

From Compliance to Contextual Legitimacy: A Sociotechnical Framework for Evaluating AI Systems

Giuditta Bellosi¹, Roberta Calegari¹

¹University of Bologna
giuditta.bellosi2@unibo.it, roberta.calegari@unibo.it

Abstract

AI systems are increasingly recognized as sociotechnical systems whose impacts emerge not only from algorithmic behavior, but also from the contexts in which they operate. While Trustworthy AI (TAI) provides a foundational framework for AI governance, several sociotechnical legitimacy conditions remain difficult to systematically evaluate within predominantly artifact- and compliance-oriented approaches. This paper proposes a relational framework for evaluating *contextual legitimacy* in AI-enabled sociotechnical systems as a complement to conventional TAI-oriented evaluations. The framework identifies and operationalizes sociotechnical legitimacy conditions that remain fragmented, implicit, or weakly addressed within existing trustworthy AI approaches, and empirically applies them to TAI-oriented Urban Digital Twin literature. The findings suggest that complementing TAI with a relational sociotechnical perspective is necessary for more comprehensively evaluating the legitimacy of contemporary AI deployments.

Introduction

AI systems are increasingly embedded within complex sociotechnical environments in which technical infrastructures, institutions, citizens, and the environment continuously interact (Kudina and Van De Poel 2024). In these contexts, evaluating AI systems predominantly through compliance-oriented and artifact-centered interpretations of requirements such as robustness, transparency, fairness, or explainability is often insufficient for understanding their broader societal implications (Selbst et al. 2019). A system may satisfy formal Trustworthy AI (TAI) requirements while simultaneously reinforcing technocratic governance (Morozov 2013), marginalizing communities (Eubanks 2018), concentrating infrastructural power (Sadowski and Bendor 2019), or generating unsustainable environmental and economic costs (Strubell, Ganesh, and McCallum 2019). For example, urban prediction systems relying on smartphone or digital payment data may systematically exclude populations lacking access to these infrastructures, rendering specific communities effectively invisible within computational representations (Van Dijk 2020). These tensions are not properties of the algorithmic artifact alone,

but emerge from the relations between technology, society, and the environment within which AI systems operate. This raises the need for evaluating what we define as *contextual legitimacy*: following organizational legitimacy theory (Suchman 1995), the socially constructed perception that an AI-enabled system’s actions are appropriate within a given normative, institutional, and material context, negotiated relationally rather than assessed solely against fixed technical criteria. Unlike artifact-centered and compliance-oriented evaluations, contextual legitimacy requires examining how AI systems redistribute power, shape participation, represent social reality, and generate systemic costs across sociotechnical environments. Ethical Guidelines of Trustworthy AI (EGTAI) (European Commission High-Level Expert Group on AI 2019) provides a foundational framework for AI governance through requirements such as robustness, transparency, fairness, accountability, and human oversight. While these principles have significantly shaped AI governance, several legitimacy conditions remain difficult to systematically evaluate within predominantly technical and compliance-oriented interpretations of TAI requirements (Mittelstadt 2019; Selbst et al. 2019; Kudina and Van De Poel 2024). Existing frameworks primarily evaluate ethical properties at the level of the artifact, often overlooking broader sociotechnical conditions — governance arrangements, participation asymmetries, representational biases, uneven distributions of costs and benefits — that human-centered AI perspectives (Shneiderman 2020) similarly foreground by situating systems within the lived contexts of the people they affect. This paper addresses this gap by proposing a sociotechnical framework for evaluating contextual legitimacy in AI-enabled sociotechnical systems. Grounded in sociotechnical systems theory (Emery and Trist 1960) and critical AI studies (Eubanks 2018; Kudina and Van De Poel 2024), the framework models AI-enabled systems through the reciprocal relations between technology, society, and environment. From these relations emerge four analytical lenses — governance, relational, representational, and contextual — operationalized through 19 sociotechnical dimensions capturing recurring legitimacy-related mechanisms identified across critical AI and sociotechnical literature. The framework is not intended to replace TAI requirements, but to complement them by making visible sociotechnical legitimacy conditions that remain insufficiently fore-

grounded within dominant TAI evaluations. To empirically assess the relevance of this perspective, we conduct a systematic literature review (SLR) of TAI-oriented Urban Digital Twin (UDT) research. UDTs were selected as an analytically relevant domain because they represent paradigmatic AI-enabled sociotechnical systems in which algorithmic infrastructures, governance processes, data practices, and citizens are deeply entangled (Nochta et al. 2021; Ruohomäki et al. 2024). Applying the proposed framework to the SLR corpus, we show that many sociotechnical legitimacy conditions remain underexamined within existing TAI-oriented evaluations.

Methodological Approach

This study adopts a three-stage methodological approach combining conceptual framework construction with exploratory empirical validation. The first two phases follow an interpretive theory-building approach grounded in sociotechnical systems theory and critical AI studies, aimed at identifying and operationalizing legitimacy-related dimensions of AI-enabled sociotechnical systems. The third phase empirically applies the resulting framework through a structured systematic literature review in order to assess how these sociotechnical legitimacy conditions remain insufficiently foregrounded within dominant TAI evaluations.

Phase 1. RQ1: *How can the contextual legitimacy of AI-enabled sociotechnical systems be systematically evaluated?* To address RQ1, the first phase focused on constructing a sociotechnical framework for evaluating contextual legitimacy in AI-enabled sociotechnical systems, grounded in sociotechnical systems theory and critical AI studies. AI-enabled sociotechnical systems were modeled as reciprocal relations between *technology*, *society*, and *environment*. From this relational model, four analytical lenses were identified through which contextual legitimacy can be examined: governance, relational, representational, and contextual (Figure 1). Framework construction followed an iterative interpretive process in which conceptual modeling and literature analysis mutually informed one another. The relational model guided the thematic synthesis of critical AI and sociotechnical literature, through which recurrent legitimacy-related mechanisms were identified and organized around four core relational concerns — governance, participation, representation, and systemic conditions — providing the theoretical grounding for the proposed analytical lenses. Each analytical lens was subsequently operationalized through a set of sociotechnical dimensions capturing analytically distinct legitimacy-related mechanisms recurrently identified across the literature. This process resulted in the evaluation framework presented in the following section, consisting of four analytical lenses and 19 sociotechnical dimensions.

Phase 2. RQ2: *How does the proposed sociotechnical framework complement TAI-oriented evaluations in identifying sociotechnical legitimacy conditions within AI-enabled systems?* The objective of this phase is to examine how the sociotechnical dimensions identified in the proposed framework intersect with and extend existing TAI-oriented evaluations. To address this, we developed a structured analyt-

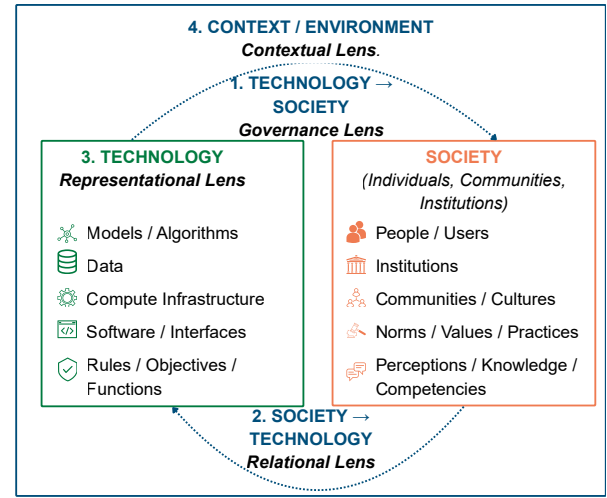


Figure 1: Sociotechnical Framework of AI-enabled sociotechnical systems and the four analytical lenses emerging from relations between technology, society, and environment.

ical mapping between the seven TAI pillars and the 19 sociotechnical dimensions composing the proposed sociotechnical framework. The mapping was developed through a systematic close reading of the EGTAI. Each paragraph associated with the seven TAI pillars was individually analyzed to identify which sociotechnical dimensions emerged as necessary conditions for the substantive operationalization of the corresponding TAI objectives within AI-enabled sociotechnical systems.

