealed-bias-found-in-ai-system-used-to-detect-uk-benefits

[23] Bowen, G. A. (2009). Document analysis as a qualitative research method. Qualitative Research Journal, 9(2), 27–40. https://doi.org/10.3316/QRJ0902027

[24] Brauneis, R., & Goodman, E. P. (2018). Algorithmic transparency for the smart city. Yale Journal of Law & Technology, 20, 103.

[25] Burrell, J. (2016). How the machine “thinks”: Understanding opacity in machine learning algorithms. Big Data & Society, 3(1). https://doi.org/10.1177/2053951715622512

[26] Calzada, I. (2021). The right to have digital rights in smart cities. Sustainability, 13(20), 11438. https://doi.org/10.3390/su132011438

[27] Caprotti, F. (2016). Eco-cities and the transition to low carbon economies. Palgrave Pivot.

[28] Cardullo, P., Di Feliciantonio, C., & Kitchin, R. (2019). The right to the smart city. Emerald Publishing Limited.

[29] Carvalho, L. (2015). Smart cities from scratch? A socio-technical perspective. Cambridge Journal of Regions, Economy and Society, 8(1), 43–60.
[30] CBS News. (2024, November 14). Cruise admits to false report in 2023 dragging of San Francisco pedestrian. https://www.cbsnews.com/sanfrancisco/news/cruise-automation-admits-false-report-sf-pedestrian-dragging/
[31] Center for Global Development (CGD). (2024). Three reasons why AI may widen global inequality. Center for Global Development. https://www.cgdev.org/blog/three-reasons-why-ai-may-widen-global-inequality
[32] Chicago Office of Inspector General. (2021). The Chicago Police Department's use of ShotSpotter technology. https://igchicago.org/2021/08/24/oig-finds-that-shotspotter-alerts-rarely-lead-to-evidence-of-a-gun-related-crime-and-that-presence-of-the-technology-changes-police-behavior/
[33] Choi, J., Jang, G., Kim, J., Choi, S., Sung, J., Hong, J., & Eom, S. (2025). 도시계획 분야 생성형 AI 활용사례 [Applications of generative AI in urban planning]. Urban Information Service, (525), 5-23.
[34] City of Amsterdam. (n.d.). Algorithm Register. https://algoritmeregister.amsterdam.nl/en/
[35] City of Helsinki. (n.d.). AI Register. https://ai.hel.fi/en/ai-register/
[36] Crawford, K. (2021) Atlas of AI, Yale University Press eBooks. https://doi.org/10.12987/9780300252392
[37] Cugurullo, F. (2020). Urban artificial intelligence: From automation to autonomy in the smart city. Frontiers in Sustainable Cities, 2, 38. https://doi.org/10.3389/frsc.2020.00038
[38] Cugurullo, F. (2021). Frankenstein urbanism: Eco, smart and autonomous cities, artificial intelligence and the end of the city. Routledge.
[39] 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
[40] Cunningham, T., Cormode, G., Ferhatosmanoglu, H., & Srivastava, D. (2021). Real-world trajectory sharing with local differential privacy. Proceedings of the VLDB Endowment, 14(11), 2283–2295. https://doi.org/10.14778/3476249.3476280
