Law v Algorithmic Governance

China's Social Credit Systems and other Data Experiments

Zhenbin Zuo

Faculty of Law

University of Cambridge

This dissertation is submitted for the degree of

Doctor of Philosophy

Hughes Hall

May 2024

Declaration

This thesis is the result of my own work and includes nothing which is the outcome of work done in collaboration except as declared in the preface and specified in the text. It is not substantially the same as any work that has already been submitted, or, is being concurrently submitted, for any degree, diploma or other qualification at the University of Cambridge or any other University or similar institution except as declared in the preface and specified in the text. It does not exceed the prescribed word limit for the relevant Degree Committee.

Abstract

This dissertation proposes an institutionalist framework and a 'scaling and layering' hypothesis to understand the emerging theoretical domain of algorithmic governance. More specifically, the dissertation is concerned with the debate on the relationship between 'law' and 'code', with law referring here to various accepted or well established forms of text-based legal governance, and code to emerging forms of algorithmic governance, using machine learning and other aspects of artificial intelligence ('AI'). The dissertation uses China's Social Credit System (SCS) and other data/code experiments as case studies to test the validity of the proposed framework and hypothesis.

The contributions made by the dissertation are twofold: (1) theoretical and (2) empirical. Theoretically, the dissertation provides a rationale for viewing law and algorithmic governance as complements of, rather than substitutes for each other. This rationale is to be found in the inherent trade-off which exists between scaling and layering in complex forms of legal and algorithmic governance across extended geographies and populations. Empirically, the dissertation presents new evidence on China's SCS, providing a more systematic and realistic picture of its development alongside various data experiments at both local and national level, and using the scaling-layering framework to revealing gains and problems from its mode of operation.

Acknowledgements

什么是纪念? 纪念无需记得,是心在偷偷下雪

‘What is remembrance? Remembrance does not require remembering.
It is the snow gently, silently falling onto the heart.’

-- <It Is a Hard Thing to Say Farewell> (Shanghai Rainbow Singers)
《道别是一件难事》 (上海彩虹室内合唱团)

This thesis would not be possible without the help and support of so many people. My gratitude first goes to my PhD supervisor, Professor Simon Deakin. I join many other students of Simon to say that one could simply not ask for a better supervisor. He encouraged me to trust in my own voice and to keep digging. He told me it is normal to feel like you are the only one working on your PhD subject, and that may mean you are onto something important. Simon also genuinely care for his students and people around, not just in work but also in life. Now I have started my first job in academia at Essex, I know I want to become someone like Simon, and pass on this genuine care to my own students. I also want to thank Simon's family, Elaine and Thomas, whose friendship and support, Christmas cards, books and good chats, made me feel warm, strong and at home in the darkest hours of my PhD.

I am deeply grateful to many other professors. I was inspired by Professor Tony Lawson and our discussions at the Cambridge social ontology group. Professor David Howarth and Dr Zoe Adams, who were my first year PhD examiners and helped me all the way through the PhD and my job applications. Professor Meng Gaofeng and Dr Wang Boya who inspired me both in theory and empirical work. Professor Jennifer Cobbe and Dr Christopher Markou, who started the cross-disciplinary law and technology group where this PhD’s many questions originated from. Professor Mireille Hildebrandt, whose work and generous comments during the Brussels conferences opened the imagination of my PhD. To Professor Moritz Hennemann who invited me to contribute my first book chapter written in English and there met many new friends on EU data law. To Professors Reinhard Zimmermann, Ralf Michaels and Knut Pißler, whose work and our brief meetings in Hamburg at the Max Planck Institute enlightened me on Roman law, comparative law, the theories of my PhD. To my new colleagues at Essex Law School: Professors Xu Ting, Lorna Woods, Stavroula Karapapa, Theodore Konstadinides,

Audrey Guinchard, Sabine Michalowski, Karen Hulme, Lorna McGregor, Onyeka Osuji, Alexandros Antoniou, Eden Sarid and Drs Soyeon Kim, Edward Mitchell, Zeng Jie, Claiton Fyock. You made my first academic job in the UK felt at home.

Before Cambridge, many professors were crucial to the inception of this PhD. I want to thank my professors at Harvard Law School. Professors William Alford, Lawrence Lessig, Yochai Benkler and Rebecca Tushnet, were the four main people that shaped my initial ideas on the theoretical framework and the empirical case studies on China's Social Credit Systems (SCS). Bill showed me how to work as a legal scholar on Chinese and comparative law, and always look across different disciplines, deep into the history and cultural fabrics of Chinese society. Larry inspired and encouraged me at the very beginning of my research on law and code, and his seminal work on code as 'law' and architecture is deep into my theoretical framing of this PhD. Yochai was always at the frontier of research methods and taught me to ground my big theoretical ideas with empirical findings. Rebecca showed me how to connect my empirical work with legal practices, and how to overcome my fears.

I also want to thank my professors at Peking University and Renmin University in Beijing, who helped form the very DNA of this PhD. Professor Wang Xixin has always inspired me in both work and life. He was with me in Shanghai and introduced me to the first group of people I had the privilege to conduct interviews, and I would not have come this far without his unwavering support and belief in me. Through Professor Wang I also had the opportunity to learn from and work with brilliant professors at Peking and Yale, including Professors Shen Kui, Dai Xin, Peng Chun, Yan Tian, Zuo Yilu, Zhang Taisu, Jeremy Daum, and Jamie Horsley.

At Renmin University, Professors Han Dayuan ignited my interest in 2016 on law and technology. Professor Wang Guisong guided my first dissertation, which also opened a gate for me to study US law. Professor Guo Rui introduced me to the world of corporate governance and Coase Theorem, and recommended me to do an LLM at Harvard. I am grateful for Renmin Law School's partnership with the ECtHR and the generous support of a King & Wood Mallesons scholarship. With their support I was able to work with Judge Sicilianos and the research division under Paul Harvey which proved to be a transformative experience. I'm grateful to Professor Jiao Shubing who opened a gate for me to intellectual giants like Herbert Simon, whose theories of bounded rationality in decision-making, administrative state and corporation as AI and expert systems, remained one of my fascinations during the PhD. Professor Wang Hongyun who taught me industrial organisations and antitrust on big-tech companies, which later led me to Clifford Chance Beijing at the Antitrust team for my first legal internship. Professor Huang Jun who taught me macroeconomics, and planted the seed in

me to work on Keynesian economics at Cambridge. Professor Paul Lewis at King's College London during my exchange in 2013, taught me game theory and introduced me to Ostrom's work on polycentric governance, which is deeply embedded in the framing of my PhD.

I am grateful for the generous financial supports from the Cambridge Law Journal, the Cambridge Political Economy Society Trust, the Cambridge School of Humanities and Social Science Fieldwork Fund, the Cambridge Law Faculty and Hughes Hall college. During the difficult times of the pandemic, the financial support and the facilities you provided remained a bedrock of my studies and everyday life.

I am indebted to all the interviewees who shared their time and thoughts with me. Among those who agreed to share their names, I'm particularly grateful for Dr Luo Peixin, who allowed for the snowballing of interviewees, and Dr Lin Junyue for his long-hour conversation on the genealogy of the SCS. I am also grateful for all those friends at Peking who helped me during my interviews in Shanghai, including Professor Cai Peiru, Dr Lu Pingxin and Huang Zhijie.

I'm grateful for Hughes Hall, my tutor Amy Khlor, law fellows Professors Lars Vinx and Markus Gehring who shared their ideas with me during lunch, and made the Hughes law community vibrant. To the Hughes Hall Boat Club, my captain Jonah Zankl; and to Hughes MCR, and our favourite IT officer Michael Tang, who was the secret life saver for all college events, and now my roommate and very dear friend. For my PhD friends. Nishant Gokhale, whose friendship since our Harvard times kept me strong in Cambridge. Dan Ward who helped with the final read through and revisions of this PhD thesis, and inspired me with both his academic and creative work. Pedro whose academic and practical work I always look up to, and every time meeting you and Pati simply makes me happy. Lily who always shares with me your kindness and the warmest conversations. Dora, Lora and Kenneth, all your energy make me feel encouraged and optimistic about the world. Bhumika your poetry and work keep inspiring me. Soyeon, so happy now we are office mates. And Rupa who is now at Oxford, we helped each other through our darkest time, and I simply cannot ask for a better friend in my life these years.

For the Cambridge rowers: Civia and Yizhou, Peng Li and Zhu Miao, Heqing and Grace Tang, see you on the river soon. Drag racers and the Cambridge queer community: Roeland, Niall, Jason, Antonio, Rohit, and all of you who gave me a new family in Cambridge. For the first time in my life I felt part of me came alive. For my friends at St Philip's Church, Lucy, Catherine, Chris and Eleanor, Ruth, Eric, Sally. You made me feel at home when family is so far away across the sea. Musical theatre and choir friends: I would not have survived

Cambridge without you. Kinky Boots and A New Brain will always be my favourite memories in Cambridge. Felix, Hugo, Jas, Joseph, Freya and the ANB cast who made me cry and laugh so much in those days. I will remember ‘you gotta have heart and music’, and ‘time’ to make a song. As Hugo said, we gotta make our own music, and I am proud we did. I want to thank Hugo Eberhard and Glen Lim for taking me to St Catharine’s chapel choir. I think it changed my life, and led me to the beauty of Cambridge music. Jamie, Alex, Daniel, Gwyn, Giles, Sam, June, Annie and everyone on the Prague tour 2022. For the first time somehow I felt so completely myself on that trip. I am grateful for Dr Edward Wickham and Dr Simon Jackson as my music directors at Catz and Peterhouse. You showed me the spirit of Cambridge choirs and how to listen to my own voice within the harmony. As engraved in St Catharine’s College Chapel, SURSUM CORDA, lift up your heart. My Catz and Peterhouse choir friends, you have lifted up my heart at times when I almost wanted to give up.

Thank you Angel for being there and inspiring me through the most difficult time of the LLM. You always believed in me and was the first to inspire me on starting the PhD. Daniel Wells you were there for my first year PhD viva, and I remember our discussion on evolution and statistics, concept drift, math history, and you helped draw the layered model as a volcano. Giles thank you for being there for my job interviews, faculty presentations, and trying to explain different types of mathematical infinities to me in set theory. Thanks Riccardo for teaching me how to paint, and explaining difficult math and category theories to me.

Finally my family. My parents, grandparents and the greater family who have supported and sheltered me all these years. You gave me the privilege and freedom to pursue my education and dreams so far away from home. You have sacrificed so much, worried too much, laughed a lot, and made me laugh, in order to make me who I am today. And all the difficulties, the conflicts, the tears and heartbreaks, never stopped us from loving each other. Happiness became the small moments that I used to take for granted. Father's desk full of books and paper, his back crouching a little under the lamp light when I got home from school, stilling working and never speaks much. Mother's eyes that always shine with such love that I cannot bear to see her cry. The steamy kitchen and smell of potato with pork ribs and spring onions.

Maybe after all, my PhD thesis was trying to tell a story about our existence in a world that is not a Markov process which has no memory. Our heart is engraved and layered with memories beyond the brains. As what the fox told the little prince: 'One sees clearly only with the heart. Anything essential is invisible to the eyes’.

Zhenbin Zuo

May 28th 2024

Table of contents

Chapter 1 Introduction **1**

1.1 Contours of the 'law v code' debate 1

1.1.1 Introducing an emerging field 2

1.1.2 Introducing the debate: law v algorithmic governance 4

1.1.3 The limiting contours of the debate 8

1.2 Research questions and summary of arguments 9

1.2.1 Sub-questions in each chapter 10

1.3 Contributions and motivations 11

1.4 Site of analysis: a literature review of China's Social Credit Systems (SCS) 12

1.4.1 Miscellaneous socio-legal approaches to the SCS 13

1.4.2 Assessments of the SCS's impacts, efficacy and legitimacy 14

1.4.3 Theorisations of state strategies and reputation systems: law & economics 15

1.4.4 Section conclusion: scoring, blacklisting, and data experiments 18

1.5 Structure 19

Chapter 2 Methodology **21**

2.1 Theory as the basis for methodology 21

2.2 Time and scope 23

2.3 Empirical methods 24

2.4 Interview data 25

2.5 Chapter conclusion tied to research question 26

Chapter 3 Theoretical Framework: Law v algorithmic governance, substitution or complementarity? **27**

3.1 Summary of the argument 27

3.1.1 Research questions 27

3.1.2 A metaphor of the 'layered' model 28

3.1.3 Hypothesis: complementarity between modes of governance 28

3.2 Framework and hypothesis: scaling and layering of governance 29

3.2.1 Governance: norms, laws, numbers and code 29

3.2.2 Algorithms: The 'algorithmic turn', scaling, and layering 33

3.3 From Norms to Laws: Institutionalism, Rule-of-law state, and Path-dependency 38

3.3.1 Legal institutionalism: law v social norms 39

3.3.2 Scaling: the rule-of-law state and modern capitalism 40

3.3.3 Layering: path-dependency, reinforcement and cooling effects 41
3.4 From Laws to Numbers: Statistics, Cybernetics, and the Erosion of Rule-of-law 42
3.4.1 Statistics: law geared to numbers 43
3.4.2 Scaling by cybernetics 45
3.4.3 Layering: statistics and the erosion of the rule-of-law state 46
3.5 From Numbers to Code: Computation, Big-data, and 'Freezing' 51
3.5.1 Computation: from statistics to 'learning' code 51
3.5.2 Scaling: big data, surveillance capitalism and 'information civilization' 54
3.5.3 Layering: cold start, concept drift, and freezing 55
3.6 Chapter conclusions 58
3.6.1 Research questions 58
3.6.2 Contributions 59
Chapter 4 (Semi)Automated Scoring Systems: an analysis of data-driven financial and social scoring in China's SCS 61
4.1 Summary of the argument 61
4.1.1 Research questions 61
4.1.2 Structure of the chapter 62
4.2 Central Bank v Alibaba: scaling and layering based on transaction data 63
4.2.1 The Central Bank's databases: Public credit referencing without scores 64
4.2.2 Alibaba's Sesame Credit: expansion and fall of private credit referencing after mid-2000s 68
4.2.3 Credit reporting in the U.S. and Europe compared to China 78
4.2.4 Assessment: scaling and layering in commercial scoring systems 81
4.2.5 Summary 85
4.3 Local (semi)automated scoring systems: Urban and rural China 86
4.3.1 Scoring in urban China 86
4.3.2 Rural Scoring: the rise and fall of Qingzhen's local scores 92
4.4 Chapter conclusion 99
4.4.1 Research Questions 99
4.4.2 Scaling and layering in the SCS scoring systems 99
Chapter 5 (Semi)Automated Law Enforcement: an analysis of code-driven blacklists and joint-sanctions in China's SCS 101
5.1 Summary of the argument 101
5.1.1 Research question 101

5.1.2 Structure of the chapter 102
5.2 Blacklists and joint sanctions in the SCS: roots, expansions and legal basis 103
5.2.1 Roots: the Supreme People's Court's (SPC) experiments 104
5.2.2 Scaling: national expansions of algorithmic governance (2016-) 107
5.2.3 Layering: legal basis for code in the SCS 112
5.2.4 Section conclusions 118
5.3 Case study: an assessment of code-driven SCS using interview evidence from Shanghai 119
5.3.1 A flowchart of civil procedure in the SCS and law-code interactions 120
5.3.2 Interview evidence from local enforcers: Perceived gains and loss of automated enforcement 128
5.3.3 Assessment: layering effects, concept drift and human agency in automated enforcement 134
5.3.4 Section conclusions 137
5.4 Chapter conclusion 138
Chapter 6 Data Experiments: scaling and layering towards a national digital economy 141
6.1 Summary of argument 141
6.2 Introduction: China's data experiments and strategies 142
6.2.1 China's data strategies in paper / policies 143
6.2.2 Practical questions and gap of understanding 146
6.2.3 Scope and structure of the chapter 148
6.3 Institutionalist theory and conceptual model of state strategies 150
6.3.1 Institutionalisation: State-market Co-production 150
6.3.2 Activation: Coding Property or Pooling Commons 152
6.3.3 Layering: Law-Tech Coevolution 152
6.4 Institutionalisation: Data Security, Protection, and Antitrust 154
6.4.1 DiDi's 2021 Cyber / Data Security Review 154
6.4.2 Data Protection, Gatekeeper, and Managerial Liabilities 160
6.4.3 Antitrust: Data Power and Interoperability 167
6.4.4 Summary: State-Market Co-Production via Institutionalisation 169
6.5 Activation: Data Property, Commons, or 'Party manages data' 170
6.5.1 Ant Group v the Central Bank—Licensing or Commons? 170
6.5.2 Beijing v Shenzhen: 'Use without Retaining' or Property Right? 173

6.5.3 Tianjin: ‘Party manages data’ 177
6.5.4 Section conclusion: Legal coding / pooling of data assets via activation 179
6.6 Layering: Data Cycles, Local Habitats, and Sustainable Design 180
6.6.1 Smart Court: National v Local Practices 180
6.6.2 Law-tech Coevolution and Layering in Design 182
6.7 Chapter conclusion 183
Chapter 7 Conclusions **185**
7.1 Research questions 185
7.2 Synopsis of the argument 186
7.3 Future research agenda and concluding remarks 195
Bibliography **197**
Books 197
Articles 202
Major laws and regulations (including drafts) 217
Cases and administrative actions 222
Others 223
List of Major Abbreviations and Acronyms **234**
General 234
European Union laws and concepts 234
U.S. laws and entities 234
Chinese laws, concepts and entities 234

Chapter 1 Introduction

1.1 Contours of the 'law v code' debate

This dissertation is concerned with the field of algorithmic governance and the theory of law. It draws on interdisciplinary perspectives from law, economics and computer science.

Code, or algorithmic tools based on data, machine learning ('ML') along with a number of other emerging computer technologies, are increasingly being used by a range of public and private actors as modes of governance (Hildebrandt, 2018, 2023; Danaher et al., 2017; Coglianese and Lehr, 2019; Engstrom and Ho, 2020a, 2020b; Cobbe, 2023). This thesis explores their interaction with modes of governance of more established or accepted legal nature, that is text-based law. These interactions are studied in the concrete context of China's Social Credit Systems (SCS) and other data experiments.

It should be noted upfront that in this thesis the word 'code' mainly refers to 'computer code' or 'algorithmic tools', and in the theoretical framing, a mode of governance based on computer code/algorithms. 'Coding' therefore simply refers to the activities of writing/making computer code. However, in a few other places, code has other specific meanings. For example, 'codification' or 'codify' is a socio-legal process that institutionally recognise or formalise social norms, concepts and physical materials into legal concepts (e.g. in Pistor's work, lawyers invent laws to 'code' capital, intellectual property, and land into legal concepts and then the global financial market, see Pistor, 2019). A narrower understanding of 'codification' is the process of systematic legislation through a body of lawmakers (e.g. members of a parliament). Here the product of 'codification' is a 'legal code', is a corpus of law that systematically addresses an area of law, e.g. the Roman or French civil code (Zimmermann, 2015). Such 'legal code' and 'coding' should not be confounded with the 'computer code' discussed in the thesis.

1.1.1 Introducing an emerging field

To start with an example, in *Surveillance Capitalism* (2019) Shoshana Zuboff challenged the vision of self-executing contracts presented by Google's Hal Varian. She hypothesised a scenario where a computer code can automatically enforce a car-hiring contract by locking the engine from afar, oblivious to whether there is a blizzard, a child in the car, or a mother waiting in a hospital.

Zuboff calls attention to the 'uncontracts' which use code[1] to substitute for legal contracts, in the process removing human dialogue, problem-solving and empathy from law enforcement. Where Varian believed that code could scale up techno-legal governance, Zuboff warns of a world in which 'the blankness of perpetual compliance' leads to the disappearance of both uncertainty and freedom, and 'the right to the future tense is endangered'(Zuboff, 2019, pp.332-336)[2].

China's Social Credit System (SCS), meanwhile, has already automatically enforced legal verdicts that have refused millions of people's booking of flights or high-speed train journeys due to, among other things, unpaid commercial and consumer debts (e.g. Li 2016; Dai, 2020)[3]. The technology includes simple information infrastructures and facial recognition code (based on machine learning), embedded in cameras at boarding gates. But what if there were a similar emergency situation to the one posed in Zuboff's question, where a single-parent mother needs to fly to see her child in a hospital across the country? Has the code embedded in the state-directed SCS, compared to the corporate 'uncontracts' of the surveillance capitalism in advanced market economies, led to similar dangers in removing human judgements?

In addition, China's SCS also has another major component, namely the (semi)automated scoring or rating systems. It uses data and algorithmic technologies to give confidence scores to individuals and organisations in order to regulate their behaviour risks, with some comparable structures to its US and European counterparts (e.g. Krause et al., 2023; Mac Síthigh and Siems, 2019). Would such 'scoring' systems create problems of fairness and

[1] Or in her words 'the positivist calculations of automatic machine processes'. Zuboff, 2019. p.333. Chapter 7.VI Executing the Uncontract.

[2] See similar concerns, e.g. Pasquale, 2019.

[3] See e.g. news report on China's railway facial recognition systems since 2016, Wang Wei, 'Beijing Station Pilots 'Face Scanning' for Entry Using Facial Recognition Technology'[北京站试运行"刷脸"进站 使用人脸识别技术] (Souhu.com, 5 Feb 2016) <https://www.sohu.com/a/58100500_119703> accessed 7 June 2023; The system was officially used since 2017, Li Jinlei, 'Four Major Changes in the 2017 Spring Festival Travel Rush: 'Face Scanning' Entry Available at Some Train Stations' [2017 年春运四大变化：部分火车站可"刷脸"进站] (People's Daily Online, 15 Dec 2016) <http://politics.people.com.cn/n1/2016/1215/c1001-28951537.html> accessed 7 June 2023

transparency when using data and machine-learning algorithms? Are there more systematic problems emerging when such metrics and automations are widely adopted?