Phase 3. RQ3: *How are the sociotechnical dimensions identified in the sociotechnical framework engaged within TAI-oriented literature on a paradigmatic AI-enabled sociotechnical system?* This phase focused on empirically applying the proposed framework through a systematic literature review of Urban Digital Twins (UDTs), selected as a paradigmatic AI-enabled sociotechnical system. The systematic literature review analyzed a corpus of 80 papers. Each paper was systematically analyzed through the proposed sociotechnical framework in order to evaluate the extent to which the identified sociotechnical dimensions were explicitly engaged within TAI-oriented UDT literature. For each sociotechnical dimension, the analysis recorded the level of engagement according to four coding categories: *ignored*, *handled uncritically*, *handled critically*, and *operationalized*. The resulting scores were subsequently aggregated at both dimension and analytical lens levels to identify recurring patterns, systematic blind spots, and forms of critical engagement across the reviewed literature.

Sociotechnical Framework for Contextual Legitimacy Analysis

Phase 1 resulted in the sociotechnical framework for evaluating contextual legitimacy presented in this section. In AI-enabled sociotechnical systems, three interdependent entities can be identified: technology, society, and environment.

Technology refers to the AI-enabled system itself; society refers to the individuals, communities, and institutions interacting with it; and environment refers to the broader political, economic, cultural, and material conditions within which these interactions occur. Within this model, AI systems are understood not as isolated technical artifacts, but as systems shaped through reciprocal relations between technological infrastructures and social actors, both of which are continuously influenced by their surrounding environment (Figure 1). From these relations emerge four analytical lenses through which the contextual legitimacy of AI-enabled sociotechnical systems can be evaluated.

Governance lens. The first analytical lens examines how AI-enabled technologies reshape power, authority, and control in sociotechnical environments. It captures dimensions related to governance structures, technocracy, surveillance, marginalization of communities, corporate influence, and legitimacy (Figure 1, relation 1).

Relational lens. The second analytical lens examines how individuals and communities perceive, engage with, and participate in AI-enabled systems. It captures dimensions related to participation, technical literacy, human reliance on algorithmic systems, and the construction of algorithmic authority (Figure 1, relation 2).

Representational lens. The third analytical lens examines how reality is represented within AI systems through data, models, and computational abstractions. It captures dimensions related to visibility, intersectionality, and biased data practices (Figure 1, relation 3).

Contextual lens. The fourth analytical lens examines the broader contextual conditions within which sociotechnical systems operate, including temporal, economic, and environmental constraints and costs. It captures dimensions related to sustainability, resource distribution, and systemic impacts (Figure 1, relation 4).

Each analytical lens is operationalized through a set of sociotechnical dimensions capturing recurring legitimacy-related mechanisms identified across critical AI and sociotechnical literature (Orlikowski and Gash 1994; Orlikowski 2007; D'Ignazio and Klein 2023; Crawford 2025; Strubell, Ganesh, and McCallum 2019). These dimensions make the analytical lenses empirically observable within real-world AI deployments. We identified 19 sociotechnical dimensions summarized in Table 1. While analytically distinct, these dimensions are empirically entangled and frequently co-occur within AI-enabled sociotechnical systems, forming a unified evaluative framework for examining contextual legitimacy. Existing frameworks have addressed related concerns from complementary angles: Selbst et al. (2019) identify the abstraction traps through which AI systems fail when decontextualized from their sociotechnical environment, while (Virk 2022) proposes a taxonomy of algorithmic harms as system outputs. The proposed framework extends both contributions by shifting the analytical focus from failure modes and output harms to the relational conditions that generate them, and by explicitly grounding each dimension within the TAI evaluative structure — enabling systematic assessment of contextual legitimacy rather than post-hoc identification of its absence. Furthermore, the

framework is intentionally open: the 19 dimensions represent a theoretically grounded but non-exhaustive starting point, derived inductively from critical AI and sociotechnical literature, and intended as a basis for future extension as new legitimacy-related mechanisms are identified across emerging AI deployment contexts.

Analytical Dimensions of Contextual Legitimacy

Governance Lens The governance lens examines how AI-enabled systems configure, redistribute, and legitimize power relations within sociotechnical environments. It addresses the following analytical question:

How do AI-enabled systems redistribute power, authority, and control within society?

Seven dimensions operationalize this lens. These dimensions emerge from more specific analytical sub-questions concerning how power is structured, exercised, and legitimized across sociotechnical systems.

Who holds power? Within AI-enabled sociotechnical systems, power can be concentrated both institutionally and economically. At the institutional level, *Technocracy* (D1) captures the displacement of political judgment and democratic deliberation by technical rationality and data-driven governance (Morozov 2013; Malleson et al. 2024). Closely related, *Corporate and neoliberalist vision* (D2) examines how private technological actors and market-oriented governance logics increasingly shape public infrastructures and political decision-making, often generating forms of vendor lock-in and infrastructural dependency (O'Neil 2017; Pasquale 2015; Sadowski and Bendor 2019).

How is this power exercised? Following sociotechnical approaches that distinguish multiple scales of social organization, power within AI-enabled sociotechnical systems can be analyzed across three interconnected levels: the micro level, concerning individuals and their interaction with technological systems; the meso level, concerning communities and collective social groups; and the macro level, concerning global geopolitical and infrastructural relations (Virk 2022). Each level makes visible different legitimacy-related dimensions. At the micro level, power can shape both imaginaries and individual behavior. *Imposition of dreams* (D3) concerns the capacity of technological systems to shape collective imaginaries, aspirations, and socially desirable futures (Wang 2021; Bina, Inch, and Pereira 2020), while *Surveillance* (D4) captures the expansion of dataveillance, algorithmic monitoring, and behavioral control enabled by pervasive digital infrastructures (Wang 2021; Krivý 2018). At the meso level, power can operate through the uneven representation and visibility of social groups within sociotechnical infrastructures. *Marginalization of Communities* (D5) concerns how data-driven governance and AI-enabled infrastructures can systematically disadvantage specific social groups, not as incidental failures, but as structural consequences of how power is exercised through technological systems (Eubanks 2018; Noble 2018). Finally, at the macro level, *Digital colonialism* (D6) concerns asymmetric technological dependencies through which external technological

