Javnost - The Public

Journal of the European Institute for Communication and Culture

ISSN: (Print) (Online) Journal homepage: https://www.tandfonline.com/loi/rjav20

Structural Logic of AI Surveillance and Its Normalisation in the Public Sphere

Yong Jin Park

To cite this article: Yong Jin Park (2021) Structural Logic of AI Surveillance and Its Normalisation in the Public Sphere, Javnost - The Public, 28:4, 341-357, DOI: 10.1080/13183222.2021.1955323

To link to this article: https://doi.org/10.1080/13183222.2021.1955323

Published online: 20 Sep 2021.

Submit your article to this journal

Article views: 397

View related articles

View Crossmark data

Citing articles: 2 View citing articles

Full Terms & Conditions of access and use can be found at
https://www.tandfonline.com/action/journalInformation?journalCode=rjav20

Javnost: The Public, 2021
Vol. 28, No. 4, 341–357, https://doi.org/10.1080/13183222.2021.1955323

STRUCTURAL LOGIC OF AI SURVEILLANCE AND ITS NORMALISATION IN THE PUBLIC SPHERE

Yong Jin Park

This study examines the fundamental logics of surveillance impetus in the rapid transition to AI-based information processing. In this paper, these logics are called axioms—three principles of (1) concentrated architectural codes, (2) constrained user psychology, and (3) peculiar characteristics of data as information. This study argues that each axiom perpetuates AI's tendency to solidify data surveillance and normalises it in newly emerged AI-driven public spheres. This is a conceptual paper structured in the following sections—(a) axioms (three principles maintaining the impetus of surveillance normalisation), (b) mutual shaping (interaction between users and institutions reinforcing surveillance), and (c) policy remedies (policy principles fixing normalisation). The thesis of this paper is the normalisation of AI—perpetuated by three axioms—is the product of mutual shaping between institutions and uses as "data-hungry" algorithms exacerbate the tendency in which users are to participate willingly in surveillance. This poses the concern that data surveillance in its pronounced normalising processes becomes an industrial structural problem, not an episodic one. This paper concludes by calling for sanguine intervention measures, collectively tackling the structural recurrence of surveillance in the U.S.-specific contexts but also touching upon even broader global policy discussion.

KEYWORDS Personal data; surveillance; AI policy; critical algorithm studies

By any measure, it has become quite a banal statement to pronounce that AI surveillance is "literally" everywhere. Consumer products, with its intelligence shifted from analogue to AI-based algorithm processing, have insatiable appetite for personal data, posing unprecedented risks to privacy. With Amazon, shopping has already changed. Personal histories of shopping records remain ontologically permanent and its AI accurately predicts what a shopper wants better than she does. In this public spheres of ubiquitous surveillance, the government has easier means to view citizens' lives by obtaining access to a single digital platform. Facebook alone, for instance, can hold a vast array of traceable digital trails of personal records surpassing that of any national government database. Simply put, AI gobbles up personal data and surveilles us "automatically."

The aim of this paper is to conceptualise the pattern of emergent AI-based automated surveillance as "normalisation" process. On a fundamental ground, this work draws upon the insight by Michel Foucault (1975) in his "Discipline and Punish" in order to establish AI surveillance as the process of "normalisation." Foucault's original idea is that social institutions, such as law, policing, medicine, and/or education, exercise disciplinary practices to control those citizens who do not conform to established standards of

© 2021 EURICOM

social practice. The essence of this hegemonic practice—or simply, the prevailing norm in given domains of social lives—is only to serve the dominant interest of established powers. In his words, Foucault (1975) noted,

> The judges of normality are present everywhere. We are in the society of the teacher-judge, the doctor-judge, the educator-judge, the social worker-judge; it is on them that the universal reign of the normative is based; and each individual, wherever he may find himself, subjects to it his body, his gestures, his behavior, his aptitudes, his achievements.

It is significant to note that the statement above does not directly touch upon personal data surveillance. However, Foucault's insight is that the practice of prevailing normality, in ways in which certain systems naturally reject and punish any individual resistance to the established norms, is becoming the structural logic on its own inertia that sustains control. From this, one can infer why the practice of AI surveillance, in its personal data collection and monitoring in digital platforms, may become the institutional form of normalisation, which sets behavioural standards, rules or expectations in governing individual users in their digital consumption. That is, AI functions effectively as a structural device that disciplines (or fails to accept) those who refuse to comply with automated surveillance by simply denying the platform access to the users who do not follow the most basic rule of data submission. On a more practical level, this study explicates three AI axioms to support the central thesis, which aims to show how such conceptualising AI surveillance as "normalisation" can serve as a useful conceptual basis to understand (and thus, intervene to fix) the impetus behind current institutional practices by digital AI industry and social media platforms.

In examining the structural and normalising logic of AI surveillance in public sphere, this paper starts with the underlying premise that the shaping of digital surveillance will never be separable from users-people as well as system producers-institutions (Giddens 1983). Pursuing this insight, one might find a reason to pause to dissect the interactive process in which individuals and institutions, not only influence each other, but also produce a different set of net outcomes as a result of their interaction (Neuman 1991, 2016; Van Dijk 2020). In this regard, the typology of three AI axioms is an effort to clarify conceptual ambiguities regarding institutional causes of perpetuating AI surveillance, ultimately aiming to (1) improve the conceptual precision in scholarly and policy discourses and (2) generate regulatory insights that can guide future policy intervention in AI-related U.S. and EU regulatory contexts.

This work is thus motivated by the deepening concern about the algorithmic well-being of public sphere (Habermas 1984; Neuman 1991, 2016; Van Dijk 2020), as the fundamental function of "self-organising" and "rational" participation in public life among citizens began to be replaced entirely by digital platforms and their commercial algorithms (Gillespie 2018; Shin 2019). Increasing dependence of public sphere on personal data-based AI-driven automatisation heralds a new era in which the algorithmic nudging of people's behaviour can be exercised in massive but subtle ways than ever before. The root of the problem is that as the citizens' relationships with public life become organised and reorganised based on AI, the relationship between the state government and individuals, which is newly emerging in public sphere, has also altered and raised the question of the platform governmentality—how commercial algorithms set out to establish the rationales, rules or techniques surrounding the data collection, retention or appropriation in their control (Foucault 1975; Törnberg and Uitermark 2020).

