

Governance Through Uncertainty: What Chinese Algorithmic Systems Reveal About the Limits of Fairness, Accountability, and Transparency

JASON C. LAU, University of Toronto, Canada

Frameworks for algorithmic governance assume that opacity is a problem to be solved through transparency. This paper argues that algorithmic systems may govern not despite their opacity but through it. Drawing on ethnographic research in China (2013–2020), I examine how social credit systems, content moderation, and platform infrastructure operate through “governance through uncertainty”: a mode of control where indeterminacy about rules, boundaries, and consequences functions as the governing mechanism itself. I develop the concept of the “microworld” to theorize how platforms construct simplified, quantified environments that users must navigate according to logics they cannot see. This produces distinctive social effects: involution (intensified striving without advancement) and lying flat (withdrawal from the game)—responses to what I call compulsory striving under conditions of epistemic deprivation. Rather than treating China as an authoritarian outlier, I argue that governance through uncertainty is an emergent property of algorithmic systems—one that arises under specifiable conditions and that the Chinese case makes unusually visible. The paper concludes by examining implications for FAccT. If opacity is integral to algorithmic governance, transparency frameworks require reconsideration. But fairness frameworks need rethinking too: governance through uncertainty produces stratification not through biased outputs but through asymmetric capacity to navigate opacity. Those with resources, connections, and political standing can bear uncertainty; those without cannot. The question is not only how to make algorithms legible, but who can afford illegibility—and who pays for it.

CCS Concepts: • **Human-centered computing** → **Empirical studies in HCI**; • **Social and professional topics** → **Computing / technology policy**; • **Security and privacy** → **Social aspects of security and privacy**.

Additional Key Words and Phrases: algorithmic governance, governance through uncertainty, microworld, opacity, fairness, China, social credit, content moderation, ethnography

ACM Reference Format:

Jason C. Lau. 2026. Governance Through Uncertainty: What Chinese Algorithmic Systems Reveal About the Limits of Fairness, Accountability, and Transparency. In *The 2026 ACM Conference on Fairness, Accountability, and Transparency (FAccT '26)*, June 25–28, 2026, Montreal, QC, Canada. ACM, New York, NY, USA, 17 pages. <https://doi.org/10.1145/3805689.3806495>

1 Introduction

One night in 2016, I was having dinner with friends in Sanlitun—one of the trendiest areas in Beijing. It was that night I first heard of “Sesame Credit” (芝麻信用). Intrigued by the unusual name, my Chinese friends told me it was a mobile app developed by Alibaba, the e-commerce giant, that had just gone viral. Introduced a year earlier as a loyalty program, Sesame Credit uses data collected from Alibaba’s internet services to compile commercial “credit scores” for its users. Roughly based on FICO scores, Sesame scores range from 350 to 950, representing different levels of “trustworthiness.” Higher scores unlock privileges: discounted bike-sharing, fast-track visa

Author’s Contact Information: Jason C. Lau, Munk School of Global Affairs & Public Policy, University of Toronto, Toronto, Ontario, Canada, x.lau@utoronto.ca.



This work is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License.

FAccT '26, Montreal, QC, Canada

© 2026 Copyright held by the owner/author(s).

ACM ISBN 979-8-4007-2596-8/2026/06

<https://doi.org/10.1145/3805689.3806495>

applications, low-interest loans, even preferential treatment on dating apps. But unlike FICO scores, which carry a sense of seriousness and are associated with formal financial institutions, Sesame Credit is designed to be playful. Like its name—referring to “open sesame” in *Ali Baba and the Forty Thieves*—the system is meant to feel fun. Users check their scores on a colorful gauge and discover perks instantly. It feels less like a financial tool than a game.

What determines these scores remains opaque. While traditional credit scores rely on payment history and debt burden, Sesame Credit draws from Alibaba’s vast ecosystem: shopping patterns on Taobao, transactions on Alipay, social media interactions, even one’s circle of friends. In one interview, a Sesame Credit executive suggested that buying diapers might improve one’s score—parenting signals responsibility—while playing video games for ten hours a day would likely lower it [30]. The confidence with which behavior is mapped to character is striking.

When I asked my friends how they felt about this, the first thing they said was “convenience.” They made fun of me for being backward in accepting new technologies. Like many young professionals in China who attended prestigious universities and held well-paid jobs, they received high Sesame scores. They were beneficiaries of the Chinese economic boom. With a few swipes on their phones, apps like Sesame Credit unlocked social perks and financial rewards. No wonder they loved these new technologies.

Then I asked unwittingly: “Will any of the information used by these apps be sent to the government?”

“这个不好说 (*zhe ge bu hao shuo*),” one of them replied, and quickly changed the topic.

Bu hao shuo. The phrase literally means “hard to say”—expressing uncertainty. But it also means “not good to say”—implying something that cannot or should not be spoken. This ambiguity, I came to realize, captures how algorithmic governance operates in China. No one knows exactly what is monitored. No one knows exactly how scores are calculated. No one knows exactly what consequences follow from what behaviors. And crucially, no one knows exactly where the line lies between what the state sees and what platforms see, between what is permitted and what is forbidden. The uncertainty is not a bug in the system. It is the system’s governing logic.

Western frameworks for understanding algorithmic systems—organized around fairness, accountability, and transparency (FAccT)—assume that opacity is a problem to be solved. But this assumption is failing even in the West. Platforms moderate content through opaque and shifting rules; recommendation systems work as black boxes; users navigate uncertainty about what is monitored and what triggers consequences. The difference is that Western technological systems maintain the *rhetoric* of transparency while becoming ever less transparent, whereas Chinese systems have more openly abandoned that pretense. In China, opacity is not hidden as failure. It is acknowledged as governance.

This paper argues that “governance through uncertainty”—a mode of control where indeterminacy about rules, boundaries, and consequences functions as the governing mechanism—is not uniquely Chinese. It is an emergent property that arises when computational systems become indispensable to daily life, when rule enforcement is discretionary, and when evaluation criteria remain opaque to those subject to them. The analysis is concerned not with whether opacity is intended but with what it enables—and under what institutional conditions that enablement can be exploited, tolerated, or contested. The analysis is diagnostic rather than prescriptive: identifying how opacity governs does not endorse it, but argues that accountability frameworks premised on achievable transparency may be insufficient for systems whose opacity is structural rather than incidental. Drawing on ethnographic research in China between 2013 and 2020, I examine how the “parallel internet” operates through managed uncertainty—and what this produces for those who must navigate it. The paper’s normative commitment is to those on whom the burden of uncertainty falls hardest—and to understanding what it means to live within systems whose rules cannot be known.

2 The Limits of Transparency

The interdisciplinary field of Fairness, Accountability, and Transparency has emerged as a critical response to algorithmic power. Scholars have documented how algorithms encode racial bias [2, 20], automate inequality [10], and operate as opaque “black boxes” [21]. This work has been essential in revealing algorithmic harms and demanding accountability.