Analytical Lens	ID	Sociotechnical Dimension	Human agency & oversight	Robustness & safety	Privacy & data governance	Transparency	Diversity and fairness	Societal & environmental well-being	Accountability
L1 – Governance Power, Legitimacy, and Control	D1	Technocracy	•			•	•	•	•
	D2	Corporate and Neoliberalist Vision		•	•			•	•
	D3	Imposition of Dreams	•	•				•	
	D4	Surveillance	•	•	•				
	D5	Marginalization of Communities					•	•	•
	D6	Digital Colonialism			•			•	
	D7	Digital Universalism	•						
L2 – Relational Human Reliance and Participation	D8	Technical Positivism	•						
	D9	Oracle Effect	•	•		•			•
	D10	Skill Gap in Citizens	•		•	•	•		
	D11	Lack of Participation	•			•	•		•
	D12	Lack of Evaluation	•	•		•	•	•	•
L3 – Representational Data Justice and Representation	D13	Lack of Intersectionality			•		•	•	
	D14	Invisible People			•		•		
	D15	Biased Data Practices			•		•		
	D16	Citizen datafication					•		
L4 – Contextual Costs and Sustainability	D17	Time Impact							
	D18	Economic Impact							
	D19	Sustainability		•				•	

Table 1: Sociotechnical dimensions operationalizing the four analytical lenses of the proposed sociotechnical evaluation framework.

infrastructures, governance models, and epistemic frameworks are imposed across different geopolitical and socio-economic contexts (Kwet 2019; Couldry and Mejias 2019).

What assumptions legitimize these power relations? Power relations within AI-enabled sociotechnical systems are also sustained by epistemic assumptions concerning what aspects of social reality can be represented, measured, and therefore governed through technology. *Digital universalism* (D7) examines the assumption that social reality can be fully represented through digital abstraction, legitimizing universal technological models while marginalizing situated and contextual forms of knowledge.

These dimensions evaluate how AI-enabled systems configure sociotechnical power relations and shape the governance conditions under which contextual legitimacy is constructed and contested.

Relational Lens This examines the relational conditions through which individuals and communities engage with, interpret, and exercise agency toward AI-enabled systems. It addresses the following analytical question:

How do people relate to AI-enabled systems, and to what extent are they able to meaningfully understand,

question, and participate in them?

Five dimensions operationalize this lens. These dimensions emerge from analytical sub-questions concerning how technological authority is perceived, how engagement with AI-enabled systems is structured, and which conditions enable or constrain meaningful participation and accountability.

How are AI-enabled systems perceived? This dimension examines how authority and legitimacy are attributed to AI-enabled systems. Two dimensions are analytically relevant. *Technical positivism* (D8) captures the assumption that data-driven systems are inherently neutral and objective, naturalizing their authority and limiting critical scrutiny (Pasquale 2015; Eubanks 2018; Jasanoff and Kim 2019). The *Oracle effect* (D9) describes the tendency to treat AI-generated outputs as authoritative and unquestionable, delegating epistemic judgment to automated processes (Selbst et al. 2019; O’Neil 2017).

How do people relate to AI-enabled systems? The relation between individuals and AI-enabled systems can be examined across both individual and collective levels, capturing how agency, participation, and the capacity to contest technological authority are unevenly distributed within STS.

At the individual level, *Skill Gap in Citizens* (D10) concerns inequalities in technical literacy that limit the capacity of certain populations to understand, navigate, contest, and meaningfully engage with the systems that govern them (Van Dijk 2020). At the collective level, *Lack of participation* (D11) captures the exclusion of affected communities from the design, deployment, and evaluation of AI-enabled systems, reproducing structural asymmetries between those who govern technological systems and those subjected to them (Selbst et al. 2019; Mittelstadt 2019).

What structural conditions constrain accountability?

The perceptual and relational asymmetries described above are reinforced by a structural condition operating across both individual and collective levels. *Lack of evaluation* (D12) concerns the absence of mechanisms and indicators capable of assessing the broader social impacts of AI systems beyond technical performance metrics (Mittelstadt 2019; Mehrotra et al. 2026). Without robust frameworks for social evaluation, these asymmetries remain difficult to identify, contest, and govern. This dimension therefore helps explain why relational asymmetries often remain unaddressed despite formal accountability mechanisms.

These dimensions evaluate whether AI-enabled systems enable meaningful participation and informed engagement, or instead reinforce asymmetries of agency and dependence.

Representational Lens The representational lens examines how social reality is represented within AI-enabled systems through data, models, and computational abstractions. It addresses the following analytical question:

Whose realities become visible, represented, and actionable within AI-enabled systems, and whose are excluded or distorted?

Four dimensions operationalize this lens through three representational conditions: omission, bias, and datafication without agency.

Representational omission. Representational omission concerns the incomplete or absent representation of individuals within AI-enabled systems. This can operate both as partial omission and complete exclusion. *Lack of intersectionality* (D13) captures the omission of overlapping dimensions of identity and social experience within computational representations (Crenshaw 1989; Benjamin 2023). *Invisible people* (D14) concerns the exclusion of individuals and communities that fall outside dominant data collection infrastructures (Crawford 2025; D'Ignazio and Klein 2023).

Representational bias. This concerns the incorporation and reproduction of social and institutional biases within computational representations. *Biased data practices* (D15) examine how data collection, aggregation, labeling, and evaluation processes reproduce existing inequalities and institutional assumptions within AI-enabled systems (Barocas and Selbst 2016; D'Ignazio and Klein 2023).

Datafication without agency. This concerns the reduction of individuals to data profiles and computational abstractions within AI-enabled systems. As data subjects, individu-

als are reduced to computational profiles detached from social context and self-determination (Pasquale 2015; Van Dijk 2014). As end users, they are positioned as passive recipients of system outputs with limited capacity for contestation or agency (Selbst et al. 2019). This dimension is captured through *Citizen datafication* (D16) (Pasquale 2015).

These dimensions evaluate whether AI-enabled systems reproduce partial, exclusionary, or distorted representations of social reality.

Contextual Lens The contextual lens examines the broader material and systemic conditions within which AI-enabled sociotechnical systems operate. It addresses the following analytical question:

What material conditions, costs, and systemic constraints shape the legitimacy and applicability of AI-enabled systems?

Three dimensions operationalize this lens. These dimensions emerge from analytical sub-questions concerning the temporal, economic, and environmental conditions that constrain the sustainability and applicability of AI systems.

Temporal conditions. *Time impact* (D17) concerns both the substantial time required to develop and maintain AI-enabled systems and the temporal mismatch between the pace of social change and the lifecycle of those systems, which may become obsolete or socially misaligned before they can be effectively operationalized (Selbst et al. 2019; Pasmore et al. 2019).

Economic conditions. *Economic Impact* (D18) examines the unequal distribution of economic costs, investments, and benefits associated with AI infrastructures and governance models, which may reinforce existing socioeconomic inequalities across communities, institutions, and geopolitical contexts (Sadowski and Bendor 2019; Hollands 2015).

Environmental conditions. *Sustainability* (D19) captures the environmental costs associated with AI-enabled systems, including energy consumption, computational infrastructures, and resource extraction (Strubell, Ganesh, and McCallum 2019; Schwartz et al. 2020; Patterson et al. 2021).

These dimensions evaluate whether AI-enabled systems remain materially sustainable and socially applicable within the environments in which they operate.