[41] Datta, A. (2020). The 'smart safe city': Gendered time, speed, and violence in the margins of India's urban age. Annals of the American Association of Geographers, 110(5), 1318–1334.
[42] Deakin, M., & Al Waer, H. (2011). From intelligent to smart cities. Intelligent Buildings International, 3(3), 133–139.
[43] Del Casino Jr, V. J. (2016). Social geographies II: Robots. Progress in Human Geography, 40(1), 149-158. https://doi.org/10.1177/0309132515618807
[44] Deloitte. (2023). NYC Local Law 144: Algorithmic bias audit requirements.
[45] Demiris, G., Hensel, B. K., Skubic, M., & Rantz, M. (2008). Senior residents' perceived need of and preferences for "smart home" sensor technologies. International Journal of Technology Assessment in Health Care, 24(1), 120–124. https://doi.org/10.1017/S0266462307080154
[46] Diakopoulos, N. (2016). Accountability in algorithmic decision making. Communications of the ACM, 59(2), 56–62. https://doi.org/10.1145/2844110
[47] Du, J., et al. (2023). Artificial intelligence-enabled participatory planning: A review. International Journal of Urban Sciences, 28(2), 183–210.
[48] Eckhoff, D., & Wagner, I. (2018). Privacy in the smart city: Applications, technologies, challenges, and solutions. IEEE Communications Surveys & Tutorials, 20(1), 489–516. https://doi.org/10.1109/COMST.2017.2748998
[49] Element AI & Evergreen. (2019). Data trusts and the governance of smart environments.
[50] Eubanks, V. (2018). Automating inequality: How high-tech tools profile, police, and punish the poor. St. Martin's Press.
[51] European Parliament & Council of the European Union. (2024). Regulation (EU) 2024/1689 (Artificial Intelligence Act). Official Journal of the European Union, L 2024/1689. https://eur-lex.europa.eu/eli/reg/2024/1689/oj/eng
[52] Fainstein, S. S. (2010). The just city. Cornell University Press.
[53] Feldstein, S. (2019). The global expansion of AI surveillance. Carnegie Endowment for International Peace.
[54] Fia, T. (2024). Governance of urban data commons as a matter of value redistribution. Sens Public.
[55] Frey, C. B., & Osborne, M. A. (2017). The future of employment: How susceptible are jobs to computerisation? Technological Forecasting and Social Change, 114, 254–280. https://doi.org/10.1016/j.techfore.2016.08.019
[56] Gabrys, J. (2019). How to do things with sensors. University of Minnesota Press.
[57] Gough, D., Oliver, S., & Thomas, J. (2017). An introduction to systematic reviews (2nd ed.). SAGE.
[58] Garvie, C. (2019). Garbage in, garbage out. Georgetown Law Center.
[59] Ghena, B., et al. (2014). Green lights forever: Analyzing the security of traffic infrastructure. USENIX Workshop on Offensive Technologies.