Going beyond the (semi)automated enforcement and scoring systems examples within the SCS, scholars around the world are concerned about other data experiments using emerging technologies. These include, among others, 'smart courts' in China (Liebman et al., 2020, 2023; Rachel et al., 2020), judgement predictions in the European Court of Human Rights (Medvedeva et al., 2020, 2023), algorithmic supply chains (Cobbe et al., 2023) and 'AI as a service' (Lewiki et al. 2023) used as public and private infrastructures. The experiments of data and algorithmic technologies are argued to be changing the modes of governance in nearly all aspects of individual and organisational activities (e.g. Hildebrandt, 2016, 2023; Yeung, 2018; Cohen, 2019).

In March 2023 OpenAI released its GPT models including ChatGPT as an AI assistant for the general public to use. This made the debate on algorithmic governance more pertinent, with increasing concern over generative AI and large language models (LLMs) that can generate human dialogues by using machine-learning algorithms (e.g. Surden, 2024). Many judicial systems have responded by providing guidelines on how judges and legal professions should use these AI assistants[4].

It should be stated at the outset that the dissertation does not cover the debate over these generative models released after 2023. The reasons are that (1) these new technologies are still in their early stage of deployment and the literature in law is only emerging, (2) the space available is limited, and (3) the research being mostly conducted before the proliferation of such models. The thesis nevertheless is even more pertinent in light of these rapid technical developments. To understand these new AI technologies used in law and society, we need a foundational framework to understand the relationship between law and algorithmic governance. This dissertation intends to provide such a framework through both theoretical work and empirical evidence. The following part will introduce the specific debate of this thesis, which identifies the tensions between two opposing views on whether code can substitute law.

[4] See e.g. Guidance for Judicial Office Holders, England and Wales 2023 https://www.judiciary.uk/wp-content/uploads/2023/12/AI-Judicial-Guidance.pdf ; Guidelines For Use Of Generative Artificial Intelligence In Courts And Tribunals, Courts of New Zealand, 2023 https://www.courtsofnz.govt.nz/assets/6-Going-to-Court/practice-directions/practice-guidelines/all-benches/20231207-GenAI-Guidelines-Judicial.pdf Also see Case in the US where lawyers use Chat GPT to argue for a case. [2023] UKFTT 1007 (TC) https://www.bailii.org/uk/cases/UKFTT/TC/2023/TC09010.html

1.1.2 Introducing the debate: law v algorithmic governance

Within the emerging field of algorithmic governance, a crucial theme is the relationship between law and code, that is, between legal and algorithmic modes of governance. The debate can be generally understood by reference to two opposing arguments on whether law is computable, that is, whether law is functionally complete and thus can be substituted with computer code. One side proposes a hypothesis of 'legal singularity', while the other side is more sceptical of the hypothesis.

Legal singularity: a substitution hypothesis

The 'legal singularity' hypothesis / argument is that 'law will reach functional completeness, in the sense that practically any legal question will have an instantaneous and just solution.' (Aidid and Alarie, 2023, p4).

This view argues that law is currently incomplete in the sense that it is expensive, inaccessible, and frequently unjust. Many problems are identified with respect to modern legal systems, including that legal procedures are over complicated for people to use, courts and lawyers are costly, incentives are misaligned, legal doctrines can be outdated, and outcomes can be frequently unfair and unjust. Aidid and Alarie hypothesise that new technologies, in particular AI, can help solve these problems for law, and make the legal systems into computer code systems that are more stable / certain, resilient and adaptable (Aidid and Alarie, 2023; Alarie, 2016). Building on Alarie's work, Eliot argues that the discussion can be enriched with parametric analysis which uses more technical standards to measure the progress of legal singularity. He proposes to use Levels of Autonomy for AI legal reasoning systems as a start (Eliot, 2020).

Other scholars have hypothesised various forms of legal singularities. They often conclude on certain degrees of retention of human judgement and not a complete substitution of law by code. Goldsworthy for example proposed that deep-learning (DL), as a branch of ML technology, has the capacity to make possible Dworkin's most controversial theory: that there is one right answer to any legal question. This is proposed as a Dworkinian legal singularity, where the right answer was the one that 'best fit'. It is imagined that Dworkin's 'seamless web' of laws can be navigated by AI to cover both the local and global needs of different legal contexts. (Goldsworthy, 2019)

Goldsworthy relates this Dworkinian legal singularity to comparative law, where Watson and Legrand argued differently on law's nature. Watson claims that law is usually borrowed

from elsewhere and from the past, operating at places very different from where it originated, and making legal transplants easy. In contrast, Legrand argues that the harmonisation of law will only diminish variety and justice, and there is no value in comparative law to identifying universal patterns of law. The possibility of detaching law from its local context, applying law at the global level to reach a 'singularity of law', is criticised as against justice. (Legrand, 2006). Goldsworthy claims however, that the DL algorithms can achieve the balance between the local and global, and the 'singularity of law' opposed by Legrand can be avoided by DL's self-learning capacities. In this sense the legal singularity is hypothesised as not necessarily presupposing a singularity of law, and the Dworkinian singularity retains legal plurality. This balance is argued to require human judges and reason to decide on morality, values and fairness, and selecting cases that 'fit' (Goldsworthy, 2019).

This Dworkinian legal singularity is similar to an idea raised by Leibniz almost 300 hundred years ago. Leibniz in his early work tried to use mathematics to represent law, and turn human law into an exercise of computation. Modern followers of Leibniz argued that computational law can be achieved by designing new symbolic discourses, such as the Wolfram language, eventually realising the Leibniz dream (Wolfram, 2016).

In addition, other forms of legal singularity arguments have been more focused on the potential for automating aspects of legal processes. Some argue that in the coming decades technological management will 'have superseded legal rules' and become a significant regulatory instrument that co-exists with legal and social norms (Brownsword, 2015). Others examine the possibility of the 'death of rules and standards' by using 'microdirectives' carried out by predictive technology which will be feasible in the near future (Casey and Niblett, 2016).

These specific predictions suggest that code can substitute law to certain degrees, but not completely. They are aligned with some computational law literature that focuses on designing legal technologies or data / code-driven 'law' (Waddington, 2020; Hildebrandt, 2020b; Diver, 2022). Such legal technologies can take the form of using computer language to directly represent legislation and regulation in machine-readable formats, which allows for fewer interpretative gaps associated with human judgement (Huttner and Merigoux, 2022). In law enforcement settings, 'smart contract' which uses blockchain and crypto-technologies can automatically enforce contracts (De Filippi and Wright, 2018). There are also emerging practices of automating tax law (Huttner and Merigoux, 2022) and antitrust enforcement across the globe (Schrepel and Groza, 2022). They all imply that code is to some extent already facilitating and / or substituting text-based law and its practices.

This line of literature above can be viewed as, to different degrees, supporting or evidencing a 'substitution hypothesis', that is computer code can make law's functioning more complete by substituting or replacing law's various modes of operations.

Against legal singularity

Many scholars however are more sceptical of the legal singularity hypothesis. They have presented views and evidence on why data and code, instead of achieving law's completeness, can be counterproductive. From different methods and traditions, the authors argue that law to certain extent is not computable.

Deakin and Markou, for example, are concerned that algorithmic governance can lead to law's ossification. This means that law, as inherently a social institution, can become increasingly rigid and less reflective of the underlying political and social changes, when being substituted by code (Deakin and Markou, 2020). They use evidence from legal evolution and institutional theories to show how legal concepts evolve not towards efficiency but suboptimal solutions embedded in history and its socio-political and economic environments. The use of analogy and legal concepts made legal reasoning and the judicial system as a whole reflexive in the sense they shape the environment law operates in, and store information about itself and its surrounding environment within legal concepts and the epistemic community of lawyers. Meanwhile the nature of algorithmic governance based on numbers and artificial language is such that it can only perform optimisation tasks with predefined objectives that cannot capture such system-dynamic, reflexive nature of law. Evolving ML algorithms, even though with the ability to adapt to new situations, in this sense cannot look into the future but only function to optimise the existing objectives of the past. This can endanger law's function as a social-institutional cognitive system that is both designed and evolved for connecting the past into the future (Deakin and Markou, 2022).

In a similar vein, Cobbe has argued that the reflexivity of law means that law inherently makes normative assumptions and priorities when exercised by lawyers and practitioners. The legal technologies being developed will inevitably be confronted by the power dynamics and relations created within the legal system as reflexive of its social environments. Therefore, it is naive to argue that law can reach completeness or singularity by being substituted or automated by technology. She argues that the depiction of a legal singularity as being more just and neutral using AI is only prioritising market and commercially oriented norms and

values, and that the legal technology industry should caution against such techno-solutionism (Cobbe, 2020).

Hildebrandt argued that the past-facing algorithms run by machines will cause 'freezing' of human futures (2020b), in a way similar to what Deakin and Markou calls the ossification of law. The main idea is that algorithms predict the future by only looking into the past datasets and social practices, and have the tendency to reinforce that past. She further argues from both legal philosophy and computer science perspectives that text-based law's operation is inherently adaptive to radical social realities. The result is that text-driven law cannot be substituted or quantified by data and algorithms, because human existence and interactions are inherently indeterminate and are uncountable in this sense (Hildebrandt, 2021).

Aligned with Hildebrandt's approach, Diver argues that law is deliberately designed to be slow. Instead of making law operate faster and even instantaneously (as proposed in the 'legal singularity' argument), Diver argues that the 'affordance of delay' by law is a crucial feature of legal systems to protect humans and deliberation efforts. The slowness of law allowed for debates from opposing opinions to emerge, develop, and be examined by legal reasoning and evidence presented. It is argued that data or code-driven solutions risk diminishing this protective space deliberately afforded by law to individuals and communities (Diver 2020, 2022).

In addition, despite not purporting to answer whether law is computable, Moses argued that there is no single singularity for law. She proposes instead a three dimensional way to map how AI in law evolves into the future, which asks three questions on (1) the availability of the institutional setting which allows for AI adoption, (2) the capacity of the technology, and (3) the normative limits of such technical adoption. This approach is proposed as a realistic and critical framework for assessing the deployment of legal technologies through time (Moses 2020).

Another effort also aimed at providing a critical typology or assessment framework of the legal technologies is developed by Diver et al. (2022). This serves as an online interactive methodology for identifying and categorising the emerging new technologies in governance practices, in particular legal governance. This typology and its related framework aims to understand how these data and code based technologies change legal effects that rely on text, in order to map how the current law (as we know it) is being influenced by algorithmic governance.

1.1.3 The limiting contours of the debate

The opposing views on the relationship between law and algorithmic governance reveals different perceptions and methodologies to the question on whether law can be computed. Some do not directly address the question of law's computability, but instead provide assessment frameworks for continuous monitoring and observations. The debate is ongoing, and the supporters of legal singularity have, for example, replied to the concerns raised by the sceptics with some new evidence and explanations (e.g. Aidid and Alarie, 2023). Continued debates from both sides provide opportunities to observe the potential trajectories of law's coevolution, as a mode of governance based on text, with new modes of governance based on data and / or code.

This dissertation stands more with the sceptics of legal singularity, and will later provide empirical evidence to support this scepticism to law's computability. The thesis however, should first identify some limiting factors of the current debates.

The first limitation identified is that of empirical testing. The question of law's computability is well defined and argued by the sceptics, and the various dangers of code are supported with initial evidence. The sceptical literature however is confined by the empirical findings, and more evidence needs to be collected in order to understand how law and code interact in daily operations and in the long-run at institutional levels. The supporters of legal singularity or law's computability, provide also limited evidence on how new legal technologies are being deployed, or whether they can be scaled up and live up to the tasks they envisioned as making law functionally complete.

The second limitation is that of the theoretical foundations, which is the underpinnings for empirical testing. The supporters of legal singularity focus more on explaining individual cases where technology works, but do not provide a systematic framing or theoretical foundation to how they work. In contrast the sceptics have generated more theoretical frameworks and analysis which are well-rooted and documented in different traditions, with good explanatory powers. They are, however, individual efforts at analysing the question of law's computability, and have not intended to provide a theory for empirical assessment in a systematic approach. The cross-disciplinary aspects of the field are also only emerging, and many insights from computer science and social sciences could be further drawn on. The debate on law's computability, and the larger field of algorithmic governance, both require a more comprehensive framework that is conducive to empirical testing.

The dissertation intends to fill these gaps on both theoretical framing and empirical testing.

1.2 Research questions and summary of arguments

The thesis asks one overarching research question: to what extent can text-based law be systematically computed, namely substituted by computer code.

It should be stated at the outset that the thesis challenges the view that law could be ultimately replaced by code, that is, the legal singularity or substitution argument (e.g. Alarie, 2016; Aidid and Alarie, 2023). To the contrary, the thesis argues that law and code complement each other in a path-dependent way, and proposes an original framework to uncover the rationale behind. This rationale is found in the inherent trade-off which exists between scaling and layering, namely, while code has the potential of governing greater scales compared to law, it is limited by various negative effects and relies on law to avoid systematic failure.

This new framework draws on institutionalist theories that focus on reflexive systems, co-evolving organisations, common-pool resources, narratives and embedded cognition emerged from communities, and non-ergodic or time-irreversible processes (Ostrom, 2010; Cohen, 2019; Xu, 2019; Cobbe, 2020; Supiot, 2017; Aoki, 2010; Deakin and Markou, 2022). This theoretical framing roots the debate of law's computability in the wider settings of institutions and the field of law and economics. Those settings provide complementary framing and hypotheses to the orthodox neoclassical economics, the new Chicago School economics, and their theory of law-code competition (e.g. Dai, 2020, Lessig, 2006).

In essence, the thesis argues that law is computable only to the extent that such computer code is (empirically) sustained and conditioned by the operations of law in its meaning and institutional contexts. This is similar to the way in which social norms (for example, customary law) can be 'codified' into laws, only to the extent that these laws are cognitively sustained by the operations of the accepted norms of that economy and community. This relationship between social norms and laws is well evidenced within the law and economics literature. This literature is drawn on to help propose a novel framing of the interactions between legal and algorithmic governance, which is conducive to empirical testing.

1.2.1 Sub-questions in each chapter

A number of sub-questions are assessed in order to formulate this answer to the overarching research question. After a discussion of the methodology of this thesis in Chapter 2, Chapter 3-6 each addresses a specific set of sub-questions.

Chapter 3 asks:

- What is the nature of algorithmic governance, compared to more established ways of text-based legal governance?
- What is the relationship between legal and algorithmic modes of governance, in relations to other established modes of governance like social norms and numbers?
- Is there initial evidence to support the hypothesis that algorithmic governance is 'layered' upon legal governance, namely that code scales up governance but is limited by its path-dependencies on laws?

Chapter 4 asks:

- What is the nature of the 'scoring' systems in China's SCS?
- How do data and code work in the scoring systems, and what are their institutional roots?
- Is there evidence of scaling and layering effects, where code expands governance scales but is limited by its path-dependency on legal institutions?

Chapter 5 asks:

- What is the nature of the 'blacklists' and 'joint-enforcement' systems in China's SCS?
- How do data and code function in these systems?
- Is there evidence to show scaling and layering of algorithmic governance upon legal governance?

Finally, Chapter 6 asks:

- What are the emerging approaches, namely 'strategies', that the Chinese state adopts, to the governance of massive amounts of public and private data and code in the contexts of wider local and national experiments?

- How do data and code, as modes of governance, work in practice in these experiments?
- Is there evidence of scaling and layering in these data experiments?

Chapter 7 draws overall conclusions from these strands, returning to the overarching question of law's computability.

1.3 Contributions and motivations

The dissertation's contributions are twofold: (1) theoretical and (2) empirical. Theoretically, it proposes an original institutionalist framework and the 'scaling-layering' hypothesis to understand why law and algorithmic governance are complements rather than substitutes. Empirically, it tests the validity of the new framework by presenting new evidence on China's SCS, providing a more systematic and realistic picture of its development alongside various data experiments at both local and national level, and using the scaling-layering framework to revealing gains and losses from its mode of operation.

More specifically, the three empirical case studies of the thesis show that negative 'layering' problems happen when governments, courts and banks over-rely on code and statistics for the operations of law and norms, and fail to adapt to dynamic social reality. These negative layering effects include, among others, the problems of (1) rigidity at the level of code, (2) perverse or counter-productive incentives at the level of human behaviour (for both code designers and users), and (3) 'institutional overload' at the level of organisations (legitimacy loss). Although code demonstrates positive scaling potentials, law is shown to be essential and useful to protect code from negative layering. Law is observed to serve as an indispensable layer of governance, between social norms, numbers and code. Legal governance affords dynamic 'breathing' windows for new social practices and futures to emerge (e.g. Hildebrandt 2020b), and temporal delays for individuals, communities and institutions to adapt with continuity (e.g. Diver, 2021).

The ultimate goal of the dissertation is to help guide law-makers, scientists and practitioners when seeking to use algorithmic modes of governance in effective and societally beneficial ways. Future studies may further explore the rationale of how numbers and code sustain large-scale governance (for example the global financial system increasingly uses code), in a similar way that modern legal governance provides a centralising force for diverse informal market norms to operate (e.g. Deakin et al., 2017; Supiot, 2007, 2017).

In its wider context, the thesis is motivated by, and hopes to motivate future understandings of, whether data and code are becoming an established institutional force in governing across large geographies and time spans. The German legal theorist Rudolf von Jhering once famously claimed in his opening words of the 'Spirit of Roman law at the various stages of its development' (1852–1865):

> 'Three times the laws of the world were dictated by Rome, three times Rome bound the nations together in unity: first when the Roman people still stood in the fullness of their power, the unity of the State; secondly after the fall of that state, the unity of the Church; thirdly as a result of the reception of Roman law in the Middle Ages, the unity of the Law. The first was achieved by force of arms and compulsion, the latter two by the force of mind and reason.' (Translation to English by Zimmermann, 2015)[5]

If ancient Rome conquered or unified Europe first through violence, then through religion, and finally the law, could data and code be the next forces that govern humanity on a global scale? Could they be sustainable in the long term like law? If Romans were the main agents, who would be the new governors? The public governors in Brussels, Washington D.C., Beijing, the new transnational corporate 'governors' from London, California, and Shenzhen, or something co-produced by the networks they form at the global level? If von Jhering's words may be borrowed, it is hoped by the thesis to start finding the 'spirit' of Human laws at its various stages of development, in the age of automata and machines.

1.4 Site of analysis: a literature review of China's Social Credit Systems (SCS)

The SCS has been an emerging field of research since 2010, and provides abundant real-life evidence that can shed light on the research question of the thesis: to what extent can text-based law be computed systematically, namely replaced / substituted by computer code? A literature review on the SCS is provided below to identify the gap of understanding in the field. It also demonstrates how this fast-growing field provides rich empirical evidence to the understanding of law's computability.

[5] Zimmermann, R. (2015) 'Roman Law in the Modern World', in D. Johnston (ed.) The Cambridge Companion to Roman Law. Cambridge: Cambridge University Press (Cambridge Companions to the Ancient World), pp. 452–480. Also see Max Planck entry on Roman Law by Zimmermann, 2012. https://max-eup2012.mpipriv.de/index.php/Roman_Law When I first met Professor Zimmermann in person at the Max Planck Institute (Hamburg) 2023, I was wearing a T-shirt that said 'law is more than code' on the front, which Zimmermann saw and said 'I disagree'. He understood 'code' here as 'codification', which means the formalisation and unification of the law into a comprehensive legal code. To some extent, one hypothesis of the thesis is that today's computer code is in the process of formalising and unifying global governance, just as Roman law once was with legal 'codification' in Europe.

Existing bodies of literature on the SCS have mainly focused on (1) various socio-legal concerns over the SCS; (2) SCS's general and potential impacts, efficacy and legitimacy; and (3) theorisations of state strategies and the reputation mechanisms in terms of law and economics. The literature has, however, not fully examined the SCS systematically as a new model of governance using emerging technologies (big-data, machine learning and other aspects of AI). Empirical evidence on the SCS, in particular field-work based narratives, is also lacking and only emerging in literature.

1.4.1 Miscellaneous socio-legal approaches to the SCS

Many scholars have focused on aspects of socio-legal concerns arising from the SCS. For example, Meissner (2017) focuses on the SCS's potential enabling effect with respect to China's economic regulatory capabilities and the resulting impact on domestic and foreign business compliance practices. Hoffman (2018) discusses the SCS's compliance implications in jurisdictions outside China.

Ahmed (2017a, 2017b) examines some cases of FinTech adoption in the SCS such as Sesame Credit, although her main focus was on its data transfer and privacy implications. Chen and Cheung (2017) also zoom in to identify data privacy flaws in China's privacy regulations concerning the SCS. Chen and Engelmann et al. continue this trail of work on SCS and privacy issues and identify its relations to social media governance (2023). Chen and Engelmann et al. also examine the digital role model narratives in the SCS, and how these online narratives produce certain social norms by casting ordinary peoples as moral heroes and foes (2022).

Engelmann and Chen et al. focus on analysing how sanctions and rewards define good and bad behaviours in the blacklists and redlists of the SCS (Engelmann et al., 2019). They further examined these lists in terms of the diversity, flexibility and comprehensiveness of the rules set within (Engelmann, et al. 2021).