Relating TAI and Contextual Legitimacy

To examine how the proposed framework complements conventional TAI-oriented evaluations, we developed a mapping between the seven TAI pillars and the 19 sociotechnical dimensions of the framework. Table 1 summarizes the resulting associations, with superscript numbers referring to the corresponding EGTAI paragraphs.

Human Agency & Oversight AI systems supporting human autonomy and decision-making⁽⁶²⁾ require meaningful forms of human oversight. From a *fundamental rights perspective*⁽⁶³⁾, TAI mandates impact assessments aimed at protecting individual freedoms, directly engaging *D4 – Surveillance* and *D3 – Imposition of Dreams*. On the dimension of

human agency⁽⁶⁴⁾, users must receive adequate knowledge and tools to interact with AI systems, directly engaging *D10 – Skill Gap in Citizens*. Preserving user autonomy further requires avoiding both the structural delegation of decision-making authority to technical systems (*D1 – Technocracy*) and the uncritical acceptance of AI-generated outputs (*D9 – Oracle Effect*). Finally, *human oversight*⁽⁶⁵⁾ requires meaningful participation across the three recommended modalities — human-in-the-loop, human-on-the-loop, and human-in-command — directly engaging *D11 – Lack of Participation* and *D12 – Lack of Evaluation*. The requirement further intersects with *D7 – Digital Universalism* and *D8 – Technical positivism*, where technological authority remains insufficiently contestable.

Technical Robustness & Safety Technical robustness is grounded in harm prevention⁽⁶⁶⁾, which TAI explicitly extends to physical and psychological integrity. This directly engages both *D3 – Imposition of Dreams* and *D4 – Surveillance*. On *resilience to attack and security*⁽⁶⁷⁾, TAI requires protection against adversarial exploitation and system misuse, intersecting with *D2 – Corporate and Neoliberalist Vision*. On *fallback plan and general safety*⁽⁶⁸⁾, the requirement that systems do not harm individuals or the environment directly engages *D19 – Sustainability*, while risk assessment presupposes evaluative mechanisms capable of capturing broader systemic effects through *D12 – Lack of Evaluation*. On *accuracy*⁽⁶⁹⁾, the obligation to mitigate incorrect predictions and communicate uncertainty intersects with *D9 – Oracle Effect*. Finally, *reliability and reproducibility*⁽⁷⁰⁾ require evaluation metrics capable of specifying the contexts in which system performance remains valid, directly engaging *D12 – Lack of Evaluation*.

Privacy & Data Governance Privacy is established by TAI as a fundamental right in the context of AI⁽⁷¹⁾. On *privacy and data protection*⁽⁷²⁾, systems must not be used for unlawful purposes, directly engaging both *D2 – Corporate and Neoliberalist Vision* and *D4 – Surveillance*. On *data quality and integrity*⁽⁷³⁾, TAI emphasizes the relevance of data to the deployment context. This directly engages *D15 – Biased Data Practices*, as well as *D13 – Lack of Intersectionality* and *D14 – Invisible People*. The requirement further intersects with *D6 – Digital Colonialism*, particularly in cross-context deployments of AI models and data practices. Finally, on *access to data*⁽⁷⁴⁾, meaningful access to and contestation of personal data presuppose the technical literacy required to understand and exercise those rights, directly engaging *D10 – Skill Gap in Citizens*.

Transparency Transparency requires that the relevant elements of AI systems be accessible and interpretable⁽⁷⁵⁾. On *traceability*⁽⁷⁶⁾, comprehensive documentation constitutes a prerequisite for meaningful accountability and system evaluation. On *explainability*⁽⁷⁷⁾, TAI requires that individuals be able to understand system decisions, directly engaging both *D10 – Skill Gap in Citizens* and *D11 – Lack of Participation*. The right to request explanations further intersects with *D9 – Oracle Effect*. The requirement that systems communicate their degree of influence on decision-making processes di-

rectly engages *D12 – Lack of Evaluation*. Finally, on *communication*⁽⁷⁸⁾, TAI requires that AI systems be identifiable as such and that their limitations and levels of accuracy be communicated in accessible terms. These requirements further intersect with *D1 – Technocracy*.

Diversity, Non-Discrimination & Fairness TAI requires that inclusion and diversity be ensured across the entire AI lifecycle⁽⁷⁹⁾, directly engaging *D11 – Lack of Participation* and *D1 – Technocracy*. The requirement further intersects with *D5 – Marginalization of Communities*, *D13 – Lack of Intersectionality*, and *D14 – Invisible People*. On *avoidance of unfair bias*⁽⁸⁰⁾, particular attention must be paid to *D15 – Biased Data Practices*. On *accessibility and universal design*⁽⁸¹⁾, systems must remain user-centered rather than reducing individuals to computational objects, directly engaging *D16 – Citizen Datafication*. The requirement additionally intersects with *D10 – Skill Gap in Citizens*. Finally, *stakeholder participation*⁽⁸²⁾, requires mechanisms capable of evaluating and incorporating stakeholder feedback over time, directly engaging *D12 – Lack of Evaluation*.

Societal and Environmental Well-Being TAI establishes that AI systems should generate benefits that are broadly distributed across society⁽⁸³⁾, directly engaging *D5 – Marginalization of Communities* and *D13 – Lack of Intersectionality*. On *environmental sustainability*⁽⁸⁴⁾, TAI recognizes ecological impact as a constitutive dimension of responsible AI, directly intersecting with *D19 – Sustainability*. On *social impact*⁽⁸⁵⁾, TAI acknowledges the capacity of AI systems to shape social relations and collective imaginaries, engaging *D3 – Imposition of Dreams*. Meaningful assessment of these impacts further presupposes evaluative mechanisms extending beyond technical performance metrics, directly engaging *D12 – Lack of Evaluation*. Finally, on *society and democracy*⁽⁸⁶⁾, TAI emphasizes the effects of AI on democratic institutions and social structures, directly engaging *D1 – Technocracy*, *D2 – Corporate and Neoliberalist Vision*, and *D6 – Digital Colonialism*.

Accountability Accountability operationalizes the preceding TAI requirements through mechanisms of evaluation, contestation, and redress⁽⁸⁷⁾. On *auditability*⁽⁸⁸⁾, TAI requires evaluative procedures capable of assessing AI systems beyond technical performance metrics, directly engaging *D12 – Lack of Evaluation*. The requirement additionally intersects with *D2 – Corporate and Neoliberalist Vision*. On *minimisation and reporting of negative impacts*⁽⁸⁹⁾, accountability presupposes the capacity to contest and respond to harmful system outputs, directly engaging *D9 – Oracle Effect* and *D11 – Lack of Participation*. On *trade-offs*⁽⁹⁰⁾, TAI acknowledges that conflicts between requirements may emerge and must be resolved based on ethical principles and fundamental rights rather than technical efficiency alone, engaging *D1 – Technocracy*. Finally, on *redress*⁽⁹¹⁾, TAI emphasizes the need to protect vulnerable individuals and groups, directly intersecting with *D5 – Marginalization of Communities*, where structural disadvantage limits access to redress mechanisms.