[60] Gizmodo. (2025). People are backflipping off Waymo's robotaxis. https://gizmodo.com/waymo-san-francisco-2000653614
[61] Goodman, E. P., & Powles, J. (2019). Urbanism under Google. Fordham Law Review, 88(2), 457–498.
[62] Graham, M. (2020). Regulate, replicate, and resist–The conjunctural geographies of platform urbanism. Urban Geography.
[63] Greenfield, A. (2013). Against the smart city (The city is here for you to use Book 1). Do projects.
[64] Guardian Staff. (2024). GM cruise self-driving taxi unit to pay $500,000 fine. The Guardian.
[65] Hao, K. (2020). Live facial recognition is tracking kids suspected of being criminals. MIT Technology Review. Retrieved June 10, 2026 from https://www.technologyreview.com/2020/10/09/1009992/live-facial-recognition-is-tracking-kids-suspected-of-crime/
[66] He, M.-L., and Jin, J. (2021). The Negative Aspects of the Sharing Economy: The Relationship between Airbnb and Housing Rents. Journal of Korea Planning Association, 56(6), 160-175.
[67] Hong, H. (2024, December 18). '빨간불'인데 횡단보도 건너던 배달 로봇, 승용차와 충돌 ['Red light' delivery robot collides with car while crossing crosswalk]. Yonhap News. https://www.yna.co.kr/view/AKR20241218160000065
[68] Incheon Metropolitan City. (2023). Incheon City, strengthening the safety net for caring by expanding AI safety checks for single-person households.
[69] Incheon Metropolitan City. (2026). Strengthening the collaborative framework to address social isolation through AI and ICT-based care services. Incheon Metropolitan City. https://www.incheon.go.kr/en/EN020401/3065447
[70] Jang, S.-G. (2026). From efficiency to adaptability: Comparing the characteristics of smart city and AI city concepts. Journal of Korea Planning Association, 61(2), 208–221. https://doi.org/10.17208/jkpa.2026.03.61.2.208
[71] Jin, H., Sandoval, M., & Roy, A. (2024). Crowd sets Waymo self-driving vehicle ablaze in San Francisco. Reuters. https://www.reuters.com/world/us/crowd-sets-waymo-self-driving-car-ablaze-san-francisco-2024-02-11/
[72] Kitchin, R. (2014). The real-time city? Big data and smart urbanism. GeoJournal, 79(1), 1–14.
[73] Kitchin, R. (2016). The ethics of smart cities and urban science. Philosophical Transactions of the Royal Society A, 374(2083).
[74] Kitchin, R., Cardullo, P., & Di Feliciantonio, C. (2019). Citizenship, justice, and the right to the smart city. In P. Cardullo, C. Di Feliciantonio, & R. Kitchin (Eds.), The right to the smart city (pp. 1–24). Emerald Publishing Limited.
[75] Kitchin, R., & Dodge, M. (2019). The (in)security of smart cities: Vulnerabilities, risks, mitigation, and prevention. Journal of Urban Technology, 26(2), 47–65.
[76] Korea Agency for Infrastructure Technology Advancement (KAIA). (n.d.). KAIA official website. Retrieved March 21, 2026, from https://www.kaia.re.kr
[77] Korea Policy Briefing. (2025). 세계 3대 AI 강국을 향해 [Toward the World's Top 3 AI Powerhouses][Press release]. https://www.korea.kr
[78] Kuffer, M., Wang, J., Thomson, D. R., Georganos, S., Abascal, A., Owusu, M., & Vanhuysse, S. (2021). Spatial Information Gaps on Deprived Urban Areas (Slums) in Low-and-Middle-Income-Countries: A User-Centered Approach. Urban Science, 5(4), 72. https://doi.org/10.3390/urbansci5040072
[79] Latour, B. (2004). Politics of nature: How to bring the sciences into democracy. Harvard University Press.
[80] Lecher, C. (2018). What happens when an algorithm cuts your health care. The Verge. https://www.theverge.com/2018/3/21/17144260/healthcare-medicaid-algorithm-arkansas-cerebral-palsy
[81] Lee, S. K., Kwon, H. R., Cho, H., Kim, J., & Lee, D. (2016). International Case Studies of Smart Cities: Anyang, Republic of Korea. https://doi.org/10.18235/0007013
[82] Lefebvre, H. (1968). Le droit à la ville. Anthropos.
[83] Ienca, M., Fabrice, J., Elger, B., Caon, M., Pappagallo, A. S., Kressig, R. W., Wangmo, T., & Benini, L. (2018). Ethical design of intelligent assistive technologies for dementia: A descriptive review. Science and Engineering Ethics, 24(4), 1035–1055. https://doi.org/10.1007/s11948-017-9976-1
[84] Lincoln, Y. S., & Guba, E. G. (1985). Naturalistic inquiry. SAGE.
[85] Lim, Y., Edelenbos, J., & Gianoli, A. (2023). Dynamics in the governance of smart cities: Insights from South Korean smart cities. International Development Planning Review, 45(2), 183–205. https://doi.org/10.1080/12265934.2022.2063158
[86] Lighthouse Reports. (2023). Suspicion machines: Unprecedented experiment on welfare surveillance algorithm reveals discrimination. https://www.lighthousereports.com/investigation/suspicion-machines/

