

CHAPTER 1

Automating Governance and Data-Driven Scoring in China: A Critical Introduction

Haiqing Yu and Rogier Creemers

Abstract

This chapter offers a critical introduction to key concepts, central questions, and the scope of this volume. It reviews how China has encouraged and employed automation technologies and toolkits to enhance social governance. Central to the discussion about automation and governance is the role of data in tracking, sorting, dividing, categorizing, and scoring practices. The chapter discusses examples and discontent in China's data-driven governance and its datafication and dataveillance practices in its digital innovations. China offers a reference point for rethinking the politics of automation and data-driven systems being used in governance in the 21st century. The global dimension of China's informationization, digitization, and automation in transforming and reshaping our political systems, socioeconomic fabrics, and cultural ethos warrants nuanced, contextual, and comparative analyses from multidisciplinary researchers. By not merely contrasting China to Western expectations but considering it as a starting point in approaching questions of technology and society, this book intends to contribute to a more inclusive and culturally diverse approach to scholarship on automated decision-making and society.

Keywords: Automating governance; Social governance; Regulatory state; Data; Scoring; Global digital China

Introduction

Perhaps more than any other government in the world, the Chinese government is embracing the potential of artificial intelligence (AI), blockchain, and other data-driven technologies in governing its society and economy. Over the past decade, it has released a series of major policies and plans on using big data (State Council 2015), artificial intelligence (State Council 2017), and blockchain (Yang, Yang and Liang 2022) for constructing a “digital China.” (Central Committee and State Council 2023). In the 14th Five-Year Planning cycle, a dedicated document provides a blueprint for the overall digitization of government services (NDRC 2021). This development has seen a terminology change in China's official documents written since the 18th congress of the Chinese Communist Party (CCP) in November 2012.

When referring to social control and management the language has shifted from “social management” (*shehui guanli* 社会管理) to “social governance” (*shehui zhili* 社会治理). More recently, “digital” has been added as a modifier to social governance, as in “digitized governance” (*shuzihua zhili* 数字化治理) and “digital transformation of social governance” (*shehui zhili shuzihua zhuanxing* 社会治理数字化转型) (Li and Wu 2022; Xiang 2021). It is required that the digitized social governance system is led by the CCP, coordinated by the government, participated in by social actors and the public, protected by law, and sustained by science and technology. In other words, social governance is centrally controlled and coordinated—by what Xi has termed “top-down design” (*dīngcēng shèjì* 顶层设计)—and enabled by an evolving digital, smart, new type of infrastructure. Science and technology, known a century ago as “Mr Science” (*sai xiānshēng* 赛先生) during the May Fourth Movement in 1919, is again called upon to usher in a new era of digitization (*shuzihua* 数字化) and intelligentization (*zhīnénghuà* 智能化).

AI is a key component of efforts to build the new type of infrastructure, networked and data-centred, to enhance China’s productivity and prosperity (China Information Centre 2021). The judiciary is experimenting with AI and big data to enhance its efficiency and reduce the workload of its overburdened and under-resourced staff (Papagiannenas 2022). Increasingly powerful algorithmic sorting and facial recognition tools are incorporated into growing surveillance networks, as China’s ever more innovative and sophisticated tech companies continuously improve their products in the hope of attracting lucrative government procurement contracts (Batke and Ohlberg 2020). Tech giants such as Huawei, DJI, and Hikvision supply a range of hardware and software to equip China’s massive surveillance networks, while Baidu, Alibaba, and Tencent (BAT) are contractors and operators for e-government services (from city services to the health code), as well as for consumer-oriented third-party businesses of public and private suppliers (including of services such as payment, travel and transportation, shopping, and entertainment). The private sector has enthusiastically adopted data-enabled tools for a range of purposes, from targeted advertising and content recommendations to work assignment algorithms for delivery drivers. In short, Chinese individuals are increasingly confronted by a context in which both governmental and corporate actors are engaging in automated decision-making (ADM).

The nature and applications of ADM tools and technologies are highly diverse. They differ in technological terms. Some applications require very sophisticated forms of AI, machine learning technologies, and blockchains, as well as very large sets of training data; others—such as some functions of the social credit system—are far less technological, employing traditional bureaucratic data collection and entry methods with the help of Weixin (China’s all-in-one super app). They differ in terms of functionality. Some systems, such as the health code system, are aimed

at scoring and sorting individuals based on sets of behavioral parameters, relational, temporal, and spatial, with consequences flowing from that evaluation as in China's COVID-19 pandemic control (2020–2022). Others, such as the application of facial recognition technologies in policing and surveillance, are intended to facilitate citizen identification and community safety. Apart from technology and functionality, they also differ in their purposes. Government agencies tell us that they use ADM tools and systems to improve efficiency and thus achieve more coordinated social and economic governance. Companies are reported to have used these technologies to streamline workflows and maximize revenue.