Structural Patterns Across the Mapping

The mapping reveals four structural patterns that highlight the complementary value of the proposed framework with respect to TAI requirements. First, *D12 – Lack of Evaluation* emerges as the most transversal dimension across the mapping, appearing as a necessary condition in nearly every TAI pillar. This suggests that requirements such as auditability, explainability, and impact assessment presuppose evaluative mechanisms capable of assessing broader sociotechnical effects beyond technical system performance. Second, *D17 – Time Impact* and *D18 – Economic Impact* receive no explicit treatment within TAI requirements, despite substantially shaping whether those requirements can be met in practice. This absence suggests a structural gap between the normative formulation of TAI principles and the material conditions required for their implementation across different institutional and socioeconomic contexts. Third, *D6 – Digital Colonialism* surfaces only weakly and indirectly across the mapping, primarily in relation to data governance and democratic integrity. The geopolitical asymmetries through which AI infrastructures, models, and governance frameworks are transferred across different social and cultural contexts therefore remain weakly operationalized within conventional TAI evaluations. Fourth, *D1 – Technocracy* and *D2 – Corporate and Neoliberalist Vision* intersect with multiple TAI pillars — particularly those related to human oversight, transparency, accountability, and democratic governance. This highlights how questions concerning the concentration and legitimization of power remain implicit rather than explicitly addressed within conventional TAI-oriented evaluations. Overall, the mapping suggests that several sociotechnical legitimacy conditions remain fragmented, indirectly addressed, or weakly operationalized within conventional TAI evaluations. Rather than introducing entirely external concerns, the proposed framework reorganizes and systematizes evaluative conditions that often remain implicitly dispersed across existing TAI requirements.

Empirical Evaluation

Empirical Research Design

Operational Definition of AI-Enabled Urban Digital Twins For the purpose of this study, an Urban Digital Twin (UDT) is defined as a digital representation of an urban environment integrating data-driven models, simulation capabilities, and decision-support functionalities connected to urban processes (Dembski et al. 2020; Digital Twin Consortium 2020; Shao and Helu 2022). To distinguish UDTs from static urban models or offline simulations, only systems explicitly supporting monitoring, prediction, simulation, or decision-making through continuously updated urban data were considered. A paper was classified as addressing an AI-enabled UDT when AI techniques constituted an explicit functional component of the UDT architecture. This included, for example, machine learning, predictive analytics, optimization algorithms, computer vision, generative AI, or AI-based decision-support systems integrated within urban digital twin workflows (Huang et al. 2021; Ferré-Bigorra, Casals, and Gangolells 2022; Xu et al.

TAI pillar	Additional query terms
Human agency and oversight	("human agency" OR "human oversight")
Technical robustness and safety	("technical robustness" OR "safety")
Privacy and data governance	("privacy" OR "data governance")
Transparency	("transparency")
Diversity, non-discrimination, and fairness	("diversity" OR "non-discrimination" OR "fairness")
Societal and environmental well-being	("societal well-being" OR "environmental well-being" OR "sustainability" OR "environmental impacts")
Accountability	("accountability" OR "responsibility")

Table 2: Queries used for the Phase 3 SLR. All queries shared the common base ("urban digital twin" OR "city digital twin") AND ("AI" OR "artificial intelligence").

2024). Static 3D city models, descriptive dashboards, and systems without explicit AI-enabled functionalities were excluded. City-scale digital representations for urban decision support were included even when not fully satisfying strict synchronization or bidirectional coupling criteria.

SLR Protocol and Inclusion Criteria The literature search was conducted across Google Scholar, Scopus, and DBLP in September 2025 and updated in May 2026. To identify TAI-oriented AI-enabled UDT literature, we defined one query per TAI pillar using UDT-, AI-, and TAI-related keywords. The complete query set is reported in Table 2. Papers were included if they *i*) described an AI-enabled UDT according to the operational definition provided in the previous Section; *ii*) explicitly addressed at least one TAI-related requirement or evaluative concern; *iii*) provided sufficient conceptual or technical detail to evaluate sociotechnical legitimacy conditions through the proposed framework; and *iv*) were written in English. Review papers and surveys were used exclusively for backward reference analysis and excluded from the empirical coding corpus. After deduplication, screening, and backward snowballing, the final corpus consisted of 80 papers¹.

Analytical Coding Procedure For each dimension, the coding assessed whether the corresponding legitimacy condition was treated as an explicit evaluative concern within the paper, rather than merely acknowledged implicitly. As a preliminary step, each paper was associated with the TAI pillar through which it had been retrieved — as identified by the search query in Table 2 — and the sociotechnical dimensions mapped to that pillar in Table 1 were used as the primary analytical entry points. Evidence was identified through close reading of introductions, conclusions, and passages addressing social, governance, or representational aspects relevant to the framework. Four levels of engagement were identified:

¹The full corpus of papers included in the systematic literature review is will be made available on arXiv.

- *Issue ignored*: the legitimacy condition is not explicitly mentioned;
- *Issue handled uncritically*: the issue is present, whether explicitly or implicitly, but is accepted as given without examining its broader implications;
- *Issue handled critically*: the issue is explicitly acknowledged and critically discussed;
- *Issue operationalized*: the issue is translated into explicit evaluative, methodological, or governance mechanisms.

Examples from the reviewed corpus illustrate the four levels of engagement. At the *operationalized* level, Villanueva-Merino et al. (2024) explicitly addresses the impacts of urban AI systems on elderly populations by proposing concrete evaluative mechanisms and participatory design approaches (*D11*, *D12*). At the *handled critically* level, Nechesov and Ruponen (2024) acknowledges the need to include diverse citizen groups in AI civic participation systems and recognizes unequal technical literacy as a barrier to meaningful engagement (*D10*), without fully formalizing a systematic methodological response. At the *handled uncritically* level, Nanthasamroeng et al. (2026), despite pursuing sustainability objectives, does not examine whether the computational costs of the deployed AI models may partially offset the environmental benefits pursued (*D19*). Finally, at the *ignored* level, Xiao et al. (2024) makes no reference to intersectionality or to the exclusion of populations absent from dominant data infrastructures (*D13*, *D14*).

Empirical Findings

Findings are presented at two complementary levels. First, the analysis is conducted at the pillar level, examining whether papers associated with a given TAI pillar also engage the sociotechnical dimensions identified in the previous mapping as necessary conditions for the substantive operationalization of that pillar. Second, the analysis is conducted at the analytical lens level, examining the overall treatment of the four analytical lenses of the proposed framework in order to identify broader patterns of sociotechnical engagement, systematic blind spots, and differences in critical treatment across the reviewed corpus.

Pillar-level Analysis At the pillar level, the analysis examines whether papers associated with a given TAI pillar also engage the sociotechnical dimensions identified in the previous mapping as necessary conditions for the substantive operationalization of that pillar. For each TAI category, we analyzed the subset of associated papers and evaluated the level of engagement of each mapped sociotechnical dimension according to the four coding categories introduced above. Figure 2 shows the distribution of papers for each TAI pillar across the four levels of engagement: ignored, handled uncritically, handled critically, and operationalized.

