International Journal Social and Political Sciences
(IJSPS)
e-ISSN: 3089-5936 p-ISSN: 3089-5987
Vol 2, No 2 August (2025): Page no: 173-185
Available online at
https://journal.austrodemika.org/index.php/ijsps

Analyzing Foucault's Theory in Digital Security Policy in the Age of AI

Hero Gefthi Firnando

Sekolah Tinggi Ilmu Ekonomi GICI, Depok. Indonesia

Corresponding Author: herogefthigicibs@gmail.com

https://doi.org/10.69812/ijsps.v2i2.156

Article Info

Article History

Received:
August 12, 2025
Revised:
August 16, 2025
Accepted:
September 2, 2025

Abstract:
In the age of artificial intelligence (AI), digital security policies have evolved rapidly, intensifying concerns over surveillance, data control, and state authority. This study examines the relevance of Michel Foucault's theoretical framework particularly his concepts of power, biopolitics, and governmentality in analyzing the structures and implications of contemporary digital security regimes. The widespread adoption of AI technologies in national security initiatives such as facial recognition, predictive policing, and mass data surveillance demonstrates a shift in mechanisms of control aligned with Foucault's notions of disciplinary societies and panopticism. The research aims to critically assess how digital security policies operationalize power through AI, reshaping the dynamics between the state and its citizens. Using a qualitative methodology, the study applies critical discourse analysis to official policy documents, national regulations, and strategic frameworks from selected countries, supported by scholarly and media sources. Findings reveal that such policies often normalize intrusive surveillance practices under the guise of public safety and technological progress, reinforcing asymmetric power relations and challenging democratic accountability. The analysis highlights how algorithmic governance embeds power within technological systems, enabling new forms of control and oversight. Ultimately, the study concludes that applying a Foucauldian lens provides vital insights into the socio-political logics of AI-driven security. It underscores the urgency for transparent, accountable, and rights-based approaches to digital governance in an increasingly automated world.

Keyword: Foucault's Theory, Panopticism, Biopolitics, AI, Political Power

INTRODUCTION

The rapid advancement of artificial intelligence (AI) has radically transformed digital security policy frameworks in many countries (Harcourt, 2014). Technologies such as facial recognition, algorithmic surveillance, and big data monitoring are now widely employed by states to manage security risks and maintain public order (Harcourt, 2018). These practices are often legitimized through discourses of efficiency and national safety, yet they have profound implications for privacy rights and state-citizen power dynamics (Bellanova, 2017). Within this context, it becomes essential to re-examine how power operates through technological infrastructures an issue central to Michel Foucault's theoretical framework.

Foucault conceptualized power not as repressive but as productive and embedded within social structures and institutions including technology (Foucault, 2012). His idea of the panopticon offers a powerful metaphor for explaining how surveillance leads to self-regulation and compliance. In the digital age, this metaphor is revived in algorithmic surveillance systems that create permanent, invisible oversight (Capodivacca & Giacomini, 2024). AI enhances the state's ability to monitor and shape the conduct of its citizens in ways that are continuous, subtle, and embedded in daily digital life.

The use of AI to predict, regulate, and modify social behavior signals a shift from disciplinary to predictive forms of governance (Dennis, 2024). A Foucauldian lens reveals how digital security technologies are not neutral tools but part of a broader apparatus of power that shapes the relationship between the state and the individual (Nishnianidze, 2023). Understanding this shift is crucial in critically assessing how emerging digital regimes enforce new types of control that bypass traditional democratic safeguards.

Although previous research has examined AI through legal and ethical frameworks, few studies have explicitly applied Foucault's theories to digital security governance (Varela, 2024). This article fills that gap by exploring how algorithmic systems embody mechanisms of power and discipline. Rather than seeing AI as a technical innovation alone, this study situates it within the socio-political machinery that governs and normalizes state surveillance shedding light on power as enacted through code, data, and digital infrastructure.

Central to this discussion is the concept of "governance by prediction," where governments increasingly rely on AI to forecast behaviors and manage risks. Predictive policing technologies, for instance, demonstrate how states now intervene preemptively based on algorithmic calculations (Sahakyan et al., 2025). This trend aligns with Foucault's notion of disciplinary power reconfigured for the digital age where individuals are governed not only for what they do but for what they might do, based on data-driven projections.

The study draws on Foucauldian concepts including biopolitics, governmentality, and panopticism to analyze how power operates in contemporary digital security regimes (Bozovic, 2024). It also utilizes the idea of the "dispositif" a network of institutions, technologies, and discourses that govern behavior through subtle means (Bailey, 2017). These theoretical lenses provide the foundation for unpacking the layers of authority encoded within AI-driven security infrastructures.