The argument below departs from this, with critical eyes on three principles perpetuating the impetus of surveillance normalisation and its shaping that comprises institutions, on the one hand, and users on the other hand. This is a conceptual paper, drawing the interdisciplinary insights from information economics (data object), behavioural economics (user), critical studies (perspective) and policy analysis (problem-solution). Henceforth, this study is not a discourse analysis; neither does it aim to be an empirical inquiry. Instead, it takes a holistic approach to lay conceptual foundations upon which the dynamics of automated AI surveillance in various future AI-based digital platforms can be better understood and channelled.

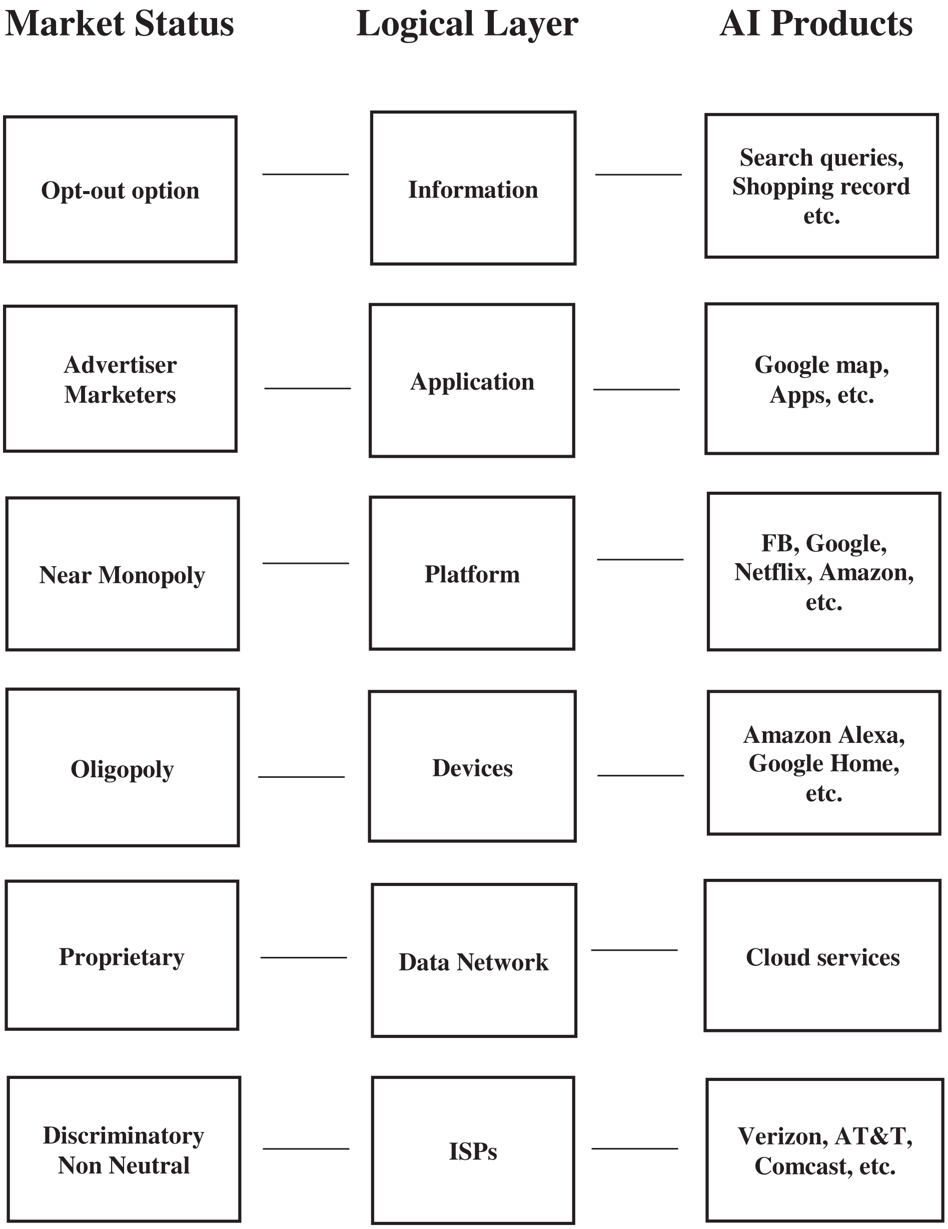
When AI Meets Old Surveillance—Three Axioms

Axiom 1: Concentrated Surveillance

Lessig (2009) suggested that "code is a law"—an elegant point that illustrates how a new technology is organised in particular ways can effectively serve as a set of structural constraints regulating people's behaviour. The code in this context of personal data surveillance is the concentrated consumption condition under which people and their behavioural data are calibrated to be monitored, collected and appropriated with a particular set of algorithmic principles. The precise nature of these architectural constraints can vary from one software or hardware to another; however, Lessig's pointed insight may well speak to the fact that the unregulated industry of algorithmic surveillance, as in AIs of Facebook and Google, remains unfettered and entirely free to set up, design and modify the technological-architectural condition of electronic public spheres, controlling people's behaviour, their interaction and data flow.

The elements of the architectural condition of surveillance concentration in public spheres are described in Figure 1. First, we see the potentially vast flow of data from the top layer, where a person or a device user is located, to the very bottom layer, where the physical backbone of digital infrastructure, namely Internet, is rooted. One might characterise this integrated flow of personal data across different levels of data ecosystem as the vertical integration of surveillance. Second, personal data are constantly "pushed and pulled" by a digital platform, whose databases can interconnect sets of information that users generate with those of other platforms or/and affiliated third parties. This can be said as the horizontal integration of surveillance in which personal data possibly travel among different platforms in the same area of business interest. Combined, that is to note unregulated data flow in AI-based digital ecosystem.

It is not hard to see the ways in which the horizontal and vertical concentration of digital ecosystems will deepen the vulnerability of personal data misuse, unwarranted access, security breach, potential exploitations or manipulation based on a person's detailed private record, to the extent that collection, appropriation, transfer and algorithm processing of personal data remain largely unregulated. For example, Netflix (the most dominant subscription-based digital platform for viewing) can pass on personal viewing data to in-house or affiliated third-party app developer, which might better programme automated video suggestion (vertical integration), while Amazon Prime, in a (hypothetical) strategic alliance, may want to share with YouTube or access its customer data, such as



Market Status	Logical Layer	AI Products
Opt-out option	Information	Search queries, Shopping record etc.
Advertiser Marketers	Application	Google map, Apps, etc.
Near Monopoly	Platform	FB, Google, Netflix, Amazon, etc.
Oligopoly	Devices	Amazon Alexa, Google Home, etc.
Proprietary	Data Network	Cloud services
Discriminatory Non Neutral	ISPs	Verizon, AT&T, Comcast, etc.