Human Agency & Oversight (11 papers) Papers associated with this pillar show a predominantly uncritical treatment of the mapped sociotechnical dimensions. In particular, *D7 – Digital Universalism*, *D3 – Imposition of Dreams*, and *D8 – Technical positivism* are present in more than 90%

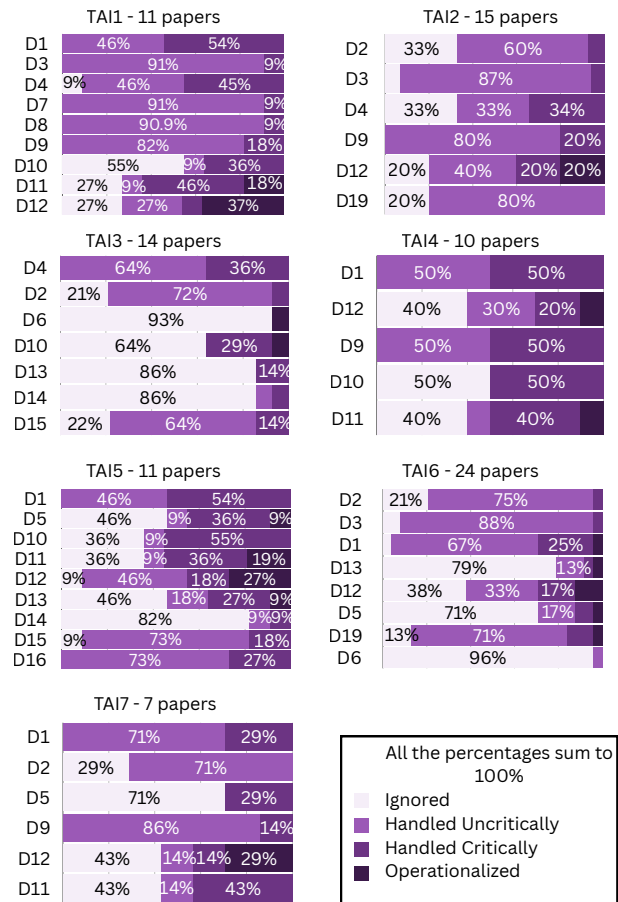


Figure 2: Engagement distribution across sociotechnical dimensions within the corresponding TAI pillar.

of the corpus without an explicit evaluative concern, suggesting that assumptions regarding the neutrality and desirability of AI-enabled urban systems remain largely unquestioned. More critical treatment emerges around *D1 – Technocracy* and *D4 – Surveillance*, where critical handling reaches 55% and 45% respectively, indicating partial awareness of power concentration and behavioral control dynamics. *D11 – Lack of Participation* and *D12 – Lack of Evaluation* record the highest shares of critical and operationalized handling within this pillar, suggesting growing attention toward participatory governance and broader mechanisms of social evaluation. The findings indicate that literature addressing Human Agency & Oversight still operationalizes agency primarily at the level of technical interaction and user control, while only partially engaging with the broader structural conditions shaping participation, contestability, and oversight.

Technical Robustness & Safety (15 papers) Papers associated with this pillar display a predominantly uncritical treatment of the mapped sociotechnical dimensions. In particular, *D3 – Imposition of Dreams* and *D9 – Oracle Effect* are handled almost exclusively without critical reflection,

consistent with the strongly technical framing characterizing this pillar. *D4 – Surveillance* shows the most contested profile, with papers distributed almost evenly across ignored, uncritical, and critical handling, suggesting emerging awareness of monitoring and data governance concerns. *D12 – Lack of Evaluation* presents the most differentiated distribution, including the highest share of operationalized cases within this pillar. At the same time, *D19 – Sustainability* is treated predominantly uncritically: environmental sustainability is invoked almost exclusively as a goal or intended outcome of AI-enabled urban systems, while the environmental costs associated with AI infrastructures are rarely examined. Overall, robustness and safety are framed primarily in technical and infrastructural terms, with limited engagement with their broader sociotechnical implications.

Privacy & Data Governance (14 papers) Papers associated with this pillar present the highest concentration of ignored sociotechnical dimensions across the TAI analysis. In particular, *D6 – Digital Colonialism*, *D13 – Lack of Intersectionality*, and *D14 – Invisible People* are almost entirely absent from the reviewed corpus, indicating that privacy and data governance are predominantly framed in technical and procedural terms. More critical treatment emerges around *D4 – Surveillance*, suggesting partial awareness of monitoring and data exploitation concerns, while *D10 – Skill Gap in Citizens* records moderate levels of critical handling related to inequalities in data access and understanding. Overall, the findings suggest that existing privacy-oriented literature remains weakly engaged with the broader sociotechnical conditions shaping exclusion, representation, and power asymmetries in data governance.

Transparency (10 papers) Papers associated with this pillar display a comparatively more balanced distribution between uncritical and critical handling across several sociotechnical dimensions. In particular, *D1 – Technocracy*, and *D9 – Oracle Effect* show evenly distributed levels of uncritical and critical treatment, suggesting more consistent engagement with transparency-related sociotechnical concerns than in other TAI pillars. *D10 – Skill Gap in Citizens* presents a distinctive profile, being either ignored or handled critically, with limited uncritical treatment. This suggests that technical literacy is recognized as a substantive condition for transparency rather than taken for granted. *D11 – Lack of Participation* and *D12 – Lack of Evaluation* present the most differentiated profiles, including the only operationalized cases observed within this pillar. Overall, the findings indicate that transparency-oriented literature engages more directly with issues of explanation, contestability, and evaluation, while still providing limited operational responses to the broader sociotechnical challenges associated with transparency in AI-enabled urban systems.

Diversity, Non-Discrimination & Fairness (11 papers) Papers associated with this pillar display the most internally heterogeneous profile across the TAI analysis. *D1 – Technocracy* records the highest rate of critical handling observed for this dimension across all pillars, while *D12 – Lack of Evaluation* reaches the highest share of operational-

ized cases in the corpus. In contrast, dimensions directly related to social exclusion remain weakly addressed: *D14 – Invisible People* is ignored in the large majority of papers, and *D13 – Lack of Intersectionality* shows similarly limited engagement. More moderate levels of critical treatment emerge for *D5 – Marginalization of Communities* and *D16 – Citizen Datafication*. Overall, the findings reveal a structural tension within this pillar: while diversity-oriented literature engages critically with governance and evaluation concerns, it remains comparatively weak in addressing the sociotechnical conditions most directly related to exclusion and representational inequality.

Societal & Environmental Wellbeing (24 papers) Papers associated with this pillar display a predominantly uncritical treatment of the mapped sociotechnical dimensions. In particular, *D3 – Imposition of Dreams* and *D2 – Corporate and Neoliberalist Vision* are rarely addressed as explicit evaluative concerns, while *D6 – Digital Colonialism* is largely ignored across the corpus. *D12 – Lack of Evaluation* presents the most differentiated profile, including moderate levels of critical and operationalized handling, while *D5 – Marginalization of Communities* records limited operational engagement. At the same time, *D19 – Sustainability* is treated predominantly uncritically. Overall, despite its explicit focus on societal and environmental concerns, this pillar reproduces many of the structural blind spots observed across the broader corpus, particularly regarding colonial, representational, and systemic dimensions.

Accountability (7 papers) Papers associated with this pillar display a predominantly uncritical treatment of the mapped sociotechnical dimensions. In particular, *D9 – Oracle Effect* and *D1 – Technocracy* are rarely addressed as explicit evaluative concerns, while *D2 – Corporate and Neoliberalist Vision* remains partially ignored across the corpus. More differentiated profiles emerge for *D5 – Marginalization of Communities* and *D12 – Lack of Evaluation*, the latter recording the highest share of operationalized cases within this pillar. Overall, despite the limited sample size, the findings suggest that accountability within urban digital twin literature is framed primarily in terms of technical evaluation and system performance, with comparatively weaker engagement with broader questions of social responsibility, contestability, and structural inequality.