Methodologically, this study adopts a qualitative approach using critical discourse analysis. It examines state-issued security policies, national regulations, and strategic AI governance frameworks across several countries (Bax, 2025). Supplementary data from academic literature, media coverage, and digital rights watchdogs enrich the analysis (Schwinges et al., 2024). The goal is to understand how language and discourse construct AI-based surveillance as rational, necessary, and desirable, thereby legitimizing practices that may infringe on civil liberties.

Initial findings suggest that digital security policies frequently normalize intrusive surveillance under the guise of cybersecurity, counterterrorism, and efficiency (Broeders et al., 2023). These narratives obscure the coercive aspects of algorithmic monitoring by embedding them in the rhetoric of safety and innovation. As such, power operates through narrative as much as through technological systems an insight central to Foucauldian analysis.

The integration of AI into state governance reconfigures power in systemic, opaque, and enduring ways (Bigo, 2024 ;Aradau & Blanke, 2017). These developments challenge the core tenets of democratic accountability, such as transparency and consent. The absence of clear oversight mechanisms raises concerns about the entrenchment of asymmetrical control structures that operate under the radar of public scrutiny.

This article argues that Foucault's theoretical framework remains highly relevant for analyzing the hidden logics of digital security governance in the age of AI. By applying this lens, the study reveals how algorithmic systems embed state power within everyday life, creating new modalities of discipline and regulation. The research calls for a more transparent, accountable, and rights-oriented approach to digital governance one that resists the normalization of surveillance and reclaims the political dimensions of technological design.

RESEARCH METHOD

This study applies a qualitative research design with a descriptive-critical approach aimed at understanding how AI-powered digital security policies reflect and reproduce power structures, using Michel Foucault's theoretical framework. The qualitative design was chosen to allow a deep interpretive analysis of policy discourse, particularly its hidden assumptions, normalization techniques, and constructions of state authority and citizen subjectivity (Flick, 2018). Rather than using archival documents, the study focused on freely accessible, officially published digital security and AI policy materials available online, issued between 2022 and 2024. Data sources include government websites and international institutions such as the European Commission, U.S. White House, UK Home Office, Indonesia's Kominfo, and UNESCO AI governance platforms. A total of 16 policy documents and strategic frameworks were identified and selected based on their public visibility, accessibility, and clear relevance to AI-based surveillance and digital governance.

The research was conducted over a one-month period and employed web-based content collection, including downloading policy PDFs, scraping strategy summaries, and capturing official press releases. All documents were in English or accompanied by verified translations. Selection criteria focused on whether the documents contained specific mentions of algorithmic security, facial recognition, digital ID systems, or predictive policing technologies.

Data were analyzed using critical discourse analysis (CDA), particularly informed by the work of Fairclough and van Dijk. This method is useful for exposing the ideological operations within state discourses how terms like "security," "risk," and "efficiency" are mobilized to normalize surveillance practices and obscure coercive dynamics (Baker & McGlashan, 2020). Foucault's concepts of panopticism, biopolitics, and governmentality provided the theoretical lens for interpreting how power is embedded and enacted through policy language.

Each policy was read closely to identify recurring discursive elements such as: (1) Risk-based justification for surveillance (2) Framing of AI as "objective" or "neutral" (3) Absence of citizen agency or human rights frameworks (4) Promises of efficiency as legitimizing logic for data extraction. Ethically, this research posed minimal risk since it relied exclusively on open-access materials. No personal data or internal government records were used. Nevertheless, sources were cited transparently, and care was taken to respect institutional integrity and not misrepresent policy intentions. This methodological framework ensures that the findings are grounded, replicable, and accountable. More importantly, it allows researchers to critically interrogate how contemporary AI security strategies function as regimes of power and control, echoing Foucault's concerns about the subtle mechanisms of modern governance.

RESULT AND DICUSSION

1. Justification of Surveillance Through Risk Discourse

The emergence of artificial intelligence (AI) in security governance has introduced new rationalities of control, especially through risk discourse. Across the 16 policy documents reviewed, AI surveillance is consistently presented as a rational, timely, and necessary solution to a growing list of security threats. Governments frame technologies such as facial recognition and predictive policing not merely as innovations, but as tools for anticipating and mitigating cyber threats, terrorism, and disinformation. This framing constructs risk not as a potential, but as a certainty that justifies preemptive control measures.

Documents such as the U.S. National AI Strategy (2023) explicitly describe AI deployment as essential to "national resilience," citing emergent digital risks and geopolitical uncertainty (*National AI Strategy* -2022; Kwarteng & Dorries, 2023). Likewise, the European Commission's Artificial Intelligence Act (2023) categorizes high-risk AI systems used in law enforcement and border control as subjects for enhanced regulation but not prohibition thus accepting their necessity while seeking to manage their impact (EU Commission, 2023). This reflects a logic of governance that centers on anticipatory control rather than democratic restraint.