Yet the translation of this scholarship into policy intervention often relies on an unexamined assumption: that transparency is both achievable and desirable, that making systems visible will make them governable. This shapes not only academic research but policy interventions—from the EU’s “right to explanation” to proposals for algorithmic auditing.

Ananny and Crawford [1] challenged the “transparency ideal,” noting that visibility alone cannot secure accountability. Algorithmic systems are dynamic, distributed, and entangled with human practices in ways that resist simple visibility. Even when we can see inside an algorithm, we may not understand what we are seeing; even when we do understand, we may lack power to change it. Selbst et al. [25] developed a related critique by identifying “traps” that undermine fairness interventions—technical definitions can conflict with social understandings of justice, and solutions that appear valid in one setting may fail when ported across social contexts.

These critiques are valuable, but they stop short of a more unsettling possibility: that opacity is not merely difficult to overcome but integral to how algorithmic systems govern. Gillespie [14] documents how platforms make millions of moderation decisions daily through processes that remain invisible even as they shape public discourse. Recommendation algorithms determine what billions see, yet their criteria remain opaque even to the companies deploying them—not always because companies refuse to explain, but because the systems themselves resist interpretation. Users everywhere increasingly navigate uncertainty about what is monitored and what triggers consequences. The question is whether this is a flaw or an operating logic of algorithmic governance as such.

3 Methodology

This paper draws on ethnographic research in China between 2013 and 2020, with focus on Beijing, Shanghai, Huizhou, and Shenzhen. The research involved participant observation, semi-structured interviews, site visits to technology companies and electronics manufacturers, and ongoing conversations with entrepreneurs, developers, platform users, and industry experts. Following Seaver’s [24] argument that algorithms should be studied as “culture”—enacted through practices, discourses, and social relationships—I approached Chinese algorithmic systems ethnographically. This methodological choice is not incidental. The dynamics this paper analyzes—generalized vigilance, recursive uncertainty about what can be said, calibration to illegible rules—are difficult to detect through audits or metrics because they operate at the level of lived navigation rather than measurable outputs. A distinctive element was my role as a jury member for the China Mobile Competition (中国手机设计大赛) in 2015 and 2016, a nationwide competition organized by a research subsidiary of the Ministry of Industry and Information Technology. This provided access to industry insiders, government-affiliated researchers, and informal channels through which policy signals circulate between state and private sectors.

My positionality shaped both access and interpretation. As a Chinese-speaking researcher trained in Western institutions, I occupied an insider-outsider position enabling certain conversations while foreclosing others. The phrase “*bu hao shuo*” itself—offered by friends when I asked about government data access—indexes the limits of what can be said. These limits are not obstacles to ethnographic knowledge but part of its object: uncertainty about what is speakable is itself a datum about governance through uncertainty.

Post-2020 developments discussed in this paper—including COVID health codes and recent content moderation campaigns—are drawn from documentary sources, news reports, and informal conversations with my

informants rather than systematic fieldwork. This material illustrates how the dynamics observed during my primary research period have intensified rather than dissipated, but should be read as contextual extension rather than ethnographic evidence of the same order.

4 The Chinese Internet as Parallel Microworld

4.1 The Parallel Internet

The methodological challenge of studying algorithmic governance in China is also its analytical opportunity: the infrastructure of uncertainty is unusually visible. Over the past two decades, China has built not a restricted copy of the “open internet” but a different system with its own logic.

With homegrown platforms like WeChat, Taobao, and Toutiao, and innovations like mobile payments and social credit systems, China has created digital infrastructure rivaling or exceeding Western counterparts in convenience—while separating it from the global network. The result is what I call a “parallel internet” (the ecosystem of interconnected platforms, regulatory infrastructure, and governance practices): a massive digital network sharing most functions of the global internet yet running apart from it. The Great Firewall renders global services like Google and WhatsApp inaccessible without circumvention tools; Chinese users are accustomed to blocked search results and posts vanishing without explanation.

Yet framing the Chinese internet as a restricted version of the “real” internet misses something crucial. Whenever I discussed absent global services with people in China, they usually shrugged. Without Google, there was Baidu; without Facebook, there was WeChat Moments (朋友圈). The Chinese internet is not experienced by its users as deprivation. For many, it represents progress.

This does not mean information is completely blocked. The practice of *fanqiang* (翻墙)—“climbing over the wall,” using VPNs to bypass the Great Firewall—is widespread among certain populations. “All of my students would *fanqiang*,” a media professor at an elite Beijing university told me at an academic conference, as matter-of-factly as if describing the weather. The practice is technically illegal but widely tolerated—for certain populations.

Here lies a crucial distinction. When I spoke with food delivery riders—most were migrant workers—they had heard of *fanqiang*, but nearly all of them had never used it. The tools, the knowledge, the sense that there was something beyond the wall worth accessing—none of this was part of their digital experience. The parallel internet is not uniformly experienced. Its boundaries are differently visible depending on where you stand.

This stratification became clearer to me through Zhao (a pseudonym), a well-connected industry expert I interviewed in Shenzhen. Zhao advised the China Mobile Competition where I served as jury member. His business card—which he presented with both hands in the formal Chinese manner—carried a long list of titles: CEO of a consulting firm, adjunct professor at two universities, committee member of various government advisory boards. In China, the line between government and private sectors is rarely clear; they are usually crisscrossed. People like Zhao inhabit this interstitial space, translating between state priorities and market opportunities.

What made Zhao unusual was his directness. When I asked about the apparent contradiction between the firewall’s existence and the widespread practice of circumventing it, he laughed. “I have a Twitter account too,” he said. “You should know that Twitter is banned in China. But there is no problem for me. I call it ‘party-elite *fanqiang*’ (党精英翻墙).”

By “party,” he meant the Chinese Communist Party. Those who contribute to national development and maintain loyalty to the regime are quietly permitted to access what others cannot. “Because China needs to continue to develop, we need information,” Zhao explained. “Some of us need to get information from the outside.” The parallel internet is not sealed; it is managed. Its boundaries are porous in deliberate and differentiated ways.

4.2 Microworlds

How should we understand this managed ecosystem theoretically? I propose that the Chinese algorithmic systems function as what computer scientist Paul N. Edwards [8] called a “microworld”—a concept that illuminates how algorithmic platforms operate as spaces of governance.

Edwards developed the microworld concept to describe simulated environments within computers: spaces that are “partial, internally consistent but externally incomplete” (p. 109). Computer simulations create “worlds without irrelevant or unwanted complexity” (p. 109). A microworld echoes the real world without sharing its full complexity; it operates by simplified, rule-governed logic. Users must navigate according to the microworld’s rules, which may not correspond to rules outside it.