FIGURE 1
Vertical-Horizontal Data Concentration

demographics and viewing patterns/preference, in order to better customise Amazon Prime offerings and to increase their market share in digital video platform (horizontal integration).

The overall structure of un-regulated AI-based data industry surely indicates: (1) the de facto status of the near monopoly in digital platforms whereby a single company controls most of market shares (e.g. Facebook in social network and Google in search engine industry); (2) the oligopoly of few manufacturers that dominate in Internet of Things, wearables, mobile and smartphone devices (e.g. Google Home, Alexa, Apple and Samsung); (3) the closely-guarded proprietary standard of big data cloud services (e.g. Amazon, Google, Netflix, etc.); and (4) the virtual monopoly of ISPs (e.g. Verizon and Comcast) in most cities, where citizens have only limited choices over access to digital infrastructure and online (see Dahlgren 2018). Particularly, this concentrated surveillance structure, under which users' personal data can be transmitted and shared across different commercial platforms, raises a grave concern about numerous platforms respectively and jointly developing a deeper knowledge of a person that would have been impossible to obtain via a single platform.

Axiom 2: Data Peculiarities in AI

Another layer of complexity can be found in unique characteristics of personal data as information. First, on top of the decreasing cost of digital data processing in general (Neuman 1991), the collection, as well as the retention, transfer and appropriation of personal data, have become cheaper than ever before. Moore's Curve in 1975 predicted processing capability doubling roughly every 18 months, along with the drastic decrease in cost per bit. As dramatic as this prediction was, however, today's exponential growth of the digital capacity in data storage is even more drastic, driving the cost down even further. This overwhelming trend of lowering costs of data collection has critical consequences. For it makes the AI harvesting of extraneous data points far less costly than any human-to-human information collection and thus, unthreatening to any business model based on AI-algorithm. Even when the massively surveilled data turn out to be erroneous for the purpose of data collection, the cost of data collection is minimum.

Second, the monetary value, created based on data, is disproportionately larger than that of raw data. In other words, the return of investment on personal data processing and appropriation tends to be far greater than the cost that has to be initially "sunk" for the obtainment of the data upfront. This is different from mass media product, of which the production of "the first copy" is enormously costly with "high sunk" fee, such as talent salaries and above-the-line production cost like scriptwriting, directing and editing (see Couldry and Mejias 2018, for the abundance of data that are freely available for appropriation). Rather, the sunk cost, as in the case of online tracking of an individual person's web browsing via bots, cookies, web beacons, etc., is so low that any AI-generated value will be likely to be high enough to recuperate the upfront investment for data surveillance.

Third, the increase in value will be exponential in both scale and scope, once an initial value is created based on data. In terms of scale, the cost of reproducing the value-added data profiles will amount to almost zero, by linking to the full scopes of databases (i.e. copying and transferring digital databases interconnected vertically and horizontally). In terms of scope, the depth of exploitation can be even deeper, given a person's digital traces in AI-driven ecosystem that is vastly interconnected can virtually capture every facet of the person's life. The lesson is that selling or transferring these values compiled on comprehensive ranges of data points to the third parties like marketers, non-profits, government agencies and campaign fundraisers, does not (1) diminish the value of personal data and (2) exclude the continuous future uses by the AI in its algorithmic processing of the already constructed databases. Simply put, the irreversible trend of lowering costs of data collection and processing by AI-algorithms cannot be even comparable to that of any other information activities by traditional institutional actors, such as print news media or broadcasting journalism—for instance.

Axiom 3: The Surveillance "Nudge" in AI

Gandy and Nemorin (2018) described a persistent pattern in which the omnipresent AI-based digitisation can exacerbate the vulnerability of individuals, whose behavioural principle is often hindered by cognitive constraints as well as structural conditions. By drawing upon a critical perspective of political economy, their argument builds the problematic nature of asymmetrical power between individuals and institution and the extent to which their interaction plays a role in shaping the performance of unregulated markets. Their argument is helpful, given their insight that privacy (and the lack of it) is a much product of the mutual function of persons and institutions (Webster and Ksiazek 2012). In this context, the "nudge" is a necessary policy intervention without which the ontological status of individuals can be reduced to objects or "things" for the purpose of institutional algorithmic calculations (Sellar and Thompson 2016).

Prisoner's dilemma (Dutton et al. 2005; Hamilton 2000) extends the argument in an interesting way. Consider the diagram in Figure 2. A and B indicate respective values for an institutional actor (digital AI platform) and an individual actor (user data as an object of AI). When they decide to cooperate, both obtain values (A + B). In this case, cooperation means

	No Data	Data
No Data	**–A – B (no AI service)**	**A – B (CB)**
Data	**–A + B (CA)**	**A + B (full AI service)**

A = value for digital platform, such as advertising
B = value for person such as automatic suggestions

CA = cost for AI, opportunity to micro-target a person
CB = cost for person, AI-generated automation

FIGURE 2
Data Submission (Person) and Use (AI)

date submission or agreement that data will be used for the AI. Values that an AI platform obtains are the data and potential databases, while the values for users are (1) (allegedly) better, customised or personalised service and (2) AI-automated recommendations as well as (3) the access to the service. When both do not trust each other and decide to defect, all lose values (−A − B). If one decides to cooperate but the other defects, we begin to have complications. To be more precise, if the defector is a user, we have A − B. The trouble is that when a user decides to defect, that means de facto exclusion from, not a particular platform, but the same type of AI-based service, because a digital platform given one market is often a monopoly.