There are different types and levels of ADM in practice. Ulrik Roehl (2022), for example, offers a three-level, six-level typology in administrative ADM, from “no automation” and “semi-automation” (acquisition and presentation of data, suggested procedural steps, supported decisions) to “full automation” (automated decisions and autonomous decisions). He cautions against technologically deterministic understandings of ADM, calling for contextualizing of technology in user experiences and cultural milieus. We heed his advice about contextualizing and localizing ADM systems including their failures (see Chapter 2). That means the term “automation” will be used in a broad sense in this book. On the one hand, it can include applications in social governance systems that do not necessarily contain sophisticated digital decision-making tools. On the other hand, it may also include highly targeted price discrimination algorithms used by e-commerce platforms (see Chapter 9).

Moreover, the deployment of ADM tools and technologies is taking place within rapidly changing socioeconomic and geopolitical contexts. Concerns around surveillance, for instance, have been a major impetus for justifying the imposition of US export sanctions against China, and for foreign governments to review the domestic use of China-sourced surveillance technologies. Both Chinese citizens and its government are increasingly aware of data-related privacy concerns reflected in the promulgation of the Personal Information Protection Law. This law seeks to navigate the tension between ensuring autonomy over the collection and use of data pertaining to individuals while disciplining data-handling activities within government departments on the one hand and enabling police and security services' operations on the other. Lastly, it is also feeding into learning processes in government itself.

In sum, China offers a reference point for rethinking the politics of automation and data-driven systems being used in governance in the 21st century. The country has embraced AI, blockchain, and big data technologies as key drivers of economic growth, technological innovation, and digital governance. It offers excellent case studies through which to examine the technological, sociocultural, and geopolitical dimensions of digital strategies in governance innovation. It also

offers a comparative framework for examining the role of automation technologies in reshaping state-society relations. In this introduction, we highlight two key concerns that run throughout this volume, concerns regarding the automation of governance in China and, central to it, the role of data-driven scoring practices.

Automating Governance in China

China's journey toward digitizing and automating governance started as early as the 1950s with its attempts at simplifying and automating Chinese characters in typing, printing, and computer inputting (Tsu 2022). As Tsu has pointed out, through the technological revolution of the Chinese script China has not only caught up with developed countries in technological and economic development but has also restored its confidence. The only dream it is now chasing is the digital future that it envisions as the "Chinese dream" and "community of common destiny" (buzzwords of Chinese President Xi Jinping), often portrayed in the West in dystopian terms (by Hayes 2020, for example). Although China has a long history in its experimentation with digital and automation technologies in social governance, in the Xi era it has taken a great-leap-forward approach to digital governance.

For decades, the CCP has cultivated a technocratic approach to governance. From the 1980s onwards, it has married the Marxist principle—that the historical evolution of human societies progresses along a scientifically predictable path—with an eclectic set of ideas derived from cybernetics, systems theory, and social management theory (Knight 2025). Governing, defined as implementing a science-based programme for human progress, thus became intimately intertwined with social science research on the assumption that, if more data was collected and better analyzed, scientifically optimal policies could be derived. Similar to the Webb Space Telescope enabling greater knowledge about the universe, the Chinese government believes that more data and greater computational power provide the instrumentation for deepening its science-based governance capabilities and realizing more of the promise of its theoretical framework for achieving progress. In the government's view, digitization carries many benefits. Automating governance may reduce, for instance, much of the discretion (and thus corruption) of local officials, either by removing decision-making power from them or by exposing them to greater scrutiny.¹

¹ This is also a goal of the "rule of law" reforms implemented over the past decades. In official texts, "rule of law" (*fazhi*) is often juxtaposed with "rule of man" (*renzhi*), suggesting that the leadership intends it to reduce human inconsistency rather than to impose meaningful legal constraints on state action as the Western reading of this term implies.

Yet the transformation of these broad ideas and ambitious goals into actual, concrete policy, regulation, and governance practice raises many questions about both ADM systems and their inclusion in the state's operations, and about the broader reforms and transformations in the relationship between the state and citizens that they enable. Moreover, such transformations are part of dynamic processes in which lessons learned from earlier measures, new technological advances, and developments in private industry continually create new feedback loops and impetus for change. It would therefore be beside the point to try to identify a specific "Chinese" approach to ADM at this stage. Instead, many insights can be gained by taking a more granular approach, focusing on specific ADM practices, applications, or discourses, and connecting this micro-perspective to macro-level questions both about China and about the evolution of ADM technologies worldwide.