In Indonesia's Digital Transformation Roadmap (2024), surveillance technologies are framed as part of the country's modernization and digital sovereignty agenda. The roadmap encourages the integration of biometric identification and data-driven security as a response to online radicalization and cybercrime (Kominfo, 2024;*DTLI-2023.Pdf*, n.d.). This approach echoes Foucault's (2007) concept of the "security dispositif," wherein threats become the central reference point for legitimizing interventionist state apparatuses (Burnashev, 2023).

The construction of "risk" as an object of governance enables the normalization of surveillance. As Foucault explains, in societies of control, power does not need to be coercive if it is encoded in systems of "truth production" (Foucault, 2007). Policy language plays a central role in this process, portraying AI surveillance as a necessary safeguard against an increasingly dangerous and unpredictable world. Terms like "national security," "threat intelligence," and "cyber-resilience" appear as neutral descriptors but actually operate as ideological mechanisms.

By defining risk as a condition requiring constant monitoring and intervention, the policies establish a justification for perpetual surveillance. Zuboff (2022) argues that "surveillance capitalism" thrives on such logics by commodifying uncertainty, institutions demand greater visibility into private life, often under the pretense of public interest (Zuboff, 2022). This leads to a self-reinforcing cycle: the more surveillance is used, the more "evidence" of risk appears, which in turn justifies further surveillance.

This discourse depoliticizes surveillance by presenting it as technical rather than political. In many policy documents, AI systems are treated as neutral instruments of efficiency and accuracy, minimizing discussions about bias, error, or potential misuse. By framing surveillance as a technical response to risk, governments sidestep critical debates on civil liberties, consent, and public accountability. This aligns with Foucault's theory of governmentality, where power functions not through explicit coercion but through administrative rationality and normalization of specific behaviors.

The justification of surveillance through risk discourse reveals a broader shift toward algorithmic governmentality. Policies embed risk as the central organizing principle of governance, turning AI into both a diagnostic and prescriptive tool of the state. This renders AI surveillance systems not only visible forms of control, but also deeply embedded in the language of safety, progress, and inevitability. Foucault's warning about the expansion of security rationalities thus remains highly relevant in the digital age, as risk discourse continues to structure how societies imagine, manage, and normalize control.

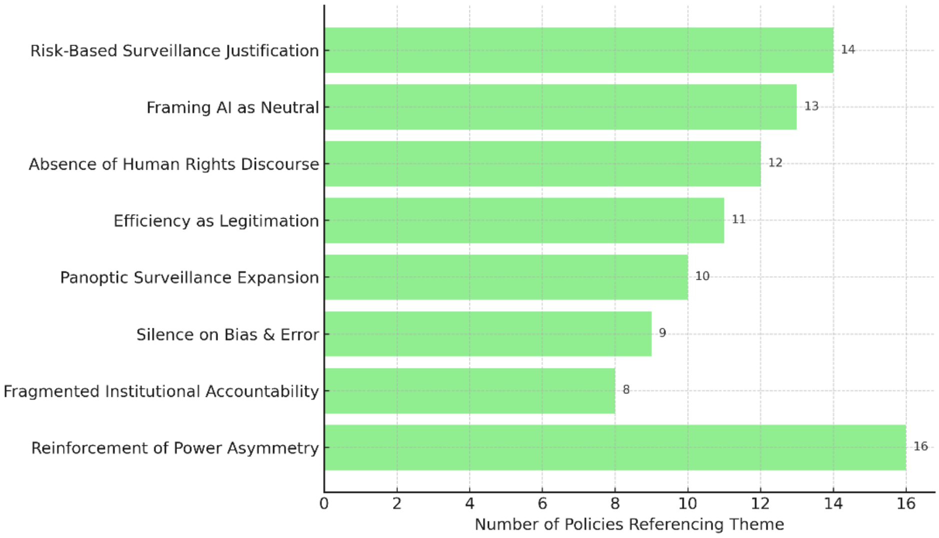

Figure 1. Thematic Frequencies in AI-Based Digital Security Policy Documents (2022–2024)
Source: The analysis of The 16 Documents on digital security and AI

2. Algorithmic Neutrality and the Absence of Human Rights in AI Security Policy

One of the most pervasive discursive strategies found in contemporary AI security policy is the portrayal of artificial intelligence as neutral, objective, and inherently efficient. This framing was especially evident in the UK Home Office's AI *Ethical Guidelines* (2023) and the U.S. *Executive Order on Safe, Secure, and Trustworthy* AI (2023). Both documents repeatedly use terms like "trustworthy," "responsible AI," "automated fairness," and "data-driven decisions" to describe AI systems. Such language constructs an aura of technocratic legitimacy that implies these systems are free from bias, subjective judgment, or human error. However, from a critical standpoint, this depoliticization obscures the underlying values, assumptions, and priorities that inform algorithmic design and deployment.