This line of literature provides some discrete analysis on various aspects of the SCS, often focusing on policy and the socio-legal implications of the SCS. Its evidence, mainly drawn from online archives of specific aspects of the SCS, was not gathered with the intention of providing a comprehensive understanding or framework, nor explained the workings and impacts of technological governance in detail.

1.4.2 Assessments of the SCS's impacts, efficacy and legitimacy

In contrast, a number of studies have provided more general discussion on the SCS's impacts, efficacy and legitimacy. Some claimed that the SCS is a technology-driven tool to exert social control and surveillance. For example, Creemers (2016, 2018) understands the SCS as a practice of control to stabilise the state. Loubere and Brehm (2018) theorise the SCS as a 'coercive surveillance infrastructure' and the 'financialisation of governance' with a repressive logic. Orgad and Reijers (2021) examined how China's SCS is a state-run cybernetic engineering of citizenship compared to the more bottom-up citizenship building efforts in other countries. Cheung and Chen (2022) argued that the SCS is a transformation of China's datafication practices towards a 'Data State', posing dangers to the rule-of-law state.

On the other hand, Liang et al. (2018) have argued that the SCS is a state surveillance infrastructure, but one which 'focuses primarily on financial and commercial activities rather than political ones.' This is echoed by Mac Síthigh and Siems (2019), who understand China's SCS as a 'specific instance in a wider phenomenon' of the global rise of algorithmic 'grading and ranking' systems. They compare the SCS with similar practices of reputational and quantitative scoring systems around the world, asking if China's SCS could be a model for other countries. In addition, Backer (2018) argues that the SCS is part of the global shift of the focus of public law 'from constitution and rule-of-law to analytics and algorithm', foreshadowing a future evolution in global normative orders. Relatedly, Kostka's empirical study (2018) shows that Chinese public opinion predominantly supports the SCS, seeing it as an instrument for improving the 'quality of life' and upgrading governance.

Notwithstanding these different views and approaches to the SCS, the existing literature has one feature in common: generalising on the roles and rationales of the SCS, without sufficiently examining its working mechanisms and consequences. Moreover, a significant omission is close analysis of how the data and code-driven technologies work in the institutional and organisational contexts of the SCS.

This gap also exists in Chinese scholarship. Since the late 1990s, Chinese scholars in economics, finance and computer science have been studying financial credit-rating techniques and institutions in advanced economies, with the aim of building the SCS. Their analyses, however, were focused on the technology and on the economic rationales for the SCS, within specific disciplinary boundaries (Yu et al., 2015; Lin, 2012). Only in the past few years have Chinese legal scholars joined the debate on the SCS. Peng (2021) analysed the legal dilemma of joint sanctions mechanisms in China's SCS and provided certain solutions.

These legal discussions are mostly concerned with offering doctrinal analysis with a view to preventing the SCS from giving rise to potential abuses of the law (Shen, 2019). They offer some general analysis of the efficacy and legitimacy of SCS, but have not sufficiently examined its detailed institutions—how different branches of the SCS work and what their empirical limits are (Shen, 2019; Wang, 2018; Wang and Huang, 2021; Peng, 2021). This leaves an opening for an in-depth theorisation and analysis of the working mechanisms of the SCS, paying close regard to its relationship to code as a mode of governance, some involving machine learning and other aspects of AI technologies.

1.4.3 Theorisations of state strategies and reputation systems: law & economics

One of the most comprehensive examinations of the SCS is provided by Dai's paper, 'Toward a Reputation State' (2020). He argues that the 'reputation society' is heading towards the 'reputation state': a more state-centred interventionist model of governing societies. Four major state strategies were identified in the SCS (illustrated in a 2x2 matrix) to formulate this reputation state, in which laws and regulations compete with other measures such as private ranking and scoring.

Dai's framework was influenced by Strahilevitz (2008, 2011), who proposes a 'reputation nation' theory, using information microeconomics, to analyse law from an economic perspective. This understanding follows a Chicago School 'law and economics' tradition—assuming that state-enforced laws are like 'products' competing with other normative measures on a 'market' of norms.

Writing within the 'law and economics' tradition, Dai assumes that the different branches of the SCS have the same natures and working mechanisms, and are therefore comparable to one another in a 2x2 matrix (Table 3.1, p. 13).

	Private sector information	**Public sector information**
Private actors' decision-making	regulation	searchlight
Public authorities' decision-making	incorporation	institutionalization

Figure 1. Dai's 2x2 matrix for the theorisations of state strategies in the SCS

He regards financial credit rating systems as equivalent in this sense to blacklist regulations, and see them as competing reputational instruments in governance. This approach ignores very real differences between FinTech-based techniques, on the one hand, and traditional state power of legal enforcement and punishments, on the other; these are distinct modes of governance with very different institutions, logics and evolution trajectories. Dai (2020, p. 61) frankly admits that his chapter is focused on sketching a general analytical framework, leaving unresolved the empirical evidence needed to fully support his claims, along with and their normative implications for future studies.

Another theorisation of the SCS in terms of state-adopted reputation systems is that of Mac Síthigh and Siems (2019). They compare Western ranking and grading systems with the Chinese SCS, seeking to understand them in a shared analytical framework with eight key factors— 'drafter', 'user', 'aim', 'scoring', 'application', 'algorithm', 'enforcement' and 'accountability'. These factors were mapped against the level of state intervention in different reputation systems: from low, to medium to high. These two sets of variables form the basis of the 8 x 3 matrix as a conceptual framework (2019, Table 1, p. 1046).

	Low	*Medium*	*High*
1)Drafter	Private	Co-drafting	State
2)User	Choice	Strong incentive	Mandatory
3)Aim	Specific	Socio-economic	General
4)Scoring	Multiple	Main and sub-indicators	Single
5)Application	Flexible	Comply or explain	Uniform
6)Algorithm	Transparency	Controlled transparency	Protected
7)Enforcement	Market	Stages of enforcement	State
8)Accountability	Oversight body	Review possible	Immunity

Figure 2. The 8x3 matrix by Mac Síthigh and Siems, 2019

Utilising this framework, they argue that China's Social Credit System provides a useful reference point for other countries. As such it should be studied by the US and EU legislators, not as a template or counter-model, but as 'illustrations of the implications of today's emphasis upon quantification and reputation across a range of domains, personal and official.' (p. 1070; Table 2, p. 1047; Table 3, p. 1054).

	Credit ratings in many countries	*UK research excellence framework*	*Sharing economy platforms*
1)Drafter	Low	High	Low
2)User	Medium	Medium	Low
3)Aim	Low / Medium	Low	Low
4)Scoring	Medium	Medium	Medium / High
5)Application	Medium / High	High	High
6)Algorithm	Medium	Medium	Low / Medium
7)Enforcement	Low	Medium	Low
8)Accountability	Low	Low	Medium

Figure 3. Mac Síthigh and Siems use the matrix for a comparison of 3 rating systems

	China-wide blacklists	*Pilot cities*	*Financial institutions*
1)Drafter	High	High	Low
2)User	High	High	Medium
3)Aim	Medium	Medium	High
4)Scoring	Low	High	High
5)Application	High	Medium	Medium
6)Algorithm	Low	High	High
7)Enforcement	High	Medium	Low
8)Accountability	High	Low	Low

Figure 4. Mac Síthigh and Siems use the matrix for a comparison of 3 SCS branches

Compared to Dai's reputation state theory, Mac Síthigh and Siems do not resort to an explicit law and economics approach. Nevertheless, they also assume that laws and regulations are on a par with market-based reputation systems—with the major difference being their 'interventionist' level.

In addition, Krause and Chen et al. (2023) explain China's Corporate Credit Reporting Systems (CCR) by comparing it to the U.S. and Germany. A similar approach was taken with Mac Síthigh and Siems in that certain matrices were created by comparing the indicators within the different rating systems. Here Krause et al. focused on transparency, clarity and accuracy. The theoretical framework looks at how the Chinese state, by adopting technology-enhanced regulations, wields both centralised powers and decentralised, instrumental powers of like a regulatory state. It is argued that the CCR is adopted to help China overcome the middle-

income trap. It aims to help build institutional trust from the state, saving China from the lengthy bottom-up processes of organic institutional trust-building in modern market economies. The paper provided a good institutional framework to understand the corporate rating part of the SCS, and generated a narrative to understand the SCS in economic development.

However the evidence remained in the corporate rating, and the framing is still focused on how reputation and trust is built within the regulatory environment. Similar to Mac Síthigh and Siems, and Dai's work, this piece provides a more realistic picture of the SCS, but still based on individual decision-making processes. They are not aimed at revealing the practical workings of the data and algorithms in their institutional and organisational contexts, or uncovering the gains and losses in the operations of algorithmic governance.

These theorisations are good starting points to understand the workings of China's SCS in relation to its western counterparts and the global adoptions of reputation and trust-building mechanisms in governance. However, they still put to one side the nature of the SCS as a site of complex interactions between law and algorithms, a point of view which raises wider questions than can be addressed through the logic of reputation, trust, transparency or norms competition alone. A more flexible and comprehensive theory of algorithmic governance, embedded in institutional history and organisational operations, can enrich the debates on the SCS and contribute to the understanding of algorithmic governance.

1.4.4 Section conclusion: scoring, blacklisting, and data experiments

A literature review shows the limiting contours of the emerging studies on China's SCS. Most studies have focused on one or two aspects of the systems, without providing a comprehensive understanding or a framing on law and technology. Some studies that tried to create general theoretical frameworks of the SCS did not provide enough evidence of the detailed workings of the systems in operation. A few studies have made comparisons of the Chinese systems with other countries which generated matrices to identify various regulatory patterns and helped demystify the SCS. They are often based on orthodox law and economics approaches, including reputation and information economics theories, which see law as competing horizontally with other forms of technology-enabled regulations. Their theoretical focus is on how individual decision-making and behaviours are produced under different regulatory settings. There is less discussion on the system-institution and organisational aspects of the SCS, and empirical evidence is lacking in general.

From the literature review it can also be drawn that the SCS has two distinctive components: (semi)automated scoring systems and (semi)automated law enforcement systems like blacklisting. The dissertation thus discusses them in two separate case studies respectively in Chapters 4 and 5. Using Hildebrandt's distinction between code-driven and data-driven systems (Hildebrandt 2016, 2020b, 2021), the thesis categorises the social scoring systems as data-driven while the blacklists and joint-sanctions more as code-driven systems. The former essentially uses data modelling and profiling for predictions, often involving machine-learning algorithms. The latter is built on computer infrastructures and generally follows the 'if-then' logic which does not necessarily require heavy computation or complex algorithms like ML. Sometimes the latter code-driven systems are layered or built on top of the data-driven systems such as scoring. Hence the thesis chooses the following chapters titles for the two case studies on the SCS (and the natural progression from scoring to blacklisting/sanctioning): Chapter 4 (Semi)Automated Scoring Systems: an analysis of *data-driven* financial and social scoring in China's SCS; and Chapter 5 (Semi)Automated Law Enforcement: an analysis of *code-driven* blacklists and joint-sanctions in China's SCS.

The literature review also demonstrates that China's wider contexts of data experiments are important to understand the SCS and its functioning, including the adoption of machine learning in courts, government decision-making, and how governments and firms share and use public and private data sets. The final case study looks at these experiments, in order to provide a comprehensive understanding of law and code interactions where the market and new technological developments evolve rapidly.

1.5 Structure

The remainder of the thesis is structured as follows.

Chapter 2 provides the methodology, the justification of the choice of case studies, and the empirical fieldwork conducted between 2020 and 2021 in China.

Chapter 3 examines the law v code debate and proposes a complementarity hypothesis, namely law and code are complements rather than substitutes. It develops an original theoretical framework that draws on institutionalist framings, and reveals a 'layered' model where social norms, laws, numbers and code form a stratification of governance modes. It is suggested that the nature of these governance modes is that of being layered or imprinted, with path-dependencies. This theoretical framework paves way for the empirical analysis in the coming chapters.

Chapter 4, 5 and 6 then use three empirical case studies to test and examine the validity of the framing and hypothesis. Each case study respectively demonstrates how law and code complement, rather than substitute each other.

Chapter 4 is a case study of the (semi)automated 'scoring' system, which is the first major institutional constituent of China's SCS. It is shown in this chapter that the institutional origins of these systems are derived from activities of the Chinese Central Bank, the credit-reporting industry and technology companies that learnt from Anglo-American practices, local governments in cities and villages, and regulatory agencies. The empirical evidence shows how these scoring projects, based on data and algorithms, initially scaled up (rose) but eventually faced setbacks (fell). Data and code, as new modes of governance, are seen to demonstrate complex and layered interactions with law and social norms when expanding across geographies and populations.

Chapter 5 provides a case study on the (semi)automated blacklists and enforcement systems, which form the second major component of China's SCS. The institutional actors are revealed as courts, regulatory agencies, and private agents including banks and companies at local and national levels. A more focused case study on contract and debt is provided here on Shanghai's debtors list, where new code-driven technologies and infrastructures were used by institutional actors to automate enforcement. Gains and losses perceived by local enforcers such as judges, civil servants, and program managers / scientists, are analysed. These analyses uncover the challenges faced by the local enforcers from the perspective of daily operations within their respective organisations and professional expertise. Main setbacks include (1) the rigidity of code, (2) the counter-productive behavioural changes of human enforcers, and (3) institutional overload and risks to system legitimacy.

Chapter 6 is a case study of the wider practices of China's data experiments involving automation and algorithmic governance. It is focused on the emerging and systematic approaches adopted by the Chinese state, namely 'strategies', in terms of data and algorithmic governance. Evidence shows that three main strategies emerged, namely institutionalisation, activation and layering. They respectively addressed (1) data security and privacy concerns in, for example, the public listing of DiDi (that is a ride-hailing service provider like Uber), on the New York Stock Exchange, (2) the data property or public resources debate, manifested in for example, Shenzhen's failed attempt to establish 'data property rights', and (3) how to increase the valuable and sustainable use of data, for example in 'smart courts' and digital governments. This case study further demonstrates that code's operations are layered on top of social, legal and statistical governance. An overall conclusion follows in Chapter 7.

Chapter 2 Methodology

There are three main methodological decisions which warrant further explanation in this chapter. The first decision is to draw on governance and institutionalist theories as the basis for proposing a novel framework. The second decision is to complement doctrinal analysis with qualitative field-work evidence in the form of interview data. The third is the choice of method applied to conduct semi-structured interviews and during the field trip to China.

2.1 Theory as the basis for methodology

In this section, the relationship between theory and methodology is spelled out. In developing a framework, this dissertation adopts the method of theoretical synthesis. It aims to contribute to a theory of law and algorithmic governance by synthesising a series of interdisciplinary works from law and economics that share a legal (neo-)realist and institutionalist tradition[6].

This tradition goes back to Commons (1924, 1932, 1950) and Williamson (1979, 1998, 2002, 2005) on institutionalist governance, Macaulay (1936) on social norms and non-contractual relations, Pound (1922, 1923) on an engineering interpretation of legal history and Llewellyn (1947, 1960) on law as 'craft'. This tradition is updated here by drawing on Aoki (2010) and Hodgson (2015) on evolutionary law and economics, Deakin (2015, 2017, 2023) on legal institutionalism and evolution, Howarth (2013) on law as engineering, Supiot (2007, 2017) on legal anthropology and governance by numbers, and Hildebrandt (2016, 2018, 2020a, 2020b) on law as computation.

Work by Ostrom on polycentric governance and common pool resources (CPR) will also be drawn to understand the complex and emergent nature of property rights, which are important to uncover the nature of data and algorithms as local institutions evolved from their socio-legal habitats through time (Ostrom, 2010). Ostrom's polycentric governance will be tied to Deakin and Meng's work on China's governance of Covid-19 where public and private

[6] As to the latest explanation of this tradition, see e.g. Deakin, 2023.

institutions co-produced local and global knowledge (2020), and on China's household responsibility systems that deals with North's puzzle on property rights by shifting the concept of 'legally enforced private property rights' to a more diverse, emergent and community-based conceptualisation (2022). This is further drawn together with a line of Xu's work on property rights and the governance of Covid-19. This includes her work on evolutionary (Xu, 2017) and community-based theory of property rights in China (Xu and Gong, 2020), her argument for a law-and-community approach to understand compensation of takings in the European Convention of Human Rights (Xu, 2019), and her critique of mathematical modelling in governance decision-making and the quantification of law (2021, 2023).

The underlying ontological assumption being made here is one common to legal realist and institutionalist approaches, namely that certain social referents, termed 'institutions', have sufficiently stable properties to enable them to be the focus of systematic study (Buchanan, Chai, and Deakin, 2012). More specifically, as in the modern evolutionary tradition, institutions are seen to have properties of emergence, self-organisation, and holistic or systemic 'complexity' (Aoki, 2010). Such institutions are cognitively open, and operate by interrelating and evolving with external environments, achieving an approximate, asynchronous and often fragile degree of 'fit' with their context (Deakin, 2015, 2017).

The term 'governance' is understood to refer to an organised mode or practice, associated with institutions of varying degrees of formality, through which modern capitalist societies and market economies are constituted. Thanks to the emergent, systemic and co-evolutionary features of contemporary modes of governance, institutions should be understood not merely as rules, but as practices which define the cognitive foundations and prevailing values of a society. Governance in this sense is not merely about efficiency, but also about the norms, legitimacy, and sustainability of social order—or in Supiot's words, the 'habitat' of a society (Supiot, 2007).

2.2 Time and scope

The thesis chooses the two major components of China's SCS as sites of analysis, namely the (semi)automated scoring and law enforcement systems. The third case study expands from the SCS to consider China's other wider data experiments. These three interrelated sites of analysis emerged as early as the 1990s, and the empirical research conducted by this dissertation had a cut-off point at April 2023, when the author started to tidy up data and write up.

In particular, the interview data is collected on a six-month field trip to China between December 2020 and May 2021. It was amidst the global pandemic but travelling in China remained possible for that time. However the author could not make another field trip to China after large-scale lock downs were introduced in 2022, starting from Shanghai's Covid-19 emergency. The cut-off point for interview data is therefore May 2021, and the dissertation does not have sufficient field-based evidence to address the state of emergency and changing circumstances that took place during China's Covid lockdowns between January 2022 and January 2023.

However, such interview evidence collected before China's intensive lockdowns already have shown certain aspects of the pandemic's influence on the SCS and other data experiments. They also provide data from the field in a more stable period rather than a 'state of exception'.

In addition, as mentioned at the beginning of the thesis, the pervasive use of ChatGPT and other generative AI models since March 2023 is also not covered by the dissertation. The reasons being (1) these new technologies are still in early deployment, and the literature in law is scarce, (2) the available space of the thesis, and (3) the main body of the research is done before the release of these new AI models. The thesis nevertheless is only made more pertinent by these new AI developments, which show that we need a more foundational framework to understand the law-code relationship, evidenced in pre-existing algorithmic systems deployed in law and societies.

This deliberate choice of sites of analysis is based on the emerging literatures on the SCS in the global discussion of algorithmic governance, and the lack of detailed understanding of the SCS with regard to empirical evidence. The choice is also both motivated and limited by the interests and skills of the author, who hopes to understand and describe a more realistic picture of these systems in China, by utilising a combination of linguistic and legal skills, as well as access to interview networks.

2.3 Empirical methods

The empirical case studies combine doctrinal, archival and interview methods. The legal-doctrinal materials and archives used were publicly available online and mostly in the Chinese language. All the interviews were semi-structured and conducted in Mandarin Chinese (the local accent's impact was minimal). The author is responsible for most of the translations of Chinese into English and the potential errors thereof, except for the public or commercial translations that were specifically credited in the dissertation.

Interviews provide in-depth and thick descriptions of real practices and behavioural interactions between actors within socio-technical systems (Poteete, Janssen and Ostrom, 2010, pp.33-35; Buchanan, Chai and Deakin, 2012, p13). Although limited numbers of interviews conducted for this study are non-replicable data points, the transcripts provide rich information on perceptions and narratives which throw light on causality. Causal processes are hard to infer from quantitative statistical associations alone. The interview evidence, triangulated with legal-doctrinal materials and online archives, can help reveal how social norms, law, and code interact in reality (Poteete, Janssen, and Ostrom, 2010, pp.9, 11-12, 33-37)[7]. It thus can help to answer the research question of whether code and law are substitutes or complements as governance practices.

However, the scope for triangulation between doctrine, online archives and interview evidence may be restricted by the limited amount of existing empirical evidence and research. Establishing the veracity of self-reports by interviewees is inherently difficult and requires the interviewer to exercise their own judgement on certain points. The interview evidence is presented with regard to the social, cultural and spatio-temporal contexts of each of the interviews. As suggested by Elinor Ostrom and her co-authors, interview-based fieldwork of this kind, by directly accessing the perceptions of actors including their understandings of causal processes and sequences in their local time-space, provides rich qualitative evidence which is not available from surveys or data analysis (Poteete, Janssen and Ostrom, 2010, pp.11-15, 33-35). It is relevant to note in this context that publicly-available and reliable statistical evidence on the effects of SCS is currently lacking. Once it becomes available the findings of the present qualitative study could be tested using quantitative approaches.

[7] See advantages of Small-N case studies by qualitative fieldwork and interviews, and the advantages of mixed methods.

2.4 Interview data

A total number of 26 people were interviewed in 21 settings, in China between December 2020 and May 2021. The interviews were mainly conducted face-to-face in Beijing and Shanghai, while a few interviews were made online with interviewees based in Qingdao (Shandong Province), and Hefei (Anhui Province). The Shanghai field trip, in particular, was intensively done in April 2021, with 15 interviewees involved.

