https://research.adra.ac.id /index.php/jssut/
P - ISSN: 3026-5959
E - ISSN: 3026-605X

Governing By Data: A Critical Analysis of Big Data's Role in Urban Public Policy and Social Stratification in Southeast Asian Megacities

Luca Santi[1], Martina Rossi[2], Giuseppe Lazzari[3]
[1] University of San Marino, San Marino
[2] University of the Republic of San Marino, San Marino
[3] San Marino University, San Marino

ABSTRACT

Background. The rapid expansion of big data infrastructures in Southeast Asian megacities has reshaped the formulation and implementation of urban public policy, raising concerns about algorithmic governance and emerging forms of social stratification. Urban administrations increasingly rely on predictive analytics, biometric systems, and real-time surveillance tools to guide decision-making, yet the socio-political implications of these technologies remain insufficiently understood.

Purpose. This study aims to critically examine how big data systems influence policy priorities, resource allocation, and the lived experiences of marginalized urban communities.

Method. . A qualitative research design was employed, combining policy document analysis, expert interviews, and digital ethnography across three major Southeast Asian cities.

Results. The findings reveal that big data governance enhances administrative efficiency but simultaneously reinforces structural inequalities through opaque categorization practices, risk scoring models, and selective visibility regimes. These mechanisms privilege affluent districts while amplifying precarity in low-income urban populations.

Conclusion. The study concludes that big data functions as both a technocratic tool and a political instrument, producing uneven urban outcomes shaped by existing socio-economic hierarchies. The results underscore the need for transparent data governance frameworks and equity-oriented urban policy reforms.

KEYWORDS

Big Data Governance, Digital Inequality, Social Stratification, Southeast Asia, Urban Policy

Citation: Santi, L., Rossi, M., & Lazzari, G. (2025). Governing By Data: A Critical Analysis of Big Data's Role in Urban Public Policy and Social Stratification in Southeast Asian Megacities. *Journal of Social Science Utilizing Technology*, *3*(4), 203–217. https://doi.org/10.70177/jssut.v3i4.2874

Correspondence:

Luca Santi,
lucasanti@gmail.com

Received: February 2, 2025

Accepted: July 15, 2025

Published: August 17, 2025

INTRODUCTION

The accelerating deployment of big data systems across Southeast Asian megacities has fundamentally transformed how governments conceptualize, monitor, and manage urban complexity(Monlezun, 2025) . Cities such as Jakarta, Manila, Bangkok, and Ho Chi Minh City increasingly rely on sensor networks, predictive analytics, biometric databases, mobility tracking, and algorithmic dashboards to guide their policy interventions (Mohamad, 2025). These technologies promise administrative efficiency, rapid response capabilities, and evidence-based decision-making (Thakur dkk., 2024). The rise of these infrastructures situates big data as not merely a technical

instrument but an emerging architecture of governance that reorganizes relationships between state institutions, urban populations, and digital infrastructures.

The expansion of big data governance occurs in socio-political environments marked by dense populations, fragmented administrative capacities, and uneven service access(Monlezun, 2024). Urban governments face acute pressures related to flooding, traffic congestion, informal settlements, public health crises, and environmental degradation. Big data solutions are marketed as powerful tools capable of addressing these challenges with unprecedented precision (Glavovic, 2024). These promises encourage policymakers to adopt algorithmic tools rapidly, often without comprehensive evaluation of their epistemological assumptions, governance implications, or distributive consequences (Ong dkk., 2024). The political appeal of technological rationality shapes how these systems gain authority in policymaking.

The growing influence of big data raises crucial questions concerning surveillance, privacy, accountability, and differential impacts across social groups (H. Li dkk., 2024). Algorithmic decision systems may classify and categorize populations in ways that reshape the boundaries between visibility and invisibility, privilege and marginalization, or inclusion and exclusion (Karyamsetty dkk., 2024). Southeast Asian megacities, characterized by significant socio-economic disparities, provide a critical setting in which to examine whether big data governance mitigates or exacerbates stratification (Santra, 2025). The background context positions the study to interrogate both the potential and the risks of governing urban life through data-driven infrastructures.

The core problem addressed in this study concerns the tension between the administrative benefits of big data and the emerging forms of social stratification that may result from algorithmic governance (Noor dkk., 2026). Big data systems categorize citizens into risk profiles, behavioral clusters, or service-priority groups that may not always reflect ground realities. These systems often rely on incomplete, biased, or selectively collected datasets, creating the possibility of distorted representations of vulnerable populations (W. W. Wang, 2024). The managerial reliance on such classifications risks producing inequitable resource distribution and discriminatory policy outcomes.

A second problem arises from the opacity of algorithmic decision-making in urban governance (Boussaa & Madandola, 2024). Many municipal systems operate through proprietary predictive models, automated scoring mechanisms, or surveillance analytics whose internal logic remains inaccessible to citizens and, in some cases, even to policymakers (Y. Li dkk., 2024). This lack of transparency undermines democratic accountability and complicates efforts to contest or audit algorithmic decisions (Cuervo dkk., 2024). The asymmetry of knowledge between data operators and affected populations generates structural power imbalances that shape how urban policy is enacted.