[87] Lincoln, Y.S., Guba, E.G. and Pilotta, J.J. (1985b) 'Naturalistic inquiry,' International Journal of Intercultural Relations, 9(4), pp. 438–439. https://doi.org/10.1016/0147-1767(85)90062-8
[88] Lim, Y., Edelenbos, J., & Gianoli, A. (2022). Dynamics in the governance of smart cities: insights from South Korean smart cities. International Journal of Urban Sciences, 26(4). https://doi.org/10.1080/12265934.2022.2063158
[89] Logan, J. R., & Molotch, H. L. (1987). Urban fortunes: The political economy of place. University of California Press.
[90] Luusua, A., Ylipulli, J., Foth, M., & Aurigi, A. (2023). Urban AI: understanding the emerging role of artificial intelligence in smart cities. AI & Society, 38(3), 1039–1044. https://doi.org/10.1007/s00146-022-01537-5
[91] Little, R. G. (2002). Controlling cascading failure: Understanding the vulnerabilities of interconnected infrastructures. Journal of Urban Technology, 9(1), 109–123.
[92] MacArthur Justice Center. (2021). ShotSpotter generated over 40,000 dead-end police deployments in Chicago.
[93] Marx, J. (2020). Smart streetlights are now exclusively a tool for police. Voice of San Diego.
[94] Mattern, S. (2021). A city is not a computer: Other urban intelligences. Princeton University Press.
[95] Ministry of Land, Infrastructure and Transport (MOLIT). (2025). 도시공간에 AI를 폭넓게 적용하는 'AI 시티' 본격 추진 (보도자료) [Official Rollout of 'AI City' to Broadly Integrate AI into Urban Environments] (Press release). https://www.molit.go.kr
[96] Ministry of Land, Infrastructure and Transport (MOLIT). (2026). AI가 시민 일상으로 빠르게, 우리 도시에서 시작됩니다 [AI spreads quickly into citizens' daily lives, starting in our cities] (Press release). https://www.molit.go.kr
[97] Ministry of Science and ICT (MSIT). (2020). 사람이 중심이 되는 인공지능(AI) 윤리기준 [National guidelines for AI ethics] (Press release). Korea Information Society Development Institute. https://ai.kisdi.re.kr
[98] Mignolo, W. D. (2011). The darker side of Western modernity: Global futures, decolonial options. Duke University Press.
[99] Mitieka, D., Luke, R., Twinomurinzi, H., & Mageto, J. (2023). Smart Mobility in Urban Areas: A Bibliometric Review and Research Agenda. Sustainability, 15(8), 6754. https://doi.org/10.3390/su15086754 (Highlighted: Replaced placeholder)
[100] Mistry, I., Tanwar, S., Tyagi, S., & Kumar, N. (2020). Blockchain for 5G-enabled IoT for industrial automation: A systematic review, solutions, and challenges. Mechanical Systems and Signal Processing, 135, 106382.
[101] Molotch, H. (1976). The city as a growth machine: Toward a political economy of place. American Journal of Sociology, 82(2), 309–332.
[102] Morozov, E., & Bria, F. (2018). Rethinking the smart city: Democratizing urban technologies. Rosa Luxemburg Stiftung.
[103] Mohamed, S., Png, M. T., & Isaac, W. (2020). Decolonial AI: Decolonial theory as sociotechnical foresight in artificial intelligence. Philosophy & Technology, 33(4), 659–684. https://doi.org/10.1007/s13347-020-00405-8
[104] Muldoon, J. (2022). Platform socialism: How to reclaim our digital future from Silicon Valley. Pluto Press.
[105] National Fair Housing Alliance. (2024). From policy to practice: Essential competencies for federal Chief AI Officers.
[106] New York City Council. (2021). Local Law 144 of 2021: Automated employment decision tools.
[107] New York Times. (2021). 'Dangerous stuff': Hackers tried to poison water supply of Florida town.
[108] Ngata, W. M., Bashir, N., Westerlaken, M., Liote, L., Chandio, Y., & Olivetti, E. (2025). The cloud next door: Investigating the environmental and socioeconomic strain of datacenters on local communities. Proceedings of the ACM SIGCAS/SIGCHI Conference on Computing and Sustainable Societies (COMPASS '25), 769–774. https://doi.org/10.1145/3715335.3736324
[109] Nixon, R. (2011). Slow violence and the environmentalism of the poor. Harvard University Press.
[110] NJCM et al. v. The State of the Netherlands, ECLI:NL:RBDHA:2020:1878 (District Court of The Hague, February 5, 2020). https://uitspraken.rechtspraak.nl/details?id=ECLI:NL:RBDHA:2020:1878
[111] National Transportation Safety Board. (2018). Collision between a pedestrian and an Uber vehicle operated in automated driving mode, Tempe, Arizona, March 18, 2018 (Report No. NTSB/HAR-19/03). https://www.ntsb.gov/investigations/accidentreports/reports/har1903.pdf
[112] National Transportation Safety Board. (2019). Collision between a sport utility vehicle operating with partial driving automation and a crash attenuator, Mountain View, California, March 23, 2018 (Report No. NTSB/HAR-20/01). https://www.ntsb.gov/investigations/accidentreports/reports/har2001.pdf
[113] Organisation for Economic Co-operation and Development [OECD]. (2019). OECD principles on artificial intelligence. OECD. https://oecd.ai/en/ai-principles
[114] OECD. (n.d.). AI Incidents Monitor. https://oecd.ai/en/incident-monitor