Most interviews were not recorded for confidentiality reasons. Some materials might be sensitive and many interviewees preferred to remain anonymous. But notes were taken almost verbatim and written up as transcripts right after the interviews, and later translated into English by the author. The interviewees were recruited using a snowballing method, in which personal and professional networks helped secure the initial few interviewees in order to start snowballing.

The interviewees are mainly government officials, judges, court administrators, local legislators, public data centre managers, state-affiliated corporate managers and employees. These are the people who are in supervisory positions or at the front line of operations for China's SCS and other data experiments at local levels. Some academics were also interviewed for their expertise and critical policy engagements in the field.

The deliberate choice to conduct an intensive set of interviews in Shanghai, instead of other large cities, is because of its more well developed civil and criminal justice systems in general (including more independent courts and regulators), and a more developed SCS in particular. Shanghai was one of the earliest cities to trial the debtors list, an early example of algorithmic governance in China, and also China's first SCS pilot city in general since 1999 (Lin, 2003, 2019)[8]. It was also the first city in China to pass local legislation on the SCS (in 2017)[9]. Views from Shanghai can thus provide evidence over a longer time frame, in a more stable environment of economic development and rule-of-law safeguards, than is possible elsewhere. This distinct socio-economic and legal environment also makes Shanghai's SCS a suitable comparative reference to the algorithmic systems developed in Western countries / cities (cf. Mac Síthigh and Siems, 2019. Krause et al., 2023).

[8] Lin Junyue is one of the intellectual and policy founders of the SCS, who was also one of the interview respondent of this thesis.

[9] Regulations of Shanghai Municipality on Social Credit (Shanghai Municipal People's Congress, [2017] No. 54.

2.5 Chapter conclusion tied to research question

This chapter has explained the three methodological choices of the thesis. Firstly, it has explicated the institutionalist foundations of the thesis, and the ontological claims on governance and institutions. The theoretical synthesis is a method used to draw together various threads of theories following a (neo-)realist and institutionalist tradition.

Secondly, the chapter has introduced the bounded time and scope of the thesis, and noted the advantages and limitations of empirical methods. The main limitation is that the qualitative fieldwork provides non-replicable data points, but they also have a greater potential to reveal causal processes compared to other more quantitative methods. Interview data is suitable to be combined with traditional legal interpretative methods. Together they can provide a more evidence-based approach to understanding how law and code interact in real-life practices. Despite the limitation in scaling-up and triangulating these data points, they are conducive to providing thick descriptions and in-depth understanding of the research question. They are also more suitable for testing the institutional framing and hypothesis proposed by this thesis on law-code complementarity which requires evidence on causality.

Thirdly, interview data and methods have been explained in detail. This includes the numbers of interviewees, how they were recruited through snowballing, the time and space of the interviews in China, how data is collected confidentially and transcribed later, and the choice of the intensive fieldwork in Shanghai, among others.

By explaining the methodological choices of the thesis, this chapter has already started laying out the theoretical foundations of the thesis which is essentially (neo-)realist and institutional. The following chapter applies theoretical analysis as a method to generate a novel framing on law and algorithmic governance, and proposes the scaling and layering hypothesis.

Chapter 3 Theoretical Framework: Law v algorithmic governance, substitution or complementarity?

3.1 Summary of the argument

In this chapter, an original institutional framework of algorithmic governance is formulated to argue for law-code complementarity. The argument draws evidence from existing literature of governance theories and institutional law and economics. Rather than law being subject to meta-competitive forces like social norms, numbers and code, an institutionalist framing follows the view that law provides the preconditions of possibility for these forces to operate (e.g. Deakin et al, 2017; Pistor 2019). A 'scaling and layering' hypothesis is proposed to the uncovering of why law provides such conditions, namely the inherent trade-off between scales and layering, or imprinted path-dependencies. This institutionalist framework, compared to those of orthodox law and economics theories, is argued to be more comprehensive and realistic, with greater explanatory power, and to pave the way for empirical testing using case studies.

3.1.1 Research questions

To answer the overarching research question, governance theory and institutional economics are consulted and developed in order to generate new framing and hypotheses. The research questions of the chapter include: what is the nature of algorithmic governance, compared to more established ways of text-based legal governance? How does law interact with other modes of governance, like social norms, numbers and code?

3.1.2 A metaphor of the 'layered' model

In response to the questions, governance is first defined in this thesis as consisting of four distinct but interrelated modes of existence: social norms, laws, numbers and code, namely algorithmic governance. These four modes of governance, for better visualisation purposes, can be displayed as vertically situated on top of each other. They are like the earth layers, metaphorically, of an erupting volcano (the 'layered' model). At the bottom there are the social practices called 'magma', which are fundamentally indeterminate that can break the exterior modes of governance with 'heat' (see e.g., Castoriadis, 1994; Supiot, 2007, pp. 13-29)[10]. Stemming up from magma, the governance layers are respectively social norms, laws, numbers and code. Each governance mode or layer, in contrast, gets 'colder' as they become more stationary and determinate with increased level of abstraction and formalisation, moving away from the 'heat' of the magma.[11]

This layered model is helpful in depicting the four modes of governance as operating and interacting simultaneously in a vertical direction. They are layered on a geographical and populated space / habitat, which is the horizontal axis where the inhabitants or population generate real-time social practices (namely, magma).

3.1.3 Hypothesis: complementarity between modes of governance

It is then hypothesised by the thesis that, in essence, algorithmic governance is a more abstract and unified mode of governance compared to law. It promises larger governance scales, but is also sustained and path-dependent on the daily operations of legal governance. This hypothesised relationship between code and law, or between algorithmic and legal governance, is based on the observation that laws are 'codified' and more abstracted forms of social norms. Laws provide scaling effects but are still preconditioned by the operations of social norms.

The full hypothesis is two-fold. (1) First, the scaling hypothesis, defined as the more formalised / abstract modes of governance emerged from the lesser ones, and with this higher abstraction levels came along the capacity to govern territories and populations at larger scales. (2) Second, the layering hypothesis, defined as that the more abstract modes of governance are sustained by the lesser ones, with institutional path-dependencies of a kind associated with

[10] Borrowing from Castoriadis on magma. Castoriadis (1994).

[11] *Ibid.* In Castoriadis' framework they would be called 'ensemble', or 'sets' in French, namely the determinate beings.

potential negative effects. The reason for such negative effects is that governance through code will decay / fail without the framing effects of other, anterior and less abstract modes of governance, including law and norms.

The next section of the chapter proposes the institutionalist framework and sets out the scaling and layering hypothesis in detail (3.2). Then the following three sections respectively provide preliminary studies on how scaling and layering occur when modes of governance move from social norms to laws (3.3), from laws to numbers (3.4), and from numbers to code (3.5). The final section concludes the chapter (3.6).

3.2 Framework and hypothesis: scaling and layering of governance

3.2.1 Governance: norms, laws, numbers and code

(1) Recent attempts to define governance

The term 'governance' has been widely adopted in various disciplines since the 1990s.[12] These uses include 'corporate governance' (since the 1990s)[13], 'global governance' (since the 2000s)[14] and 'algorithmic governance' (since the 2010s; B. Williamson, 2014; Danaher et al., 2017; Coglianese and Lehr, 2019; Engstrom and Ho, 2020a, 2020b). But the concept of 'governance' itself has not been well-defined or consistently applied across these different disciplines (Boyer 1990; Rhodes, 1996, 1997, 2007; Kooiman, 2003; Rodrik, 2008; Offe, 2009; Fukuyama, 2013; Bevir, 2013).

Some international organisations have provided practical definitions of governance. For example, the UN Global Compact (2020) defines governance as 'the systems and processes that ensure the overall effectiveness of an entity–whether a business, government or multilateral institution'. With specific reference to public governance, the OECD (2011) refers to 'the formal and informal arrangements that determine how public decisions are made and how public actions are carried out, from the perspective of maintaining a country's constitutional values when facing changing problems and environments.'

[12] *See*, google Ngram, the usage of the word 'governance' started to grow since the 1980s, and showed exponential growth since the early 1990s.
[13] *See*, google Ngram, growth in 1980s, and exponential growth since 1990s.
[14] See, google Ngram, exponential growth since 1990s.

More theoretical definitions of governance were also recently offered: 'all processes of governing, whether undertaken by a government, market, or network, whether over a family, tribe, formal or informal organization, or territory, and whether through laws, norms, power or language.' (Bevir, 2013, p. 1). These definitions tend to have a decentralising philosophy with an emphasis on network theory as the core approach to understanding social organisation (Bevir, 2013, Ch. 7-8).

It should also be clarified upfront that the dissertation chooses not to delve into the governmentality literature by Foucault and others, despite its high importance and influence in governance theory. Foucault has addressed the modern transformation of state government (take life or let live) to 'governmentality' and biopolitics under the rule of numbers and surveillance (make live and let die), from a critical theoretical approach in his lectures in the 1970s-80s (Foucault, 2008, 2011, 2012). This is highly relevant to the thesis's framing on governance, but is not the approach the thesis decides to take (despite the thesis indeed gains lots from the French critical-historical literature by discussing Supiot's *Governance by Numbers*, 2017, which is important for law, and some ideas in the *Politics of Large Numbers* by Desrosières, 2002). Similarly, recent governance literature from management, government and social science have been vast (see e.g. Jackson, 2011; Davis et al, 2012; Siems et al, 2017; Twining, 2015; Hansen et al, 2012). Nevertheless, they do not directly address a *theory* of governance (despite focusing on various issues of statistical governance), and the thesis can only cover the most relevant theories to the discussions on the different modes of governance, including social norms, laws, statistics and code, which follows a law and institutional economics tradition. Supiot's work (2017) on governance is highly relevant also in the sense it reveals the different layers of modes of governance including law, norms and statistics, and in addition, how international organisations like the World Bank and UN promoted the new modes of statistical governance to the globe.

This choice of theoretical traditions is also more conducive to the empirical examinations of this thesis, using the case studies from China which started in the 1990s. China's Social Credit Systems focus on credits, enforcement of debts, and modern socio-financial institutions sustained by law and courts in China, which sit at the centre of the law and economics discussion on governance. In addition, the SCS were introduced in China in the 1990s around the same time when governance theories (*治理 zhili*) were introduced to China by international organisations such as the OECD and World Bank (see e.g. Yu, 1999, 2002, 2018; Lin, 2003). This also makes Supiot's work on the global diffusion of numerical governance more relevant

(and thus can be examined empirically) in this thesis, compared to, say, previous Foucauldian discussions on governmentality and surveillance.

(2) Modes of governance proposed: social norms, laws, numbers and code

By 'mode' is to be understood as the 'mode of existence' (Latour, 2013), so that 'modes of governance' are synonymous with the ways in which governing processes operate or exist (Hildebrandt, 2016, 2018). An institutionalist approach can be used to identify four principal modes of governance: governance by (i) informal social norms, (ii) state-enforced laws, (iii) statistical calculations and (iv) computational code.

Norms, laws, numbers, and code, have been recognised as four distinct ways of regulating individual behaviours since early debates on law and technology (Cohen, 2012; cf Lessig, 2006). They function as four principal modes of governance, and each medium has self-reinforcing properties that create independent systems of governing and related modes of existence.

(i) *Norms.* In this thesis, 'governance by norms' refers to the use of social norms as the main medium for governing in societies. The literature on social norms see them as informal rules and customs developed through repeated interpersonal interactions in communities. They are the customary or private arrangements without the aid of state or other central coordinators (Hayek, 1973; Ellickson 1994; Ostrom, 2005, 2010).

(ii) *Laws*. By contrast, the thesis defines 'governance by laws' as the use of formal institutions, underpinned by the modern state, as the main mode of governance of a society. This implies a realm of public ordering and centralised design. Literature on formal and informal institutions has pointed out that even though law relies on and works in conjunction with social norms and private ordering, it requires law-makers to act as the institutionalising agents and final arbiters under the general ambit of the rule-of-law state (Deakin et al., 2017; Supiot, 2007).

(iii) *Numbers*. 'Governance by numbers' is defined as using quantified measures and statistical calculations as major forms of governing. While relying on mathematical techniques rather than the interpretation of texts, governance by numbers is still derivative of governance by laws in the sense of inheriting the basic institutions and structures of the state, including a formal legislature and judiciary (Desrosières, 2002). It has a particular focus on administration (public and private), formal regulation and planned policy-making; it is often characterised by the transplanting of the formal logic of the market, above all the price mechanism, into public

institutions including the legislature, judiciary and public administration. In its ubiquitous use of statistical modelling, employing the 'artificial' language of numbers, it is distinct from laws and social norms, which are based on the use of natural language for communication (Supiot 2007, 2017).

(iv) *Code.* 'Governance by code' is defined as an extension of governance by numbers, but with qualitatively higher sophistication both in the 'learning' capacity of computations and calculations (Hildebrandt, 2020a, 2020b, 2021; Deakin and Markou, 2020), and in code's 'hyper-reflexivity' (Zuboff, 2015, 2019). Code is hyper-reflexive in the sense that it can be used to shape social environments in real-time. For example, a Facebook algorithm can nudge individual behaviours based on their real-time interactions online. Such code, when manipulated by engineers or self-adjusted by a machine-learning algorithm, can be used to experiment with human emotions and behaviours, or even trigger certain predetermined feedbacks, which reinforce the code's own preconditions (Zuboff, 2019). Such processes require a higher level of formalisation or abstraction compared to governance by numbers, and a faster, sometimes real-time, feedback-enforcement infrastructure. Governance by social norms, laws, and numbers all have reflexive capacities to various extent in the sense they can react to their environment and be adaptive. Code however is not as reflexive as the other modes at a system or institutional level. Code's hyper-reflexivity at monitoring and nudging individual's behaviours paradoxically made it the most self-reinforcing mode of governance which lacks a systematic adaptive nature. In essence, algorithmic governance is the use of advanced computational algorithms written in binary numbers (0-1) and operated on computers, often involving deep-learning and other machine-learning techniques in which models self-adjust or 'learn' through repeated iterations with the data on which they are 'trained'.

It should be stated upfront that governance has local meaning and context. Most relevant to this thesis, the theoretical concept of governance (治理 *zhi li*) in modern China was largely introduced by western governance theories since the 1990s through academia and international organisations like the UN, Work Bank and OECD (see e.g. Yu, 1999, 2002, 2018). The thesis acknowledges the different meanings and depths of the concept 'governance' or *zhi li* in the Chinese context, where debates often focus on whether informal institutions or social norms like *guanxi* are more important than state laws for economic growth in China compared to western economies (cf. Chen and Deakin, 2015; Deakin and Meng, 2022). However, this difference does not affect the above categorisation of the four modes of governance into social

norms, laws, numbers and code. In fact, these differences in governance practices between China (e.g. social credit) and the west (e.g. surveillance capitalism) make the proposed theoretical framing of the thesis more interesting, as it potentially can reveal shared functions and structures of these different layers of governance. The validity and usefulness of the proposed framing will be later tested and examined by the empirical case studies in Chapters 4, 5 and 6.

The thesis therefore provides an opportunity to examine theories that predominantly originated from the west (or common law and/or continental legal traditions) using practices from China's legal and technological systems. The goal is to formulate a preliminary general framework that can potentially transcend geographical boundaries and jurisdictions. It also poses new questions to challenge the preconceptions on whether governance and computational law theories originating from the west are relevant/applicable to China, and in reverse, whether theories generated from Chinese practices can be useful to understand and solve problems in the common law and/or continental legal traditions.

3.2.2 Algorithms: The 'algorithmic turn', scaling, and layering

(1) An institutional approach to governance

By contrast, the seminal studies of governance by John R. Commons from the 1930s (1932, 1950) did not have such a strong decentralisation emphasis. Commons' early work (1932, p. 4) defined economic governance in terms of its 'three principles of conflict, mutuality and order'. His definition emphasised the continuity of mutual exchange relationship, and the rules that governed transactions.

However, the paradigm of orthodox economics of the time was increasingly taking shape as a 'science of choice', which studied how equilibrium is reached as a result of the choices of utility-maximising consumers and profit-maximising firms (Reder, 1999, p. 48). Commons (1950, p. 21) criticised the orthodox paradigm on the grounds that economics should instead centre on 'transactions and working rules, ... and on the ... [ways] the organization of activity is ... stabilized.'

This institutionalist tradition and the triple principle of 'conflict, mutuality and order' was taken on by Oliver E. Williamson from the 1970s. He (2002, p. 439; 2005, p. 3) proposed that governance should be defined as 'the means by which to infuse order, thereby to relieve conflict and realize mutual gain.' Williamson (1979; 1998, p. 58) argued that rules, including laws, are

essential for stabilising transactions in order to form reciprocal and continuous exchanges through time.

Moreover, this institutionalist approach to governance was enriched by Elinor Ostrom's study of polycentric governance (2005, 2010). 'Polycentricity' here refers to governing decisions that are made through multiple loci, each operating with some degree of autonomy. Ostrom (2010) claimed to develop a model of governance beyond the dichotomy of market and state, in which complex economic systems are governed by a hybridised framework of diversified rules, established by both state and non-state institutions.

In summary, an institutionalist tradition understands governance in terms of (1) its functions to infuse social order and cooperation, and (2) the various means or institutions by which these functions come to operate, that is the different modes of governance, including law.

(2) Defining algorithms and computation

An 'algorithm', generally put, is a 'step-by-step procedure for solving a problem or accomplishing some end.'[15] The term is often associated with the rise of computer science. 'An algorithm is thus a sequence of computational steps that transform the input into the output.' (Cormen et al., 2009, p. 5).

'Computational' denotes 'involving the calculation of answers, amounts, results', and a computational procedure means it must be readable and operable by a physical computing machine (namely, 'using computers')[16]. This restrains the operation of computational steps within the material hardware of the computer, that is, the transistors placed on silicon chips that process binary electric signals following Boolean algebra (Hildebrandt, 2020b; Markou and Deakin, 2020).

This thesis adopts a simple definition of an algorithm as follows: any process can be called 'algorithmic', or a process decided by algorithms, if it (1) has 'problem-solving' as the goal, and (2) provides a 'step-by-step' or 'computational' procedure. More computational modes of governance require a higher level of abstraction and formalisation.

[15] Merriam-Webster (2019). *Definition of ALGORITHM*. [online] Merriam-webster.com. Available at: https://www.merriam-webster.com/dictionary/algorithm

[16] Cambridge Dictionary (2022), online at: https://dictionary.cambridge.org/dictionary/english/computational.

(3) Identifying an 'algorithmic turn' in governance

Applying this definition of algorithm to governance, it can be seen that all four modes of governance – norms, laws, numbers and code – are 'algorithmic' to various extents. The four modes of governance have varying degrees of algorithmic features situated in their institutional history. As the mode of governance moves from norms to laws, to numbers and to code, the ways societies govern themselves become more algorithmic, and at the same time more explicitly computational and abstract. The thesis conceptualises this evolutionary movement as the 'algorithmic turn' in governance.

To start with, (i) governance by norms and (ii) governance by laws can be algorithmic, but only to limited extents. They are partly algorithmic, because they fulfil the definition of involving some 'step-by-step' or computer-readable procedures in order to solve some well-defined goals of governance. Nevertheless, they both maintain ambiguous concepts, principles, and goals that are intrinsically impossible to be expressed in purely numerical or binary computational terms. Laws and social norms maintain a non-calculable and non-algorithmic element that cannot be programmed into machine-readable code (Supiot, 2007, 2017; Markou and Deakin, 2020). On the other hand, as between laws and social norms, law could be deemed as more algorithmic because it is often by definition more formalised with better-defined concepts and instructions (Deakin et al., 2017; Chen and Deakin, 2015).

In contrast, (iii) governance by numbers and (iv) governance by code are much more algorithmic, or that is to say computational, in that they both follow a highly mathematical step-by-step procedure to solve problems. For example, regulatory ranking and credit-rating systems operate on step-by-step instructions to evaluate compliance performances and risks using statistical calculations. Some of these numerical calculations could thus be easily programmed into, for instance, FinTech or RegTech algorithms that are run on computers (Hurley and Adebayo, 2016; Sun 2017; Dai, 2020). Meanwhile, code in the form of ML / DL algorithms that decide credit scores or criminal sentencing and probation periods are readily operated on computers to solve computational problems following computational steps (Hurley and Adebayo, 2016; Kehl et al., 2017). These codes are purely computational and the most abstract.

(4) Proposing a 'scaling and layering' hypothesis

Having proposed the movement of the four modes of governance (from norms to laws, to numbers and to code), this chapter further proposes the hypothesis that this 'algorithmic turn'

has two defining qualities—scaling and layering. First, 'scaling' denotes that the more algorithmic modes of governance become, the greater their potential for governing populations and territories on an ever-larger scale at a smaller or similar social cost. This can be explained by their higher levels of abstraction / formalisation which provide greater capacities in processing larger amounts of data to facilitate complex division of labours and dispute resolutions.

The concept of scaling is taken from the idea of 'economies of scale' in microeconomics, which refers to the cost advantages that enterprises gain as a result of their increasing scale of production (Alfred, 1977). It can be traced back at least to Adam Smith's idea of gaining higher production returns by adopting the division of labour (Smith, 1776). In the context of governance, scaling is not purely about producing more cheaply, but rather about expanding the general scale of governed territory and population in a sustainable way. Such sustainable scaling can achieve mutual gains and order, without inducing disproportionate social costs in resolving new conflicts.