In my formulation, the term captures a paradox: a simplified system that has become an inescapable world. The *micro* refers not to physical scale but to epistemic reduction—a rule-governed environment whose complexity is radically less than the social world it mediates, rendering behavior computationally manageable even when users number in the billions. The *world* refers to mundane indispensability—the fact that this simplified environment has become the infrastructure of daily existence: payments, transit, communication, social life. Neither condition alone produces governance through uncertainty. A micro system without world (a game, a simulation) can be exited. A world without micro (the unquantified social world) resists the continuous, personalized, opaque evaluation that algorithmic governance requires. The microworld is where both conditions meet—and that intersection is made possible only by computational infrastructure mediating mundane daily life.

A microworld¹, as I use the concept, has three defining features. First, it is *rule-governed*: behavior within it is structured by logics that may not correspond to rules outside. Second, it is *consequential*: the microworld’s categorizations produce real-world effects on mobility, opportunity, and life chances. Third, it is *epistemically opaque*: users know the microworld has rules but cannot fully see what those rules are or when they might change. This combination makes microworlds spaces of governance, not merely participation.

More precisely, the “rules” governing a microworld operate at multiple registers: evaluation criteria (what factors determine scores or status), access conditions (what thresholds enable or restrict participation), and enforcement thresholds (what triggers penalties or sanctions). Each of these can be independently opaque—a user may know the access conditions for a platform but not the evaluation criteria determining their standing within it, or may understand the general scoring logic but not the enforcement thresholds that trigger consequences. These distinctions become concrete in the Chinese case. It is this layered illegibility that makes the microworld epistemically opaque in the relevant sense.²

WeChat evolved from a simple messenger into a comprehensive ecosystem: communication, social media, payments, shopping, government services. With over one billion monthly active users, it has become what Tencent calls “a way of living” (微信, 是一个生活方式). To live without WeChat is nearly impossible. In major cities, hailing a taxi on the street has become difficult because drivers take orders through WeChat. Exchanging

¹The concept differs from adjacent notions in digital studies. Unlike “platform,” which emphasizes economic and social interdependencies, microworld foregrounds the *experiential* dimension: what it feels like to navigate a rule-governed space whose rules you cannot fully see. Unlike “digital life” [3], which implies a boundary between online and offline, microworld emphasizes that the boundary itself is indeterminate—users may not know where the dividing lines are or what lies beyond them. Unlike “infrastructure” [22], which emphasizes essentiality and ubiquity of the tech systems, microworld emphasizes epistemic closure: the difficulty of knowing the rules even while being subject to them.

²Not all consequential, rule-governed systems constitute microworlds in this sense. Consider consumer credit scoring under the U.S. Fair Credit Reporting Act: the scoring algorithm itself is proprietary and complex, but the regulatory framework mandates disclosure of factors affecting scores, provides formal dispute mechanisms, and requires independent adjudication of contested decisions. The algorithm is opaque, but subjects have institutional recourse against its categorizations. What distinguishes a microworld is not opacity alone but opacity *combined with* the absence of effective mechanisms for contesting or appealing the system’s judgments—the condition under which epistemic opacity becomes a governing mechanism rather than merely a technical limitation.

business cards is pure ritual; people simply ask to “scan” (扫一扫) the QR code on their WeChat. It is almost socially reclusive to say, “I don’t have a WeChat account.”

The microworld logic is reinforced by China’s “internet real-name registration system” (网络实名制), implemented from 2017. All mobile users must register state-issued IDs with telecommunications providers. This extends to platforms: without identity verification, core functions like WeChat Pay are disabled. The smartphone-enabled microworld ties digital identity directly to authenticated, state-legible identity.

This transforms the microworld into a governance space. By managing rules of a simplified computational environment essential to daily life, authorities effectively manage behavior without directly intervening in each action. The platform’s rules become rules citizens must navigate daily, yet these rules are internally consistent but externally opaque: users know the microworld has rules, but cannot fully see them.

4.3 Differentiated Permeability

The parallel internet is not uniformly sealed. As Zhao’s “party-elite *fanqiang*” suggests, the system tolerates selective circumvention. Even Hu Xijin, editor-in-chief of the nationalistic *Global Times* and a staunch regime defender, openly acknowledged in a 2019 interview with Hong Kong media: “With VPNs... anyone who wants to bypass the Firewall can bypass it... Everyone can do it. I do it just like everyone else.”

Margaret Roberts [23] theorizes this as “porous censorship”—a system that increases the cost of accessing certain information without making it unavailable. I extend this analysis: the porosity is not uniform but stratified. Who can easily *fanqiang*, who struggles, and who would not dare attempt it—these distinctions map onto hierarchies of class, education, political loyalty, and geography. While millions of Chinese users employ VPNs, for the vast majority, the parallel internet is the only internet.

This stratification reveals something important about fairness—a dimension FAccT has approached primarily through algorithmic bias. Here, unfairness operates not through biased outputs but through asymmetric access to the microworld’s boundaries. Zhao can *fanqiang* without consequences; a rights advocate attempting the same risks trouble. The system does not treat them unequally through discriminatory algorithms; it stratifies them through differentiated permeability to uncertainty itself. Those with resources, connections, and political standing navigate opacity more easily—a form of structural asymmetry I analyze as a fairness problem in Section 7.

5 Governance Through Uncertainty

5.1 The Uncertainty Mechanism

Under what conditions does uncertainty emerge as a governing mechanism rather than merely a byproduct of bureaucratic complexity? The Chinese case suggests three enabling conditions. First, *platform-state entanglement*: commercial platforms and state agencies share data, enforce each other’s categories, and blur the line between market convenience and political compliance. Second, *discretionary enforcement*: rules exist but their application is selective, unpredictable, and often retroactive. Third, *absence of independent adjudication*: there is no effective mechanism for contesting categorizations or appealing decisions. Under these conditions, uncertainty is not incidental friction but productive power.³

In 2014, the Chinese government announced plans for “social credit” databases that would “ensure that the trustworthy benefit everywhere, while the discredited find themselves unable to take even a single step” (使守

³“Productive” is used here in the Foucauldian sense: power that *produces* subjects, behaviors, and dispositions, rather than merely repressing or prohibiting [11]. This is a descriptive claim about what the mechanism does, not a normative endorsement of its effects. The difference from Foucault is that while disciplinary power operates through uncertainty about *surveillance* while keeping norms legible, the systems examined here operate through uncertainty about the *norms themselves*. I develop this distinction in Section 5.5.

信者处处受益，失信者寸步难行). The entanglement between state and commercial systems became formalized: In 2015, the Supreme People's Court signed a memorandum of cooperation with Sesame Credit for credit punishment of blacklisted individuals [26]. The government's blacklist (失信被执行人名单) operates alongside commercial scoring systems such as Sesame Credit, producing convergent regimes of compliance by translating legal noncompliance into market-facing consequences rather than through direct data integration [18]. By 2018, the blacklist had resulted in 17.5 million bans on plane tickets and 5.5 million bans on train tickets [19]. The logic is explicit: "Once discredited, limited everywhere" (一处失信，处处受限).

What remains uncertain is nearly everything else. Which behaviors trigger which consequences? How do commercial scores relate to government databases? What information flows between platforms and state agencies? When my Beijing friends responded with "*bu hao shuo*," they were not evading—they were accurately describing the system's epistemology. They did not know. No one knows exactly.