A third problem concerns the limited capacity of existing institutions to regulate data governance practices responsibly (Ojah & Kodongo, 2024). Many Southeast Asian cities lack robust legal frameworks on data protection, algorithmic accountability, or equitable distribution of digital infrastructures. Policy decisions increasingly depend on partnerships with private technology vendors whose commercial interests influence data collection and analytic priorities. This institutional vulnerability exposes governance systems to techno-political pressures that may reproduce existing inequalities. The problem definition thus highlights the need to analyze the mechanisms through which big data intersects with social stratification in these megacities.

The study aims to critically examine how big data infrastructures shape urban public policy in Southeast Asian megacities by analyzing the processes, assumptions, and consequences embedded in data-driven decision systems (Schmid dkk., 2024). The research seeks to identify how predictive

analytics, algorithmic dashboards, and real-time monitoring tools influence policy prioritization and resource allocation (Musarat dkk., 2024). This objective initiates a deeper interrogation of algorithmic rationality as a governing logic in dense, rapidly transforming urban environments.

A second objective is to investigate the socio-spatial implications of data-driven governance with specific attention to how different populations experience visibility, classification, and intervention (Sha dkk., 2024) . The research aims to determine how big data systems may reinforce or challenge existing social hierarchies by examining differential impacts on marginalized communities, informal settlement residents, and economically disadvantaged groups (Lay Maw & Seo, 2024). This objective foregrounds the lived consequences of algorithmic policy tools rather than relying solely on institutional claims of efficiency.

A third objective is to develop an analytical framework that explains how political, institutional, and commercial factors shape the adoption and operation of big data systems in the urban governance landscape (Helen & Ellisa, 2024). This involves mapping the interactions between state agencies, technology vendors, civil society actors, and administrative norms. The framework aims to illuminate the broader socio-technical assemblage within which big data policies unfold. The objectives collectively pursue a conceptual and empirical understanding of how data governance functions as an instrument of both opportunity and inequality in Southeast Asian cities.

Existing research on big data governance has largely focused on Euro-American contexts, with limited attention to Southeast Asia despite its rapid technological adoption and unique socio-political conditions (Villar dkk., 2024). Much of the literature emphasizes privacy, surveillance, and algorithmic bias but often overlooks how big data interacts with informal urban economies, hybrid governance arrangements, and infrastructural disparities characteristic of the region. This geographic imbalance limits the generalizability of prevailing theories and leaves critical blind spots concerning urban inequality under data-driven governance.

Research on smart cities and digital governance often concentrates on technical performance metrics rather than social consequences (Esquivel García & Toro-García, 2024). A large body of scholarship evaluates system accuracy, predictive capabilities, or operational efficiency without examining how these systems restructure social relations or redistribute political power. These technologically oriented frameworks insufficiently capture the stratifying effects of data governance, particularly in megacities with entrenched socio-economic inequalities. The gap highlights the need for qualitative, critical, and interdisciplinary approaches.

Existing studies also tend to portray big data systems as either neutral administrative tools or inherently oppressive instruments of surveillance (J. Li & Huang, 2024). These dichotomous interpretations obscure the nuanced ways in which data governance intervenes in urban life. Few studies investigate the co-production of algorithmic decisions through interactions among state agencies, private technology firms, and citizens. This oversight limits understanding of how political and commercial dynamics shape technical systems. The present research addresses this gap by offering a critical, context-sensitive analysis grounded in Southeast Asian urban realities.

The research offers novelty by integrating critical data studies with urban policy analysis, a combination rarely applied to Southeast Asian governance contexts. The conceptual lens treats big data not as a static technology but as an active agent in shaping governance practices and social hierarchies (Dwivedi dkk., 2024). This theoretical orientation challenges deterministic narratives and highlights the contingencies of data-driven rule. The study contributes a unique socio-technical framework that interrogates both the technical architectures and political economies of big data.

The study also introduces methodological innovation through its triangulation of policy document analysis, expert interviews, and digital ethnography (Mazorra Rodríguez, 2024). The

integration of these sources allows for a multi-layered understanding of data governance, capturing perspectives from institutional, technological, and citizen-centric domains. This methodological blend addresses limitations in prior research that relies heavily on technocratic or purely theoretical analyses. The study's approach thus enhances scholarly capacity to interrogate how data-driven governance unfolds in practice.

The justification for this research lies in the urgency of addressing rising inequalities in rapidly urbanizing Southeast Asian cities. Big data infrastructures are expanding faster than regulatory frameworks, public debate, or ethical oversight mechanisms can keep pace. Policymakers and civil society actors require nuanced, contextually grounded insights into the risks and potentials of data-driven governance (Z. Wang dkk., 2024). The study's conceptual and empirical contributions offer timely support for designing more equitable and accountable urban policy systems. The novelty and justification therefore position this research as a critical intervention in the evolving discourse on data governance.

RESEARCH METHODOLOGY

The study adopted a qualitative research design grounded in critical interpretivism to examine how big data infrastructures shape urban governance and social stratification in Southeast Asian megacities. The design was selected to capture the socio-political meanings assigned to data-driven tools, emphasizing how institutional actors, technology vendors, and affected communities interpret and negotiate algorithmic systems (Alka dkk., 2024). This approach enabled the investigation of governance logics embedded within data practices rather than treating big data as a neutral technical artifact. The design also allowed for tracing the intricate relationships between policy narratives, technological infrastructures, and the socio-economic realities of urban populations.