The scaling hypothesis can be initially supported by some historical cases. For example, a variety of informal social norms like repeated trading, reputation and interpersonal trust first assisted the transition from smaller tribal transactions to larger impersonal trade in the west (North and Thomas, 1973; North, 1990). The later rise of modern contract law and that of the market economy based on numerical calculation are widely identified as among the most important defining institutional developments that led to the dominance of the global north (Ferguson, 2012; Cooter and Schaefer, 2013).

In addition to 'scaling', this chapter presents the hypothesis that governance is characterised by 'layering effects'. Layering is defined as: a process through which the more algorithmic / computational modes of governance are cognitively layered and path-dependent on their less algorithmic predecessors.

Some legal scholars have already pointed out that numbers and code can reinforce or lock-in negative features of existing governance modes with path-dependency (Supiot, 2007, 2017). Deakin argued in particular, that legal concepts like employment have been historically 'imprinted' in legal evolution with path-dependencies. Similarly to the evolution of technologies such as the computer keyboard arrangement (QWERTY), these juridical concepts are like 'frozen accidents' which are stable but suboptimal in terms of efficiency. They are 'frozen' in the sense that these concepts are societally costly to change once they are imprinted and fixed into a 'punctuated equilibrium' (Deakin 2003, 2011, 2015; Deakin and Markou 2022).

This chapter draws these existing arguments together to propose a version of the 'layering effect', according to which code and numbers are layered on laws, in a similar way that law is layered on social norms. When it comes to numbers and code, the negative reinforcement of these more abstracted forms of governance can, for example, 'cool down' or 'freeze' the otherwise open and dynamic interactions in societies (e.g. Hildebrandt, 2020). 'Freezing' here is another metaphor which indicates the high cost of change of institutions once they are socially embedded in cognition. It is a concept that aligns with the 'layered' model which uses 'heat' as a metaphor to represent the level of indeterminacy of a mode of governance. Here numbers and code are 'cooled' in the sense they are more determinate compared to social norms and laws, and more likely to reinforce negative social patterns and practices derived from the past.

Furthermore, the thesis observes two broad types of negative layering which are useful to the framework and the analysis. The first one involves the more abstract mode of governance such as code, reinforcing the negative path-dependencies of the less abstract modes such as laws. This is similar to the idea of 'freezing' by Hildebrandt (2020) or 'frozen accidents' by Deakin and Markou (2022). The second type of negative layering refers to the gradual decay of the more abstract mode of governance such as code, due to its misalignment with the less abstract modes such as laws and social norms over time. In this case, code loses its efficacy and legitimacy as it extends beyond the cognitive 'support' by law and social norms which sustain code's social functioning (see e.g. Chapter 5 Section 3.3 on 'institutional overload').

The chapter presents the layering hypothesis with the aim to provide a preliminary but novel approach to governance that helps explain why claims such as those involved with the 'legal singularity' (e.g. Aidid and Alarie 2023; Alarie, 2016) cannot stand. These claims hypothesise that law can reach so-called 'completeness' where code will function fully on its own in providing personalised, instantaneous and just legal solutions. Layering defeats these claims by arguing code is sustained by law just as law is sustained by social norms.

Layering explains why over-reliance on numbers and code in governance can lead to negative reinforcements. Just as laws once reinforced, and may continue to reinforce, outdated social norms and practices, code can work similarly. The following three sections of the chapter each examines this institutionalist framework with more detail. The chapter explains how scaling and layering happen between the modes of governance, following its movement from social norms to laws, to numbers and to code.

It is also worth noting here that the proposed scaling and layering hypothesis gains a lot from the critical algorithms/code literature as backgrounds (e.g. Lessig, 2006; Grimmelmann,

2005; Lewkowicz Gregory et al, 2020; Katzenbach et al, 2019; Amoore, 2020; Bucher, 2018; Cobbe et al, 2023), despite not delving into the details of these debates. For example, Lessig's framing of the four modes of regulations (i.e. social norms, laws, markets and code) inspired the four modes of governance of this thesis, which follows the debate in law and economics (e.g. Lessig's New Chicago School framing, 2006). However, as mentioned earlier this thesis distinguishes from Lessig's framing by seeing the modes of governance not as competing on a flat, horizontal field, but on a vertical space with path-dependencies upon one another, defined by their nature of existence (e.g. abstraction/formalisation level) and by highlighting social practices and norms at its vertical centre/core. In addition, the four modes of governance defined in this thesis are not only about the micro-regulations of individual's behaviours (as in Lessig's framework), but also about their own ways of existence as macro social-technical systems operating on human-machine hybrid networks. This means the regulated individuals and organisations also constitute the modes of governance that regulate them. In this sense, the thesis intends to enrich existing law and economics literature on regulations and governance (e.g. Lessig) by adding this institutional and organisational approach focusing on systems, networks and cognitive/behaviour feedback loops.

Nevertheless, the thesis cannot cover all relevant discussions when proposing a theory of algorithmic governance and law. This means some more nuanced debates in the critical algorithms/code literature are implicitly embedded in the framing of 'scaling and layering', but not discussed in detail or explicitly at times, such as Grimmelmann's criticism of Lessig's framing of code by pointing out that software itself is a different mode of existence/regulation between law and architecture/code (2005). Instead of introducing these various theories and debates of algorithmic governance and critical data studies in the theoretical framing, these theories are discussed more in-depth within the case studies in the later chapters (such as the claim 'All Data are Local', Loukissas, 2022). The hope is to propose a 'scaling and layering' hypothesis that maintains a delicate balance between simplicity of form (thus useful for analysis) and complexity of meaning (thus have rich explanatory power in diverse contexts).

3.3 From Norms to Laws: Institutionalism, Rule-of-law state, and Path-dependency

The chapter first examines the scaling and layering effects that occur when modes of governance turn from social norms to laws. This turn can be demonstrated via the literature on 'legal institutionalism' (e.g. Deakin et al., 2017, Pistor 2013).

3.3.1 Legal institutionalism: law v social norms

The importance of informal social norms in governance has long been evidenced. Repeated trading, reputation and interpersonal trusts have been effective in, for example, the governance cases of non-contractual relations in businesses (Macaulay, 1963), informal property disputes resolutions between neighbours in California (Ellickson, 1994), and the medieval long-distance commerce among religious communities like the Jewish Maghribi traders (Greif, 1993).

Nevertheless, some scholars claimed that (1) such norms without the aid of a state or other central coordinator are sufficient to support complex modern governance, and (2) the law is a mere epiphenomenal expression of or could be reduced to these customary norms or private ordering (for discussion of these views see Deakin et al., 2017, p. 189).

Legal institutionalism rejects these views, in the process presenting two ontological claims. First, it argues that law institutes modern capitalist societies by involving the state, and qualitatively turning away from sole reliance on social norms. A key reason for this turn is the scale effect acquired through governance by laws. The 'high degree of complexity and uncertainty in large, developed economies' requires standardised legal institutions involving the state, a realm of public ordering, and enforcement that supports division of labour at greater scale (*id*, pp. 189, 198). In contrast, businesses in societies under a 'rule-of-norm' often resort to means such as 'clan or family ties, shared religion or ethnicity, bureaucratic co-option and corruption …' (*id*, p. 189). These social norms are often formed on interpersonal trust and local networks that work at a micro scale but become costlier when the scale and complexity of governance increase. Hodgson (2015, 2023) goes so far as to claim that the *most* important factor in the historical emergence of capitalism is the constitutive role of law and the state.

(2) Secondly, law must always involve both public / designed (state intervention) and private / spontaneous ordering (social norms) working in a relation of mutual complementarity. State-enforced law must align to at least an approximate (if precarious and somewhat asynchronous) extent with social norms to achieve hybridity, in order to fulfil its instituting and scaling effects (Deakin et al., 2017, p. 198; Pistor, 2013). At the same time, the institutionalist thesis of the rule-of-law state does not overlook the importance of social norms in modern governance. This point of view follows the early argument of Commons (1924) that law's enforcement must be founded on beliefs—in particular, a wide perception of law's reasonableness, appropriateness and fairness.

Law essentially needs to conform with, or at least not depart too fundamentally from, social norms in order to be accepted and enforced. For example, Chen and Deakin (2015) argue that one of the most important social norms that upholds the state and legal enforcement is the concept of a 'rule-of-law state' itself, understood as both a social practice and a formal, constitutional principle. Basing their argument on evolutionary game theory, they claim that with repeated games at large scale, social agents collectively recognise law's capacities such as accessibility, neutrality and stability, that facilitate sustainable economic development (pp. 132-134). This shared recognition and respect becomes a focal point or common knowledge among actors in a society, which is a cognitive meta-convention supporting the functioning of the state and laws (pp. 134, 135).

3.3.2 Scaling: the rule-of-law state and modern capitalism

Scaling of governance can be observed in the historical turn from governance by social norms to laws. First, many scholars relate the rise of capitalist societies to the emergence of rule-of-law states. Hodgson (2015) explains how the state instituted modern capitalism in eighteenth-century Europe with the early development of financial institutions in England, particularly concerning the law of property as collateral and the basis of transactions based on debt. Hodgson (2023) further demonstrates that external pressure at the time, namely military conflicts overseas, forced the British state to radically reform internal legal and financial institutions, which not only equipped the British war machine, but also fostered the British Industrial Revolution.

Pistor (2013, pp. 322, 323) argues that the way in which law institutes modern financial systems gives rise to an 'essential hybridity' between both public (state) and private (market) ordering. She claims that: 'Indeed, the scale of today's transnational financial markets would not be feasible without their legal backing'. Law's use of legal language and fictions makes it possible to 'code' material assets like land and debts into capital, thereby constituting 'a code for the globe' (Pistor, 2019, Ch. 6).

Second, these scale effects are also recently evidenced in emerging economies like China, where the state and law gained increasing centrality in governance replacing informal social norms (guanxi). This contributed to China's rapid industrialisation and growth since the 1990s. As Chen and Deakin argue (2015, p. 11), social norms such as guanxi in China provided valuable interpersonal trust networks for businesses to access financial resources, secure property rights, and enforce informal contracting. Nevertheless, these social norms of guanxi

can also limit the scale of governance and hence exchange, because they are often formed in tacit and common understandings among business parties and local niches. For example, China's foreign direct investment (FDI) in the 1990s was highly concentrated in the coastal areas where dense guanxi networks enabled the local economy to grow, in the process generating inequalities between regions. Moreover, guanxi-based norms confined benefits to members in closed networks and led to local corruption (pp. 11-13).

Chen and Deakin (2015, pp. 12-13) argue that when these deadweight costs of the guanxi system began to outweigh the benefits, state-enforced laws started to grow as the dominant mode of governance. Formal legal institutions like new commercial laws securing contract and property rights enforced by courts have helped China's economy grow at scale since the 1990s and 2000s. Their claims are supported by quantitative analysis (2015, pp. 137-141) providing evidence that legal transplants of shareholder protection have supported financial developments in emerging economies. Qualitative fieldwork in China also provides evidence that in product markets, 'business relations are increasingly characterised by a mix of trust-based transacting and legal formality' (Chen et al., 2017, pp. 18-22). In short, both in advanced and emerging economies, formal legal institutions have brought scale effects to governance by functioning together with informal social norms.

3.3.3 Layering: path-dependency, reinforcement and cooling effects

The second ontological claim of legal institutionalism is concerned with the path-dependence of laws on social norms, which this thesis calls 'layering'. The centrality of law in modern capitalist governance depends not only on the public state but also on private social norms, especially in the case of legal transplants and reforms (North, 1990; Chen and Deakin, 2015).

Alain Supiot's works on legal anthropology can be used here to further support the layering hypothesis of law with a more philosophical approach. He argues (2007, pp. 79-83; 2017, pp. 52-67) that every civilisation or society has certain founding beliefs, and that the modern nation-state, with its particular understanding of law, is no exception. These beliefs are the widely accepted and respected social norms, which may have been taken for granted or implicit. Laws are like surface-level rocks that have 'cooled down' from the magma—the social norms and practices. Laws are thus layered upon social norms and reinforce their basic structures.

For example, Supiot claims (2007, pp. 27-29, 95) that the modern and western formal laws and states require an epistemic foundation of an essential belief in a 'third party as guarantor of identity'. When applied to governance, these collective social norms become a search for an impersonal arbiter, who gives perfect judgements without human arbitrariness (2007, pp. 86-100; 2017, pp. 19-25, 67-73). Related notions like justice, trust and equality, are all parts of the 'magma' that forms the epistemic foundations of law (2007, pp. 13-29; 2017, p. 1).

This kind of path-dependency can lead to negative reinforcements due to 'concept drift', a term borrowed from computer science (Gama et al., 2014; Žliobaitė et al., 2016). While social norms change in social interactions rapidly, the more static legal forms sometimes lag behind. Law's inter-temporal and asynchronous qualities often reinforce past social norms which might have already changed, and thus can 'cool down' the evolving social dynamics in communities. Such cooling effects bring incompatibilities and conflicts between the rather stationary law and the more dynamic social norms (Supiot, 2007; Deakin, 2015).

Therefore, the emergence of the rule-of-law state as the exterior layer of governance that is more abstract compared to norms does not *necessarily* equate to social progress. Laws are not intrinsically more effective at governance than social norms in every situation. Law must often fade into the background (Chen and Deakin, 2015), provide room for local private autonomy, and keep adapting to evolving social practices at appropriate times to avoid cooling effects[17]. Coevolution and complementarity between laws and social norms are crucial for their functioning as different but layered modes of governance. This understanding has the potential to provide a blueprint for understanding the code versus law debate.

To summarise, in the shift from social norms to laws, governance becomes more formalised and step-by-step (that is, somewhat more algorithmic) in order to achieve scale effects. However, the layering and path-dependence of laws upon norms could bring negative reinforcements and cooling effects. As we shall see in the next part, Supiot's historical account of *Governance by Numbers* (2017) further unearths the evolution of governance by laws to governance by quantification and statistics.

3.4 From Laws to Numbers: Statistics, Cybernetics, and the Erosion of Rule-of-law

[17] E.g. in the case of margin of appreciation. Benvenisti, 1998.

The chapter now moves on to test the scaling and layering hypothesis in the turn from laws to numbers. Governance by numbers relies on statistical analyses, which are themselves defined by reference to non-numerical frames of reference, including those of law and social norms. Societies' over-reliance on numbers may erode the rule-of-law state and cause negative lock-in effects or 'imprinting'. This dynamic is similar to the way in which over-reliance on law's formal and technical features can harm the informal social foundations of both law and state.

3.4.1 Statistics: law geared to numbers

(1) Numbers from law

Existing literature has argued that governance by numbers is essentially an institutional extension of governance by laws, where number-based statistical analysis becomes the main medium for governing, adopted by either public or private actors (Supiot, 2017). Its functioning is dependent on the basic structures of the state and formal legal institutions (Desrosières, 2002). At the same time, the main medium of governing shifts from legal interpretation to statistical analysis. Statistical techniques are increasingly used by nation-states in legislative, public administrative and judicial decision-making. They are also adopted by corporations in governing online transactions, disputes and moreover, carrying out state administrative functions like compliance ranking and grading (Supiot, 2017, pp. 78-102; Mac Síthigh and Siems, 2019). Similarly, in global governance, statistics is adopted by decision-making bodies in international organisations, and is directly applied by private entities (Supiot, 2017, pp.160-163).

(2) Numbers v. law

Governance by numbers differs from governance by laws in using statistical analysis in place of legal interpretation. Statistics is 'the science of using information discovered from collecting, organizing, and studying *numbers*' (Romeijn, 2017). These numbers are, in essence, mathematical representations of (a) empirical facts and (b) hypotheses on the relations between such facts. The empirical facts are measured and structured into numerical data sets, and the hypotheses are 'formulated in terms of probability distributions over possible data sets' (Romeijn, 2017). Statistics is thus also defined (Upton and Cook, 2008) as 'the science of

collecting, displaying, and analysing *data*.' It uses probability for decision-making in the face of uncertainty (Moore, 1992).

These features of statistical analysis, which draw on mathematical abstractions of empirical facts and their related hypotheses, imply an increase in the computational content of governance. This arises from the use of number-based mathematical language, with its high levels of disambiguation and completeness. Problem-solving by statistics closely follows the binary logic on which computers operate. Calculations using statistical methods can thus be automated by machines. In contrast, the law relies on text-driven legal concepts expressed using natural human language. Although these concepts are also abstractions of social entities and their relations, the use of natural language enables them to take advantage of the ambiguity and openness of legal texts, and thereby create room for re-interpretation and innovation in the law.

(3) Numbers replace Law

The history of statistics begins with probability theory in the seventeenth century, and culminated in twentieth-century applied econometrics (Desrosières, 2002). Desrosières's conceptual exploration of statistics in *The Politics of Large Numbers* is deepened in Supiot's *Governance by Numbers* (2017). Supiot argues that statistical methods based on numbers have dethroned laws and became the predominant medium for governance in their place.

He first explains how numerical quantification has already been put to normative use in accounting, managing, judging and legislating (2017, pp. 78-102). He then shows how they were transformed into statistical methods that achieved the subjection of law to numbers in a range of societies (pp. 103-143). In advanced capitalist societies, in particular, law lost its isonomic function as a 'shared reference valid for everyone', and was downgraded to becoming merely 'legislative products which compete on a global market of norms.' (p. 121). In the process there was a loss of the belief that state and laws are the precondition of the calculations which, according to neoclassical economics, are the basis for achieving social harmony in governance (pp. 103-104). Supiot claims that (pp. 121-143) various branches of neoclassical law and economics contributed to this process. He argues that game theory, agency theory, the Coase Theorem (1960), and the new comparative analysis (La Porta et al., 1998) all constituted a concerted effort to transform governance into a statistical calculation process.

The result is that statistical analysis further dominated governance in various national and global contexts. Governance by numbers aimed at bringing about beneficial scale effects, but

according to Supiot, did not last for long before undermining its own foundations—state and laws (Supiot, 2017).

3.4.2 Scaling by cybernetics

The scaling of governance by numbers could be described in a logic of 'cybernetics', according to Supiot (2017). Cybernetics aims to enlarge the scale of social cooperation through managing material and human resources together through real-time feedback mechanisms. Yet it sees each social agent as atomised and indifferent to others. In the cybernetic conception of society, each agent, driven by subjective calculation of personal interests, instantly reacts to stimuli transmitted by statistical indicators in feedback loops to self-adjust their behaviour in order to achieve systematic stability and efficiency (2017, pp. 1-13, 25-31).

Supiot (2017, pp. 153-156, 193-196) argues that the New Public Management Movement, in vogue since the 1990s, demonstrated how the statistical-cybernetic knowledge originating in corporate management quickly penetrated national public administration practices across the globe. In addition, numbers used in financial markets and the accounts of corporations have become one of the main media of corporate governance since the 1980s.

Moreover, for global scaling, Supiot claims that (2017, p. 160) the IMF has used quantified indicators in its financial assistance programmes, namely the structural adjustment plans, in Latin America, Africa and the Eurozone. These quantified criteria include macro-economic variables which measure the success of the countries receiving financial assistance in attaining agreed statistical objectives in order to achieve the overall global governance goals of the IMF.

The EU and the World Bank have used similar numerical indicators for other large-scale global governance projects. They have adopted the same 'conditionality' mechanisms as the IMF. They have used quantified objectives as the conditions for their aid to poor countries. Some of these statistical indicators had evolved into the 'human development index' by 2011. In another example, quantitative targets and tradable rights for national CO2 emissions also became the dominant means of governance of global warming issues (Supiot, 2017, pp. 161-162).

In addition to the public sector, private actors such as credit-rating agencies also use statistical measurements to scale up global governance. These agencies generate risk indicators of 'a country's actual or projected financial performances, and the trustworthiness of its debt securities.' They are regulated only by a private body, the International Organisation of

Securities Commissions (IOSC). It provided its member with a non-legally binding code of conduct, or 'Methodology', in 2003 (Supiot, 2017, p. 162), and complemented this Methodology in 2017 with and 'Objectives and Principles of Securities Regulation'[18].

Nevertheless, Supiot argues that this scaling of governance by numbers is unsustainable in the long-run in the financial market. The 2008 financial crisis reveals its limits. This is so because statistics and cybernetics gave a disproportionate amount of power to those who construct the numbers. The attempt to use quantified analysis to provide a neutral façade for political decision-making is eventually self-defeating. Since the 2008 financial crisis, the EU for example, has gone back on its self-regulatory attitude towards the credit-rating industry and established a European Securities and Markets Authority in 2010 to tighten its legal checks on private rating-agencies (Supiot 2017, pp. 162-163).

In addition, after the financial crisis the private international body for banking regulations, that is the Basel Committee, also managed to produce the Basel III standards agreed by the central banks of major jurisdictions. These standards however were only finalised in 2017 and 2019, and are still awaiting to be implemented due in 2025 by the central banks of, for example, the EU, UK and USA (European Central Bank, 2023; Bank of England, 2023; Federal Reserve, 2023).

These setbacks for governance by numbers in the global financial markets, including both credit-rating and banking regulations, demonstrate the importance of identifying negative layering effects and the risks thereof.

3.4.3 Layering: statistics and the erosion of the rule-of-law state

(1) The ontology of statistics

As we have seen, governance by numbers largely inherits the basic structures of the nation-state and formal legal institutions (e.g. Desrosières, 2002) and expands them to international institutions (Supiot, 2017). After all, statistics, as the core medium for governance by numbers, originated as a 'science of the state' or 'political arithmetic', derived in the 18th century from the German word *Staatistik* referring to a 'descriptive and not quantitative knowledge of the state' (Supiot, 2017, p. 78; Desrosières, 2002). Governance by numbers in this sense, has an inherent non-quantified or non-calculable element, which must be defined by reference to laws and social norms.