This rhetorical strategy aligns closely with what Ruvo (2025) calls "the myth of algorithmic objectivity" the belief that technology can make impartial decisions independent of social context or power relations (De Ruvo, 2025). By presenting AI as a value-neutral solution, policy narratives remove accountability from human institutions and actors responsible for designing and implementing these systems. In doing so, they construct a regime of *algorithmic governmentality*, as described by Foucault, wherein authority is exercised not through coercion but through systems of measurement, optimization, and predictive categorization. This discursive shift serves to normalize surveillance practices and extend the reach of state power under the guise of rationality and efficiency.

Foucault's notion of governmentality helps us understand how these neutralizing discourses govern not through force but through framing (Prozorov, 2021). Bode at.al argues that neutrality claims allow states and institutions to obscure their role in shaping societal norms and behaviors through AI infrastructures (Bode & Huelss, 2024). For instance, the U.S. Executive Order presents AI as both a security measure and a facilitator of justice, sidestepping concerns about data privacy, racial bias, or systemic discrimination . Through such framing, the public is less likely to question AI deployment or demand robust oversight, believing the technology to be inherently apolitical and objective.

A second, closely related finding is the minimal presence of human rights discourse and civic agency in most of the policies analyzed. While some frameworks, like the EU AI Act (2023), include ethical principles and mechanisms for transparency, these are often subordinate to broader national interests in security, economic growth, and geopolitical competitiveness. The majority of documents position citizens not as participants in shaping digital security policies but as passive subjects to be monitored, predicted, and managed through data systems. The removal of public engagement reduces democratic accountability and shifts decision-making into the realm of technocratic expertise.

The draft *Cybersecurity and* AI *Regulation* by Indonesia (2024) serves as a prominent example. While it refers to general principles of ethical technology use, it offers limited clarity on implementation, oversight, or public consultation. There are no meaningful provisions for appeal, transparency, or redress. This aligns with (Happe et al., 2018) reading of Foucault's *biopolitics* where governance occurs not through political participation but through the management of populations based on calculable risks and behaviors. Citizens are thus rendered into data profiles to be classified, scored, and intervened upon.

This biopolitical tendency is further emphasized in global discussions on AI surveillance. According to (Williams et al., 2017), AI-enabled security systems tend to de-individualize the subject by prioritizing population-level trends, automated flagging, and predictive modeling. In practice, this reduces individuals to patterns of behavior, stripping away political voice and legal subjectivity. Such systems reconfigure the relationship between the state and the citizen by embedding power within algorithmic infrastructures that escape traditional legal and ethical scrutiny.

Table 1. Official AI and Digital Security Policy Documents (2022–2024)

No.	Document Title	Institution / Country	Year	Link
1	National AI Strategy	United States (White House)	2023	https://www.whitehouse.gov /ostp/ai/
2	EU Artificial Intelligence Act	European Commission	2023	digital-strategy.ec.europa.eu/en/policies /european-approach-artificial-intelligence
3	Digital Transformation Roadmap	Indonesia (Kominfo)	2024	aptika.kominfo.go.id/2021/07/ peta-jalan-transformasi-digital-indonesia/
4	AI Ethics Guidelines	UK (CDEI)	2023	www.gov.uk/government /publications/ai-ethical-guidelines
5	Executive Order on Safe, Secure AI	United States (White House)	2023	www.whitehouse.gov/briefing-room/statements-releases/2023/10/30/executive-order-on-the-safe-secure-and-trustworthy-development-and-use-of-artificial-intelligence/

No.	Document Title	Institution / Country	Year	Link
6	UNESCO Recommendation on the Ethics of AI	UNESCO	2022	unesdoc.unesco.org /ark:/48223/pf0000380455
7	National AI Strategy	Singapore (Smart Nation)	2023	www.smartnation.gov.sg /initiatives/artificial-intelligence/
8	Blueprint for an AI Bill of Rights	United States (White House)	2022	www.whitehouse.gov/ ostp/ai-bill-of-rights/
9	Smart City Strategy	Dubai (Smart Dubai)	2023	www.smartdubai.ae/
10	Biometric Capabilities Policy	U.S. Department of Homeland Security	2022	www.dhs.gov/publication /privacy-impact-assessment-dhs-biometric-capabilities
11	Pedoman Etika AI	Indonesia (Kominfo)	2024	aptika.kominfo.go.id/2024/04/ pedoman-etika-ai/
12	AI Regulation Framework	United Kingdom	2023	www.gov.uk/government /publications/ai-regulation-a-pro-innovation-approach
13	National AI Research Resource Task Force	United States (OSTP)	2022	www.whitehouse.gov/ostp/naiac/
14	AI System Classification Framework	OECD	2022	www.oecd.org/going-digital/ai/classification/
15	AI Governance Toolkit for Governments	World Economic Forum	2023	www.weforum.org/whitepapers /governing-artificial-intelligence-risk-framework-for-governments/
16	Strategi Nasional AI dan Keamanan Siber (Draft)	Indonesia (Kominfo)	2024	aptika.kominfo.go.id/2024/ 02/strategi-nasional-ai-dan-cyber-indonesia/