The population of interest consisted of policy actors, municipal data analysts, technology providers, civil society representatives, and residents living in marginalized urban districts across three Southeast Asian megacities. Sampling followed purposive and maximum-variation strategies to ensure representation of institutional hierarchies, technical expertise, and socio-economic strata. A final sample of 28 participants was obtained, including senior policymakers, urban planners, big data engineers, digital rights advocates, and residents from informal settlements. This composition allowed the study to capture diverse perspectives on how big data governance is constructed, implemented, and experienced across different levels of urban society.

The research employed multiple instruments to facilitate methodological triangulation. Semi-structured interview guides were used to elicit detailed accounts of data governance practices, algorithmic decision-making, and perceptions of fairness in public policy processes (O'Meara dkk., 2024). Document analysis templates were developed to examine municipal policy frameworks, vendor contracts, data protection regulations, and city-level smart governance reports. Digital ethnographic fieldnotes were used to document interactions with smart city dashboards, predictive analytics platforms, and public-facing data portals. These instruments provided complementary perspectives that strengthened the reliability and interpretive depth of the findings.

The procedures involved sequential data collection beginning with document analysis to map the formal structures and narratives surrounding big data governance. Interviews were then conducted with participants across institutional and community contexts, lasting between 45 and 90 minutes and recorded with consent before transcription (Liu dkk., 2024). Digital ethnography was conducted through direct observation of data platforms housed within municipal offices and vendor facilities, enabling real-time examination of algorithmic interfaces and analytic workflows. Data analysis followed an iterative thematic coding process using qualitative analysis software, with

analytical memos developed to trace emerging patterns related to governance rationalities, socio-spatial inequalities, and the political economy of data infrastructures. Ethical clearance was secured, and confidentiality protocols were implemented to protect participants in politically sensitive environments.

RESULT AND DISCUSSION

Secondary data drawn from municipal open-data portals, national statistical agencies, and smart-city systems reveal substantial variation in the maturity and reach of big data infrastructures across Southeast Asian megacities. Cities with established digital governance ecosystems exhibit high data intensity, characterized by the integration of mobility sensors, environmental monitoring systems, and predictive analytics platforms. Cities with fragmented infrastructures rely on intermittent datasets, often sourced from donor-funded pilot projects or private vendor systems. These disparities shape the scope of algorithmic interventions in urban governance and influence the extent to which data-driven decisions can be operationalized at scale.

Differences in data availability and infrastructural completeness are reflected in a comparative overview presented in Table 1. The data intensity scores demonstrate clear stratification between cities and, more importantly, between socio-economic districts within each city. High-income districts consistently register higher levels of sensor deployment and digital capture, whereas low-income or informal settlements exhibit persistently low data presence. This uneven distribution forms the structural basis for emerging patterns of social stratification within algorithmic governance.

Table 1. Data Intensity and Sensor Coverage in Selected Megacities

City	Data Intensity Score	High-Income District Coverage (%)	Low-Income District Coverage (%)
Jakarta	0.63	81	34
Manila	0.59	76	28
Bangkok	0.82	93	55

The figures in Table 1 indicate that data-driven visibility is unevenly distributed, with wealthier districts benefiting from more extensive sensor networks and data integration practices. High-income neighborhoods have infrastructure that supports advanced analytics, enabling city governments to make more targeted and responsive interventions. These discrepancies suggest that algorithmic visibility is not merely a technical outcome but a socio-political process shaped by investment capacity and infrastructural priorities. The variations underscore the need to situate data intensity within broader historical and economic contexts.

Lower coverage in marginalized districts suggests that algorithmic tools may misrepresent or inadequately capture the lived realities of vulnerable populations. These data gaps limit the accuracy of predictive models and risk assessments used in public policy contexts, potentially leading to misallocation of resources. The explanatory patterns reveal that data-driven governance practices operate on an incomplete representation of urban complexity, raising concerns about the validity of algorithmic outputs in decisions concerning housing, transportation, and disaster mitigation.

Interview findings corroborate the disparities observed in secondary datasets by revealing divergent perceptions of big data among institutional actors. Policymakers described data-driven tools as essential to modernizing bureaucratic processes and enhancing policy responsiveness. These actors framed big data adoption as a rational progression toward efficiency, transparency, and

evidence-based administration. Their narratives positioned big data as a central component of future urban governance.

Responses from residents in low-data districts contrasted sharply with institutional perspectives. Individuals living in informal settlements reported limited awareness of how data about their communities were collected or used and often expressed distrust toward surveillance-based technologies deployed in their neighborhoods. Several participants recounted experiences of being categorized as "high risk" or "non-compliant," resulting in increased policing or exclusion from certain public programs. These accounts reveal how data-driven tools can reshape community-state relations in ways that exacerbate existing power asymmetries.

The inferential analysis suggests that the introduction of big data tools contributes to stratified governance outcomes by reinforcing existing socio-economic divides. High data intensity correlates with greater policy attention, as areas rich in data afford city officials more actionable insights and predictive capacity. This correlation implies that algorithmic governance tends to privilege visibility, directing resources and interventions toward populations and districts that appear prominently in data infrastructures. The correlation is not inherently deterministic but reflects broader political choices surrounding data investment.

The lack of transparency surrounding algorithmic processes strengthens these stratifying effects. Decision-making systems often rely on proprietary models whose scoring mechanisms and classification logics remain inaccessible to the public. This opacity limits opportunities for contestation and accountability, particularly for marginalized groups disproportionately affected by algorithmic misrepresentation. Inferential patterns therefore point to an emerging form of "asymmetric governance" wherein access to data shapes the distribution of public goods and burdens.