This uncertainty is not a design flaw. It is the productive core. Consider the phenomenology of living within it: citizens know monitoring exists, scores matter, consequences can follow, but cannot know precisely what is watched, how it is weighted, when penalties materialize. The effect is generalized vigilance—self-regulation without explicit rules. The system requires only that subjects believe monitoring is possible and consequences unpredictable.

The phrase "*bu hao shuo*," discussed in the introduction, captures this linguistically: the uncertainty propagates recursively, because one cannot know the rules and one cannot safely ask what the rules are.

5.2 The Weather of Censorship

In my fieldwork, people in China repeatedly described censorship as "like the weather"—unpredictable, ever-present, something one learns to live with rather than understand. The metaphor is apt beyond unpredictability. Weather is ambient: it affects everything yet recedes from attention until it disrupts. It creates shared conditions experienced differently: the same rain falls on everyone, but some have umbrellas. It cannot be argued with; you can only adapt.

Internet users in China are accustomed to seeing articles mysteriously removed, or reappearing later without explanation. Almost any topic—education policies, land acquisitions, public health—can become sensitive without warning. A phrase that circulated freely yesterday may be unsearchable today. A WeChat group that discussed current events for months may suddenly vanish. Everyone I spoke with had learned to second-guess the continuously shifting political climate, developing intuitions about what might be sensitive even when nothing explicit had been announced.

King, Pan, and Roberts [16] provide evidence for why this weather is so difficult to predict. Analyzing millions of social media posts, they found that censorship targets not criticism but "collective action potential"—content that might mobilize coordination. This criterion is inherently contextual, which means the rules cannot be collectively mapped: any attempt to coordinate knowledge about what is permissible is itself the kind of collective action the system targets. Users must navigate the weather alone.

Even Hu Xijin, despite running one of the regime's most reliable propaganda outlets, has had his Weibo articles removed multiple times. In China, no one can be absolutely certain what is acceptable to say. People constantly calculate what can and cannot be spoken, at risk of censorship, reporting, or worse. The safest strategy, as many told me, is to "sing along with the main melody" (主旋律)—to echo official discourse and avoid topics that might attract scrutiny.

This atmospheric quality is central to governance through uncertainty. Because rules are not fixed or legible, everyone must continuously interpret signals, assess risks, calibrate behavior. The system operates not through clear prohibitions but through cultivating a disposition—habituated alertness to shifting conditions. As a software engineer at a Chinese tech startup described it casually to me: when you cannot search for something on

your smartphone, you learn that you had better not talk about it. When a topic suddenly becomes unsearchable, you adjust without being told. The Chinese internet has become not just a tool but a mode of governance—one that works through the production and management of uncertainty itself.

What makes this powerful is precisely its indeterminacy. A system with clear rules can be navigated, tested, gamed, resisted. A system where rules are uncertain produces subjects who govern themselves, who internalize vigilance, who anticipate prohibitions before they are announced. Governance through uncertainty does not need to watch everyone continuously. It needs only to make everyone feel continuously assessed, and to ensure that the consequences of transgression remain unpredictable.

5.3 Health Codes: Infrastructure Materialized

The weather of censorship is atmospheric—pervasive but intangible. The following account draws on documentary sources and informal conversations rather than direct fieldwork, as noted in Section 3; it illustrates how the dynamics observed during my primary research period took material form. The COVID-19 pandemic revealed what happens when uncertainty’s logic materializes into infrastructure that governs physical movement itself.

By March 2020, entering even a convenience store in Beijing required scanning a QR code. The code—green, yellow, or red—appeared on your phone within seconds, determining whether you could proceed or be turned away. Within weeks, these health codes (健康码) became mandatory for navigating daily life across China: public transportation, office buildings, residential compounds, restaurants, hospitals. Hangzhou launched one of the first health codes on February 11, 2020—just four days after the city proposed it. By late March, over 200 cities had implemented similar systems with Alibaba and Tencent [5].

The system connected databases that had previously operated separately—health records, telecommunications data, transportation logs, identity verification—into a unified algorithmic layer. A person could wake with a green code and discover hours later it had turned yellow or red—without notification, explanation, or apparent change in circumstances. Because the system calculated overall “risk” rather than detecting specific exposures, the codes were “basically unfalsifiable” [5]: users could not verify or contest their categorization. During the pandemic, citizens I spoke with described checking codes anxiously before leaving home, unsure whether yesterday’s green status would remain.

This exemplifies microworld logic in its most consequential form. Unlike earlier access control—ID checks, permits, physical barriers—the algorithmic system operated through criteria that were dynamic, opaque, and unknowable, simultaneously defining what constitutes “risk,” identifying who is “risky,” and enforcing consequences—all in real-time, without explanation. The system’s fungibility—its capacity to be repurposed for different governance objectives without building new infrastructure—is a distinctly computational property.

The 2022 Henan bank protests revealed how easily this infrastructure could be repurposed beyond its stated purpose. When depositors traveled to Zhengzhou to protest frozen accounts at rural banks, their health codes turned red upon arrival—immobilizing them. According to the Zhengzhou Discipline Inspection Commission, 1,317 depositors received red codes [28]. Some received them without even traveling to Henan; they had merely scanned a location QR code that other depositors had shared in a WeChat group [12]. The infrastructure built for pandemic management proved immediately adaptable to deterring political coordination. No new system needed to be built; the existing algorithmic layer simply received different inputs.

Local officials were later punished for what state media called “abuse” of the health code system [28]. But the punishment was for this particular misuse becoming visible—not for the system’s capacity to enable such manipulation in the first place. The infrastructure remains. The uncertainty about whether any given code change is legitimate health policy, bureaucratic error, or political targeting persists. Individual abuses can be sanctioned, but the systemic condition of uncertainty that enables them remains intact.

Governance through uncertainty does not require a conspiracy. It requires infrastructure. When consequential algorithmic systems are built to categorize populations in real time, governance through uncertainty arises as an emergent property—available for exploitation whether or not anyone planned for it. The analytical question is not whether opacity is intended but whether the sociotechnical system enables this form of governance.

5.4 Emergence and Algorithmic Specificity

The cases examined above—social credit’s convergent scoring regimes, censorship’s atmospheric unpredictability, health codes’ infrastructural fungibility—reveal a common dynamic. Three features of this dynamic require articulation.

First, governance through uncertainty is not reducible to intention. What it requires is a configuration: the interaction of computational infrastructure, discretionary enforcement, and absent adjudication that produces a system whose opacity has governing effects—effects that persist whether or not anyone intended them. The Henan case illustrates this precisely: infrastructure built for pandemic management proved immediately available for political repurposing, not because anyone designed a tool of political control but because the emergent properties of the system made such use effortlessly possible.