In this sense, this book follows the dominant tendency in the literature to analyze the Chinese deployment of technologies through the perspective of authoritarian governance, focusing on discourses and applications in AI, blockchain, and other ADM technologies and systems in surveilling, profiling, categorizing, and servicing the world's second largest population. China is, of course, not a liberal democracy, and no other government in the world is as ambitious in automating its society as is China's—it is unique in that sense. However, many topics and approaches are missed through the spotlight beam on authoritarian governance in the current literature. That focus tends to foreground an antagonistic state-versus-society perspective, for instance, paying far less attention to the complexities within the CCP or the Party-state itself, or to emergent techno-political landscape and multipolarity in modes of articulations via blockchain technologies (see Chapter 3). Even if it is the goal of the CCP to retain its hold on power (and there are few incumbent political organizations or individuals who do not), that tells us little about the evolving ideas and practices underpinning the realization of that goal; these are the subject of relatively diverse debates, particularly on how to configure and manage new data-driven ADM systems at national and local levels.

China's approach to automating governance is state led, driven by a long-term strategic vision for national development and a desire to maintain information control and social stability. Most of the research on digitization in China has focused on questions of political risk-related social control, including surveillance and stability maintenance. This reflects growing anxiety about the "perfect dictatorship" (Ringen 2016) of the Chinese Party-state and its increasing digital capabilities in social and opinion management (Creemers 2017). Indeed, digitization in the judicial reforms, for example, reinforce the Party-state's authority over the judiciary while seemingly contributing to professionalism in the courts (see Chapter 4).

Lastly, regime type-based studies tend to overestimate the degree of centralization and coordination while underestimating the messiness that characterizes

realities on the ground. It bears remembering that the CCP counts more members than Germany has citizens, and that it governs over thousands of large cities and small towns. Activity at the local level is, in many ways, at least as impactful as that driven by the central government; experiments and pilot projects take place at that level and are later expanded across the entire nation—it is where central directives must be transformed in the light of local realities.

In a People's Daily article in response to the CCP's 20th congress (held in October 2022), Chen Yixin (Minister of State Security) laid out three core objectives in "improving the social governance system": *political security* (referring to terrorism; foreign interference, and sabotage), *social stability* (*weiwén* 维稳, such as collective incidents and petitions), and *public safety* (including crimes and workplace safety). He emphasized local and grassroots social governance—so-called "bottom-up logic" in response to "top-down design"—to promote the modernization of social governance (Chen 2023). Such a coordinated approach to social governance must be supported by "intelligence", that is, data-centered, networked information databases, such as the Integrated Joint Operations Platform, a massive database combining personal data automatically harvested online from public and private platforms together with information that is entered manually by on-the-ground "grid" members.² Also known as China's "Big Brother App" (Wang 2019), the platform is one of many policing and social governance platforms that require "boots on the ground" (grid management) or other work of humans in the loop to support the automation of governance.

Data-Driven Scoring in China

Central to the discussion about automation and governance is the role of data in tracking, sorting, dividing, categorizing, and scoring practices. China has been known as a scoring society with a long history of categorizing and ranking its population through policies or decrees and conventions (Ghosh 2020; Wallace 2023). As von Galahn (2012, 39) writes, "[c]ivil registration for the purpose of social control and the mobilization of labor and other resources was a cornerstone of the Chinese imperial state." Information-gathering about, and registration of, individuals, households, clans, and their assets and property was an important tool of governance in China's long and continuous history of civil registration in

² Piloted since 2004 to improve grassroots management, China has revived and optimized the grid governance practice by introducing intelligent, automatic control and communication mechanisms to enable information-sharing and data integration. It has proved pivotal in the maintenance of China's stability and its COVID-19 lockdown management. See Mittelstaedt (2022), for example.

successive dynasties for the purposes of taxation, social reproduction, labor service, military conscription, and land allocation. It enabled rulers to manage and discipline their society and to impose military-style organization and mobilization in times of crisis.