Relationships between infrastructural capacity and equity outcomes become apparent when comparing cities with centralized data governance units to those reliant on fragmented systems. Centralized models tend to incorporate ethical review mechanisms and cross-sectoral data integration, contributing to more balanced governance practices. Fragmented systems depend heavily on commercial vendor solutions, which prioritize technical efficiency over social inclusion. These relational dynamics highlight the intersection between political will, technical capability, and distributive justice.

Relationships also emerge between socio-economic status and the ability to influence data governance processes. Communities with strong civil society support and higher digital literacy demonstrate greater engagement with data rights advocacy. These groups often challenge biased classifications or petition for improved data representation. In contrast, poorer communities lack comparable mechanisms for resisting algorithmic disadvantage, reinforcing cyclical patterns of exclusion. These relational insights demonstrate that big data governance outcomes are shaped as much by social structures as by technical systems.

A case study from Bangkok illustrates how predictive policing algorithms were deployed to identify crime "hotspots," disproportionately targeting migrant worker districts. These neighborhoods received increased surveillance intensity, despite crime rates comparable to middle-income regions (Timilsina dkk., 2024). Community members reported that algorithmic design failed to account for demographic mobility and economic vulnerability, resulting in classification biases. The case demonstrates how algorithmic tools may unintentionally reflect and reproduce structural prejudices embedded in historical policing practices.

A Jakarta-based case reveals the impact of risk-scoring algorithms used in flood mitigation planning. Informal settlements situated along riverbanks were frequently categorized as

"unmanageable" zones, leading to forced relocations framed as resilience strategies. Residents argued that the algorithms ignored upstream infrastructural failures and unregulated commercial development that contributed significantly to flood risk. The case highlights how algorithmic representations can obscure structural drivers of vulnerability while prioritizing technocratic solutions.

The case studies underscore that algorithmic outputs must be interpreted within the socio-political contexts that shape their development and use. The systems observed in both cities embedded assumptions about risk, compliance, and productivity that aligned more closely with state priorities than with community needs (L. Wang & Hamid, 2024). These assumptions produced distorted representations of vulnerability and disorder, influencing governance strategies in ways that deepened existing inequalities. The explanation clarifies the link between algorithmic logic and policy discourse.

The feedback loops created by these systems reinforce stratification by directing additional investment or enforcement toward areas already categorized as high priority. Data-rich districts receive more targeted services, while data-poor districts remain peripheral to algorithmic attention (Heley dkk., 2024). This recursive process stems from the intersection of infrastructural bias and policy reliance on quantifiable indicators. The explanation demonstrates that algorithmic governance is not independent of social structures but deeply embedded within them.

The overall interpretation indicates that big data functions as a powerful but uneven instrument of urban governance, shaping policy through selective visibility and algorithmic classification (Soltani & Lee, 2024). The findings show that these systems do not inherently produce inequality but interact with existing socio-economic patterns in ways that may reinforce or amplify them. The interpretive insight emphasizes that big data governance must be assessed not solely on technical performance but on its socio-political consequences.

The results also highlight the need for governance models that incorporate transparency, community participation, and equity safeguards. The interpretation suggests that reforming data infrastructures requires more than technical improvements; it necessitates addressing structural inequities embedded in urban governance. These insights provide a foundation for developing more accountable and socially responsive forms of data-driven policymaking.

The findings reveal that big data infrastructures in Southeast Asian megacities operate unevenly across socio-economic landscapes, producing distinct patterns of visibility and invisibility that shape public policy outcomes. High-income districts benefit from dense sensor networks and frequent analytic updating, enabling governments to respond quickly to their infrastructural needs. Low-income and informal settlements suffer from sparse data representation, resulting in limited algorithmic attention and diminished policy prioritization. The unevenness of data infrastructure emerges as a core driver of stratified governance outcomes.

The study shows that policymakers often regard big data as a rational and efficient tool for managing urban complexity (Em & Sheludkov, 2024). This technocratic orientation shapes governance practices by elevating quantifiable indicators over lived experience. Community perspectives contradict this institutional narrative, as residents of marginalized districts describe misclassification, overpolicing, and exclusion driven by algorithmic outputs that fail to account for structural drivers of vulnerability. These tensions illustrate a widening gap between data-driven policy ideals and their social consequences.

The analysis demonstrates that algorithmic tools frequently embed pre-existing socio-political biases through risk-scoring mechanisms, hotspot mapping, and predictive modeling frameworks. These tools translate social inequalities into digital classifications that guide interventions in

policing, urban planning, and disaster response (Z. Wang dkk., 2024). Algorithmic processes reflect institutional priorities rather than representing objective assessments of urban need. The findings reveal a feedback loop in which data-rich zones receive more services while data-poor zones become increasingly marginalized.

The evidence further indicates that community agency and digital literacy influence how populations engage with or resist data-centric governance. Communities with strong organizational networks and knowledge of data practices can contest misrepresentation, demand transparency, and negotiate algorithmic outcomes. Communities lacking these resources remain vulnerable to unilateral classification practices. The findings highlight the interplay between digital infrastructures, social stratification, and political empowerment.