Second, it is not reducible to any single component. It is not opacity alone, not scoring alone, not censorship or platform-state entanglement or the absence of adjudication taken individually. It is the interaction among these elements—each insufficient on its own—that produces the governing effect. This is why the microworld concept is analytically necessary: it names the enclosed environment within which these elements *interact* to produce governance, rather than merely coexisting as separate features of a political system.

Third, once these conditions are in place, uncertainty persists as a governing effect, regardless of design intent. Officials were punished for the Henan exploitation. But the infrastructure that made repurposing possible remains intact, available for the next exploitation—or for governance effects that require no deliberate exploitation at all.

The analytical value of the Chinese case lies precisely in its explicitness. Because the state’s role is overt and the political context is acknowledged, governance through uncertainty is visible in ways that make it available for theoretical specification. But the mechanism this paper identifies does not depend on that explicitness. It does not require an authoritarian designer or a deliberate strategy of opacity. It requires computational infrastructure that is indispensable, evaluation criteria that are opaque, enforcement that is discretionary, and adjudication that is absent or ineffective. Where these conditions converge, governance through uncertainty can emerge—whether the system was designed with that purpose or not. The Chinese case makes the mechanism legible. It does not make the mechanism unique. This moves beyond intention as an analytical category for understanding algorithmic harm: governance through uncertainty names a class of effects that emerge from how systems operate, not from what anyone intended them to do.

Specifying what is distinctly *algorithmic* about this mode of governance is essential. The mechanism is not new—authoritarian states have long governed through discretionary enforcement and opaque decision-making. But pre-computational uncertainty operates differently: norms are broadly knowable even when enforcement is unpredictable, the system engages subjects episodically rather than continuously, and uncertainty is confined to political life rather than penetrating the mundane. What computational infrastructure makes possible is the microworld as a *form*—an enclosed, always-on evaluative environment mediating everyday functions through continuously shifting, individually opaque criteria. This requires properties unavailable to pre-digital governance: real-time categorization at individual granularity, criteria that update without deliberate policy change, and integration into infrastructure so essential that non-participation is practically impossible. The result is not merely intensified political uncertainty but a qualitatively different epistemic condition: governance operating not at

the margins of political life but at the center of daily existence, through ambient, continuous evaluation whose criteria cannot be known.

5.5 From Discipline to Uncertainty

This extends beyond Foucault's [11] panopticon. The panopticon model also involves uncertainty: the prisoner cannot see whether the guard is watching. But this is uncertainty about *surveillance*, not about *norms*. The prisoner knows what norms require; the uncertainty concerns only whether deviation will be observed. In Foucault's broader analyses—from discipline through biopolitics and governmentality—modern power operates through a double legibility: subjects are made knowable to power through examination, statistics, and classification, while norms are made available to subjects for self-regulation. Governance through uncertainty retains the first half and dismantles the second. Subjects remain maximally legible to power—real-name registration, platform data collection, algorithmic surveillance render them more knowable than Foucault could have anticipated. But the norms, the criteria, the thresholds are made illegible to subjects. The direction of conformity is perceptible—one knows that one should be compliant and politically cautious. But the precise requirements remain unknowable: what counts as compliant enough, which behaviors affect which scores, where exactly the red line (红线) falls. The system governs not because subjects cannot perceive the direction of acceptable conduct, but because they cannot determine its precise coordinates. Self-regulation of unclear rules still produces conformity, but it also produces something more: because the threshold of sufficiency is unknowable, striving and complying have no natural stopping point. The system does not only produce subjects who follow rules. It also produces subjects who never stop trying to find them.

6 Living Under Uncertainty: Exhaustion and Withdrawal

The governance mechanisms analyzed above do not operate without producing effects on those who live within them. This section examines what it feels like to live under governance through uncertainty—and what responses become thinkable when the game's rules cannot be known.

The enabling conditions described above specify the institutional configuration from which governance through uncertainty emerges. To describe how that governance is *experienced* by those subject to it, I formalize four interlocking properties: *outcome visibility* (users can see scores, status changes, categorizations), *rule illegibility* (users cannot reliably map actions to outcomes), *dynamic shifting* (evaluation criteria change continuously, outdated even partially successful inference), and *epistemic asymmetry* (the system can infer far more about its users than users can infer about the system's logic). These properties help explain both the forms of exhaustion discussed in this section and the mechanism's possible appearance beyond the Chinese case. Two terms have emerged in Chinese discourse to name this experience: 内卷 (*neijuan*, involution) and 躺平 (*tang ping*, lying flat). Together, they articulate the social consequences of systems that make life more game-like while keeping the game's rules hidden.

6.1 Involution and the Quantified Microworld

Involution, a term borrowed from anthropologist Clifford Geertz's [13] analysis of Javanese agriculture, describes intensifying competition that produces no advancement—running faster to stay in place. The term exploded in Chinese online discourse around 2020, crystallizing a generation's sense of exhausting competition without advancement. In a widely shared interview, anthropologist Xiang Biao analyzed involution as a condition of spinning intensely while remaining in place—what he called a “gyroscope-like” (陀螺) existence [27]. It names the grueling competition among students and workers: longer hours, more credentials, ever-higher stakes, yet no improvement in life prospects.

Most discussions of involution emphasize over-competition and structural scarcity—too many people chasing too few opportunities in education, employment, and housing markets. This is accurate but incomplete. Involution's relationship to governance through uncertainty runs deeper than scarcity alone.

The sentiment preceded the keyword. Lily, a longtime informant in my research on Chinese-made mobile phones, had migrated from an inland city to Shenzhen in her twenties to sell mobile phone parts. She put it quite simply when I asked how she sees her future: “读书不如当网红”—studying is not as good as becoming an internet celebrity. The calculation was not cynical but pragmatic: if advancement through education and credentials felt increasingly arbitrary, why not pursue visibility through equally arbitrary means? This disillusionment with traditional paths—the sense that the game's rules had become illegible—is what “involution” would soon name. It is also, notably, precisely the narrative that the 2025 “negative emotions” campaign targeted for removal: “读书无用论,” the view that studying is useless.

The microworld logic does more than produce opacity; it transforms the texture of social life itself. Platforms like Sesame Credit do not simply monitor behavior—they render existence into quantified form. Scores, metrics, rankings translate human action into comparable, consequential numbers. The gamification is deliberate: colorful gauges, instant feedback, unlockable perks. Users are recruited into environments where social life takes on game structure—where everything is measured, advancement is promised, falling behind becomes visible and punishing.

This intensifies the imperative to strive. In a world without scores, social standing remains ambiguous—difficult to measure, open to interpretation, and resistant to direct comparison. But once life is quantified, the imperative becomes absolute. Scores can be compared. Rankings can be seen. The distance between oneself and others becomes a precise, measurable gap. Every interaction, transaction, behavior potentially affects numbers that determine opportunities.

Yet the clarity is partial. While the microworld quantifies outcomes, the rules generating those quantifications remain opaque. A user can see their Sesame score but not how it is calculated. They know the score matters—it determines access to loans, rentals, dating prospects—but cannot know which behaviors produce which effects. The executive who suggested buying diapers might improve one's score was not explaining the system; he was gesturing toward its fundamental illegibility. The weights, the variables, the algorithmic logic—all remain hidden.