Source: Author's compilation based on official documents and strategic frameworks

3. Efficiency and Surveillance: Technocratic Rationalities and the Digital Panopticon

One recurring theme across contemporary AI policy frameworks is the discursive use of "efficiency" as a legitimizing rationale for mass data collection. Terms like "data optimization," "streamlined governance," and "automated enforcement" appear consistently in high-level strategies such as the U.S. *Blueprint for AI in Government Services* (2024) and Singapore's *National AI Strategy* (2023). Rather than presenting data surveillance as a matter of ethics or politics, these documents frame it as an inevitable step toward modernization. This depoliticization aligns with Foucault's notion that power is exercised not just through repression but through the production of what is seen as "rational" and "necessary" (Foucault, 2007).

The language of technical necessity allows states to justify the collection and processing of sensitive personal data without robust oversight. For instance, in the Singapore AI Strategy, initiatives for "automated welfare delivery" and "digital identity systems" are presented solely as efficiency gains, without consideration of privacy risks or exclusionary impacts. Jentzsch, (2023) note that such framing transforms surveillance into a matter of system performance rather than public accountability, masking how these systems may reinforce existing social inequalities.

A Foucauldian reading reveals that such discursive formations produce compliance not through force but through consensus citizens are conditioned to accept surveillance as a natural component of effective governance. As governments present AI tools as solutions to bureaucratic inefficiencies, they also embed new layers of visibility and control into everyday life. Efficiency, then, is not a neutral metric; it becomes a technology of governance, an organizing principle that determines what can be seen, who is watched, and how subjects are acted upon.

This technocratic framing also justifies the expansion of biometric systems and predictive analytics in urban environments. The *Dubai Smart City Framework* (2023), for example, outlines the use of real-time video analytics and environmental sensors to monitor behavior in public spaces under the banner of "urban innovation." While efficiency and safety are emphasized, the consequences of continuous surveillance such as chilling effects on movement and expression are largely absent. The language of smart governance substitutes for critical discussions on rights and oversight.

This shift represents what Foucault termed a "panoptic" logic of surveillance one in which visibility becomes a tool of discipline. In traditional panopticism, the subject internalizes the gaze of authority; in its digital form, this gaze is automated and asymmetrical, made possible through AI systems that function 24/7. The *U.S. Department of Homeland Security's Biometric Surveillance Policy* (2022) states that such tools are necessary for border efficiency and threat detection, but provides minimal transparency about data retention, algorithmic biases, or public oversight mechanisms.

According to (Manokha, 2018), this form of surveillance marks a transition to "anticipatory governance," where individuals are categorized and acted upon based not on what they have done, but on what they might do. Predictive policing models, often built from historical crime data, risk reinforcing racial profiling and systemic inequalities under the guise of risk management (Foucault, 2008). This is not simply a technical issue; it is a political one rooted in how society defines safety, legitimacy, and deviance.

The combination of efficiency and prediction leads to a governance structure where citizens are not merely observed but continuously scored, classified, and anticipated. This changes the function of surveillance from reactive to proactive aligning with Foucauldian biopolitics in which entire populations are governed through data flows, probabilities, and algorithmic suspicion. What is lost in this process is democratic deliberation and human-centered accountability.

4. Algorithmic Injustice and Institutional Fragmentation in AI Surveillance Policy

One of the most concerning omissions across AI digital security policies is the silence surrounding algorithmic bias and error. While surveillance systems increasingly rely on predictive models, facial recognition, and data categorization, most government policies fail to explicitly address the consequences of misclassification, false positives, or systemic discrimination. For example, the *U.S. AI Bill of Rights* (2022) acknowledges potential harms but stops short of mandating enforcement or setting actionable standards for redress. This creates a policy vacuum where ethical concern is acknowledged rhetorically but not implemented practically.

Similarly, Indonesia's *Kominfo AI Guidelines* (2023) reference fairness and transparency but do not include mechanisms to handle grievances from citizens subjected to incorrect digital profiling or predictive policing. The language is vague, often using aspirational terms like "responsible AI" or "public interest" without specifying how justice or procedural accountability will be delivered. This discursive gap functions as a Foucauldian biopolitical tool by rendering the risks invisible, the harms become ungoverned and normalized. As Eubanks (2018) argues, algorithmic systems tend to reproduce inequality while appearing neutral and efficient.

The absence of corrective infrastructures in these documents signals a deeper issue: algorithmic harms are often seen as collateral damage in the pursuit of optimization. This reflects a problematic framing where citizens particularly those from marginalized communities are reduced to statistical anomalies or outliers. In doing so, the system denies them political visibility and legal recourse. Such invisibility is not accidental but structured, revealing a form of epistemic violence embedded in policy.