This non-symmetrical relationship between users and institutions can be represented numerically. Assuming that A or B can be simply assigned with an equal value of 1, we thus do have 0 for user cost (A − B) and also 0 for AI cost (−A + B) in case of non-symmetrical participation. Nonetheless, we already know that the value cannot be exactly same. The platform AI deals with numerous users (not a single user); thus losing one user amounts to *true* zero to AI. On the other hand, a user deals with only one platform; losing one platform amounts to *true* 100 to that individual. This is also different from the situation in which a digital AI platform decides to defect from data use (−A + B) or deviate from at least the way it is "explicitly or implicitly consented." For instance, the AI might fail to predict and recommend accurately what a customer prefers. Or there can be a case in which the AI does not deliver automatised service or it may divulge data to third parties, exposed to illegal data breach. The fact that the defect by a user will be forever known to the AI, whereas the defect by the AI can be hardly known to a user attests asymmetrical nature of digital transaction (Sadowski and Pasquale 2015).

The opportunity cost of defect on the part of AI will stay minimal, given the user data have been already taken. After all, the monetary benefit from data might not come in a short term but in the long term, and will be based on not a single data point, but multiple data points of repeated situations. Notice in the above Figure 2: even in the best optimal cooperation of A + B, the equilibrium and the balance between cost and benefit is not evenly distributed because the cost for the user participating in AI always remains—that is, privacy, a status of being surveilled forced upon a user, or a sense of dignity reduced to sets of data points (Cheney-Lippold 2017; Lupton 2013; Van Dijk 2020). This is the nature of a dilemma that is "coded" for the AI never to lose.

Three Axioms, Combined

What has been highlighted above is this: First, the structural condition, under which an individual person is situated in AI industries, remains deeply concentrated and interconnected as very few dominant institutional actors are in the best position to exploit the full dimensions of personal data. Second, as people become be subject to the ubiquitous AI surveillance, personal data as information do carry key unique characteristics (of non-rivalry and non-exclusivity), in that the potentially huge margin of return as a market product is propelling the intensification of massive data collection. Third, in the performance of unregulated AI industry, the purported "rational" dynamics of the interaction between an individual and AI platforms, as seen in a prisoner's dilemma-like scenario, tend to situate a person in a position to be unable to exercise privacy but to accept data surveillance as an entry condition of almost every digital participation. Tying all these is that AI surveillance can be interpreted from those multiple angles, as (1) the informational issues characteristic of data object from a perspective of information economics, as (2) the constraint of individuals' behaviour from a point of view by behavioural economics, and finally, as (3) the problem of institutional concentration of data power from a critical viewpoint. These seemingly disjointed rationales, as this study argues, are the fundamental impetuses driving AI surveillance in its own respective and distinctive ways. As the whole of different parts is not the same as an additive sum of each of them, however, their combination exerts its critical capacity for shaping AI-driven public sphere, conditioning how individuals and institutions mutually influence, shape and construct the normalisation of data-based surveillance.

Individual-Institution in Mutual Shaping

Here an important point about this is that AI is built into harvesting as much personal data as possible so as to produce "intelligence" and to make certain decisions or recommendations—thus, regulating behaviour. The logic is the more data, the better—a spurious logic of "big data" (boyd and Crawford 2012), but a much practiced one in the AI industry. This way, the unregulated AI industry, by design, will be always "data-hungry" and its algorithm will be ready to gobble up any data that may later turn out to be contextually irrelevant or even of no use (Nissenbaum 2009). Pinpointing actual harm or malicious intent of data surveillance is of less point, as the violation of privacy has already occurred with digital consumption presuming user consent (whether explicit or implicit) of data collection. On an even more fundamental level, the issue at stake is the "normalisation" of digital surveillance in personal data-based AI system. By this, we must not simply mean the intensification of data collection in Web 1.0 era. Nor does it refer to the digital transformation of interpersonal face-to-face surveillance that has been prevalent since human interaction ever existed in public arenas (Goffman 1967; Westin 2003).

On the user side, the fact that users often decide to give up data—or simply succumbed to personal data submission—is well documented (Park 2018a, 2018b, 2021). A study by Turow, Hennessy, and Draper (2015), for instance, characterised this privacy control behaviour as helplessly "being resigned to" the request for personal data, because individuals in their limited cognitive capacity cannot fight or resist the influx of data request—notably, in exchange of free access to platform services. Gandy and Nemorin (2018), on the other hand, noted that the user's decision to divulge data might reflect typical behavioural patterns among ordinary people, when being nudged by algorithms. This is alarming, but not surprising to understand people's vulnerability to manipulative AI, as user rationality is habitually bounded with the convenience of inaction (also see Neuman 2016; cf. Park and Oh 2021). Understanding these user-side dynamics, the critical concern of this paper is about the inertia behind the mutual shaping of AI surveillance—with the fundamental impetus of AI surveillance underlying the institutional behaviour, on the one hand, and the individuals' behavioural habits and cognitive constraints, on the other hand. One of the central thesis in this work is that the prior debates on AI surveillance have focused on the dystopian algorithmic trends alone. But that is the description of a symptom, not the root-structural determinants, which help us better identify (thus, give us a clue on how to redesign) the structural conditions that are inductive of normalisation of surveillance.

The continuous trend toward AI-based surveillance will surely introduce elite scholarly debates, policy discussion, and occasionally, public resentment, for instance, about Facebook's yet another data breach of its billions users, in October 2018—less than six months since Cambridge Analytica scandal broke out. However, as AI will make its way into the entirety of our "normal" digital experiences, the best prediction will be less and less of the public shock and thus, little resistance, and this trend of no change will be deepening, combined with no "nudge" from policy intervention in Washington DC (Gandy and Nemorin 2018). For instance, Edward Snowden's revelation about NSA PRISM programme in 2013 dissipated largely from public discourse—such a critical moment in mass surveillance by the U.S. government that involved every citizen's digital experience on the Internet, digital newspapers, Google, Amazon, mobile phone, emails, etc. (Evens and Van Damme 2016; Lyon 2014). In fact, we found very little evidence suggesting that Edward Snowden's revelation about NSA surveillance ever changed individual behaviour. Nor did Facebook users migrate into other social network services after a series of Facebook data breach incidents.

This is not the same as claiming that people remain unconcerned or do not care about privacy (Hargittai and Marwick 2016; Park, Chung, and Shin 2018; Park 2018b; Park and Shin 2020). Ordinary people, as AI continuously processes personal records, may be dealing and struggling with the complexity of "normalisation" of surveillance in which the type and intensity of personal information exposure remain deeply ingrained in people's mundane digital experience. Noble (2018), in her critique of high-tech industry, chastised algorithm as a tool of oppression. Focusing on our experiences with Google search engine and Yelp, Noble correctly rejected the optimism that search engines like Google through its algorithms can offer an equal playing field for different forms of ideas. Instead, algorithm perpetuates bias against women of colour and marginalised populations, as Google searches are designed based on people who are white and male. Likewise, the new forms of AI-based data collection that normalises a routine of surveillance can potentially create further inequalities by translating seemingly banal data and defining and classifying people into a "sort," or a type and a target, under the guise of bringing posited consumer benefits of personalisation and customisation.