This combination—compulsory quantification with opaque rules—produces involution's distinctive phenomenology. Because life is gamified, the pressure to optimize feels inescapable. Scores matter. Falling behind is a numerical fact. Yet because rules are hidden, strategic optimization is impossible. The rational response is diffuse intensification: work harder, study longer, accumulate more of everything, optimize across all dimensions simultaneously. If any behavior might count and none is known to count more than others, the only safe strategy is total effort.

Users do develop informal heuristics—folk theories about what the algorithm rewards or punishes [9]. But when criteria shift unpredictably, folk theories become unreliable maps that intensify rather than resolve striving: failed inference produces not resignation but redoubled effort to infer again.

Involution, understood this way, is not simply competition under scarcity. It is compulsory striving under conditions of epistemic deprivation—the exhausting imperative to optimize without knowing what optimization means.

Involution has multiple contributing causes—economic scarcity, demographic pressure, cultural expectations around achievement. The epistemic dimension this paper isolates—compulsory quantification combined with opaque criteria—intensifies these other pressures: when the rules of advancement are illegible, economic scarcity becomes harder to navigate strategically, and competitive pressure becomes more exhausting precisely because effort cannot be targeted.

The structurally linked counter-response is withdrawal. 躺平 (*tang ping*, lying flat) names the refusal to strive under conditions perceived as unwinnable—not exit from the microworld itself, which is structurally impossible when platform infrastructures are prerequisite for daily life, but withdrawal from optimization within it.⁴ That the Chinese state has moved to censor such “negative discourse” reveals what threatens governance through uncertainty most directly. In 2025, the Cyberspace Administration announced a campaign targeting content promoting narratives like “effort is useless” (努力无用论) and “studying is useless” (读书无用论) [7]—not because these expressions mobilize collective action, but because they articulate non-participation as a legitimate response. Governance through uncertainty requires subjects who remain in the game. Those who withdraw are no longer governable through uncertainty.

How involution and lying flat distribute across class, geography, generation, and platform context—and under what microworld conditions one response predominates over the other—are empirical questions this framework opens but this paper does not resolve.

6.2 Beyond the Chinese Case

This paper does not argue that Western algorithmic systems operate in the same political or institutional context as those examined in China, nor that the consequences of opacity are equivalent across regimes. The Chinese case clarifies dynamics that remain analytically under-specified elsewhere—and the four-property framework introduced above provides a way to examine whether those dynamics appear in other settings.

Content moderation on major platforms in the Western context offers a test case far removed from the Chinese political context. Creators and users can see when content is removed or reach suppressed—outcomes are visible. But the criteria generating those outcomes remain opaque: community guidelines are stated at a level of generality that provides little actionable guidance, enforcement frequently diverges from users’ reasonable interpretations, and moderation policies change without advance notice. The platforms, meanwhile, can infer far more about user behavior than users can infer about the platform’s logic. All four properties appear in structurally analogous form (outcome visibility, rule illegibility, dynamic shifting, and epistemic asymmetry). What differs is not the mechanism but its weight—and, crucially, the institutional conditions surrounding it. Platform users in liberal democracies face economic precarity and expressive constraints; they do not, in general, face restrictions on physical mobility or liberty. And institutional recourse, however imperfect, exists. State monopoly on legitimate violence transforms epistemic uncertainty from a market condition into one with existential stakes. The mechanism is structurally analogous; the consequences and available contestation differ.

Involution and lying flat, seen in this light, name responses recognizable wherever participation is compulsory, evaluation consequential, and the rules governing advancement unknowable. The framework developed here—microworlds, differentiated permeability, epistemic burden as a dimension of fairness—is grounded in the Chinese case but not limited to it. Whether and how governance through uncertainty operates in other algorithmic environments are questions that require their own empirical investigation, whether ethnographic, computational, or policy-analytic. This paper’s contribution is to make the mechanism visible and the analytic vocabulary available.

⁴The term 躺平 (*tang ping*) gained traction after a viral post on Baidu Tieba in 2021 by a user calling himself “Kind-Hearted Traveler,” who described living on occasional work and minimal expenses: “Lying flat is my wise movement... Only by lying flat can humans become the measure of all things” [29]. The post was quickly censored, but the concept spread; “lying flat” has become one of the most popular terms on the Chinese internet.

7 Implications for FAccT

7.1 Rethinking Transparency

The FAccT framework treats opacity as a problem to solve: algorithmic systems are “black boxes” to be opened, decisions are to be explained, evaluations disclosed. This assumes transparency is achievable.

The Chinese case suggests a different possibility: opacity is not incidental to algorithmic governance but integral to its operation. The systems analyzed here do not merely fail to be transparent; they govern *through* their opacity. Uncertainty about what is monitored, what triggers consequences, what rules apply—this is the mechanism recruiting subjects into self-regulation. Making such systems “transparent” would not simply fix them; it would transform them into something different, governing through other means.

This does not render transparency worthless. But transparency is not a master solution applicable to all algorithmic harms. Some systems may be rendered more accountable through disclosure. Others operate precisely through illegibility. The analytical task is to distinguish these—to ask not only “how can we make this system transparent?” but “what work does opacity do in this system’s operation?”

This position extends critiques already developing within the FAccT community. Corbett and Denton’s [6] call to “interrogate the T in FAccT” challenges simplistic approaches to transparency—arguing that transparency interventions can be superficial, performative, or even counterproductive when they privilege visibility without addressing underlying power and information asymmetries. The analysis here goes further: not only can transparency interventions fail, but for systems that govern *through* opacity, transparency is not merely insufficient—it is structurally ill-suited to the governing logic at issue.⁵

7.2 Rethinking Fairness

FAccT’s fairness discourse has focused primarily on algorithmic bias—disparate impact across protected categories, unintended unfairness in model weights.

The Chinese case reveals a different fairness problem: the differentiated permeability discussed in Section 4 is not a bug or unintended bias—it is architecture. The system stratifies access to information by political loyalty, class, education, and geography.

But the asymmetry runs deeper. Governance through uncertainty produces stratification not through biased outputs but through differential capacity to navigate opacity. The platform can infer a lot about users, but users can infer very little about how rules are determined. This epistemic asymmetry does not fall equally. Those with resources, connections, and political standing navigate uncertainty better—they have informal channels, can absorb mistakes, know people who know things. The “weather” of censorship is the same weather, but some people have umbrellas.

This reframes fairness from “does the algorithm have hidden bias?” to “who can bear uncertainty and who cannot?” Transparency interventions, if they came, would benefit those already disadvantaged by opacity. But they might also threaten the informal advantages of those who navigate through connections and insider knowledge. Uncertainty is the medium; asymmetric capacity to navigate it is the mechanism of stratification.