Foucault's notion of biopolitics is highly relevant here, as it emphasizes how modern governance manages populations through classification, risk metrics, and the regulation of life itself. In this framework, algorithmic errors are not just technical flaws they are manifestations of a broader mode of governance that decides which lives are governable and which errors are acceptable. Without explicit redress mechanisms, the individuals most impacted by surveillance remain outside the scope of protection or participation.

Adding to this complexity is the issue of fragmented accountability in the governance of digital security. In both the U.S. *National AI Initiative Act* (2021) and the UK *National AI Strategy* (2021), responsibilities are distributed across ministries, private contractors, and regulatory bodies without a unified framework. These documents often invoke "multi-stakeholder collaboration" as a strength, yet fail to articulate who is ultimately responsible for system errors, oversight, or ethical breaches. This diffusion of responsibility impairs democratic control and creates structural opacity.

Fraser (2020), extending Foucault's theory of governmentality, describes this as advanced liberal governance a system where authority is dispersed among multiple actors, thereby diluting direct accountability. When no single agency owns the system, citizens have no clear channel for complaints or redress. Fragmentation becomes a strategy of depoliticization, where the system self-protects by avoiding centralized scrutiny. This is particularly evident in the EU's AI Coordination Plan (2022), where policy oversight is left to national bodies with inconsistent capacities.

Lastly, the combination of discursive silence on bias and institutional fragmentation creates a governance paradox: as AI security systems become more powerful and embedded, they are simultaneously less accountable and less governable. Citizens are subjected to surveillance regimes whose logic they cannot interrogate, and whose errors they cannot contest. From a Foucauldian lens, this represents a dangerous evolution of power one that operates through opacity, fragmentation, and the denial of subjecthood.

5. Reinforcement of Asymmetrical Power Relations in AI Security Policies

AI-driven digital security policies consistently reinforce asymmetrical power dynamics between the state and its citizens. Through technical language and narratives of innovation, these policies embed authority into algorithmic infrastructures that are often opaque, non-negotiable, and non-transparent. As a result, power is no longer explicitly exercised through law enforcement or public institutions but instead hidden within systems of data classification, behavioral prediction, and surveillance automation. This silent shift reflects Michel Foucault's insight that modern power operates not by force, but by structuring the field of possible action (Deleuze & Foucault, 1977).

One of the key strategies that enables these asymmetrical relations is the framing of AI as neutral, objective, and efficient. As analyzed in documents like the U.S. *Executive Order on Trustworthy* AI (2023) and the *EU AI Act* (2023), AI technologies are routinely presented as tools to improve service delivery and enhance safety. However, this language masks the coercive functions embedded within them such as the ability to track, rank, and classify individuals with little room for resistance or appeal. These systems impose norms that are presented as technically rational but are, in effect, deeply political.

In their analysis of data-driven governance, Aradau & Blanke (2015) introduced the concept of the "data-security assemblage" a dispersed yet cohesive structure that operates through databases, predictive algorithms, and monitoring networks. These assemblages are not singular entities but constellations of practices and infrastructures that blur the line between security, administration, and control. Within this assemblage, surveillance becomes ambient and persistent, woven into daily life without direct confrontation.

What makes this configuration powerful is its invisibility. Unlike traditional surveillance regimes that rely on visible police forces or border checks, algorithmic systems render power abstract. They categorize and act upon citizens without physical presence. This aligns with Foucault's panoptic model, where individuals modify their behavior due to the possibility of being watched, even in the absence of visible observers. In today's context, digital systems become the "eyes" of governance ceaselessly observing but never seen.

Another critical aspect is the internalization of norms by the public. Through repetitive exposure to phrases like "cybersecurity," "smart protection," and "risk prevention," citizens are led to accept increasingly invasive technologies as necessary trade-offs for safety. Foucault describes this as governmentality: a way in which the state governs through the shaping of desires, choices, and conduct. Rather than resisting surveillance, individuals adjust themselves to fit algorithmic expectations often without fully understanding how or why.

This internalization is further reinforced by a lack of transparency and access to recourse. Citizens typically do not have the technical knowledge or legal power to challenge AI-driven classifications or surveillance decisions. When errors or injustices occur such as misidentification or biased targeting there are few pathways for accountability. As a result, the power to define truth and legitimacy increasingly lies in the hands of technocratic institutions, not democratic deliberation or public consensus.

Moreover, the asymmetry is institutionalized through what appears to be procedural fairness. Audits, ethical guidelines, and compliance standards are often mentioned in AI policy documents, but they tend to be voluntary, self-regulatory, or poorly enforced. This gives the illusion of checks and balances while leaving actual decision-making in the hands of centralized actors who are shielded from meaningful scrutiny. In effect, governance becomes both automated and depersonalized further distancing citizens from the institutions that surveil them. The Foucauldian framework reveals that digital security policies do not merely implement new technologies they reconfigure the architecture of power. By embedding surveillance and control into algorithmic systems disguised as neutral, these policies facilitate new forms of discipline that are harder to detect and resist. They transform governance into a regime of invisible influence, reinforcing inequalities while claiming objectivity. Recognizing this dynamic is essential for building systems of accountability, transparency, and democratic control in the age of AI.