Foucault's Discipline and Punish—in AI Context

Here Michel Foucault (1975) in *Discipline and Punish* offered yet another critical insight that can be applied to the "normalisation" of AI-based surveillance in which new ways of collecting, retaining and organising personal data become routinely engrained in everyday life. To Foucault, this type of a force, in the form of a machinery in the AI case, would be a disciplinary device for exercising social power and control, as it rewards people conforming to its dominant rule, but punishes those who do not accept the rule or follow the emerging norm of a proper conduct—thus, effectively reinforcing power.

On the most rudimentary sense, an individual user, who refuses to the request of data submission by a platform, is immediately punished by access denial and even disciplined (taught) (in Foucault's sense) to be compliant with the entry condition of digital platform services. Repeated situations with the constant influx of such demands are in fact the normalising power in a circle of learned submission behaviour from the part of users, given the surveillance (via data submission) is de facto normality against which to measure a user's willingness to go against the very condition of the platforms. That is to say, the crucial power of AI lies in its organising intelligence that can exclude those not conforming to the routine norm of surveillance from any digital participation. What's critical to realise in this process is that the normalisation of digital surveillance is a product of the "mutual" contribution from institutions, motivated to design the optimal digital platform, and individuals, not necessarily coerced to do give up privacy but only to choose data submission "voluntarily" so as not to create "noise" and thereby, safely avert the immediate risk of being excluded from digital participation. Put bluntly, the differentials in the cost (1) between data submission and non-submission and (2) between being surveilled and refusal to be surveilled remain algorithmically too large in the AI ecosystem to disrupt the normality of digital experience in order to undertake privacy protection—and create "noise."

We can imagine how easily we perceive the data surveillance as "normal" part of human–machine interaction as we even consider AI-driven digital platforms, such as Facebook and Alexa, as another social actor or at least, the place via which we can develop intimate societal relationships through self-disclosure (Jiang, Bazarova, and Hancock 2011; also see Evens and Van Damme 2016, in the context of developing digital newspaper access and readership). This dynamics of interdependence, between two otherwise distinct and separate entities of institutions and individuals, naturally obtains the status of normalisation, with the interplay of two forces constituting the system of digital surveillance in the future of AI-based data ecosystem. The key point is that this mutual shaping by the two separate but interconnected forces of individuals and institutions jointly sustains the AI ecosystem and its surveillance practices embedded in social structure (Giddens 1983; Neuman 1991, 2016). In other words, AI-based ecosystem is designed to produce the structural logic and condition that encourage data surveillance, which in turn helps clearly define (and limit) people's digital activities in their scopes and types, and thus, construct identities according to the maximum utility value of the system needs (Gandy and Nemorin 2018; Sandvig et al. 2016).

Note Foucault's well-known premise of surveillance—Panopticon. This is still a useful analogy describing surveillance in a totalitarian society as a prison-like architecture system, in which one prison guard on top of a control power watches over those surveilled below, thus governs the power of the system. Nevertheless, this study argues that the Panopticon-perspective might be perhaps better retired in favour of Foucault's own alternative notion of normalisation as in his "Discipline and Punish." By simply teaching a user to submit data and how it is beneficial to her/him, AI surveillance in digital platforms exercises its demand more subtly and perhaps more effectively than a central towering figure standing above all of us. Namely, AI surveillance, in obtaining the automated standpoint of observation over user participation, invokes (1) a more covert process than an overt exertion of control; (2) a more interactive shaping between users and platform-institutions than a one-way totalitarian dictation of what to do; and (3) a more self-disciplinary decision to cooperate with the request of data submission than a threat or coercion. While this interpretation may present a somewhat alternative perspective on surveillance, the notion of normalisation better describes the structural condition in which AI surveillance occurs in indirect, subtle and nuanced ways—thus explaining why the reversal of such trends in public sphere will be impossible without looking at what is beneath in the digital platforms and the three axioms fuelling their algorithmic systems.

dannah boyd (2018) made this point in her sharp critique of how the algorithmic world perpetuates hate on the Internet. The logic of boyd's critique is that the banality of unintended evil is actually more problematic than intended evil with a specific purpose. It is because structural banality would push AI surveillance into a normalised routine of digital participation whereby any observed negative effects (in defining who we are) may be simply the result of differences calculated according to particular AIs and their purposes, rather than any individual, wilful control or choices in how we present ourselves (Goffman 1967; May and Finch 2009). In other words, unintended harm may be more serious precisely because it makes even more difficult any efforts to differentiate the effects from the causes and underlying motivations of surveillance, and thus to repair the damage itself (see Gillespie 2014, 2018). The evil, in this case of digital surveillance, is not Google, any Silicon Valley corporations, or even the U.S. government tapping on its vast databases. Rather, it is systemic, lying in (1) bottom-line driven model of profit imperatives and (2) reticent psychological nature of privacy behaviour. These forces combine to perpetuate, reinforce and reward the continued growth of AI-based surveillance system. The threat of harmful results of surveillance is real because its end output, as in the case of Google search results, might eventually introduce oppressive effects, such as perpetuations of misperceptions of women of colour. Nevertheless, we must recall that the privacy loss and other "oppressive" consequences of the application of algorithms might be far from the intended effects.

Look Ahead for Institutional Remedies: From Explainability to Axiom Principles

This invites effective regulatory intervention, which should be innovative enough to be devised outside the two forces of individual agencies and institutional imperative in marketplace. On a policy level, we can focus on U.S.-specific regulatory contour as an exemplary case, but because the issues raised above—especially in regards to three axioms—touches upon personal data surveillance that are pervasive in AI and algorithm-driven platforms in general, the fundamental regulatory lessons should apply equally to EU and global contexts. Here it is important to highlight that the key debate regarding AI policy solutions often evolves around AI explainability—the idea that suggests that exposing information process specific to each individual algorithm decision will prevent error, or increase trust over AI decision by enabling the public to gain knowledge into AI (Doshi-Velez et al. 2017). This principle of AI explainability, though not directly, encompasses the notion that the AI collection and use of personal data too can be better governed by introducing a measure by which the data subject (under the purview of specific AI decisions) can see inside AI. EU General Data Protection Regulation (GDPA)'s Article 22 is a reflection of this belief on "right to explanation," providing a legal basis of why one should be granted a right to obtain information about the logic of AI decisions.