[18] IOSC, 2017.

This layering is revealed in two notions crucial to statistics: data and measurement. 'Data are characteristics or information, usually numerical, that are collected through observation.' (OECD. 2008. p. 119). Numerical observations, or data, must involve measurement, defined as 'the assignment of a number to a characteristic of an object or event, which can be compared with other objects or events' (Pedhazur et al., 1991, pp. 15–29; International Bureau of Weights and Measures, 2008, p.16).

Statistical analysis, by its very nature, is the use of numbers to describe social phenomena and make comparisons of these numbers. The social phenomena themselves must be pre-defined as categories in natural language before they can be counted and compared. This can be seen in concepts such as 'employee' which have a shared juridical and statistical foundation in certain time-space (Deakin, 2015). Statistics, through its use of numbers, achieves mathematical completeness and disambiguation only by taking conceptual categories which have origins in natural language and roots in laws and, ultimately, in social norms and practices. This implies that even statistics have a layering quality.

The complete and disambiguated quality of mathematical language is not all positive. It means that statistical concepts are more closed by nature and less open to reinterpretation in the light of changing social circumstances than is the case with legal concepts. Text-based legal concepts, such as 'employee', can drift through time with new interpretations, and thus include new samples and populations which were not previously defined within these categories. An example of this is the way in which, in the 2021 Uber judgment, the UK supreme court expanded the legal concept of 'worker' to incorporate new kinds of employment relations of the gig-economy (*Uber BV v Aslam* [2021] UKSC 5). This natural linguistic ambiguity provides law with an open texture which permits re-interpretation (Deakin, 2015; Markou and Deakin, 2020).

In contrast, statistical concepts like p-values, confidence intervals or normal distributions have to be stable in order to maintain mathematical completeness. These statistical concepts cannot themselves deal with problems like the 'conceptual drift' involved in legal evolution (Deakin, 2015). They can only absorb new data once they have been manually reconstructed and reinterpreted with new measurements of empirical facts. Statistics requires humans to consciously and diligently update the mapping of legal categories into numbers ('counting') according to the new time-space, similarly to the way that laws need to be re-made through the flexible operation of interpretive legal techniques or new legislation.

In addition to the temporal limitation, space is also a limiting factor of governance by numbers. In the emerging literature it has been argued that 'all data are local', and that data and

statistics are indexes to local knowledge. This means that data sets are often not complete or portable, but dependent on the understanding of their data settings, namely specific contexts and meaning in a certain place (Loukissas, 2022).

Governance by data / statistics, in this sense, emerges from local community practices in a way similar to how governance of common property resources (CPR) emerges as polycentric modes of collective behaviour. Ostrom's field-work based studies describe how the state and market co-produce community knowledge and governance in the case of CPR. Her case studies show that forests, fisheries, and other 'things' that trigger collective action problems, can hardly be governed by a dichotomy of state monopoly or private property rights. Thus, the formal legal coding and arrangements of CPR need more dynamic and diverse considerations of the natures of the CPR and their specific community contexts. These context-specific formal legal practices require, while co-producing, informal local knowledge in the communities (Ostrom, 2010). Numbers, or data and statistics, as new modes of governance, can be seen as co-produced by local poly-centric agents and collective behaviours, and thus have complex attachment to that geographical space and community, namely the data habitat.

In addition, Pistor describes how the law has historically 'coded' (or categorised) land, technology, financial instruments, and other tangible / intangible things into property, patents and / or capital. These standardised legal assets then become quantifiable in the eyes of the global financial market. She argues that legal institutions and lawyer's innovative practices are always in the process of creating or contesting property rights and new forms of transactions, which are the foundations of quantified capital. The legal community in this sense helps the state authorities to code / decode shared understandings and practices on coercive measures and exchange informal knowledge that allows for large-scale market transactions, global cooperation, as well as wealth and inequality (Pistor, 2019)

Deakin and Meng, in addition, used the case of the 'household responsibility system' that governs China's rural lands to argue that ownership should move beyond the concept of 'legally enforced private property rights'. Instead it should be understood as an 'emergent, diverse and complex institution'. Deakin and Meng's arguments draw from and emphasise Ostrom's theories of CPR governance and legal institutions' capacity to code / contest property (Deakin and Meng, 2022). The governance of individual property in this sense is never straightforwardly about 'private ownership', but always has to be based on the local understanding and meaning of the property that is further quantified into data and statistics of the market.

This line of work on emergent and institutional understanding of property can also be drawn together with the work of Xu on evolutionary (Xu, 2017) and community-based theory of property rights in China (Xu and Gong, 2020), and Xu's argument for a law-and-community approach to understand compensation of takings in the European Convention of Human Rights (2019). Her work shows that if property should be 'coded' into national and global capital as data and numbers in the modern market's eyes, then the local community and social norms are the foundations of quantification processes.

(2) The erosion of the rule-of-law state and lock-in effects

Layering effects give rise to path dependencies which can lock-in the negative aspects of pre-existing laws and social norms based on local data habitats. In particular, the over-reliance on numbers by public and private actors can become too self-referential and erode the foundational rule-of-law state.

For example, Supiot claims (2017, pp. 150-152) that for financial markets, state deregulation allowed exponential expansions of new financial products through self-referential securitisation and credit derivatives at national and global level, where stock price statistics dominated corporate behaviours. Around the same time, major countries including those in the EU, deregulated accounting standards by delegating them to a private body, the International Accounting Standard Committee, in order to achieve greater scale effects. The Committee inverted the logic of accounting from state regulation under the principle of prudence to the concept of 'fair value'. This transformed the meaning of the numbers in the account from an expression of the liability and responsibility a corporation has towards third parties ('accountability') to 'an instrument for measuring a company's market value in comparison with its competitors' (2017, pp. 152-153).

As mentioned above, Supiot (pp. 84-85, 193-203) argues that the deregulation of financial markets and accounting standards, as well as excessive reliance on the rating agencies, were ultimately unsustainable as ways to achieve efficient governance at large scale; instead they led directly to the Enron accounting scandal and the 2008 financial crisis. The logic of governance by numbers eroded the public-private divide and thus undermined the foundation of the legal institutions on which effective market governance depended.

In contrast to this process, Supiot demonstrates that historically, governance by numbers worked well when it was coordinated with governance by laws. The rise of modern capitalism, for example, combined both the adoption of formal legal institutions and novel statistical

methods in governance. The invention of the double-entry system of book-keeping by medieval Italian merchants, which were later stipulated as accounting rules by Italian city authorities, facilitated the expansion of credit and debt in an emerging capitalist market through the state plus statistical method (pp. 80-83). A basic private-public divide to provide state institutions, and conscious updates of socio-legal concepts in the statistical analysis of governance by numbers, were essential to achieve beneficial complementarity between numbers, laws and social norms.

Deakin and Meng's work on Covid-19 governance (2020) aligns with Supiot's critique of governance by numbers in a Chinese context. They argue that the governance of public health risks requires a hybrid approach to the co-production of institutional knowledge, in a way similar to Ostrom's polycentric governance. They argued that public institutions like local states and international organisations are at the core of this polycentric governance. Meanwhile neoliberal theories and practices which favour self-adjustments to external shocks in the past decades, have weakened public international institutions in dealing with this global pandemic, and are now under pressure testing. They thus cautioned against the over-reliance on mathematical modelling and quantification to policy and law-making in time of the Covid-19 crisis.

This can be further drawn together with Xu's critique on the mathematical modelling and quantification of law in decision-making related to uncertainties in Covid-19 governance (2021, 2023). She claims that the essence and legitimacy of quantifying law lies in a 'Cartesian dream' that mathematical modelling provides certainty in the face of uncertainty and ignorance in governance decision-making. Nevertheless she warns of the pursuit of excessive certainty about the future, and argues with case studies in Covid-19 governance that three challenges remain in the quantification of law. First, (1) uncertainty is the foundation of knowledge itself, and ignorance is in fact the 'refinement' or 'twin' of knowledge (p.25). A Cartesian dream of perfect prediction and control by quantification of law can thus be counter-productive, and can hide the real context and biases of the modeller's own choices. (2) Second, rule-makers can use or manipulate mathematical models as 'truth' or evidence for pre-determined agendas, which makes it more difficult to accommodate local knowledge from and / or local public opinions or experts in the field, such as public health professionals. Quantification of law in this sense can undermine law and public knowledge production (3) Third, rule-makers are not aware or find it hard to embrace that ignorance is the condition to knowledge, and should intentionally invent alternative strategies, and 'communicate the unknown' in governing uncertainties, rather than over relying on numbers and statistics (pp. 26-27).

These more recent case studies on the failures in global pandemic governance using quantification support Supiot's critique on governance by numbers as detailed above, and manifest the layering effects between law and numbers.

In short, initial evidence shows numbers and statistics have helped governance to scale from local to global levels, but are still deeply embedded, imprinted, and sometimes locked-in by their socio-legal foundations. To avoid negative complementarity, societies need to avoid over-reliance on governance by numbers and embrace the fact that this more abstracted mode of governance is sustained by less abstracted modes like social norms and laws.

This pursuit of hybridised governance for positive complementarity and scaling is further complicated by governance by code. As emerging technologies of ML algorithms become increasingly powerful and adopted by private and public actors in governance, they add another level of scaling and layering. The next section addresses this last pair of governance modes where algorithmic governance, or code, is postulated as being layered on numbers.

3.5 From Numbers to Code: Computation, Big-data, and 'Freezing'

Governance by code, as the most purely computational mode of governance, completes the 'algorithmic turn' from norms to laws, laws to numbers, and numbers to code. This highest level of abstraction achieves greater scale effects than all preceding modes of governance, largely thanks to big data: the unprecedented amount of data that computers can process. But the data which code depends on are still numerical measurements of empirical facts which cannot be entirely freed from linguistic constructions. Just like numbers, governance by code has to be cognitively founded on laws and social norms.

3.5.1 Computation: from statistics to 'learning' code

Computation is becoming a major medium of governance with the recent breakthroughs in 'artificial intelligence' (AI), which in practice means Machine Learning (ML) including Deep Learning (DL) algorithms. Through these 'learning' algorithms, aspects of human decision-making and problem solving can now be automated. Computations carried out at scale have features which are not readily comprehensible to humans. DL is the subset of ML that is responsible, according to some, for giving machines a form of 'intelligence' (LeCun et al., 2015). As DL claims to model or replicate how the human neurons work in the brain, it is

'regarded as the best model of perceptual similarity judgments in primates' (Bucker, 2019, pp.1-2).

(1) Code from numbers

ML and DL algorithms, when examined more closely, are 'learning' algorithms that originated from statistical analysis, that is, the logic of governance by numbers. ML is defined as 'a family of statistical techniques that uses an algorithm to "learn" over time through the iterative adjustment of mathematical parameters and optimise performance at a task' (MacKay, 2003; Alpaydin, 2016). It 'learns' by using mathematical models of artificial neural networks (NNs) that are shallow, in that they are often no more than three to four layers deep (Schmidhuber, 2015).

In contrast, DL, as part of a broader family of ML methods, 'learns' through more layers of NNs—the 'deep convolutional neural networks' (DCNNs) which can contain several hundred layers. DL algorithms have more computational stages than shallower ML algorithms, enabling them to perform more complex optimisation tasks (Schmidhuber, 2015).

Therefore, in essence, ML and DL algorithms still use certain 'statistical techniques' that are based on numbers and calculation. They must use 'mathematical parameters' to construct and adjust the NNs in order to optimise task performance. Thus 'learning or credit assignment is about finding *weights* that make the NN exhibit desired behaviour, such as driving a car.' (Schmidhuber, 2015, p. 4). This does not go beyond the basic logic of using statistical analyses in governance to achieve optimised outcomes.

In addition, historically, code came from numbers and statistical techniques. The basic mathematical models for ML and DL had already been developed in the 1960s and 1970s. Later techniques such as *backpropagation* (BP) for error-correction in supervised learning (SL) were already adopted in ML and DL in the 1980s and have stayed popular since then (Schmidhuber, 2015, sec 5.5, 5.6; Buckner, 2019, p. 2). In the 1990s, practical applications of DL were already feasible with the breakthrough of unsupervised learning (UL) (Schmidhuber, 2015, sec 5.10).

(2) Code and big-data replace numbers and laws?

It was however only in the 2000s that the applications of ML and DL algorithms proliferated in certain business settings. This was because, by then, the 'learning' code finally had enough data work on. Data generated on the internet and personal devices, which could

be stored and processed by increasingly powerful computers with greater calculation capacities, became available at scale (Buckner, 2019). Traditional statistical analysis is most helpful when data scaling is limited. It provides sampling methods and hypothetical probability distributions that harness the limited data at hand for predictions using mathematical techniques, such as the linear regressions widely used in the quantitative social sciences. In contrast, when data scaling is large enough, ML and DL algorithms no longer need to define certain probability distributions a priori in their data modelling. The capacity to infer regularities through iterative error correction meant that ML and DL became more flexible in developing complex mathematical models to solve problems (Breiman, 2001). These processes had the 'ability to discover novel solutions directly from problem data', with the use of backpropagation in SL a major use case (Buckner, 2019, p. 2). ML / DL research and applications are now considered part of a distinct discipline of 'data science', which combines statistics and computer science in practical problem-solving settings using large amounts of data (James et al., 2013; McQuillan, 2018).

When the problem-solving capacities of ML and DL algorithms were fully realised with big data, governance by code emerged as qualitatively different from other modes of governance, and seemed to have the capacity to replace them. Some philosophers argued that since DL applications, in particular, are qualitatively distinctive from the shallower ML, a new understanding in what computers could achieve in replicating human decision-making processes is required (Buckner, 2019). Natural Language Processing (NLP) uses a DL algorithm that can 'read' natural language, and thus replicate aspects of human decision-making, including legal ones (Young et al., 2018; Deng and Liu, 2018). NLP algorithms can carry out legal tasks such as predicting legal judgments by analysing case law (Aletras et al., 2016; Alarie et al., 2016; Xiao et al., 2018), even though, at present, this facility essentially takes the form of predicting one part of a single judgement text from another (Medvedeva et al., 2020, 2023).

(3) Code's hyper-reflexivity

In addition to its growing problem-solving capacity, governance by code also has a distinctive quality of 'hyper-reflexivity' (Lessig, 2006; Zuboff, 2015, 2019). Granted, statistics, laws and social norms already shape their realities when describing them; but they lack the reflexive capacity of the ML and DL-based predictive algorithms. As Zuboff shows (2015, 2019), these 'learning' codes are increasingly being used to shape social environment and

interactions in real-time by monitoring and nudging individual behaviours. Hildebrandt argues (2016, 2020a) that the increasingly 'smart' environment built by ubiquitous 'learning' technologies will create an 'onli*f*e world' (original italics by Hildebrdant), where people's everyday decisions and social interactions are steered real-time by infrastructures 'saturated with artificial, data-driven agency'.

Such use of big-data through 'learning' and hyper-reflexive code can achieve greater scale effects in governance than those available through law and numbers. Yet the layering effects are also magnified. Such layering can bring excessive 'freezing' to societies that are over-reliant on code as a mode of governance.

3.5.2 Scaling: big data, surveillance capitalism and 'information civilization'

Learning and hyper-reflexivity as features of governance through ML and DL code afford the advantages of scaling. For example Zuboff (2015, pp. 78, 85) unearths how the adoption of 'learning' algorithms in everyday governance has led to unprecedented scale effects in advanced economies. She argues that global platform companies, by developing and deploying ML and DL based forms of algorithmic governance, are potentially creating a new 'information civilisation' in the century to come.

Scaling by code and big-data differs qualitatively from scaling induced by numbers. It scales up governance not only by disrupting the sovereign states and territorial boundaries, but also by reshaping fundamental human nature and agency, according to Zuboff (2019, Ch12, 15, 16).

Zuboff traces (2015, p. 89) this development by analysing two key papers of Google's chief economist, Hal Varian (2010, 2014). Varian claims that computer code and the Internet provide 'new general technology' for 'combinatorial innovation' and 'recombinant growth'. Computer-mediated transactions based on code enable 'new forms of contract due to better monitoring', and these transactions in turn provide new tools for governance. These tools can be used to modify users' behaviours in real-time, which, Varian argues (2014), opens new possibilities for Google and other tech companies like Uber and Airbnb to realise the full potential of big-data in managing daily life from travel to accommodation.

Although Varian's papers focus on the economies of scale provided by code and big-data, Zuboff argues that they have more general implications for governance. She claims that Google and other tech companies are in fact creating a new information civilisation in which

governance by that reinforces a capitalistic logic of accumulation and calculation. She terms this 'surveillance capitalism', which in its roots governs by financial calculus, but goes beyond numbers in the way it uses ML / DL code to modify users' behaviours and reinforce a predetermined capitalistic calculus (2015).

In addition, Julie Cohen emphasises the legal foundations, or in her words, 'legal constructions' of information capitalism. She argues that today information-age harms and conflicts over ownership of information-age resources are transforming legal institutions. This process is similar to the tectonic shifts that the U.S. legal system went through due to the emergence of pervasive harms in the industrial age like population and poor labour conditions, from the late nineteenth century through the mid-twentieth century. She argues that new types of business models, code architectures and control over information flows are reshaping legal institutions, the regulatory state, and international networks and transnational governance institutions. These legal institutions in turn, also help in scaling up an information economy (Cohen 2019, chapters 1, 6, 7).

3.5.3 Layering: cold start, concept drift, and freezing

Zuboff criticises the *capitalistic version* of governance by code, and argues for a 'hacking' of it: liberating and redistributing existing computational architectures. She calls for an alternative version of governance by code and a new information civilisation that can overcome the exploitative logic of surveillance capitalism (2015, p. 85; 2019). This chapter understands Zuboff's call as one of resistance to negative layering, that is, the reinforcement of path-dependencies. To further test the layering effect of governance by code, the computer science literature on 'cold start' and 'concept drift' is examined, with examples drawn from Hildebrandt's work on code and law.

(1) The ontology of code within physical and mathematical bounds

The narrow definition of algorithm refers only to computer code: 'a tool for solving a well-specified *computational problem.*' 'The statement of the problem specifies in general terms the desired input / output relationship. The algorithm describes a specific *computational procedure* for achieving that input / output relationship' (Cormen et al., 2009, p. 5)

This narrow definition of 'algorithm' requires both the procedure and the problem being solved to be computational. This means that the governance problem must be so well-defined that it could be written in binary code and operate on computers. Computer algorithms can only

solve the subsets of problems in governance that are computable within the physical limits of computers, following a definitive logic of 'if this then that' (IFTTT) (Hildebrandt, 2020b).

Two ontological claims follow from this observation. First, no matter how 'intelligent' ML / DL algorithms are, they must still fall under the narrow definition of computational algorithms, in other words, their goals in governance terms are confined within their material limits of the physical transistors wired in computers.

Second, even if these ML / DL algorithms could replicate some human decision-making processes with unprecedented accuracy, they still cannot escape their nature as computational algorithms and 'statistical techniques'. These techniques are rooted in quantification, measurement and numerical calculation, but now enhanced by big data. In this sense, governance by code involving ML and DL is layered on governance by numbers, which is layered on governance by laws and social norms.

(2) ML / DL layered on statistics: cold start and concept drift

These ontological claims can be evidenced in two major limitations of ML and DL algorithms so far: 'cold start' and 'concept drift'. These limitations reveal the layering that is inherent in code being based on laws and social norms.

First, 'cold start' is a well-researched problem in algorithmic recommendation systems where new communities, items or users are introduced and the ML / DL code cannot draw any inferences from existing information. 'New community' refers to the situation where almost no users are present and thus the recommendation algorithms have very little information on potential interactions. The 'new item' situation happens when an item is newly added into the recommendation system, which might have content information but no interaction information. The 'new user' problem refers to the situation when a person is recently registered without any interactions yet in the community. Cold start problems manifest the limitations of 'learning' algorithms in dealing with new situations in the future. Their mitigation strategy often is to rely on hybrid recommenders, such as combining collaborative filtering and content-based filtering (Lam et al., 2008; Bobadilla et al., 2012; Wei et al., 2017; Li et al., 2019).

Another major problem of ML / DL predictive algorithms is 'concept drift'. This occurs when there is a change in the relationship between input and output data at the point when a model is applied to solve a new problem or perform a new task. As the meaning of data changes over time, the performance and accuracy of predictive models based on historical data will decay. This is because while the concepts and categories of the phenomena being observed can

shift and evolve with environments through time, the mathematical models which represent them are fixed, assuming a static relationship between input and output variables. In novel real-life situations where new data sets are generated, predictive algorithms must be updated in order to work, but this may not be straightforward (Gama et al., 2014; Žliobaitė et al., 2016). Even more difficult is the situation where the concepts at stake are dependent on a 'hidden context' not explicitly provided for in the statistical or predictive features of the model. The opacity of DL models, in which outputs can be observed but an element of processes and calculation in neural nets remain hidden, is a potential cause of such concept drift (Widmer and Kubat, 1996; Tsymbal, 2004).

Similarly to governance by numbers, governance by code has to deal with the gap between its mathematical stationarity or closure, and the open nature of social interactions. When social entities and their relations change, these changes can be addressed via norms and law through novel interpretations, using text-based natural language in non-numerical frames.