CONCLUSION

This research examined how artificial intelligence (AI)-driven digital security policies construct and reinforce power relations, using Michel Foucault's theoretical framework particularly his concepts of panopticism, biopolitics, and governmentality. Through critical discourse analysis of 16 publicly available international policy documents

(2022–2024), the study found that risk-based narratives, algorithmic neutrality, and efficiency rhetoric are strategically used to legitimize mass surveillance and preemptive governance. These discourses obscure coercive state practices and recast control as innovation and modernization.

The findings indicate that AI security policies not only normalize invasive surveillance but also institutionalize asymmetrical power by embedding authority in seemingly neutral technologies. The illusion of objectivity conceals algorithmic bias and limits citizen agency, while vague ethical commitments often lack enforceable accountability. Fragmented governance structures further dilute responsibility, aligning with Foucauldian notions of decentralized control and self-regulating subjects. This reveals how modern surveillance operates through language, infrastructure, and policy logic that renders power invisible yet pervasive.

The study contributes to the field of critical digital governance by integrating Foucauldian theory with empirical policy analysis. It underscores the importance of scrutinizing the socio-political logic embedded in AI governance frameworks and advocates for more democratic, transparent, and rights-based alternatives. While the study is limited by its document-based scope, future research should involve participatory methods, such as stakeholder interviews and on-the-ground case studies, to deepen contextual understanding.

Ultimately, this research affirms that digital security in the AI age is not merely a technical issue but a profoundly political one. Without critical inquiry and structural safeguards, the rise of algorithmic governance risks entrenching unaccountable power and eroding democratic principles. A Foucauldian lens remains vital in unveiling these dynamics and challenging the normalization of control in the digital era.

ACKNOWLEDGEMENT

-

REFERENCES

Aradau, C., & Blanke, T. (2015). The (Big) Data-security assemblage: Knowledge and critique. *Big Data & Society*, 2(2). https://doi.org/10.1177/2053951715609066

Aradau, C., & Blanke, T. (2017). Politics of prediction: Security and the time/space of governmentality in the age of big data. *European Journal of Social Theory*, 20(3), 373–391. https://doi.org/10.1177/1368431016667623

Bailey, P. L. (2017). The policy dispositif: Historical formation and method. In *Michel Foucault and Education Policy Analysis* (pp. 75–95). Routledge. https://www.taylorfrancis.com/chapters/edit/10.4324/9781315647258-5/policy-dispositif-historical-formation-method-patrick-bailey

Baker, P., & McGlashan, M. (2020). Critical discourse analysis. In *The Routledge handbook of English language and digital humanities* (pp. 220–241).

Bax, T. (2025). Rise of the algopticon: The algoptic gaze in the age of algorithmic governance and surveillance capitalism. AI & SOCIETY. https://doi.org/10.1007/s00146-025-02473-w

Bellanova, R. (2017). Digital, politics, and algorithms: Governing digital data through the lens of data protection. *European Journal of Social Theory*, 20(3), 329–347. https://doi.org/10.1177/1368431016679167

Bigo, D. (2024). The Future Perfect of Suspicion and Prediction as a Dispositive of Security Today? The Legacy of Foucault (1977). *Foucault Studies*, 36(1), 73–106.

Bode, I., & Huelss, H. (2024). Artificial Intelligence Technologies and Practical Normativity/Normality: Investigating Practices beyond the Public Space. *Open Research Europe*, 3, 160.

Bozovic, V. (2024). *Biopolitics and Digital Surveillance in the United States during the COVID-19 Pandemic*.

Broeders, D., Cristiano, F., & Weggemans, D. (2023). Too Close for Comfort: Cyber Terrorism and Information Security across National Policies and International Diplomacy. *Studies in Conflict & Terrorism*, 46(12), 2426–2453. https://doi.org/10.1080/1057610x.2021.1928887

Burnashev, R. (2023). *Michel Foucault's Concept of 'Dispositif of Security': Central Asian Weak States*. https://edizionicafoscari.it/media/pdf/books/978-88-6969-667-1/978-88-6969-667-1-ch-10.pdf

Capodivacca, S., & Giacomini, G. (2024). Discipline and Power in the Digital Age: Critical Reflections from Foucault's Thought. *Foucault Studies*, 36(1), 227–251.

De Laat, P. B. (2019). The disciplinary power of predictive algorithms: A Foucauldian perspective. *Ethics and Information Technology*, 21(4), 319–329. https://doi.org/10.1007/s10676-019-09509-y

De Ruvo, G. (2025). Algorithmic Objectivity as Ideology: Toward a Critical Ethics of Digital Capitalism. *Topoi*, 44(1), 175–186. https://doi.org/10.1007/s11245-024-10117-9

Deleuze, G., & Foucault, M. (1977). Intellectuals and power. *Language, Counter-Memory, Practice*, 205–217.