However, what is unclear about AI explainability is whether any form of explanation of specific AI decisions and its data collection practices can be plausible, and in fact helpful, for individuals who cannot do much against surveillance to enter a commercial platform. In other words, the critique here is that it is highly uncertain that people take effective actions against data surveillance even in a very specific circumstance that is publicly known for AI data monitoring and surveillance. Instead, the shift toward explainabiltiy, despite its good regulatory intention, rather creates a burden to an individual who might not even have enough cognitive resources to understand and take actions (Wachter, Mittelstadt, and Russell 2017). This point is well in line with the scepticism by Ananny and Crawford (2018), who saw pragmatic limits of instilling transparency in algorithm-driven digital environment, given their assessment of bounded user practices and knowledge.

As a matter of a fact, the central focus in this paper is to argue that the natural inclinations of institutions and individuals—either interactively or respectively—have never been (more importantly in the future, will never be) inductive of privacy than surveillance, if at all. To be clear, the task then is to recognise that it is the policy roles that should nudge the private enterprises to better shape marketplace so that a person's digital participation is not to be subject only to surveillance. As noted above, this broad insight may not be only concerning the U.S. regulatory environments where almost no or little direct governance on AI-based industry exists, but also touch upon regulatory principles of AI in the context of EU. Given the significance of GDPR that affects the data activities of U.S. Silicon Valley-based firms as well as EU's AI industry, it is important to advocate the fundamental changes in broad global principles (Wachter, Mittelstadt, and Russell 2017). Vogel (1998) argued that the freer market needs the more regulations—his thesis is that the function of market in fact depends upon how to condition its operational terms of freedom, while acknowledging that there is also a room for marketplace to fulfil. Accordingly, this paper insists that we stay with high-level principles that can be flexible regardless of specific technologies of which the unique threats can be hardly foreseen. Importantly, the top-level regulatory attributes must stay beyond a perennial moot point enshrouded in the ideological debate about government vs. market in order to bring meaningful solutions to the table.

This study suggests the following elements of specific regulatory proposals, which can be useful in addressing each of three axioms explained earlier. Note that these suggestions directly concern U.S. regulatory reforms; however, they will also apply to a global policy level, potentially guiding EU policy-makers in their effort to update GDPR.

- First, interface design of personal data-based AI applications, like social media platforms, smart home devices and smartphone or wearables, are required to contain a function that restricts third-party data access, appropriation and retention of personal records.
- Second, horizontal and vertical integrations within and across the AI-based industry need serious periodic oversight from the FTC, which can be enacted by Congress to prevent the concentration of databases and surveillance capacity.
- Third, along with the change in the effort to educate consumers, particularly in the U.S., the FTC should enact and enforce an opt-in model regarding surveillance in personal data-based AI platforms (Park 2011).
- Fourth, on the side of consumption, a long-term local-level public campaign for promoting relevant digital skills must be instilled involving various stakeholder groups.

It is important to note that these first and second proposals are to address the first axiom concerning concentrated AI surveillance structure. These two proposals will also tackle the second axiom regarding peculiarities of data as information, in that low cost of data collection and reuse can be in part restricted or at least discouraged by limiting third-party access to data collected under specific platforms. The third and fourth proposals are to tackle the problem of the third axiom concerning AI nudge (or its manipulation) on the side of individual users, as the current opt-out option is hardly engineered for users to take any meaningful actions for privacy. On the other hand, the last two proposals are also to be applied to AI-based sectors in dealing with personal data concerning children, elderly, and other vulnerable populations, as well as digital records related to personal finance, health and related genome data.

Collectively, the suggested measures above are not to naïvely suggest one approach at a state national-level to eradicate all the root causes of surveillance in algorithm-based public sphere—the problems recognised as three axioms in this paper. Still, a set of regulatory principles, based on which policy-makers devise oversight to ease the loss of our abilities to control the flow of data in AI platforms, are possible, especially considering practical natures of what is suggested above. In fact, the full scope of Fair Information Practice Principles (FIPs), though the most comprehensive working guidelines yet, remain difficult to be incorporated into an AI system design, with its regulatory emphasis premised on user-side actions such as access and transparency. This concern is a reasonable and practical one and, when it comes to the specifics of AI-driven platforms, the above macro-level regulatory solutions must be mixed with another layer of regulatory measures at a micro-level. Particularly helpful, yet again, is Lessig's premise that "code is law" (2009; also, Bowker and Star 1999), but here in a more literal sense of computer code in that algorithmic programming can be a powerful tool for regulating user behaviour as well as the operation of digital platforms. From this standpoint, it is telling that AI-based digital systems like Facebook, Amazon or Google can be engineered or reverse-engineered for fulfilling other societal purposes such as privacy.

In this sense, one of the attractive policy alternatives is to take a more normative than technical approach, and a more consequential than purely instrumental point of view about personal data. As Napoli (2015) succinctly pointed out, algorithms can be regarded as institutions, like mass media, in the way that their effects are obliged to the public interest. If such normative reconceptualisation is possible, then, the algorithmic need of personal data, as in any AI industry sector, can be re-conceptualised—not just as Weberian rational motivations or phenomena (for instance, based on three axioms reviewed above), but as the extension of public control over the use of personal data as well as institutional data management processes. In imaging such possibilities of next Internet-AI worlds, it is in fact fascinating to see the insightful suggestion by Mosco (2018) who recommended the reconceptualisation of technological convergence as carrying public utility functions, such as water and electricity. This insight resonates with the normative concern that this paper started based on the premise of public sphere. In fact, (1) if public sphere is dominated by AI, and (2) if such domination is based on personal data, then, one might ask about a strong rationale about why AI surveillance is not regulated for public benefits. It naturally follows that the state government has active stakes, not only to protect or promote citizens' rights to access, but also to maintain certain levels of quality, let's say of water, as much as of personal data.