Analytical Lens-Level Analysis Aggregating results across the four analytical lenses highlights systematic differences in how sociotechnical legitimacy conditions are treated within TAI-oriented UDT literature. Within the governance lens, the dominant pattern is one of predominantly uncritical treatment of structural power relations and their underlying legitimacy assumptions. *D1 – Technocracy*, *D2 – Corporate and Neoliberalist Vision*, *D3 – Imposition of Dreams*, and *D4 – Surveillance* are frequently present across the corpus, but are treated predominantly uncritically, suggesting limited reflexive engagement with the ways AI-enabled urban systems concentrate authority, shape behavior, and normalize particular forms of governance and authority. More structurally oriented dimensions remain even

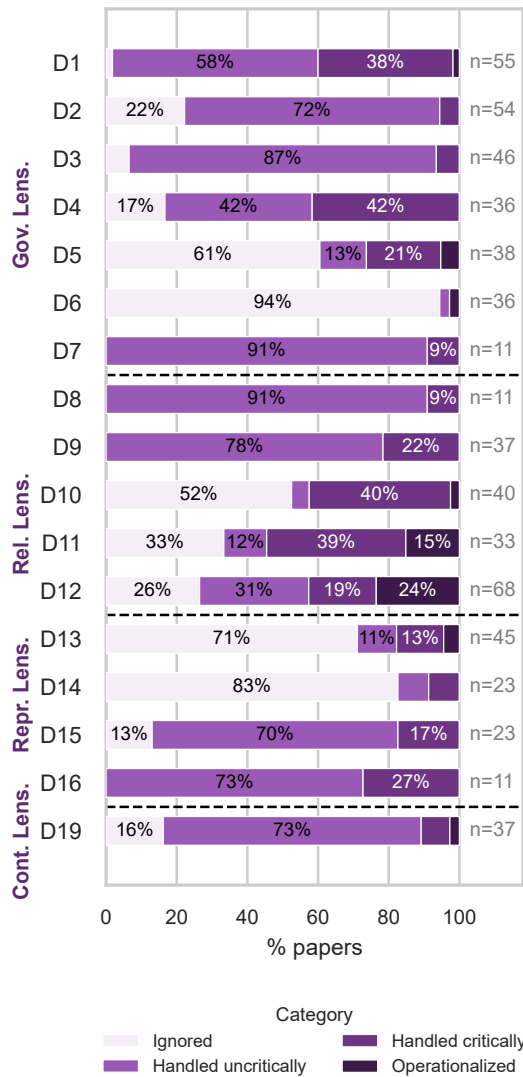


Figure 3: Distribution of engagements of ST dimensions for paper TAI related

less developed: *D5 – Marginalization of Communities* receives only limited critical treatment, while *D6 – Digital Colonialism* is almost entirely ignored across the corpus. Finally, *D7 – Digital Universalism* is handled almost exclusively uncritically, indicating that assumptions regarding the universality and neutrality of computational representations remain largely unquestioned. Overall, the findings suggest that broader geopolitical asymmetries, structural inequalities, and legitimacy assumptions remain weakly addressed within existing TAI-oriented UDT literature. The relational lens presents the most internally differentiated profile across the four analytical lenses. *D8 – Technical Positivism* and *D9 – Oracle Effect* are treated predominantly uncritically, indicating limited reflexive engagement with assumptions of technological neutrality and algorithmic authority. In contrast, *D10 – Skill Gap in Citizens*, *D11 – Lack of Participation*, and especially *D12 – Lack of Evaluation* display com-

paratively higher levels of critical and operationalized handling. The findings suggest that existing literature increasingly addresses participation and social evaluation mechanisms, while paying less attention to the relational asymmetries and legitimacy assumptions through which AI systems acquire authority and social acceptance. Within the representational lens, the dominant pattern is the limited treatment of dimensions related to representational exclusion. In particular, *D13 – Lack of Intersectionality* and *D14 – Invisible People* are largely ignored across the reviewed corpus, indicating limited attention to which individuals, identities, and social experiences remain excluded from computational representations. *D15 – Biased Data Practices* is more frequently acknowledged, though predominantly handled uncritically, indicating that bias is often treated as a technical data-quality issue rather than as a broader sociotechnical condition shaping exclusion and inequality. Similarly, *D16 – Citizen Datafication* is acknowledged more often than operationalized. Taken together, these findings suggest that representational concerns extensively discussed within broader AI ethics literature — particularly intersectionality and representational exclusion — remain only weakly integrated into domain-specific UDT evaluation and governance practices. Finally, the contextual lens is represented solely by *D19 – Sustainability*, as *D17 – Time Impact* and *D18 – Economic Impact* are not associated with any TAI pillar and therefore fall outside the pillar-based analysis. At the same time, sustainability functions primarily as the motivating objective driving the development of AI-enabled urban solutions, which partly explains the comparatively large size of this pillar's corpus. However, the environmental costs associated with AI infrastructures themselves are rarely made explicit, creating a structural blind spot in the evaluation of the sustainability profile of these systems.

Conclusion

This paper proposed a sociotechnical framework for evaluating contextual legitimacy in AI-enabled sociotechnical systems. The analytical mapping and empirical evaluation show that several legitimacy conditions remain only partially operationalized within conventional TAI-oriented evaluations: while participation, transparency, and evaluation receive increasing critical attention, structural conditions — power concentration, representational exclusion, digital colonialism, and systemic sustainability — remain comparatively underexamined. These findings suggest that artifact-centered and compliance-oriented evaluations alone are insufficient for capturing how AI legitimacy is constructed and contested in practice, motivating the proposed framework as a complementary, systematic lens for assessing these conditions. Future work should extend its empirical application beyond the UDT domain to assess which observed blind spots generalize across AI-enabled sociotechnical systems and which remain domain-specific.

Declaration on the Use of AI

The authors used generative AI tools exclusively for (i) graphical refinement and completion of author-created fig-

ures and sketches, and (ii) editing and polishing of author-written text for clarity and readability. No AI-generated scientific content, analyses, interpretations, or conclusions were directly included in the manuscript without substantial human authorship, revision, verification, and full responsibility by the authors.