FAccT has developed sophisticated tools for detecting bias in algorithmic outputs. We have fewer tools for analyzing how opacity itself distributes advantage. The Chinese case suggests this is urgent. Governance through uncertainty is never equally uncertain. What matters is not only whether the algorithm treats people fairly, but who is positioned to survive when no one can fully see the rules.

The paper’s normative stake is here: not in endorsing or condemning opacity, but in insisting that its burdens be made visible and its distribution be contested. Accountability under governance through uncertainty

⁵The issue is not whether existing transparency mechanisms are performative or sincere, but whether full transparency is structurally achievable in systems whose opacity is a technical and epistemic condition rather than merely a political choice. I develop this distinction further in Section 7.3.

cannot wait for full transparency. It requires contestability, recourse, and procedural safeguards—mechanisms that protect those subject to consequential algorithmic categorization even when the system’s criteria cannot be rendered fully legible.

7.3 Structural Opacity

I have argued that governance through uncertainty is not uniquely Chinese but an emergent property of algorithmic governance that the Chinese case makes visible—one that arises when computational systems become indispensable, enforcement is discretionary, and evaluation criteria remain epistemically opaque to those subject to them. This claim requires careful articulation. It does not rest on the assertion that Western systems are the same as Chinese ones—the political stakes of transgression differ dramatically—but on the observation that the *conditions* under which governance through uncertainty emerges are not confined to any single regime type.

Three sources of opacity must be distinguished. The first is *political*: the deliberate withholding of information for purposes of control, as in much of the Chinese case analyzed here. The second is *commercial*: the protection of proprietary algorithms and trade secrets, as in the opacity of Western platform recommendation systems—what Zuboff [31] has characterized as a defining asymmetry of surveillance capitalism. The third is *technical*: the resistance of algorithmic systems to human interpretation not because someone chose to hide the logic but because it cannot be expressed in human-legible form. Burrell [4] distinguishes these as intentional secrecy, opacity from technical illiteracy, and a deeper opacity arising from the characteristics of machine learning itself, especially where high-dimensional optimization does not align with human semantic explanation. The third form is particularly relevant here, because it suggests that some algorithmic systems are not made legible by disclosure alone: even when access is granted, their operative logic may remain difficult to render in the explanatory terms that transparency frameworks often presume. As such systems are retrained, tuned, and scaled across multiple features and components, the effective criteria of classification may shift in ways that are difficult for users—and sometimes even for designers—to fully interpret.

The framework proposed here applies across all three sources because the *mechanism* does not depend on the *origin* of opacity. Whether rules are hidden for political reasons, protected for commercial ones, or technically illegible for architectural ones, the phenomenological condition is the same: subjects navigate consequential systems whose criteria they cannot know. The behavioral responses—intensified striving, anxious self-regulation, eventual withdrawal—follow from the epistemic condition, even though their form and stakes remain shaped by political context.⁶

Because the microworld’s power works through the *interaction* of technical systems, dynamic enforcement practices, and user behavior in ways that resist decomposition into auditable components, transparency interventions—however well-designed—may be structurally insufficient for systems that govern through the construction of inescapable computational microworlds.

The challenge for FAccT, then, is not only whether algorithms are sufficiently transparent, but whether the field’s frameworks adequately grasp systems that govern through the construction of opaque, consequential, self-enclosed environments—and what accountability looks like when full legibility cannot be assumed.

8 Conclusion

This paper began with a phrase: “*bu hao shuo*.” Hard to say; not good to say. When my friends in Beijing offered this response to my question about government access to Sesame Credit data, they were not evading. They were

⁶The literature on strategic classification [15, 17] examines what happens when agents can observe or infer evaluation criteria and optimize strategically. Governance through uncertainty describes the inverse condition: high-stakes evaluation under persistent rule illegibility, where the dominant response is not targeted evasion but diffuse intensification. Both frameworks concern the relationship between rule legibility and behavior, but the directionality differs. Strategic classification asks how systems should respond when subjects can see the rules; the framework developed here asks what happens to subjects when they cannot.

accurately describing the epistemological condition of the system they inhabited. They did not know what was monitored, what was shared, what consequences might follow. And they could not safely pursue the question further.

The analysis developed here has taken that phrase seriously—as a description of governance through uncertainty’s epistemic condition rather than a failure of transparency. The microworld concept specifies how computational infrastructure constructs enclosed, consequential environments whose criteria resist legibility—and whose governance operates through that resistance. The social consequences are legible in the Chinese case: involution names compulsory striving under epistemic deprivation; lying flat names the refusal. That the state has moved to censor lying flat discourse reveals how threatening non-participation is to a system that depends on subjects who continue to play.

The field of fairness, accountability, and transparency has developed powerful tools for analyzing algorithmic systems from the perspective of their designers—fairness metrics, audit frameworks, strategic classification models. It has fewer tools for analyzing what it means to live inside those systems: the phenomenology of navigating high-stakes environments whose rules cannot be known. The analysis developed here—grounded in years of fieldwork within one such environment—suggests that this experiential dimension is not supplementary but essential. If governance through uncertainty is an emergent property rather than a design flaw, then accountability cannot wait for transparency. It requires contestability mechanisms that function even when criteria are structurally illegible, and methods for measuring epistemic burden—not only whether algorithms produce biased outputs, but who bears the cost of not knowing.

“*Bu hao shuo*” is not only a Chinese condition. It is an emergent property of algorithmic governance—one that arises under different conditions, takes different forms, and meets different institutional responses, but that no system producing consequential, opaque, inescapable evaluative environments can be assumed to avoid. The question facing researchers, policymakers, and those who live within these systems is not simply how to achieve transparency. It is how to govern, and how to live, when the rules cannot be fully known—and how to ensure that the burden of not knowing does not fall, as it so often does, on those least able to bear it.

Generative AI Usage Statement

I used ChatGPT 5 (OpenAI) and Claude Opus 4 (Anthropic) for editorial assistance (grammar, formatting, organization, and language clarity). Neither tool was used to generate research claims, evidence, analysis, or interpretations. All substantive content, theoretical arguments, and ethnographic material were authored by me.

Ethical Considerations

This research involved ethnographic fieldwork with human participants in China. All interviews were conducted with informed consent, and participants were informed of their right to withdraw at any time. Pseudonyms are used throughout to protect participant identities. Given the politically sensitive nature of some topics discussed, additional care was taken to obscure identifying details that could place participants at risk. The study received approval from the relevant institutional research ethics board.

Acknowledgments

I thank Hugh Raffles, Ackbar Abbas, Anthony Dunne, and Janet Roitman for their guidance on the broader project from which this paper develops. I thank Tong Lam and Janice Stein for valuable comments on the theoretical framework and related work. I am grateful to Karina Vold and the Periscope Research Lab for the opportunity to present an earlier version of this paper, and to Rachel Silvey and Rick Halpern for their support from the Munk School of Global Affairs and Public Policy. This research was supported in part by the National Science Foundation and the Wenner-Gren Foundation for Anthropological Research. Any opinions, findings,

and conclusions expressed in this paper are those of the author and do not necessarily reflect the views of these organizations.