(3) Freezing of the future as layering effect

The layering hypothesis, at a systematic level, can be further tested via Hildebrandt's argument of 'freezing' (2020b). Compared to text-driven laws that are structured on natural language and speech acts, Hildebrandt argues that 'code-driven law' (law written in computer code) has certain constraints in its computational architecture, leading potentially to its ineffectiveness in governance. These constraints include the requirement of formal deduction ('if this then that', IFTTT), disambiguation in terms and rules, and the incompleteness and inconsistency of computation. These limitations of code are related to the 'uncertainty that inheres in the future'(Hildebrandt, 2020b).

Hildebrandt argues (2020b, III.B) that ML / DL algorithms embrace a false assumption to the effect that the distribution of past training data can be a close approximation of the distribution of future. In reality, the future distribution of data can only be *predicted* rather than *learned* by algorithms. Hildebrandt claims (2020b, III-C) that 'radical uncertainty' defines the future instead. This does not mean the future is entirely random or arbitrary, but that it can be underdetermined, especially considering the 'double contingency' raised by Parsons and Luhmann (Vanderstraeten, 2002; Parson 1951; Luhmann, 1976, 1995). The double contingency embedded in natural language makes humans always in the process of anticipating how others anticipate our actions. This is done through the use of linguistic concepts. For example a person X anticipates how their actions are *'read'* by Y, as if Y is reading a string of

texts, while Y attaches their interpretation of a string of *texts* (stemming from X's action) to their *interpretation* of X's *action*, and vice versa. This mutual anticipation based on interpretable texts—conceived and expressed in natural language concepts—generates radical uncertainty, which demands institutionalisation of specific patterns of behaviours that form points of consolidation and stabilisation, making it possible to reduce complexity and uncertainty in social systems.

Not just legal certainty, but also open texture and interpretability, are crucial in this stabilisation process. Laws to some extent provide stable expectations in social interactions, but are more than mechanical applications of disambiguated rules. This hybrid of certainty and open-interpretative nature enables law to achieve a balance between radical uncertainty and the need to avoid freezing the future (Hildebrandt, 2020b, IV-A). Hildebrandt argues that code-driven law freezes the future. Code's temporal specificity, expressed in the disambiguated nature of mathematical modelling, means that it cannot adapt to unforeseen circumstances in the same way that human language can. Code will always lag behind an open future; legal concepts, with their inter-temporal, reflexive and adaptive nature, can better capture future scenarios (Hildebrandt, 2020b, IV).

Hildebrandt's freezing argument supports the layering hypothesis set out in this chapter. Governance by code ultimately depends on the open texture of laws and social norms, just as law is sustained by evolving social norms and practices. The layering of code on law and norms defeats the teleological claim that code is necessarily superior to laws or social norms as a mode of governance. The layering hypothesis provides a preliminary basis for understanding how code and law might work together via complementarity as opposed to substitution. Governance by code does not necessarily or automatically equate to social progress, any more than governance through law is an advance on governance through social norms.

3.6 Chapter conclusions

3.6.1 Research questions

The thesis answers the overarching research question: to what extent can text-based law be computed systematically, namely replaced / substituted by computer code.

In this chapter, this overarching question is assessed through answering a set of sub-questions. First, what is the nature of algorithmic governance, compared to more established ways of text-based legal governance? Second, how do laws interact with other modes of

governance, like social norms, numbers and code? Third, what is the underlying model of law-code interactions?

3.6.2 Contributions

By answering these questions, this chapter makes two contributions to the field of algorithmic governance. It first proposed an institutionalist framework, in contrast to orthodox law and economics theories, to understand the nature of code, namely algorithmic governance. It compared algorithmic governance to other modes of governance based on social norms, laws and numbers, and argued that code is a distinct mode of governance with the highest level (compared to the other modes) of abstraction and formalisation.

Secondly it proposed a 'scaling and layering' hypothesis to understand how different modes of governance interact. The argument is that as there is a movement from social norms to laws, from laws to numbers, and from numbers to code, the governance of larger territories and populations by means of higher level of formalisation and abstraction becomes possible; but that, at the same time, the higher level abstractions are path-dependent and cognitively reliant on the lower-level abstractions to sustain their operations.

These two arguments together reveal that algorithmic governance is sustained by law in the layering effects. They show how law and code are complements rather than substitutes.

It follows that societies must constantly balance and (re)align these four modes of governance, if negative complementarities are to be avoided or mitigated. For example, the proper functioning of ML and DL algorithms requires timely updates on their data inputs. Here, data do not, in themselves, produce more data. Data is in essence constructed by non-computational concepts. In the same way, governance by code requires reference to laws and social norms. Even the more purely computational forms of governance by code are historically and logically dependent on anterior modes of governance. This is so no matter how formally complete 'learning' algorithms might be, or become, in terms of replicating human decision-making.

Complementarity between law and code needs to be turned to beneficial ends if negative interdependencies are to be avoided. Law-makers, therefore, must be aware of the different nature and limitations of each mode of governance, and consciously avoid potential freezing or lock-in effects induced by over-reliance on code and numbers. The Enron accounting scandal and the 2008 global financial crisis serve as warning examples.

The following chapters use three detailed empirical case studies to test the validity of the framework and hypothesis proposed in this chapter. It aims to showcase how beneficial complementarities between law and code can be achieved, and negative layering avoided. It starts with the 'scoring' systems of China's SCS.

Chapter 4 (Semi)Automated Scoring Systems: an analysis of data-driven financial and social scoring in China's SCS

4.1 Summary of the argument

This chapter provides a case study of the (semi)automated 'scoring' systems within China's SCS. Historically it is the first major institutional constituent of the SCS to be developed. It is a prelude to discussing (semi)automated law enforcement mechanisms in the next chapter (Ch 5), which constitute the second major component of the SCS.

Two sets of SCS scoring systems are examined: the financial-commercial scoring systems, and the local social scoring systems. New empirical evidence (in the form of online documents and interviews) is analysed to demonstrate how scaling is limited by layering, when it comes to scoring across large territories and populations using data and code. The case studies evidence supports the overarching argument of the thesis that algorithmic governance is layered upon legal governance, namely that law is computable only to the extent that computer code is sustained by law's operations within institutions.

4.1.1 Research questions

As we have seen, the thesis asks one overarching research question: to what extent can text-based law be systematically computed, namely substituted by computer code.

This chapter assesses the overarching research question by answering a set of sub-questions related to China's SCS, which include:

- What is the nature of the 'scoring' systems in China's SCS?
- How do data and code work in the scoring systems, and what are their institutional roots?

- Is there evidence of scaling and layering effects, where code expands governance scales but is limited by its dependence on legal institutions?

It is shown in this chapter that China's (semi)automated scoring systems are essentially data-driven socio-technical systems that are rooted in the US and European credit-reporting practices and public management scores. The SCS scoring branch is therefore part of the ongoing global scaling of governance by numbers and code led by advanced economies. The SCS scoring systems are nevertheless bound by Chinese local data practices and habitats. They are institutionally rooted in the activities of the Chinese Central Bank, local governments in cities and villages, regulatory agencies, private credit-reporting industries and big-tech companies like Alibaba, which learnt from American and European practices.

The new empirical evidence provided by this chapter shows how these SCS scoring systems, based on data and algorithms, initially scaled up (rose) but eventually faced setbacks (fell). Data and code, as new modes of governance, are evidenced to demonstrate complex and layered interactions with law and social norms when expanding across geographies and populations.

Evidence demonstrates how layering effects limited the scaling of scoring mechanisms. First there are 'concept drifts' that happen over time, when social reality changes and the numbers and code fail to capture such change. Second, the condensed information units, namely individual scores, are layered upon complex contextual knowledge embedded in specific, local interactions. They become institutionally over-capacitated when applied to general social systems, and wider geographical regions, resulting in the community's change of perception and reducing the score's legitimacy.

4.1.2 Structure of the chapter

The chapter tells two sets of stories of the SCS scoring systems, that is the two subsystems of scoring mechanisms, distinguished by the data they use and the agents / organisations that run them. First, section 4.2 examines the financial-commercial scoring systems through which the Central Bank and big-tech companies utilise private transaction data. Second, section 4.3 studies the local public credit information scores through which local governments harness law enforcement / compliance data. A chapter conclusion follows in section 4.4.

Each of these two scoring systems, respectively discussed in section 4.2 and 4.3, can be further bifurcated into two subsystems: i. the Central Bank's financial databases (4.2.1), which is in complementarity and competition with, ii. the private scoring systems, including Alibaba's

Sesame Credit (4.2.2); and iii. the urban municipal scores (4.3.1), which are intertwined with the iv. rural village scores (4.3.2).

The chapter tells the respective story of how each of these four subsystems of scoring expanded at first (scaling) but later revealed their limitations (layering).

The structure of the chapter is as shown below:

4.2 Financial-commercial scoring systems for citizens and businesses

- 4.2.1 Central Bank's qualitative databases
- 4.2.2 Alibaba's AI-driven Sesame Credit

4.3 Local public scores for citizens and businesses

- 4.3.1 Urban scores
- 4.3.2 Rural scores

4.2 Central Bank v Alibaba: scaling and layering based on transaction data

This section examines the rise and fall of the financial-commercial scoring systems of China's SCS. It starts with two descriptive parts. Section 4.2.1 and 4.2.2 tell a tale of two subsystems within the financial-commercial scoring systems: (a) the public financial credit reference systems (without scores), run by the Central Bank and its local branches, and (b) the private credit reporting systems, run by corporate credit-reporting agents and / or tech companies including Alibaba's Sesame Credit.

Section 4.2.3 provides a brief comparison between credit-reporting in China, US and Europe. Section 4.2.4 assesses the evidence presented in the previous parts. A sectional summary follows in section 4.2.5.

It is evidenced in this section that there are tensions between the two subsystems, namely: (1) the Central Bank's credit-reporting system which provides data for commercial bank's risk scores, and (2) the private scoring systems developed by large platform companies.

It is revealed how layering occurred respectively in each one of them. Evidence shows, first, that the Central Bank's financial credit-reporting database, which provides contextualised qualitative information, establishes the social and legal governance on which commercial banks and lender's quantified scoring and algorithmic governance are dependent. Second, tech companies' scores that uses overparameterised big-data and ML / DL algorithms for multipurpose predictions, failed in effectiveness when deployed at large scale and settings, and faced antitrust concerns.

4.2.1 The Central Bank's databases: Public credit referencing without scores

The Chinese government first began to realise the benefits of modern credit systems in the late 1980s, a decade into Deng's economic reform that invited the modern market economy into China. In order to help improve lending decisions to private enterprises, the Central Bank started setting up financial credit information systems both inside and outside the state-controlled banks. In the 1990s, the Central Bank further operated an internal online system of credit information for commercial banks.[19] By 1997 the Central Bank had established a *Banking Credit Registration and Counselling System* (银行信贷登记咨询系统) for enterprises, and by 2002, this system had already established three levels of databases (National, Provincial, and Municipal)[20].

The Central Bank also started experimenting with credit-rating (信用评级) in the 1980s. It created public-private joint ventures which only evaluated the risks of debts and payback capacities of large and medium-sized enterprises. Credit-rating is different from credit-reporting which collects comprehensive information beyond corporate bonds and payback ability. The latter also covers medium to small size enterprises and individuals. It was evidenced in finance literature that the Central Bank's early experiments of credit-rating were not very successful, and the Bank had to select only nine major credit-rating agencies by 1997 and regulate them with licences (Liu, 2017).[21]

Despite early developments of financial credit information sharing and rating, there was no concept of the Social Credit Systems (SCS) at this point. The SCS concept took off only in 2002, when then President Jiang Zemin (at the end of his second term) first raised it in the *Report of the 16th National Congress of the Communist Party of China (CPC)*, namely China's highest policy document every five years. He referred to the need to 'strengthen the SCS in a modern market economy, and break industrial monopoly and local foreclosure'. The notion of the SCS here was explicitly put forward as an integral part of the reforms to 'strengthen the modern market system and macroeconomic regulation'.[22]

---

[19] People's Bank of China, *Introduction to the central bank's release of the foundational database of corporate and personal credit information*, 2006.1.8. (online, in Chinese) . [中国人民银行，央行发布企业和个人信用信息基础数据库简介，2006] http://finance.sina.com.cn/g/20060118/17462287106.shtml.

[20] *Ibid.*

[21] For example, see a history on China's financial credit rating systems. Liu F., (2017). 'Research on regulatory issues of China's credit rating industry'. *Southwest Finance.*(8). (In Chinese).

[22] Jiang, Z. *Report of the 16th National Congress of the Communist Party of China (CPC),* 2002.

In 2003 the Hu-Wen administration set out its aim to 'accelerate the construction of credit information service systems for enterprises and individuals' (according to the principles of improving legal systems, franchising, commercial operation, and professional service). Then Premier Wen Jiabao explained that the construction of the SCS needed to 'start from the financial Credit Reference Systems', that is the CRS. He referred to the need to accelerate the establishment of a unified national Financial Credit Information Foundational Database (FCIFD) for Enterprise and Individuals, to embrace a nation-wide credit information network, accelerate legislative processes, and gradually build institutions for sanctioning incredible behaviour.[23]

Accordingly, in 2004 the Central Bank established the Centre for Banking Credit Information Service (CBCIS, 银行信贷征信服务中心) for coordination and implementations of policies on the SCS. First, for individual credit information, the first database among seven cities was set up on an experimental basis in 2004.[24] In 2005 the Central Bank enacted "Interim Measures to Administer the Individual Credit Information Foundational Database" in order to officially establish this national data-sharing mechanism.[25] In January 2006 this individual credit information system started formal operations, and by 2005 there were RMB 2.2 trillion balance of loans recorded, which covered 97.5% of the consumer loans in China. By 2006 there were already 340 million natural persons recorded in the system, and among them around 35 million had credit information.[26]

Second, for corporate credit information, by the end of 2005 the BCRCS for enterprises (established since 1997), had already recorded 4.52 million enterprises, with more than 90% of the loans covered in China, totalling RMB 1.736 trillion. From January to November 2005 there were an average of 1.44 million searches each month of credit information on this system.[27] After 2004 the Central Bank started to develop a new system in order to replace the old BCRCS, and by 2006 established the nationally unified *Corporate Credit Information*

http://en.people.cn/200211/18/eng20021118_106983.shtml ; Original Chinese texts can be found at China's official government website. https://www.gov.cn/test/2008-08/01/content_1061490.htm.

[23] People's Bank of China, 2006 (online, in Chinese).

[24] *Id.*

[25] People's Bank of China, *Interim Measures for the Administration of the Basic Data of Individual Credit Information*, [2005] No.3. http://www.lawinfochina.com/display.aspx?lib=law&id=4514&CGid=

[26] People's Bank of China, 2006 (online, in Chinese) at http://finance.sina.com.cn/g/20060118/17462287106.shtml. Also see People's Bank of China, Report on China's Credit Reporting Industry, 2013 (online in Chinese) [中国人民银行，中国征信业发展报告]

[27] *Id.*

Foundational Database for all domestic, foreign and rural credit cooperatives was established[28].

However, these two systems did not remain separate for long. In 2006, PBOC established the Credit Reference Centre (PBCCRC, 中国人民银行征信中心) to 'build, run and maintain' the national Financial Credit Information Foundational Database (FCIFD, 金融信用信息基础数据库), in order to aggregate the entire financial credit information for both enterprises and individuals in China.[29]

The new FCIFD was endowed with two official functions: first, socially promoting honesty, a spirit of contract, and compliance with law; and second, economically helping commercial banks with credit risk control, financial stability, increasing the scope of credit and consumption, and improving structures of economic growth and sustainable development.[30] The database's infrastructure was developed by a PBOC-owned enterprise, and the operation was run single-handedly by the PBCCRC.[31]

By May 2018, the FCIFD individual credit reference system had been connected to 3,347 institutions, and the corporate system connected to 3,283 institutions. It had recorded information on 962 million natural persons (69% of China's population in 2017) and 25.3 million enterprises. The average daily searches between January and May 2018 were 4.6 million times for individual credit information, and 270,000 times for corporate credit information.[32] By comparison, in 2015 US credit rating systems had covered 75% of the country's population, and in 2014 every individual credit subject was requested or reported 8.3 times on average, almost 4.75 times the number of Chinese searches in 2018.

[28] People's Bank of China, [2006] No.117. Notice of the General Office of the People's Bank of China on Issuing the 'Switching Plan for the Corporate Credit Information Foundational Database' [中国人民银行办公厅关于印发《企业信用信息基础数据库切换方案》的通知]（银办发[2006]171号）.

[29] PBCCRC official website, online at http://www.pbccrc.org.cn/zxzx/zxgk/gywm.shtml. Also see State Council, *Regulation on Credit-reporting Industry* [2013] No.631. 征信业管理条例 (www.gov.cn), and People's Bank of China, *Specification for User Management of Financial Credit Information Basic Database* [2014], JR/T 0115-2014 [《金融信用信息基础数据库用户管理规范》]

[30] State Council, [2013] No.631. Section 5.

[31] PBCCRC official website, online at http://www.pbccrc.org.cn/zxzx/zxgk/gywm.shtml.

[32] People's Bank of China, The 'government + market' two-wheel-driven development model of the credit reporting market has initially taken shape, 2018 (online in Chinese), at http://www.pbc.gov.cn/zhengxinguanliju/128332/128352/128402/3570533/index.html [征信市场"政府+市场"双轮驱动的发展模式初步形成]

Initially, when introduced in 2005-2006, the data collected from enterprises and individuals for the SCS was restricted to financial information. They include basic account information, loans and collaterals in financial institutions, and major financial indicators of enterprises and so on. The main users and data-collectors of the system were financial institutions, mostly commercial banks.[33]

However, after 2005 the Central Bank started undertaking the 'Tasks of inter-governmental cooperation and information-sharing between credit reference bureaus for enterprises and individuals that serve financial institutions and financial markets' (*the '2005 Tasks'*).[34] This state policy document required the Central Bank to cooperate with a range of governmental agencies, and to collect data beyond financial credit from enterprises and individuals. These data included information on basic personal identification (as opposed to only bank account information), enforcement of civil judicial decisions, outstanding tax obligations, payment of social security contributions, personal education, and professional certifications, such as bar association and accounting qualifications.[35]

These non-financial data were said to be required to be processed and profiled in order to provide credit reference reports. The aim of the *2005 Tasks* was to help in certifying bank accounts and eliminating fraudulent loans at their inception, and to provide comprehensive credit information on enterprises and individuals in order to help financial institutions to decide whether to lend, how much to lend and at what interest rates. The ultimate goal was to build a sanction and reward mechanism. In addition the *2005 Tasks* wanted to improve the enforcement capacity of courts, tax agencies, industry and commerce administration bureaus, and custom offices. These are socio-legal goals in terms of improving compliance with law, respect of rules and contracts, honesty and building a harmonious society.[36]

This inter-governmental, non-financial data-sharing can also be seen in the 2006 State Council Five Year Economic Plan (2006-2010), where "improving credit records of loans, *tax payment, contract fulfilment, and product safety*" are expressly set out as crucial tasks for

[33] People's Bank of China, 2006 (online in Chinese), at http://finance.sina.com.cn/g/20060118/17462287106.shtml. Also see People's Bank of China, Report on China's Credit Reporting Industry, 2013 (online in Chinese) [中国人民银行，中国征信业发展报告].

[34] *Id.* Original Document cannot be found due to the change of the Administration after 2008 into the Ministry of Industry and Information Technology (MIIT) under the State Council's ministerial structural reform.

[35] *Id.*

[36] *Id.*

establishing joint law enforcement programs in the SCS, meaning that multiple state and private agents come together to enforce for example, debts and contracts.[37]

In addition, the *2007 Instruction* emphasised the importance of building credibility in different industries, with the help of the industries' respective self-governing associations and wider governmental agencies. The key was to use 'Joint Sanction and Reward' mechanisms (JSRs), which is the core subject of Chapter 5, namely (semi-)automated law enforcement. This emphasis on inter-agency collaborations echoes the role of the *2005 Tasks* in expanding the data categories from merely financial ones toward commercial and industrial with the help of many state agents. The *2005 Tasks* also provided early signals on the transformation of the SCS from a financial-commercial logic toward a socio-political one. [38]

The Central Bank's move from purely using financial data, to also integrating non-financial data, foreshadowed the involvement of market players in the lucrative business of credit reporting and rating. Afterall, the *2007 Instructions* stated that the starting point of the SCS was 'market demand'. It described the market economy as a 'credit economy', and emphasised important tasks relating to loans, tax, contract performance and production quality. It claimed an urgent need to 'cultivate and develop the credit service market' and 'promote competition and innovation'. It required that the credit service market should provide products that satisfied 'multi-layer, diverse and professional credit service demands', and even at the time, open up to foreign investments. It was clear that the *2007 Instruction* pinpointed the main focus of the SCS at that time was on finance and the market, rather than society at large. [39]

4.2.2 Alibaba's Sesame Credit: expansion and fall of private credit referencing after mid-2000s

The previous section mapped the trajectory of the state-led financial SCS back to the early 2000s, originating in the central reforms of the banking sector during the 1980s. These were all established by the Central Bank, and were a significant part of the market economic reforms led by the state.