This way, the protection of privacy in AI-based digital platforms can and should be thought as an institutional product providing various public utility functions in public sphere, whereby their institutional interactions with users may carry public obligations under regulatory purview. Based on this perspective, policy-makers can justify regulatory oversight and interventions against a purely functional argument supporting unregulated AI. A dilemma for AI-based business model is precisely that it can never sustain with no data surveillance—henceforth, we witness its "normalisation" process of surveillance. With the three AI axioms perpetuating AI surveillance as "normality," data protection becomes a structural problem, not an episodic one. Sanguine policy measures thus must devise corresponding levels of structural solutions, which tackle the recurrence of surveillance normalisation in public sphere.

Conclusion

The readers should not leave with the impression that the study of AI surveillance and privacy suffers from the complete void of conceptual tools beyond what has been discussed in terms of the three axioms. Neither is this study's aim to conclude that there are serious shortcomings in prior debates. Like many other areas of socio-tech-policy inquiries, the aim was that newer useful conceptual grounds can be laid out so that future debates do not reach conceptual insularity whereby scholars confront only limited sets of analytical tools and choices. We might reach similar conclusions about automated AI surveillance and its normalisation tendency; but sceptical views will advance—when aided by alternative conceptual reasons and thus, will need to continuously generate clearer practical implications.

As Foucault (1975) noted, if surveillance and its institutional mechanism do mean the power to structure how to control public sphere, we in fact have the insurmountable concern over AI surveillance, because the current trend symbolises that such disciplinary control is becoming easily normalised, handing over control to the commercial entities with their dominant algorithmic power. With this, the democratic function of rational public sphere and its ideal will be, not only very unlikely to arise, but also hardly put forth even as an ideal. Put differently, the public sphere will be simply a product of AI surveillance and algorithmic processing by commercial platforms. Future works will need to borrow further insights on this future development, while borrowing insights from related fields, such as behavioural economics, and we have seen successful examples in which scholars obtaining valid experimental-survey evidence make critical policy analysis (see Acquisti, Brandimarte, and Loewenstein 2015). The complexity of measuring any change (or no change) brought by AI, automation, and surveillance requires such interdisciplinary instruments and a significant conceptual revision.

ACKNOWLEDGEMENTS

The author expresses his fullest gratitude to anonymous reviewer for very careful readings of this manuscript and enormous insights that helped improve the quality of this work.

DISCLOSURE STATEMENT

No potential conflict of interest was reported by the author(s).

REFERENCES

Acquisti, Alessandro, Laura Brandimarte, and George Loewenstein. 2015. "Privacy and Human Behavior in the Age of Information." *Science* 347 (6221): 509–514. doi:10.1126/science.aaa1465.

Ananny, Mike, and Katie Crawford. 2018. "Seeing Without Knowing: Limitations of the Transparency Ideal and its Application to Algorithmic Accountability." *New Media and Society* 20 (3): 973–989. doi:10.1177/1461444816676645.

Bowker, Geoffrey, and Susan Star. 1999. *Sorting Things Out: Classification and Its Consequences*. Cambridge, MA: MIT Press.

boyd, danna, and Katie Crawford. 2012. "Critical Questions for Big Data: Provocations for a Cultural, Technological, and Scholarly Phenomenon." *Information, Communication and Society* 15 (5): 662–679. doi:10.1080/1369118X.2012.678878.

boyd, danna. 2018. "What Hath We Wrought?" *SXSW EDU*. https://www.sxswedu.com/news/2018/watch-danah-boyd-keynote-what-hath-we-wrought-video/.

Cheney-Lippold, John. 2017. *We are Data: Algorithms and the Making of Our Digital Selves*. NYC, NY: NYU Press.

Couldry, Nick, and Ullises Mejias. 2018. "Data Colonialism: Rethinking Big Data's Relation to the Contemporary Subject." *Television and New Media* 20 (4): 336–349. doi:10.1177/1527476418796632.

Dahlgren, Peter. 2018. "Media, Knowledge and Trust: The Deepening Epistemic Crisis of Democracy." *Javnost-The Public* 25 (1-2): 20–27. doi:10.1080/13183222.2018.1418819.

Doshi-Velez, Finale, Mason Kortz, Ryan Budish, Chris Bavitz, Sam Gershman, and Alexandra Wood. 2017. "Accountability of AI under the Law: The Role of Explanation." ArXiv preprint arXiv:1711.01134. https://arxiv.org/abs/1711.01134.

Dutton, William, Geraldo Guerra, Daniel Zizzo, and Malcom Peltu. 2005. "The Cyber Trust Tension in Egovernment: Balancing Identity, Privacy, Security." *Information Polity* 10 (1/2): 13–23. doi:10.3233/IP2005-0066.

Evens, Tom, and Kristen Van Damme. 2016. "Consumers' Willingness to Share Personal Data: Implications for Newspapers' Business Models." *International Journal on Media Management* 18 (1): 25–41. doi:10.1080/14241277.2016.1166429.

Foucault, Michel. 1975. *Discipline and Punish: The Birth of the Prison*. NYC, NY: Vintage.

Gandy, Oscar, and Selena Nemorin. 2018. "Toward a Political Economy of Nudge: Smart City Variations." *Information, Communication and Society* 22 (14): 1–15. doi:10.1080/1369118X.2018.1477969.

Giddens, Anthony. 1983. "Comments on the Theory of Structuration." *Journal for the Theory of Social Behaviour* 13 (1): 75–80. doi:10.1111/j.1468-5914.1983.tb00463.x.

Gillespie, Tarleton. 2014. "Facebook's Algorithm—Why Our Assumptions are Wrong, and Our Concerns are Right." *Culture Digitally* 4. https://culturedigitally.org/2014/07/facebooks-algorithm-why-our-assumptions-are-wrong-and-our-concerns-are-right/.

Gillespie, Tarleton. 2018. *Custodians of the Internet: Platforms, Content Moderation, and the Hidden Decisions That Shape Social Media*. New Havens, CT: Yale University Press.

Goffman, Erving. 1967. *Interaction Ritual: Essays on Face-To-Face Behavior*. Garden City, NY: Doubleday.

Hamilton, James. 2000. *Channeling Violence: The Economic Market for Violent Television Programming*. Princeton, NJ: Princeton University Press.

Hargittai, Eszter, and Alice Marwick. 2016. "What Can I Really Do? Explaining the Privacy Paradox with Online Apathy." *International Journal of Communication* 10 (21). doi: 1932–8036/20160005.