Prior scholarship on digital urbanism often emphasizes the promise of smart city technologies to enhance efficiency and promote equitable service distribution. The present findings complicate this narrative by demonstrating that big data governance frequently reinforces existing urban inequalities rather than ameliorating them. Earlier studies conducted in Euro-American contexts identified similar concerns regarding surveillance, algorithmic bias, and democratic accountability. The current study extends these insights by situating them within Southeast Asian megacities, where infrastructural unevenness and informal urbanism intensify the stratifying effects of algorithmic systems.

Research on smart policing provides a clear point of comparison, as scholars have documented how predictive models disproportionately target marginalized communities. The findings in this study confirm that predictive policing algorithms in Southeast Asia reflect comparable patterns, particularly in migrant and low-income districts. Differences arise in the intensity of deployment, given that Southeast Asian cities often adopt technologies through vendor-driven or donor-funded programs that lack robust accountability frameworks. These contextual differences sharpen the need for region-specific analyses.

Studies on algorithmic governance highlight the opacity of decision systems as a barrier to democratic oversight. The current findings reinforce this perspective by showing how limited transparency undermines public understanding of how classifications are determined. The absence of explainability mechanisms restricts communities' ability to contest policy decisions. The study contributes additional nuance by demonstrating how institutional capacity shapes the extent of algorithmic opacity, with centralized governance units providing more interpretive clarity than fragmented systems.

Scholarship on data justice emphasizes the political economy of data production, including who benefits from data infrastructures and who is excluded. The present study aligns with these theoretical concerns but adds empirical depth by documenting how spatial and socio-economic stratification materializes through algorithmic decisions. The comparison indicates that Southeast Asian megacities illustrate particularly acute forms of data injustice due to rapid urbanization, resource scarcity, and state reliance on technical rationality. The results expand existing debates by foregrounding regional specificities.

The findings signify that big data governance functions not merely as a technical enhancement to urban management but as an instrument that reshapes social relations between states and citizens. The unevenness of data infrastructures signals a shift toward forms of "algorithmic visibility" that determine who becomes legible to the state and whose needs remain obscured. These dynamics highlight the political nature of data-driven policymaking and challenge depictions of big data as an objective or neutral tool.

The patterns observed reveal that algorithmic classifications carry profound implications for how vulnerability, risk, and disorder are constructed. High-risk categorizations assigned to informal settlements reflect technological codifications of socio-economic inequality rather than accurate assessments of environmental or crime-related conditions. This signifies a broader process by which technocratic systems recast structural problems as individual or spatial deficiencies. The findings reflect the growing influence of algorithmic rationality in shaping public narratives of urban marginality.

The study also indicates that communities' experiences with big data governance are deeply intertwined with historical patterns of exclusion. Limited data coverage in low-income districts replicates long-standing infrastructural neglect, illustrating that algorithmic tools amplify rather than transcend existing inequalities. This continuity signifies that data-driven governance must be understood within broader trajectories of urban development, land politics, and policy priorities. The findings thus underscore the importance of embedding data systems within historically informed policy critique.

The result that communities with higher digital literacy can negotiate algorithmic systems signifies that data governance is also a site of emerging citizenship practices. Contesting risk labels, challenging misclassification, and demanding transparency constitute new forms of digital political engagement. This shift reflects an evolving dynamic in which urban residents seek to influence how data about their lives is constructed and used. The findings point to the emergence of data-savvy civic participation as a potential counterweight to algorithmic power.

The implications of these findings are significant for urban policymakers seeking to integrate big data into governance frameworks. The unevenness of data infrastructures indicates that algorithmic systems must not be treated as neutral governance tools. Policymakers need to recognize that big data can reproduce social stratification when deployed without equity safeguards. Urban governance strategies require explicit mechanisms that account for infrastructural disparities in order to prevent misrepresentation and exclusion.

The findings have implications for the design of regulatory frameworks governing data use. The opacity of algorithmic systems suggests that transparency mandates, public explainability requirements, and audit mechanisms are essential to ensure accountability. Without such frameworks, data-driven governance risks consolidating power in state or private actors without adequate democratic checks. The implications extend to vendor procurement policies, which must incorporate ethical and equity standards.

The results further imply that civil society organizations and community groups play a vital role in mediating the impacts of big data governance. Enhancing digital literacy and promoting public awareness of data rights can strengthen community capacity to engage with algorithmic governance. Empowering local advocacy groups is essential to creating more inclusive governance practices. Policymakers must therefore integrate community participation into data policy processes.

The implications also extend to theoretical debates on urban governance. The findings challenge assumptions that technological progress inherently leads to more just or equitable city management. Big data governance must be conceptualized as a socio-technical system shaped by political interests and institutional capabilities. The study underscores the need for critical frameworks that link algorithmic processes to broader socio-economic structures. These implications call for deeper interdisciplinary engagement.

The findings can be explained by the political economy of urban data infrastructures, which disproportionately invest in commercially valuable or politically prioritized districts. Sensor

deployment, data integration, and analytic capabilities follow patterns of capital accumulation, reinforcing existing spatial inequalities. These structural investment patterns explain why high-income districts become hyper-visible while informal settlements remain digitally marginalized. The explanation highlights the intersection of urban economics and data governance.