Central to the civil registration is anchoring subjects to their registered residence. The Chinese household registration system, known as *hukou* 户口, is not a completely new paradigm of Chinese socialist modernity or an exclusive invention of the CCP. For centuries household management tactics, such as the *baojia* system, have been used “to address the informational needs that result from a relatively high labour-land ratio” in a relatively densely settled agrarian economy (Szreter and Breckenridge 2012, 24). *Hukou* has been used to divide the Chinese population into urban and rural groups. It is used like a caste system under Mao to strictly control not only population mobility but also resource allocation. In the post-Mao era, the *hukou* system has been maintained but internal migration restrictions are loosened. This has created a “floating population”, with millions of young villagers migrating into the cities to work in construction, export-oriented manufacturing, household service and other jobs that urbanites would look down upon. Meanwhile, cities like Shenzhen have used the *hukou* mechanism to attract “talents”—the so-called high-quality (*gao suzhi*) Chinese citizens—by designing a points system on the basis of a numerical assessment to “scientifically” allocate citizen rights and public goods (see further discussion in Chapter 7).

Although its purpose and practice have undergone dramatic changes over two millennia, *hukou* in the PRC era has continued its long-lasting trajectory in China’s data-driven scoring governmentality through scoring, ranking, and sorting. In the post-reform era, there have been various attempts to implement innovations in this tradition of population control by making it more scientific and accountable. This follows the global trend to adopt the data logics, technologies, and automation processes of the private sector into the practices of citizen scoring and techno-social shaping of citizenship (Dencik et al. 2019).

Citizen scoring refers to “the use of data analytics in government *for the purpose of categorisation, assessment and prediction at both individual and population level*” (Dencik et al. 2019, 3; original emphasis). It is emblematic of the logics of data-driven scoring and rating in contemporary societies, from financial and commercial industries to governmental and public services. China’s social credit system is often used as an example of data-driven governance and state surveillance infrastructure (Liang et al. 2018; Backer 2019). The system is best known for its credit scoring and ranking, with low-scoring individuals and corporates being placed on blacklists (the untrustworthy categories) and those with higher scores on the red lists (trustworthy categories), thus determining different kinds of preferential treatment or punishment and cultivating an ideal, loyal citizenry (Tsai et

al. 2021; Hou and Fu 2022). Implemented by local governments, earlier attempts to embed the social credit system (SCS) in the social and cultural fabric of everyday life failed, often resulting from a backlash of popular discontent and protest. The SCS is predominantly limited to the financial sector and used as a compliance tool (see Chapter 6).

As the social credit system is normalized as an evolving method of social control rather than as the West's imagining of dystopian authoritarianism (e.g. Creemers 2018; Yu 2023), it is easy to see its relevance to the discussion about what David Lyon (2002) calls "social sorting." Social sorting "highlights the classifying drive of contemporary surveillance" and "defuses some of the more supposedly sinister aspects of surveillance processes" that "not only rationalize[s] but also automate[s]" the process of social and personal categorization (Lyon 2002, 13). Whether it is in marketing or policing, sociotechnical surveillance systems have turned people into data subjects and data doubles who are searchable via databases, on the move, and modifiable. Using the language of "efficiency, productivity, convenience, and comfort" (ibid, 18), covert practices of social sorting appear benign and innocent. In the Chinese context, social sorting takes a "pan-moralism" tendency (Bakken, 2000); in our present discussion this pan-moralism can be described as a tendency to ground arguments in morality discourses, whether in regard to the reasons for corruption and fraud or the reasons for poverty and disorder (Yu 2008).

Social sorting shapes moral conduct, disciplines subjects, and creates social stratifications based on a wide range of vectors in human variables. It exercises power over the social body when empowered by data assemblages—including digital transactional records, video images captured via facial recognition cameras, biometrics (fingerprints, iris scans, or DNA samples), geolocation tracking technologies on mobile phones, and computerized administrative files—in the name of building a trustworthy society and good citizenship (Liu, Lin, and Chen 2019; Zhang 2020). Using legal, administrative, and technical means, huge amounts of data can be collected and processed quickly to inform decision-making in social control and governance, such as policing the LGBTQIA+ activists (see Chapter 8).

During the COVID-19 pandemic, the Chinese government made epistemological claims to scientific truth and methods in pandemic control based on its ability to mobilize all public and private resources in data collection and processing. Facilitated by popular social media platforms like Weixin and AliPay and empowered by big data, facial recognition technologies, and geolocation tracking technologies, the health code system was regarded as the magic weapon in China's "success" in combating the pandemic (Yu 2022). While the health code system is now no longer used, it has never been officially abolished—it remains in people's mobile phone apps. What has been monitored and traced is not merely where people have been and what they have done, but also with whom they interact or

are connected, and how long they engaged with particular people, products, or services. Such information feeds into the datafication machinery of surveillance capitalism and digital authoritarianism.