References

- Barocas, S.; and Selbst, A. D. 2016. Big data's disparate impact. *Calif. L. Rev.*, 104: 671.
- Benjamin, R. 2023. Race after technology. In *Social Theory Re-Wired*, 405–415. Routledge.
- Bina, O.; Inch, A.; and Pereira, L. 2020. Beyond technoutopia and its discontents: On the role of utopianism and speculative fiction in shaping alternatives to the smart city imaginary. *Futures*, 115: 102475.
- Couldry, N.; and Mejias, U. A. 2019. Data colonialism: Rethinking big data's relation to the contemporary subject. *Television & new media*, 20(4): 336–349.
- Crawford, K. 2025. *Atlas da IA*. Edições Sesc SP.
- Crenshaw, K. 1989. Demarginalizing the intersection of race and sex: A black feminist critique of antidiscrimination doctrine, feminist theory and antiracist politics. *University of Chicago Legal Forums*.
- Dembski, F.; Wössner, U.; Letzgus, M.; Ruddat, M.; and Yamu, C. 2020. Urban Digital Twins for Smart Cities and Citizens: The Case Study of Herrenberg, Germany. *Sustainability*, 12(6): 2307.
- Digital Twin Consortium. 2020. Digital Twin Consortium Defines Digital Twin.
- D'Ignazio, C.; and Klein, L. F. 2023. *Data feminism*. MIT press.
- Emery, F. E.; and Trist, E. L. 1960. Socio-technical systems. *Management science, models and techniques*, 2: 83–97.
- Eubanks, V. 2018. *Automating inequality: How high-tech tools profile, police, and punish the poor*. Macmillan+ORM.
- European Commission High-Level Expert Group on AI. 2019. Ethics Guidelines for Trustworthy AI.
- Ferré-Bigorra, J.; Casals, M.; and Gangolells, M. 2022. The adoption of urban digital twins. *Cities*, 131: 103905.
- Hollands, R. G. 2015. Critical interventions into the corporate smart city. *Cambridge journal of regions, economy and society*, 8(1): 61–77.
- Huang, Z.; et al. 2021. A Survey on AI-Driven Digital Twins in Industry 4.0. *Computers*, 10(8): 100.
- Jasanoff, S.; and Kim, S.-H. 2019. *Dreamscapes of modernity: Sociotechnical imaginaries and the fabrication of power*. University of Chicago Press.
- Krivý, M. 2018. Towards a critique of cybernetic urbanism: The smart city and the society of control. *Planning Theory*, 17(1): 8–30.
- Kudina, O.; and Van De Poel, I. 2024. A sociotechnical system perspective on AI. *Minds and Machines*, 34(3): 21.
- Kwet, M. 2019. Digital colonialism: US empire and the new imperialism in the Global South. *Race & class*, 60(4): 3–26.
- Malleson, N.; Franklin, R.; Arribas-Bel, D.; Cheng, T.; and Birkin, M. 2024. Digital twins on trial: Can they actually solve wicked societal problems and change the world for better?
- Mehrotra, S.; Huang, J.; Fu, X.; Dobbe, R.; Sánchez, C. I.; and De Rijke, M. 2026. Understanding AI Trustworthiness: A Scoping Review of AIES & FAccT Articles. *Journal of Artificial Intelligence Research*, 85.
- Mittelstadt, B. 2019. Principles alone cannot guarantee ethical AI. *Nature machine intelligence*, 1(11): 501–507.
- Morozov, E. 2013. *To save everything, click here: The folly of technological solutionism*. PublicAffairs.
- Nanthasamroeng, N.; Luesak, P.; Pitakaso, R.; Khonjun, S.; Jirasirilerd, G.; and Matitopanum, S. 2026. A machine learning-Powered digital twin framework for adaptive management of urban air quality in Chiang Mai, Northern Thailand. *Engineering Applications of Artificial Intelligence*, 163: 112597.
- Nechesov, A.; and Ruponen, J. 2024. Empowering government efficiency through civic intelligence: Merging artificial intelligence and blockchain for smart citizen proposals. *Technologies*, 12(12): 271.
- Noble, S. U. 2018. Algorithms of oppression: How search engines reinforce racism. In *Algorithms of oppression*. New York university press.
- Nochta, T.; Wan, L.; Schooling, J. M.; and Parlikad, A. K. 2021. A socio-technical perspective on urban analytics: The case of city-scale digital twins. *Journal of urban technology*, 28(1-2): 263–287.
- O'Neil, C. 2017. *Weapons of math destruction: How big data increases inequality and threatens democracy*. Crown.
- Orlikowski, W. J. 2007. Sociomaterial practices: Exploring technology at work. *Organization studies*, 28(9): 1435–1448.
- Orlikowski, W. J.; and Gash, D. C. 1994. Technological frames: making sense of information technology in organizations. *ACM Transactions on Information Systems (TOIS)*, 12(2): 174–207.
- Pasmore, W.; Winby, S.; Mohrman, S. A.; and Vanasse, R. 2019. Reflections: sociotechnical systems design and organization change. *Journal of Change Management*, 19(2): 67–85.
- Pasquale, F. 2015. The black box society: The secret algorithms that control money and information. In *The black box society*. Harvard University Press.
- Patterson, D.; Gonzalez, J.; Le, Q.; Liang, C.; Munguia, L.-M.; Rothchild, D.; So, D.; Texier, M.; and Dean, J. 2021. Carbon emissions and large neural network training. *arXiv preprint arXiv:2104.10350*.
- Ruohomäki, T.; Ponto, H.; Santala, V.; and Virtanen, J.-P. 2024. Urban digital twin as a socio-technical construct. In *Handbook of digital twins*, 308–320. CRC Press.

- Sadowski, J.; and Bendor, R. 2019. Selling smartness: Corporate narratives and the smart city as a sociotechnical imaginary. *Science, Technology, & Human Values*, 44(3): 540–563.
- Schwartz, R.; Dodge, J.; Smith, N. A.; and Etzioni, O. 2020. Green ai. *Communications of the ACM*, 63(12): 54–63.
- Selbst, A. D.; Boyd, D.; Friedler, S. A.; Venkatasubramanian, S.; and Vertesi, J. 2019. Fairness and abstraction in sociotechnical systems. In *Proceedings of the conference on fairness, accountability, and transparency*, 59–68.
- Shao, G.; and Helu, M. 2022. Credibility Consideration for Digital Twins in Manufacturing.
- Shneiderman, B. 2020. Human-centered artificial intelligence: Reliable, safe & trustworthy. *International Journal of Human–Computer Interaction*, 36(6): 495–504.
- Strubell, E.; Ganesh, A.; and McCallum, A. 2019. Energy and policy considerations for deep learning in NLP. In *Proceedings of the 57th annual meeting of the association for computational linguistics*, 3645–3650.
- Suchman, M. C. 1995. Managing legitimacy: Strategic and institutional approaches. *Academy of management review*, 20(3): 571–610.
- Van Dijck, J. 2014. Datafication, dataism and dataveillance: Big Data between scientific paradigm and ideology. *Surveillance & society*, 12(2): 197–208.
- Van Dijk, J. 2020. *The digital divide*. John Wiley & Sons.
- Villanueva-Merino, A.; Urra-Uriarte, S.; Izkara, J. L.; Campos-Cordobes, S.; Aranguren, A.; and Molina-Costa, P. 2024. Leveraging Local Digital Twins for planning age-friendly urban environments. *Cities*, 155: 105458.
- Virk, G. 2022. Identifying Sociotechnical Harms of Algorithmic Systems: Scoping a Taxonomy for Harm Reduction. *arXiv (Cornell University)*.
- Wang, B. 2021. The seductive smart city and the benevolent role of transparency. *Interaction Design and Architecture (s)*, 48: 100–121.
- Xiao, C.; Zhou, J.; Xiao, Y.; Huang, J.; and Xiong, H. 2024. Refound: Crafting a foundation model for urban region understanding upon language and visual foundations. In *30th ACM SIGKDD Knowledge Discovery and Data Mining Conference*, 3527–3538.
- Xu, H.; et al. 2024. Leveraging generative AI for urban digital twins: a scoping review.