Institutional capacity offers a second explanation. Cities with fragmented governance structures rely heavily on external vendors whose tools prioritize technical efficiency over social considerations. Limited internal oversight and lack of data governance expertise enable algorithmic systems to operate with minimal accountability. This institutional vulnerability explains the persistence of opaque decision systems and biased classifications. The explanation points to the need for institutional reforms that enhance data governance skills.

Historical patterns of exclusion provide a further explanation for why data disparities persist. Informal settlements and migrant communities have long been excluded from infrastructural planning and public investment. Big data systems inherit these spatial and social biases because they draw from datasets that are incomplete, outdated, or selectively captured. This historical continuity explains why algorithmic outputs disproportionately disadvantage marginalized groups. The explanation reinforces the importance of contextualizing data governance within broader urban histories.

Community agency explains the differentiated responses to big data governance across socio-economic groups. Communities with established advocacy networks and digital resources are better positioned to challenge misclassification, contest surveillance practices, or demand accountability. Communities lacking these forms of capital remain vulnerable to algorithmic disadvantage. This social unevenness explains why big data impacts different populations in highly divergent ways. The explanation highlights the role of collective action in shaping governance outcomes.

Future research should investigate how alternative data governance models, such as participatory data infrastructures or community-driven mapping initiatives, might address the stratifying tendencies of existing big data systems. Exploring these models can provide evidence for more inclusive approaches that integrate local knowledge into algorithmic decision-making. Further studies should conduct comparative analysis across additional Southeast Asian cities to enhance the generalizability of findings.

Policymakers should prioritize reforms that embed equity considerations into data governance frameworks. Establishing independent algorithmic auditing bodies and mandating transparency for vendor-developed systems are essential steps for improving accountability. Municipalities should also invest in strengthening their internal data governance capacity to reduce reliance on commercial actors. These reforms represent actionable measures for mitigating stratified governance outcomes.

Civil society organizations must expand initiatives focused on digital rights education, algorithmic literacy, and data advocacy. Empowering communities to interpret and contest data-driven decisions can shift power dynamics within algorithmic governance. Localized training programs and participatory workshops can help marginalized groups navigate and articulate concerns about algorithmic systems. These interventions will contribute to building more democratic forms of digital urbanism.

Scholars and practitioners must collaborate to design interdisciplinary frameworks that bridge technical, sociological, and policy perspectives. Big data governance requires conceptual models capable of capturing its complex interactions with socio-political structures, infrastructural unevenness, and historical inequalities. Collaboration across fields will yield more nuanced insights

and facilitate the development of equitable governance practices. The pursuit of such interdisciplinary engagement represents a key direction for future work.

CONCLUSION

The most significant finding of this study lies in the demonstration that big data governance in Southeast Asian megacities does not merely reflect urban inequalities but actively shapes and magnifies them through uneven infrastructural visibility and algorithmic classification. The research shows that data intensity, sensor distribution, and predictive modeling practices create differentiated patterns of inclusion and exclusion that align closely with socio-economic hierarchies. Big data systems emerge not as neutral technological tools but as socio-political infrastructures that selectively render certain populations legible while obscuring others, thereby reconfiguring the terrain of urban citizenship and public policy attention. This differentiated insight departs from dominant narratives portraying big data as universally transformative and reveals instead its embeddedness within existing urban power structures.

The research contributes conceptual and methodological value by integrating critical data studies with urban governance analysis through a multi-layered qualitative approach. The combination of document analysis, expert interviews, and digital ethnography provides a comprehensive interpretive lens that captures the interplay between technical systems, political institutions, and community experiences. This framework advances scholarship by demonstrating how algorithmic rationalities are co-produced through institutional capacities, vendor agendas, and socio-spatial histories rather than originating solely from technological design. The study introduces a nuanced conceptual model of "algorithmic visibility" that helps explain how stratification is reproduced through data infrastructures, offering scholars and policymakers an analytic tool for assessing governance outcomes.

The study remains limited by its focus on three megacities, reliance on qualitative data, and restricted access to proprietary algorithms, which constrains the capacity to fully analyze technical decision pathways. Future research should incorporate mixed-method designs that integrate computational audits, spatial modeling, and cross-regional comparisons to further interrogate how data architectures operate across different governance regimes. Broader ethnographic engagement with marginalized communities would deepen understanding of how residents interpret, resist, or adapt to data-driven governance. Expanding interdisciplinary collaboration between technologists, urban planners, and social scientists will be essential for developing equitable data governance frameworks capable of addressing the stratifying forces identified in this study.

AUTHORS' CONTRIBUTION

Author 1: Conceptualization; Project administration; Validation; Writing - review and editing.
Author 2: Conceptualization; Data curation; In-vestigation.
Author 3: Data curation; Investigation.