Dennis, J. (2024). The Architectonic Eye: Reimaging Foucault, Surveillance Theory and Platform Pedagogy. *JIS*, 8(2).

DTLI-2023.pdf. (n.d.). Retrieved July 30, 2025, from https://cgd.ibc-institute.id/wp-content/uploads/2024/10/DTLI-2023.pdf

Eubanks, V. (2018). *Automating inequality: How high-tech tools profile, police, and punish the poor*. St. Martin's Press.

European approach to artificial intelligence | Shaping Europe's digital future. (n.d.). Retrieved July 30, 2025, from https://digital-strategy.ec.europa.eu/en/policies/european-approach-artificial-intelligence

Flick, U. (2018). *Designing qualitative research*.

Foucault, M. (2007). *Security, territory, population: Lectures at the Collège de France*, 1977-78. Springer.

Foucault, M. (2008). panopticism" from" discipline & punish: The birth of the prison. *Race/Ethnicity: Multidisciplinary Global Contexts*, 2(1), 1–12.

Foucault, M. (2012). *Discipline and punish: The birth of the prison*. Vintage.

Fraser, G. (2020). Foucault, governmentality theory and 'neoliberal community development.' *Community Development Journal*, 55(3), 437–451.

García López, P. (2024). *Foucault and digital technologies of power: Studying resistance as decentralization*. https://openaccess.uoc.edu/items/357ce874-e554-4723-ab09-21357e5a0556?locale=en

Happe, K. E., Johnson, J., & Levina, M. (2018). *Biocitizenship: The politics of bodies, governance, and power* (Vol. 19). NYU Press.

Harcourt, B. E. (2014). Digital security in the expository society: Spectacle, surveillance, and exhibition in the neoliberal age of big data. *Columbia Public Law Research Paper*, 14–404. https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2455223

Jentzsch, S. M. (2023). *Surveillance and prejudice: How is bias in digital surveillance discussed in the public discourse, in the context of select English speaking* *newspapers after Snowden's leaks from 2013 to the current discussion in 2023?* [B.S. thesis, University of Twente].

Kwarteng, K., & Dorries, N. (n.d.). *National AI Strategy*.

Manokha, I. (2018). Surveillance, panopticism, and self-discipline in the digital age. *Surveillance and Society*, 16(2). https://ora.ox.ac.uk/objects/uuid:a8f5e604-0e3e-42d2-b373-a4e650b39dcb

Montasari, R. (2024). *Cyberspace, cyberterrorism and the international security in the fourth industrial revolution: Threats, assessment and responses*. Springer Nature.

National AI Strategy HTML version. (n.d.). GOV.UK. Retrieved July 30, 2025, from https://www.gov.uk/government/publications/national-ai-strategy/national-ai-strategy-html-version

Nishnianidze, A. (2023). Surveillance in the Digital Age. *ESI Preprints (European Scientific Journal, ESJ)*, 24, 80–80.

Prozorov, S. (2021). Foucault and the birth of psychopolitics: Towards a genealogy of crisis governance. *Security Dialogue*, 52(5), 436–451. https://doi.org/10.1177/0967010620968345

Sahakyan, H., Gevorgyan, A., & Malkjyan, A. (2025). From Disciplinary Societies to Algorithmic Control: Rethinking Foucault's Human Subject in the Digital Age. *Philosophies*, 10(4), 73.

Schwinges, A., Van Der Meer, T. G. L. A., Lock, I., & Vliegenthart, R. (2024). The watchdog role in the age of Big Tech – how news media in the United States and Germany hold Big Tech corporations accountable. *Information, Communication & Society*, 27(6), 1073–1094. https://doi.org/10.1080/1369118x.2023.2234972

Varela, D. T. (2024). Artificial Intelligence Law through the Lens of Michel Foucault: Biopower, Surveillance, and the Reconfiguration of Legal Normativity. *Open Journal of Social Sciences*, 12(12), 189–201.

Virtual Transparency: From the Panopticon to the Expository Society and Beyond. (2018). In B. E. Harcourt, *Transparency, Society and Subjectivity* (pp. 369–391). Springer International Publishing. https://doi.org/10.1007/978-3-319-77161-8_18

Williams, M. L., Burnap, P., & Sloan, L. (2017). Crime sensing with big data: The affordances and limitations of using open-source communications to estimate crime patterns. *British Journal of Criminology*, 57(2), 320–340.

Zuboff, S. (2022). Surveillance Capitalism or Democracy? The Death Match of Institutional Orders and the Politics of Knowledge in Our Information Civilization. *Organization Theory*, 3(3). https://doi.org/10.1177/26317877221129290