[115] Park, S. (2023). Smart Cities: striking a balance between urban resilience and civil liberties. arXiv. https://doi.org/10.48550/arxiv.2303.14597
[116] Pasquale, F. (2015). The black box society: The secret algorithms that control money and information. Harvard University Press.
[117] Pasquale, F. (2020). New laws of robotics: Defending human expertise in the age of AI. Harvard University Press.
[118] Paiva, S., et al. (2021). Enabling technologies for urban smart mobility. Sensors, 21(6), 2143.
[119] Page, M. J., McKenzie, J. E., Bossuyt, P. M., Boutron, I., Hoffmann, T. C., Mulrow, C. D., Shamseer, L., Tetzlaff, J. M., Akl, E. A., Brennan, S. E., Chou, R., Glanville, J., Grimshaw, J. M., Hróbjartsson, A., Lalu, M. M., Li, T., Loder, E. W., Mayo-Wilson, E., McDonald, S., … Moher, D. (2021). The PRISMA 2020 statement: An updated guideline for reporting systematic reviews. BMJ, 372, n71. https://doi.org/10.1136/bmj.n71
[120] Perlroth, N. (2021, February 8). Water plant hack shows weaknesses in U.S. infrastructure. The New York Times. https://www.nytimes.com/2021/02/08/us/oldsmar-florida-water-plant-hack.html
[121] Quijano, A. (2000). Coloniality of power, Eurocentrism, and Latin America (M. Ennis, Trans.). Nepantla: Views from South, 1(3), 533–580.
[122] Rahwan, I. (2018). Society-in-the-loop: Programming the algorithmic social contract. Nature Human Behaviour, 2(9), 634–636. https://doi.org/10.1038/s41562-018-0381-0
[123] Rawls, J. (1971). A theory of justice. Harvard University Press.
[124] Republic of Korea. (2025). Framework Act on the Development of Artificial Intelligence and Establishment of a Foundation for Trust (Act No. 20676). Korean Law Information Center. Korean Law Information Center
[125] Robinson, C., & Franklin, R. S. (2020). The sensor desert quandary: What does it mean (not) to count in the smart city? Transactions of the Institute of British Geographers, 46(1), 238–254. https://doi.org/10.1111/tran.12415
[126] Roy, A., & Jin, H. (2024). San Francisco Waymo arson sparks fresh debate on self-driving cars. Reuters. https://www.reuters.com/business/autos-transportation/san-francisco-waymo-arson-sparks-fresh-debate-self-driving-cars-2024-02-13/
[127] Rugkhapan, N. T., & Murray, M. J. (2019). Songdo IBD (International Business District): experimental prototype for the city of tomorrow? International Planning Studies, 24(3–4), 272–292. https://doi.org/10.1080/13563475.2019.1650725
[128] Rudin, C. (2019). Stop explaining black box machine learning models. Nature Machine Intelligence, 1(5), 206–215.
[129] Safransky, S. (2020). Geographies of algorithmic violence: Redlining the smart city. International Journal of Urban and Regional Research, 44(2), 200–218.
[130] Sanchez-Gallegos, D. D., Carrizales-Espinoza, D. E., Torres-Charles, C., & Carretero, J. (2025). Smart Cities: A Systematic Review of Emerging Technologies. Smart Cities, 8(5), 173. https://doi.org/10.3390/smartcities8050173
[131] Scholz, T. (2016). Platform cooperativism: Challenging the corporate sharing economy. Rosa Luxemburg Stiftung.
[132] Shaw, J., & Graham, M. (2017). An informational right to the city? Code, content, control, and the urbanization of information. Antipode, 49(4), 907–927. https://doi.org/10.1111/anti.12312
[133] Shin, H., Park, S. H., & Sonn, J. W. (2015). The emergence of a multiscalar growth regime and scalar tension: the politics of urban development in Songdo New City, South Korea. Environment and Planning C: Government and Policy, 33(6), 1618–1638. https://doi.org/10.1177/0263774X15614442
[134] Shwayri, S. T. (2013). A Model Korean Ubiquitous Eco-City? The Politics of Making Songdo. Journal of Urban Technology, 20(1), 39–55. https://doi.org/10.1080/10630732.2012.735409
[135] Smith, S., Barlow, G., Xie, X.-F., & Rubinstein, Z. (2013). Smart urban signal networks: Initial application of the SURTRAC adaptive traffic signal control system. Proceedings of the International Conference on Automated Planning and Scheduling, 23(1), 434–442. https://doi.org/10.1609/icaps.v23i1.13594
[136] Sovacool, B. K., et al. (2022). Energy justice and digital transitions. Nature Energy, 7, 1–10.
[137] Sonn, J. W., Shin, H., & Park, S. H. (2017). A mega urban project and two competing accumulation strategies: Negotiating discourses of the Songdo International Business District development. International Development Planning Review, 39(3), 299–317. https://doi.org/10.3828/idpr.2017.7
[138] Spielman, F. (2024, September 17). Chicago to stop using controversial gunshot detection technology. Chicago Sun-Times.
[139] Söderström, O., Paasche, T., & Klauser, F. (2014). Smart cities as corporate storytelling. City, 18(3), 307-320. https://doi.org/10.1080/13604813.2014.906716
[140] Stilgoe, J. (2018). Machine learning, social learning and the governance of self-driving cars. Social Studies of Science, 48(1), 25-56. https://doi.org/10.1177/0306312717741687
[141] Stilgoe, J. (2020). Who's driving innovation? New technologies and the collaborative state. Palgrave Macmillan.