Fourcade and Gordon (2020) argue that modern bureaucracies derive their power from information via their vast sociotechnical machinery; public and private actors and interests coningle to “mint” data and create a new form of governmentality called the dataist state (or machine learning state and artificial intelligence state). Such a dataist state exercises data-driven social, economic, and political control over individual and social bodies, often in collaboration with corporate interests (as in value extraction such as via algorithmic price discrimination), and seeks monopoly over population classification and resource distribution. The power of the dataist state is like the rhizome, with ubiquitous roots, shoots, and modalities; it is omnipresent in the social body yet lacks public accountability for its actions or for its underlying structural inequalities, stigmatization, and biases.

The examples of the social credit system and the health code system reflect a dominant ideology in China regarding the teleology of technology and the technological fix. Both systems use big data-driven social sorting and scoring technologies to modernize bureaucratic capability and to govern and organize social and economic activities. Such a trend reflects a deep-rooted desire among Chinese elites for scientific management practices rooted in “technoscientific reasoning” (Sigley 2009). Technoscientific reasoning derives from “knowledges concerned with shaping human conduct” based on modern claims to “scientific truth” and reflects “the desire within governmental and administrative projects to create certain human subjects” (Sigley 2009, 538). It derives and exercises its power from and upon the individual and social body “in order to shape human conduct and thereby forge new relations between sovereign and subject, between nation and citizen” (ibid, 542).

The devil is always in the detail, but this is not a place to discuss in detail China’s role in surveillance capitalism or its surveillance industry. The explanation of China’s obsession with data, scoring, and surveillance lies as much with Xi Jinping and George Orwell as it does in Shoshana Zuboff (2019) and Kaifu Lee (2018). It is worth pointing out, by way of Cohen (2019), that surveillance capitalism ensures that internet platform firms have control over individual data and that platforms successfully promote a culture of access-for-data and acceptance of data-driven algorithmic service methods among users by accepting cookies or terms of service for using online services. This is despite global anxieties about surveillance capitalism, which underpin current debates on the turn to the state and its role in content regulation (Flew 2019). China has led the trend for using state power to regulate digital platforms. Since 2020, Chinese regulators have imposed a series of measures intended to better protect the personal information of online platform users and limit the degree to which they can exploit data for economic gain, thus effectively

erecting significant obstacles and limitations to the data-driven “surveillance capitalist” business models of China’s digital giants (Creemers 2023).

The regulatory state worldwide increasingly adopts compliance monitoring and reporting mechanisms that are intensively managerial in orientation, outsourced to specialized professionals, and seemingly automated. The situation creates regulatory fragmentation, oversight, and even performative regulation for the sake of appearances. The ubiquity of social control and surveillance tools and systems in China—from what one can see in regular use (e.g. surveillance cameras equipped with the facial recognition technology, see Chapter 5) to what is less visible but no less controversial (e.g. queer social sorting, see Chapter 8)—does not equal total control, as many of the technologies and systems do not always work as intended, and automation is not implemented at the local and grassroots levels (see Yu 2019, for example).

Most research has focused on critiquing elite or bureaucratic enthusiasm for data-driven governance. Some have pointed out the paradox of technocracy and the aesthetics of “open” data in various smart city initiatives. It is known that after more than two decades of e-government initiatives, smart government is limited to bureaucratic efficiency and is unlikely to translate into the politics of open government (Yu and Robinson 2012). It is also known that bureaucratic knowledge and capability at grassroots levels often fall short of expectations. Local cadres and grid members are too buried in data entry and reporting to carry their jobs into the streets, and they would fake records in order to meet their targets, or resort to sending manipulated information into the big data system to con officials higher in rank (Bakken 2022). They are also prey to power struggles within the power-money-intellect iron triangle (Zhao 2008). This is seen in the Henan health code scandal in which the metadata of certain individuals were manipulated to stop them gathering in public to protest and petition, all in the name of stability maintenance and pandemic control (Yu and Zeuthen 2023). Rather than being subject to criminalization, local cadres are demoted or given demerits according to CCP’s internal disciplinary policy. They are made scapegoats of a discriminatory system designed and promoted by the state and its willing partners in the private and public sectors.

Anyone who has visited China recently will have noted the omnipotent ID card in the lives of ordinary Chinese citizens, for banking, ticketing, hotel registration, online account registration, and many other public and private services. The second-generation ID card, referred to as a “smart card” (read by any radio-frequency-identification device), has an embedded digital microchip that contains the cardholder’s information, including name, sex, birth date, address and *hukou*. Known as the e-ID card, it has gone digital, and has been piloted in more than 15 major cities since 2018; it enables people to use facial recognition technology on Weixin to verify and authenticate their identity. The government has called