References

- [1] Mike Ananny and Kate Crawford. 2018. Seeing without knowing: Limitations of the transparency ideal and its application to algorithmic accountability. *New Media & Society* 20, 3 (2018), 973–989. doi:10.1177/1461444816676645
- [2] Ruha Benjamin. 2019. *Race After Technology: Abolitionist Tools for the New Jim Code*. Polity, Medford, MA.
- [3] Tom Boellstorff. 2008. *Coming of Age in Second Life: An Anthropologist Explores the Virtually Human*. Princeton University Press, Princeton, NJ.
- [4] Jenna Burrell. 2016. How the machine “thinks”: Understanding opacity in machine learning algorithms. *Big Data & Society* 3, 1 (2016), 1–12. doi:10.1177/2053951715622512
- [5] Wanshu Cong. 2021. From pandemic control to data-driven governance: The case of China’s health code. *Frontiers in Political Science* 3 (2021), Article 627959. doi:10.3389/fpos.2021.627959
- [6] Eric Corbett and Remi Denton. 2023. Interrogating the T in FAccT. In *Proceedings of the 2023 ACM Conference on Fairness, Accountability, and Transparency (FAccT ’23)*. ACM, New York, NY, 1624–1634. doi:10.1145/3593013.3594104
- [7] Cyberspace Administration of China. 2025. 中央网信办部署开展“清朗·整治恶意挑动负面情绪问题”专项行动. https://www.cac.gov.cn/2025-09/22/c_1760258688713582.htm. September 22, 2025.
- [8] Paul N. Edwards. 1990. The army and the microworld: Computers and the politics of gender identity. *Signs* 16, 1 (1990), 102–127. doi:10.1086/494647
- [9] Motahhare Eslami, Aimee Rickman, Kristen Vaccaro, Amirhossein Aleyasen, Andy Vuong, Karrie Karahalios, Kevin Hamilton, and Christian Sandvig. 2015. “I always assumed that I wasn’t really that close to [her]”: Reasoning about invisible algorithms in news feeds. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems (CHI ’15)*. ACM, New York, NY, 153–162. doi:10.1145/2702123.2702556
- [10] Virginia Eubanks. 2018. *Automating Inequality: How High-Tech Tools Profile, Police, and Punish the Poor*. St. Martin’s Press, New York, NY.
- [11] Michel Foucault. 1977. *Discipline and Punish: The Birth of the Prison*. Pantheon Books, New York, NY.
- [12] Nectar Gan. 2022. China’s bank run victims planned to protest. Then their health codes turned red. CNN, June 15, 2022. <https://www.cnn.com/2022/06/15/china/china-zhengzhou-bank-fraud-health-code-protest-intl-hnk>.
- [13] Clifford Geertz. 1963. *Agricultural Involution: The Processes of Ecological Change in Indonesia*. University of California Press, Berkeley, CA.
- [14] Tarleton Gillespie. 2018. *Custodians of the Internet: Platforms, Content Moderation, and the Hidden Decisions That Shape Social Media*. Yale University Press, New Haven, CT.
- [15] Moritz Hardt, Nimrod Megiddo, Christos Papadimitriou, and Mary Wootters. 2016. Strategic classification. In *Proceedings of the 2016 ACM Conference on Innovations in Theoretical Computer Science (ITCS ’16)*. ACM, New York, NY, 111–122. doi:10.1145/2840728.2840730
- [16] Gary King, Jennifer Pan, and Margaret E. Roberts. 2013. How censorship in China allows government criticism but silences collective expression. *American Political Science Review* 107, 2 (2013), 326–343. doi:10.1017/S0003055413000014
- [17] Jon Kleinberg and Manish Raghavan. 2019. How do classifiers induce agents to invest effort strategically?. In *Proceedings of the 2019 ACM Conference on Economics and Computation (EC ’19)*. ACM, New York, NY, 825–844. doi:10.1145/3328526.3329584
- [18] Genia Kostka. 2019. China’s social credit systems and public opinion: Explaining high levels of approval. *New Media & Society* 21, 7 (2019), 1565–1593. doi:10.1177/1461444819826402
- [19] Joe McDonald. 2019. China bars millions from travel for “social credit” offenses. AP News, February 23, 2019. <https://apnews.com/article/9d43f4b74260411797043ddd391c13d8>.
- [20] Safiya Umoja Noble. 2018. *Algorithms of Oppression: How Search Engines Reinforce Racism*. New York University Press, New York, NY.
- [21] Frank Pasquale. 2015. *The Black Box Society: The Secret Algorithms That Control Money and Information*. Harvard University Press, Cambridge, MA.
- [22] Jean-Christophe Plantin and Gabriele de Seta. 2019. WeChat as infrastructure: The techno-nationalist shaping of Chinese digital platforms. *Chinese Journal of Communication* 12, 3 (2019), 257–273. doi:10.1080/17544750.2019.1572633
- [23] Margaret E. Roberts. 2018. *Censored: Distraction and Diversion Inside China’s Great Firewall*. Princeton University Press, Princeton, NJ.
- [24] Nick Seaver. 2017. Algorithms as culture: Some tactics for the ethnography of algorithmic systems. *Big Data & Society* 4, 2 (2017), 1–12. doi:10.1177/2053951717738104
- [25] Andrew D. Selbst, danah boyd, Sorelle A. Friedler, Suresh Venkatasubramanian, and Janet Vertesi. 2019. Fairness and abstraction in sociotechnical systems. In *Proceedings of the 2019 Conference on Fairness, Accountability, and Transparency (FAT* ’19)*. ACM, New York, NY, 59–68. doi:10.1145/3287560.3287598

- [26] Supreme People's Court of the People's Republic of China. 2015. 最高法联手芝麻信用网络惩戒失信见成效. December 20, 2015. <https://www.court.gov.cn/zixun/xiangqing/16351.html>.
- [27] Qianni Wang and Shifan Ge. 2020. 专访 | 人类学家项飙谈内卷：一种不允许失败和退出的竞争. 澎湃新闻, October 22, 2020. https://www.thepaper.cn/newsDetail_forward_9648585.
- [28] Phoebe Zhang. 2022. China officials who abused health codes to stop bank protests punished. South China Morning Post, June 23, 2022. <https://www.scmp.com/news/china/politics/article/3182742/china-officials-who-abused-health-codes-stop-bank-protests>.
- [29] Wangqing Zhang and Mengqiu Liu. 2021. Tired of running in place, young Chinese “lie down”. Sixth Tone, May 27, 2021. <https://www.sixthtone.com/news/1007589>.
- [30] Christina Zhao. 2018. “Black Mirror” in China? 1.4 Billion citizens to be monitored through social credit system. Newsweek, May 1, 2018. <https://www.newsweek.com/china-social-credit-system-906865>.
- [31] Shoshana Zuboff. 2019. *The Age of Surveillance Capitalism: The Fight for a Human Future at the New Frontier of Power*. PublicAffairs, New York, NY.