Habermas, Jürgen. 1984. *The Theory of Communicative Action: Reason and the Rationalization of Society*. Boston, MA: Beacon Press.

Jiang, L. Crystal, Natalie Bazarova, and Jeffrey Hancock. 2011. "The Disclosure–Intimacy Link in Computer Mediated Communication: An Attributional Extension of the Hyperpersonal Model." *Human Communication Research* 37 (1): 58–77. doi:10.1111/j.1468-2958.2010.01393.x.

Lessig, Lawrence. 2009. *Code: And Other Laws of Cyberspace*. NYC, NY: Vintage.

Lupton, D. 2013. "Quantifying the Body: Monitoring and Measuring Health in the Age of MHealth Technologies." *Critical Public Health* 23 (4): 393–403. doi:10.1080/09581596.2013.794931.

Lyon, David. 2014. "Surveillance, Snowden, and Big Data: Capacities, Consequences, Critique." *Big Data & Society* 1 (2), doi:10.1177/2053951714541861.

May, Carl, and Tracey Finch. 2009. "Implementing, Embedding, and Integrating Practices: An Outline of Normalization Process Theory." *Sociology* 43 (3): 535–554. doi:10.1177/0038038509103208.

Mosco, Vincent. 2018. "A Critical Perspective on the Post-Internet World." *Javnost-The Public* 25 (1-2): 210–217. doi:10.1080/13183222.2018.1418976.

Napoli, Phil. 2015. "Social Media and the Public Interest: Governance of News Platforms in the Realm of Individual and Algorithmic Gatekeepers." *Telecommunications Policy* 39 (9): 751–760. doi:10.1016/j.telpol.2014.12.003.

Neuman, W. Russell. 1991. *The Future of the Mass Audience*. London: Cambridge University Press.

Neuman, W. Russell. 2016. *The Digital Difference: Media Technology and the Theory of Communication Effects*. Cambridge, MA: Harvard University Press.

Nissenbaum, Helen. 2009. *Privacy in Context: Technology, Policy, and the Integrity of Social Life*. Palo Alto, CA: Stanford University Press.

Noble, Sophia. 2018. *Algorithms of Oppression: How Search Engines Reinforce Racism*. New York, NY: NYU Press.

Park, Yong Jin. 2011. "Provision of Internet Privacy and Market Conditions: An Empirical Analysis." *Telecommunications Policy* 35 (7): 650–662. doi:10.1016/j.telpol.2011.06.003.

Park, Yong Jin. 2018a. "Social Antecedents and Consequences of Political Privacy." *New Media and Society* 20 (7): 2352–2369. doi:10.1177/1461444817716677.

Park, Yong Jin. 2018b. "Explicating Net Diversity in Trend Assessment." *Communication Research* 45 (5): 783–809. doi:10.1177/0093650215601883.

Park, Yong Jin. 2021. "A Socio-technological Model of Search Information Divide in US Cities." *Aslib Journal of Information Management* 73 (2): 144–159. doi:10.1108/AJIM-07-2020-0225.

Park, Yong Jin, Jae Eun Chung, and Dong Hee Shin. 2018. "The Structuration of Digital Ecosystem, Privacy, and Big Data Intelligence." *American Behavioral Scientist* 62 (10): 1319–1337. doi:10.1177/0002764218787863.

Park, Yong Jin, and Donghee Don Shin. 2020. "Contextualizing Privacy on Health-Related Use of Information Technology." *Computers in Human Behavior* 105: 106204. doi:10.1016/j.chb.2019.106204.

Park, Yong Jin, and Yu Won Oh. 2021. "Effects of Smartphones on Economic and Subjective Quality of Life." *First Monday* 26 (3). doi:10.5210/fm.v26i3.10269.

Sadowski, Jathan, and Frank Pasquale. 2015. "The Spectrum of Control: A Social Theory of the Smart City." *First Monday* 20 (7), doi:10.5210/fm.v20i7.5903.

Sandvig, Christian, Kevin Hamilton, Karrie Karahalios, and Cedric Langbort. 2016. "When the Algorithm Itself Is a Racist: Diagnosing Ethical Harm in the Basic Components of Software." *International Journal of Communication* 10: 4972–4990. http://social.cs.uiuc.edu/papers/pdfs/Sandvig-IJoC.pdf.

Sellar, Sam, and Greg Thompson. 2016. “The Becoming-Statistic: Information Ontologies and Computerized Adaptive Testing in Education.” *Critical Methodologies* 16 (5): 491–501. doi:10.1177/1532708616655770.

Shin, Don Donghee. 2019. “Toward Fair, Accountable, and Transparent Algorithms: Case Studies on Algorithm Initiatives in Korea and China.” *Javnost-The Public* 26 (3): 274–290. doi:10.1080/13183222.2019.1589249.

Törnberg, Petter, and Justus Uitermark. 2020. “Complex Control and the Governmentality of Digital Platforms.” *Frontiers in Sustainable Cities* 13 (March), doi:10.3389/frsc.2020.00006.

Turow, Joseph, Michael Hennessy, and Nora Draper. 2015. “The Tradeoff Fallacy: How Marketers Are Misrepresenting American Consumers and Opening Them Up to Exploitation.” https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2820060.

Van Dijk, Jan A.G.M. 2020. *The Network Society*. New York City, NY: Sage.

Vogel, Steve. 1998. *Freer Markets, More Rules: Regulatory Reform in Advanced Industrial Countries*. Ithaca, NY: Cornell University Press.

Wachter, S., Brent Mittelstadt, and Chris Russell. 2017. “Counterfactual Explanations Without Opening the Black Box: Automated Decisions and the GDPR.” *Harvard Journal of Law & Technology* 31 (841). https://arxiv.org/abs/1711.00399.

Webster, James, and Thomas Ksiazek. 2012. “The Dynamics of Audience Fragmentation: Public Attention in an Age of Digital Media.” *Journal of Communication* 62 (1): 39–56. doi:10.1111/j.1460-2466.2011.01616.x.

Westin, Alan. 2003. “Social and Political Dimensions of Privacy.” *Journal of Social Issues* 59 (2): 431–453. doi:10.1111/1540-4560.00072.

Yong Jin Park (corresponding author) is a Professor in the Communication, Culture and Media Studies Department at the Cathy Hughes School of Communications, Howard University, Washington, DC, USA. Email: yongjinp@hotmail.com