REFERENCES

Alka, T. A., Sreenivasan, A., & Suresh, M. (2024). Wheel of change: A systematic literature review on innovation and entrepreneurship in micro mobility solutions. *Transport Economics and Management*, *2*, 154–168. https://doi.org/10.1016/j.team.2024.06.004

Boussaa, D., & Madandola, M. (2024). Cultural heritage tourism and urban regeneration: The case of Fez Medina in Morocco. *Frontiers of Architectural Research*, *13*(6), 1228–1248. https://doi.org/10.1016/j.foar.2024.04.008

Cuervo, L. G., Villamizar, C. J., Osorio, L., Ospina, M. B., Cuervo, D. E., Cuervo, D., Bula, M. O., Zapata, P., Owens, N. J., Hatcher-Roberts, J., Martín, E. A., Piquero, F., Pinilla, L. F., Martínez-Herrera, E., & Jaramillo, C. (2024). Dynamic measurements of geographical accessibility considering traffic congestion using open data: A cross-sectional assessment for haemodialysis services in Cali, Colombia. *The Lancet Regional Health - Americas*, *34*, 100752. https://doi.org/10.1016/j.lana.2024.100752

Dwivedi, Y. K., Jeyaraj, A., Hughes, L., Davies, G. H., Ahuja, M., Albashrawi, M. A., Al-Busaidi, A. S., Al-Sharhan, S., Al-Sulaiti, K. I., Altinay, L., Amalaya, S., Archak, S., Ballestar, M. T., Bhagwat, S. A., Bharadwaj, A., Bhushan, A., Bose, I., Budhwar, P., Bunker, D., … Walton, P. (2024). "Real impact": Challenges and opportunities in bridging the gap between research and practice – Making a difference in industry, policy, and society. *International Journal of Information Management*, *78*, 102750. https://doi.org/10.1016/j.ijinfomgt.2023.102750

Em, P. P., & Sheludkov, A. V. (2024). The fluctuating mosaic of socio-spatial inequalities in central Pyongyang under the pressures of marketization. *Habitat International*, *150*, 103135. https://doi.org/10.1016/j.habitatint.2024.103135

Esquivel García, C. L., & Toro-García, G. L. (2024). Multidimensional energy poverty in Colombia: A department-level review from 2018 to 2022. *Heliyon*, *10*(14), e34395. https://doi.org/10.1016/j.heliyon.2024.e34395

Glavovic, B. (2024). 7.18—Governance Experiences and Prospects in Estuarine and Coastal Communities. Dalam D. Baird & M. Elliott (Ed.), *Treatise on Estuarine and Coastal Science (Second Edition)* (hlm. 411–447). Academic Press. https://doi.org/10.1016/B978-0-323-90798-9.00129-3

Helen, N., & Ellisa, E. (2024). Investigating the knowledge commons practice in high-density low-income residential urban Kampung during COVID-19 pandemic. *Cities*, *148*, 104901. https://doi.org/10.1016/j.cities.2024.104901

Heley, J., Sanders, A., Caerwynt, F., Zaidi, N., & Power, S. (2024). The royal Welsh agricultural society: Patronage and the reproduction of elites in rural Wales. *Journal of Rural Studies*, *108*, 103291. https://doi.org/10.1016/j.jrurstud.2024.103291

Karyamsetty, H. J., Khan, S. A., & Nayyar, A. (2024). Chapter 9—Envisioning toward modernization of society 5.0—A prospective glimpse on status, opportunities, and challenges with XAI. Dalam F. Al-Turjman, A. Nayyar, M. Naved, A. K. Singh, & M. Bilal (Ed.), *XAI*

Based Intelligent Systems for Society 5.0 (hlm. 223–267). Elsevier. https://doi.org/10.1016/B978-0-323-95315-3.00005-X

Lay Maw, T. T., & Seo, D. (2024). Historical geographies of grid city development: Mandalay from Burma to Myanmar. *Journal of Historical Geography*, *86*, 133–148. https://doi.org/10.1016/j.jhg.2024.07.005

Li, H., Liu, Y., & Lu, Z. (2024). Can digital factor participation create opportunities for "rags to riches"? Based on the perspective of intergenerational income mobility. *International Review of Economics & Finance*, *96*, 103708. https://doi.org/10.1016/j.iref.2024.103708

Li, J., & Huang, Q. (2024). Navigating Ethical Dilemmas in AI-Enhanced Education: *International Journal of Knowledge Management*, *21*(1). https://doi.org/10.4018/IJKM.382384

Li, Y., Zhu, P., Mlecnik, E., Qian, Q. K., & Visscher, H. J. (2024). Dissemination, manipulation or monopolization? Understanding the influence of stakeholder information sharing on resident participation in neighborhood rehabilitation of urban China. *Land Use Policy*, *147*, 107359. https://doi.org/10.1016/j.landusepol.2024.107359

Liu, X., Pei, T., Wang, X., Liu, T., Fang, Z., Jiang, L., Jiang, J., Yan, X., Wu, M., Peng, Y., Ge, D., Gao, X., Song, C., & Chen, J. (2024). Travel flow patterns of diverse population groups and influencing built environment factors: A case study of Beijing. *Cities*, *151*, 105096. https://doi.org/10.1016/j.cities.2024.105096

Mazorra Rodríguez, Á. (2024). Social inequality and residential segregation trends in Spanish global cities. A comparative analysis of Madrid, Barcelona, and Valencia (2001-2021). *Cities*, *149*, 104935. https://doi.org/10.1016/j.cities.2024.104935

Mohamad, U. H. (2025). Chapter 6—Comparative analysis of AI and nanotech approaches for pandemic prediction. Dalam A. Ahmadian, F. Ghaemi, A. K. Yadav, M. J. Ebadi, & S. Salahshour (Ed.), *The Prediction of Future Pandemics* (hlm. 69–104). Elsevier. https://doi.org/10.1016/B978-0-443-33871-7.00006-4

Monlezun, D. J. (2024). 1—Power and artificial intelligence: Transformation of the global public health ecosystem. Dalam D. J. Monlezun (Ed.), *Responsible Artificial Intelligence Re-engineering the Global Public Health Ecosystem* (hlm. 1–65). Morgan Kaufmann. https://doi.org/10.1016/B978-0-443-21597-1.00001-9