[142] The Guardian. (2025, October 28). More than a million people every week show suicidal intent when chatting with ChatGPT. https://www.theguardian.com/technology/2025/oct/27/chatgpt-suicide-self-harm-openai
[143] Thomas, J., & Harden, A. (2008). Methods for the thematic synthesis of qualitative research in systematic reviews. BMC Medical Research Methodology, 8(1), Article 45. https://doi.org/10.1186/1471-2288-8-45
[144] Treasury Board of Canada Secretariat. (2019). Directive on automated decision-making. Government of Canada. https://www.tbs-sct.canada.ca/pol/doc-eng.aspx?id=32592
[145] United Nations Development Programme (UNDP). (2025). The next great divergence: Why artificial intelligence may widen inequality between countries. UNDP Regional Bureau for Asia and the Pacific.
[146] United States Department of Justice. (2024). Cruise admits submitting false report to influence federal investigation. US DOJ Press Release.
[147] US Department of Housing and Urban Development. (2024). AI in housing: Opportunities and challenges.
[148] van Bekkum, M., & Zuiderveen Borgesius, F. J. (2021). Digital welfare fraud detection and the Dutch SyRI judgment. European Journal of Social Security, 23(4), 323–340. https://doi.org/10.1177/13882627211031257
[149] Vanderbilt Policy Accelerator. (2024). The national security case for a public AI option.
[150] Veale, M., & Brass, I. (2019). Administration by algorithm? Public Administration Review, 79(6), 1–12.
[151] Verbeek, P.-P. (2011). Moralizing technology: Understanding and designing the morality of things. University of Chicago Press.
[152] Wachter, S., Mittelstadt, B., & Floridi, L. (2017). Why a right to explanation of automated decision-making does not exist in the General Data Protection Regulation. International Data Privacy Law, 7(2), 76–99. https://doi.org/10.1093/idpl/ipx005
[153] Wachter, S., Mittelstadt, B., & Russell, C. (2021). Bias preservation in machine learning: The legality of fairness metrics under EU non-discrimination law. West Virginia Law Review, 123(3), 735–790. https://researchrepository.wvu.edu/wvlr/vol123/iss3/4
[154] Winfield, A. F. T., Winkle, K., Webb, H., Lyngs, U., Jirotka, M., & Macrae, C. (2020). Robot accident investigation: A case study in responsible robotics. arXiv. https://doi.org/10.48550/arXiv.2005.07474
[155] WIRED. (2019). The Apple Card didn't 'see' gender and that's the problem.
[156] Xu, J., Wang, J., Leung, J., & Gu, J. (2025). GRASP: Municipal budget AI chatbots for enhancing civic engagement. arXiv. https://arxiv.org/abs/2503.23299
[157] Yang, H. J., & Park, H. J. (2024). Theoretical and practical reflections on smart urbanism: Promoting spatial inclusiveness with smart technologies. Journal of Korea Planning Association, 59(6), 55–65. https://doi.org/10.17208/jkpa.2024.11.59.6.55
[158] Yigitcanlar, T., & Lee, S. H. (2014). Korean ubiquitous-eco-city: A smart-sustainable urban form or a branding hoax? Technological Forecasting and Social Change, 89, 100–114. https://doi.org/10.1016/j.techfore.2013.08.034
[159] Ylipulli, J., Kangasvuo, J., Alavesa, P., & Ojala, T. (2023). Chasing digital shadows: Exploring future hybrid cities through anthropological design fiction. Personal and Ubiquitous Computing, 27(3), 1213-1230.
[160] Yoo, S. (2017). Songdo: The hype and decline of world's first smart city. In F. Caprotti & L. Yu (Eds.), Sustainable cities in Asia (pp. 146–160). Routledge.
[161] Zaffagnini, M. (2025). Civil liability in the age of smart cities: New legal challenges. HAL Open Science. https://hal.science/hal-05063799v1
[162] Zhang, Z., Wu, R., Ren, F., Turaga, S., & Mendel, T. (2026). CoDesignAI: An AI-enabled multi-agent, multi-user system for collaborative urban design at the conceptual stage. arXiv. https://doi.org/10.48550/arXiv.2603.16008
[163] Zuboff, S. (2015). Big other: Surveillance capitalism and the prospects of an information civilization. Journal of Information Technology, 30(1), 75–89.
[164] Zuboff, S. (2019). The age of surveillance capitalism: The fight for a human future at the new frontier of power. Public Affairs.