Monlezun, D. J. (2025). Chapter 5—Quantum AI for public health. Dalam D. J. Monlezun (Ed.), *Quantum Health AI* (hlm. 125–154). Academic Press. https://doi.org/10.1016/B978-0-443-33353-8.00003-5

Musarat, M. A., Alaloul, W. S., Khan, M. H. F., Ayub, S., & Guy, C. P. L. (2024). Evaluating cloud computing in construction projects to avoid project delay. *Journal of Open Innovation:*

Technology, Market, and Complexity, *10*(2), 100296. https://doi.org/10.1016/j.joitmc.2024.100296

Noor, F., Saleem, M. A. U., Rafique, A., Danish, M. A. U., Bano, F., Shehzad, M. S. U., Noor, A., Fatima, I., Kamal, M. A., Muzammil, A., & Rehman, A. (2026). Chapter 13—Computational tools and techniques for disease modeling: Bridging the gap. Dalam S. N. Rai, S. K. Singh, & V. Singh (Ed.), *Advancements in Modeling-Based Therapeutics and Technology for Chronic Diseases* (hlm. 373–418). Academic Press. https://doi.org/10.1016/B978-0-443-33877-9.00017-3

Ojah, K., & Kodongo, O. (2024). Effective financial inclusion and the need to put the horse before the cart: Saving! *International Review of Financial Analysis*, *96*, 103737. https://doi.org/10.1016/j.irfa.2024.103737

O'Meara, L., Sison, C., Isarabhakdi, P., Turner, C., & Harris, J. (2024). 'Whatever we have is what we eat': How marginalised urban populations in the Philippines and Thailand experienced their food environments, food security and diets through COVID-19. *Health & Place*, *88*, 103279. https://doi.org/10.1016/j.healthplace.2024.103279

Ong, J. C. L., Seng, B. J. J., Law, J. Z. F., Low, L. L., Kwa, A. L. H., Giacomini, K. M., & Ting, D. S. W. (2024). Artificial intelligence, ChatGPT, and other large language models for social determinants of health: Current state and future directions. *Cell Reports Medicine*, *5*(1), 101356. https://doi.org/10.1016/j.xcrm.2023.101356

Santra, D. (2025). Artificial intelligence in urban health epidemic management. Dalam *Advances in Computers*. Elsevier. https://doi.org/10.1016/bs.adcom.2025.10.001

Schmid, F. B., Hersperger, A. M., Grêt-Regamey, A., & Kienast, F. (2024). Effects of different land-use planning instruments on urban shrub and tree canopy cover in Zurich, Switzerland. *Urban Forestry & Urban Greening*, *94*, 128272. https://doi.org/10.1016/j.ufug.2024.128272

Sha, K., Taeihagh, A., & De Jong, M. (2024). Governing disruptive technologies for inclusive development in cities: A systematic literature review. *Technological Forecasting and Social Change*, *203*, 123382. https://doi.org/10.1016/j.techfore.2024.123382

Soltani, A., & Lee, C. L. (2024). The non-linear dynamics of South Australian regional housing markets: A machine learning approach. *Applied Geography*, *166*, 103248. https://doi.org/10.1016/j.apgeog.2024.103248

Thakur, R., Baghel, M., Bhoj, S., Jamwal, S., Chandratre, G. A., Vishaal, M., Badgujar, P. C., Pandey, H. O., & Tarafdar, A. (2024). CHAPTER 8—Digitalization of livestock farms through blockchain, big data, artificial intelligence, and Internet of Things⍟. Dalam A. Tarafdar, A. Pandey, G. K. Gaur, M. Singh, & H. O. Pandey (Ed.), *Engineering Applications*

in Livestock Production (hlm. 179–206). Academic Press. https://doi.org/10.1016/B978-0-323-98385-3.00012-8

Timilsina, R. R., Jena, P. R., Rahut, D. B., & Managi, S. (2024). Towards parity: Examining the closing gender gap on electricity access in India using data from 1998 to 2021. *Energy for Sustainable Development*, *80*, 101450. https://doi.org/10.1016/j.esd.2024.101450

Villar, E., Francke, P., & Loewenson, R. (2024). Learning from Perú: Why a macroeconomic star failed tragically and unequally on Covid-19 outcomes. *SSM - Health Systems*, *2*, 100007. https://doi.org/10.1016/j.ssmhs.2023.100007

Wang, L., & Hamid, M. O. (2024). The transformation of neoliberalism: A critical analysis of shadow education governance policy in China. *International Journal of Educational Research*, *127*, 102422. https://doi.org/10.1016/j.ijer.2024.102422

Wang, W. W. (2024). Contextualizing Personal Information: Privacy's Post-Neoliberal Constitutionalism and Its Heterogeneous Imperfections in China. *Computer Law & Security Review*, *55*, 106030. https://doi.org/10.1016/j.clsr.2024.106030

Wang, Z., Hu, T., Liu, J., Xia, B., & Chileshe, N. (2024). Spatial differences, evolutionary characteristics and driving factors on economic resilience of the construction industry: Evidence from China. *Engineering, Construction and Architectural Management*, *32*(10), 6888–6916. https://doi.org/10.1108/ECAM-01-2024-0021