[37] State Council, Report on the Five Year Economic Plan (2006-2010), [2016] No.12, approved by the National People's Congress on 2006.3.14. Online in Chinese, at http://www.gov.cn/ztzl/2006-03/16/content_228841.htm

[38] *Id.*

[39] State Council, Several opinions of the General Office of the State Council on the construction of the social credit system, [2007] No.17. Online in Chinese, at http://www.gov.cn/zwgk/2007-04/02/content_569314.htm [国务院办公厅关于社会信用体系建 设 的 若 干 意 见(2007)]. Invalidated by State Council [2015] No. 68 Online in Chinese, at http://www.pkulaw.cn/fulltext_form.aspx?Db=chl&Gid=267463 [国务院关于宣布失效一批国务院文件的决定（国发[2015]68 号）].

Changes happened after the mid-2000s when these market reforms invited private firms to complement the Central Bank in the provision of foundational credit information to an emerging market economy, as required by both the *2005 Tasks* and the *2007 Instructions on the SCS*. Evidence shows that these market players rose and fell in the business of credit referencing.

4.2.2.1 Private credit rating and its fall after the 2008 financial crisis

At first, nine credit-rating agencies emerged in China by 1997, and they prepared the way for the emergence of more than 100 credit-rating agencies after 2005 (Liu, 2017).[40] Nevertheless, after the 2007-2008 US subprime mortgages crisis (which led to the 2008-2009 global financial crisis), transnational credit-rating firms suffered reputational losses, which also influenced China's credit-rating industry. The US Financial Crisis Inquiry Commission concluded in their 2011 Report that 'failures of credit rating agencies were essential cogs in the wheel of financial destruction'.[41] The metrics and scores provided by Moody's Investors Service, Standard & Poor's, and Fitch Ratings, for example, were largely discredited by the regulators and general public. Large quantities of securities that they gave a 'triple-A' rating, namely the rating for the safest securities, turned out to be 'junk'. Moody's for example, after 2007 had to downgrade 83% of the triple-A securities ratings it gave in 2006.[42]

Following the 2008 financial crisis, China's credit-rating market only had five major agencies. Two of which had 49% shareholdings held respectively by Moody's and Fitch Rating. The credibility of these agencies remained low nevertheless (Liu, 2017).[43] The Central Bank had also established preliminary regulations on credit-rating companies since 2006, and the most comprehensive and authoritative one was the *2019 Provisional Regulations on Credit Rating Industry*[44]. This *2019 Provisional Regulation* compared to the *2013 Regulation for Credit Reference Industry*, sets much higher fines for credit-rating malpractices compared to credit-reporting, because the price and profit for credit-rating products are usually much higher.

[40] Since the 2000s, the credit-rating market further developed with more than 100 agencies and five major players.
[41] Financial Crisis Inquiry Commission, The Financial Crisis Inquiry Report, 2011, at p.xxv. https://www.govinfo.gov/content/pkg/GPO-FCIC/pdf/GPO-FCIC.pdf
[42] *Ibid.*
[43] Since the 2000s, the credit-rating market further developed with more than 100 agencies and five major players.
[44] People's Bank of China, National Development and Reform Commission, Ministry of Finance, and China Securities Regulatory Commission, *Interim Measures for the Administration of credit rating industry.* [2019] No.5. Online in Chinese, at: https://www.gov.cn/gongbao/content/2020/content_5480489.htm; Also see a local government's comparative analysis on the two laws, Jindong government, 2019. https://www.jindong.gov.cn/art/2019/12/18/art_1229701412_59318349.html

The industry is considered still immature in China according to literature and public reports (e.g. Liu, 2017; People's Bank of China, 2020)[45].

4.2.2.2 Private credit reporting: Alibaba's rise and fall

The early public credit referencing provided by the Central Bank has its main data collected, and the service provided, in the banking sector for financial purposes. In contrast, private credit referencing collected a wider range of transaction data on individuals and enterprises. They also provided credit reference services for purposes outside of finance.

China's private credit reporting industry at first was largely limited to giving scores to corporations.[46] Many private credit-reporting agents emerged after the mid-2000s. By 2022, there were 136 credit-reporting companies registered with the Central Bank that ran credit reporting for corporations.[47] Meanwhile only two firms were licenced to run credit reporting services for individuals, namely Baihang Credit and Pudao Credit, by 2024 March.[48]

Between 2015-2018, however, the Central Bank experimented with a failed expansion of private credit reporting for individual consumers. This experiment started when the Central Bank in 2015 granted eight companies provisional licences, including Alibaba's Sesame Credit. But in 2018 the Central Bank did not continue their licences[49].

(1) The rise of Alibaba and Sesame Credit

Alibaba's rise in the credit-reporting business was the most prominent in 2015. But such success was paved way by its earlier businesses in credit-reporting and ecommerce. Using its own individual and corporate financial-commercial databases, Alibaba started credit-reporting business as early as the 2000s. Its 'Honesty Pass' (*Chengxintong,* 诚信通) was launched in

[45] E.g. Pan Shenggong, vice president of the Central Bank in 2020 claimed that the problems include overrated bonds and the lack of capacities to warn of systemic risks. People's Bank of China, Press Release, 2020.12.13, Online in Chinese, at http://www.pbc.gov.cn/goutongjiaoliu/113456/113469/4142965/index.html

[46] People's Bank of China, Report on China's Credit Reporting Industry, 2013 (online in Chinese) [中国人民银行，中国征信业发展报告]

[47] People's Bank of China, The number of corporate credit reporting agencies across the country, 2022. Online in Chinese at http://www.pbc.gov.cn/zhengxinguanliju/128332/128352/2875623/index.html

[48] People's Bank of China, The approval of the personal credit reporting business licence of Pudao Credit Information Co., Ltd., 2020. 12.25 http://www.pbc.gov.cn/goutongjiaoliu/113456/113469/4153264/index.html ； People's Bank of China, 2024.1.30 http://www.pbc.gov.cn/zhengwugongkai/4081330/4081344/4081407/4081702/4081770/4081803/4089644/index.html Also see Bian, W., 'Interview with Zhao Yi Han Chairman of Pudao Credit: Complementing each other and developing in conjunction with the Central Bank Credit Information and Baihang Credit, finance.sina.com, 2021.6.7. Online in Chinese, at https://finance.sina.cn/2021-06-07/detail-ikqciyzi8197100.d.html.

[49] See, Zhang Y., 'Wan Cunzhi: The implications of eight market institutions divesting their personal credit reporting business', Caixin.com, 2018.5.27. Online in Chinese, available at https://finance.caixin.com/2018-05-27/101257470.html

2002 as a credit-reporting system for private enterprises of medium-small size. This 'Pass' looks at certifications, consumer feedbacks and transaction records within Alibaba's business network (mostly domestic B2B transactions on 1688.com[50]).

This Honest Pass later prospered with Alibaba's dominance in the B2B market and therefore, access to corporate commercial data.[51] In 2011, the Honest Pass credit database claimed a million paid subscribers. It not only allowed not only businesses to reduce information asymmetry and transaction costs, but helped Alibaba optimise operations on credit-based transactions and expand into core financial services.[52]

This system further developed into Alibaba's 'Honesty System for Enterprises' (企业诚信体系), a much less known but comprehensive credit-reporting system launched in 2016. It was constituted of four major parts: digital code, credit rating, credit report and risk cloud map, with the previous Honesty Pass now a central node. This new system reported on both commercial and financial data of the small businesses.[53]

Compared with the often-overlooked Honesty System for Enterprises, the better known financial-commercial SCS pilot is Sesame Credit, run by the Ant Financial Service Group (AFSG，蚂蚁金服). With its seemingly high-tech algorithms and state of the art user interfaces, it became the most high-profile pilot among the eight credit-reporting platforms selected by the Central Bank in 2015. Once enrolled, a customer will be automatically given a credit score between 350 and 950 points. The higher the score, the better credibility one has, according to the ML / DL empowered, automated scoring.

Much of the literature has identified Sesame Credit as a loyalty program that provides benefits for those enrolled in the Sesame Credit app and Alibaba's related services, and then zeroed in how the platform rates people's credit scores. Seemingly, the Sesame Credit score was similar to the Fair Isaac Corporation (FICO) score in the U.S. in that it only collected financial data. However, Sesame Credit actually goes far beyond financial issues by evaluating a person's consumption behaviour, social circles, scores on Sesame Credit networks, and loan

[50] Chengxin Tong official website home page. [online] Available at: https://cxt.1688.com/
[51] *Ibid.*
[52] Liu Xinhai (刘新海), Zhengxin yu Da Shuju (征信与大数据) [CREDIT INFORMATION SERVICE AND BIG DATA](2016) at 307-8.
[53] See, e.g. 'Alibaba launched the credit system to build a future enterprise financing pass', 2016.6.16. [online] finance.sina.com.cn. Available at: http://finance.sina.com.cn/roll/2016-06-16/doc-ifxtfmrp2171126.shtml

repayments through Alibaba's 'Just Spend' platform (*Huabei*, 花呗, a consumer loan platform) (e.g. Ahmed, 2017, 2018; Kostka 2018).

By 2018, Sesame Credit claimed to provide benefits to high-score users on visa application (such as getting a fast track for Singapore visa if a person was over 700 points in credit), dating websites, hospital service, non-deposits for hiring cars, riding sharing-bikes and booking hotels, and even exhibiting on-time updates of the Supreme People's Court's Debtors Blacklist ('SPC debtors list').[54]

All these benefits were in principle made available only for those who used Sesame Credit voluntarily. The results of one's high or low credit score therefore cannot be used in other platforms or settings, such as local governments' public scoring systems explained in 4.3, which are mostly also opt-in systems. Neither can the Sesame Scores influence the Blacklists enforced by the Supreme Court and other regulatory agents, which are compulsory but can only collect individual's law compliance data from judicial or administrative rulings. The consumer data collected on Sesame Credit mostly came from Alibaba's own platforms including Taobao (C2C platform), Tmall (B2C platform), and Alipay (payment platform).

The most important channel that collects consumer data is Alipay (支付宝), the third-party online payment platform also affiliated to AFSG of Alibaba and developed since 2004. Alipay started as a unified payment channel for Alibaba's e-commerce platforms (including B2B, B2C and C2C), a Chinese version of PayPal. But later it tapped into financial businesses by utilising its dominance in commercial markets. In addition to the integrated consumer loan service 'Just Spend' (which operates like a virtual online credit card[55]), Alipay is directly linked to Yu Ebao, a funds management service run by Tianhong Fund.[56] Alipay made it possible for data from Alibaba's financial and commercial operations to be used together, which provided all Alibaba services, including Sesame Credit, an advantage in providing both financial and commercial services.[57]

[54] See e.g. Sesame Credit website: https://www.zmxy.com.cn/ ; and the corporate Sesame Credit scores website, http://www.xin.xin/#/detail/4-1-0

[55] See, e.g. Alipay's official website's explanation of *Huabei*, online in Chinese, at https://help.alipay.com/lab/help_detail.htm?help_id=513629 .

[56] For a brief introduction on Yu Ebao. *See,* on Alipay's official site. https://cshall.alipay.com/lab/help_detail.htm?help_id=257914&keyword=%D3%E0%B6%EE%B1%A6&sToken=s-68296f3a0f7548e383cb8a2ee8e08e53&from=search&flag=0

[57] See, e.g. Zhong and Peng, (2014), 'The inspiration of Ali's credit reporting model on the construction of my country's financial credit reporting system', *Credit Reference,* 2.

The Central Bank later recognised that in addition to using financial data, Alipay and its related services could make use of the commercial data they collected. Such commercial data provides significant information which could be utilised by private data aggregators outside of the banking sector. The Central Bank, for example, claimed that 'these commercial data aggregators could also be guided to provide credit information.' (*2015 The People's Bank of China Instruction on Individual Credit Information Reporting*).[58]

(2) The (partial) fall of Sesame Credit

In 2017, however, the Central Bank publicly stated some differences between the Social Credit Systems (SCS) and the Credit Referencing Systems (CRS). The CRS is a broad concept used by the Chinese government to refer to systems including both traditional public and private credit-rating, credit-reporting (including the Central Bank's FCIFD), and the more recent credit scoring using alternative data and ML / DL algorithms like Sesame Credit.

The Central Bank's vice president Chen Yulu and the head of the Credit-Reference Centre Wan Cunzhi gave a few important public talks on this. For example, Wan emphasised that the CRS players (including big-techs) need to be clear of their positions in the SCS:

> 'I start with differences (between the SCS and CRS). First, the SCS requires the people in society to be credible on a moral level, which demands cultural promotion and education.
>
> Second, the SCS entails that the governments, enterprises and individuals be credible by internal credit / reputation managements…to promote good faith and curb malice of incredibility.
>
> The third layer of the SCS is information sharing, in two ways—the first way is directly sharing by open government information (OGI), and *the second is sharing by CRS* (emphasis added by author).

[58] The People's Bank of China, *Instruction on Individual Credit Information Reporting,* 2015. Online in Chinese, available at https://www.gov.cn/xinwen/2015-01/05/content_2800381.htm.

Therefore, CRS is only a constituting institution as part of the information-sharing mechanism (of the SCS), but not all of it. Consequently, some CRS institutions involved in the SCS must be clear of their positioning in the project.'[59]

In 2018, Sesame Credit as the most prominent player in the private CRS, lost some momentum. The Central Bank did not extend Sesame Credit's licence that year. Instead, the Central Bank gave only one licence of consumer credit reporting to a newly established public-private joint venture, namely Baihang Credit.

The Central Bank argued that the eight pilots including Sesame Credit since 2015 had failed due to (1) conflicts of interest and foreclosure of data, and (2) antitrust concerns, among others. Sesame Credit, for example, was viewed as providing unfair advantages for Alibaba's e-commerce business, creating its own data silos, being biased rather than an independent scoring agent in defining social credibility, and so on.[60]

From that point, Baihang Credit started collecting individuals' financial credit information as the only officially recognised system complementing the PBCCRC. Each of the eight pilot companies, including Alibaba (based in Hangzhou) and Tencent (based in Shenzhen), holds 8% of the shares, and the Central Bank has de facto control over the company with 36% of the shares through an industry association of internet finance (中国互联网金融协会) (Xue, 2018).[61]

Nevertheless, Baihang Credit only collects consumer's financial data. It does not collect the alternative or transactional data of the kind which provides business opportunities for the digital giants like Alibaba with their various platforms and data sets.[62] Sesame Credit still functions on Alipay and provides credit scores for consumers, but the data they can collect and process remains limited by the Central Bank and the use of this score is largely limited in financial-commercial settings instead of the social.

[59] Wan Cunzhi, 2017, Accelerating the construction and improvement of a Credit Referencing system that covers the whole society, (in Chinese) [万存知：加快推进建立健全覆盖全社会的征信系统] https://huzcredit.huzhou.gov.cn/art/2017/11/5/art_1229635293_4583.html.

[60] Chen, G., 'None of the 8 personal credit reporting agencies has qualified so far. The development of credit reporting agencies must strengthen the protection of personal information', www.ce.cn, 2017.4.27 [online] Available at: http://www.ce.cn/xwzx/gnsz/gdxw/201704/27/t20170427_22415553.shtml.

[61] People's Bank of China Credit Referencing Licence, [2018] No.1. Also see Xue H., 2018, 'The model boundaries and missions of Baihang Credit', Financial Economics, (4):3. [薛洪言, 2018, 百行征信的模式边界与使命,《金融经济》].

[62] *Id.*

Baihang was soon followed by a second credit-reporting agency approved by the Central Bank in December 2020, namely Pudao Credit. While Baihang is based in Shenzhen, Pudao is based in Beijing. Pudao's main shareholders include JD.com and a few other tech companies based in northern China. JD.com for example is based in Beijing and a major e-commerce competitor with Hangzhou-based Alibaba,[63].

Alibaba's individual credit-reporting business, nevertheless, is still trying to rejoin the game. According to a Central Bank's official press conference in January 2024, the Bank is again reviewing Qiantang Credit's licence application for individual credit reporting, since its first application in late 2021[64]. But this time they have the help of the Zhejiang Government (provincial level) and Hangzhou Municipality, and the aim is to establish itself as the main competitor to Baihang and Pudao. The shareholding structure of Qiantang Credit reveals Ant Financial Group (of Alibaba Group) has 35% holdings, and that the corporation will be based in Hangzhou.[65]

If successful, together these three credit-reporting companies, Baihang, Pudao, and Qiantang, will covers the three most economically developed regions of China, namely the Pearl River Delta (that is the Guangdong-Hong Kong-Macau Greater Bay Area), Beijing area (that is the Beijing, Tianjin and Hebei region), and the Yangtze Delta (Shanghai and Hangzhou region). Their businesses are, nevertheless, still limited to providing financial scoring services which does not extend to wider social settings.

(3) Comparison between Sesame Credit and FICO Score

To provide a more detailed description of Alibaba's Sesame Credit, a comparison with its US counterparts is made in this part. It also uncovers Sesame Credit's institutional and technical origins.

The Chief Technology Officer (CTO) of Sesame Credit claimed in a news report that the company learned from FICO Score in its algorithmic design. Sesame Credit actually invited the CTO of FICO and Experian experts into the data team for product designs (Hua, 2015).

[63] People's Bank of China, The approval of the personal credit reporting business licence of Pudao Credit Information Co., Ltd., 2020. 12.25 http://www.pbc.gov.cn/goutongjiaoliu/113456/113469/4153264/index.html ; Also see Bian, W., 'Interview with Zhao Yi Han Chairman of Pudao Credit: Complementing each other and developing in conjunction with the Central Bank Credit Information and Baihang Credit, finance.sina.com, 2021.6.7. Online in Chinese, at https://finance.sina.cn/2021-06-07/detail-ikqciyzi8197100.d.html.

[64] State council press release, 2024.1.25. https://www.gov.cn/lianbo/fabu/202401/content_6928397.htm

[65] People's Bank of China, 2021.11.26. http://www.pbc.gov.cn/rmyh/105208/4401151/index.html

FICO is short for Fair Isaac Corporation (NYSE: FICO[66]). It is a data analytics company that created the FICO scores as a measurement of credit risks in 1956. Since then, the FICO score has gradually become the most widely used score in the US by lenders in their daily lending decisions to consumers and businesses.[67] The FICO score operates by way of licences with major credit-reporting agencies, including the 'big three' in the U.S, namely Experian, Transunion and Equifax (Citron and Pasquale, 2014). Each of these agencies collects their own financial data from commercial banks and calculate the FICO scores. For example, Experian (LSE: EXPN[68]) is a multinational data analytics company headquartered in Dublin. It provides world-wide credit-reporting services to consumers and businesses directly. In the US, for example, Experian provides the FICO score, while in the UK it provides an Experian score calculated by its own algorithms[69].

The scoring factors of Sesame Credit and FICO are not fully revealed because the full details of their measurement metrics and algorithms are proprietary[70]. However, it is known that they both rely on five 'pillars' of information relating to consumers when developing the scores. The five pillars are shown in *figure 5* below, which compares the FICO score with Sesame Credit:

1. FICO Score considers Payment history (35%), Amounts owed (30%), Length of credit history (15%), Credit mix (10%), and New credit (10%);[71]

2. Sesame Credit's scoring considers the Payment history and performance capacity similarly to FICO score, but also very differently, it takes account of consumption behaviour and preference, personal characteristics, and social network. It was reported that the Sesame scores are calculated through 'techniques of cloud computing and machine learning etc., with logistic regression, decision tree, and random forest algorithm' (Hua. 2015).

[66] See FICO on the NYSE, https://www.nyse.com/quote/XNYS:FICO
[67] See FICO official website, https://www.fico.com/
[68] See Experian on the London Stock Exchange, https://www.londonstockexchange.com/stock/EXPN/issuername
[69] See Experian official website, https://www.experianplc.com/
[70] See criticism on the opacity of automated scoring. Citron and Pasquale. (2014). .
[71] FICO official website explanation, https://www.myfico.com/credit-education/whats-in-your-credit-score

Factors	FICO Score	Sesame (Credit) Score
1	Payment history (35%) Whether you have paid past credit accounts on time. This helps a lender figure out the amount of risk it will take on when extending credit. This is one of the most important factors.	Payment history: Whether you have paid past credit accounts on time through Alipay or Ant Financial services like *Huabei* (like an online credit card giving small consumer loans).
2	Amounts owed (30%) The available amount of credit left on the users' account. If not much is left, this may indicate that you are overextended, and banks can interpret this to mean that you are at a higher risk of defaulting.	Performance capacity: The available amount on users' Alipay account. (similar to FICO) Moreover, also on the stability of past performance on cars / house rentals, utilities payment, and consumptions.
3	Length of credit history (15%) A longer credit history will increase your FICO Scores: 1) how long your credit accounts have been established, including the age of your oldest account, the age of your newest account and an average age of all your accounts; 2) how long specific credit accounts have been established; and 3) how long it has been since you used certain accounts.	Consumption behaviour and preference: The extent to which an account is actively used, how many purchases are done, the stability of these consumptions and transactions etc.
4	Credit mix (10%) The mix of credit cards, retail accounts, instalment loans, finance company accounts and mortgage loans.	Personal Characteristics: Basic public information including the educational background, real-name registrations etc. (shared by the courts, Public Security Department, and civil bureaus etc.)
5	New credit (10%) Research shows that opening several credit accounts in a short period of time represents a greater risk-especially for people who don't have a long credit history.	Social Network: how many online Sesame/Alipay contacts/friends one has, one's influential power in contact list, interactions between user and friends, etc.

Figure 5. Comparison between the FICO score and Sesame Credit Score

4.2.3 Credit reporting in the U.S. and Europe compared to China

In addition to comparing Sesame Credit with FICO, a brief comparison of China with American and European practices also reveals certain similarities in (semi)automated scoring systems. It reveals that this approach to credit scoring is by no means an isolated Chinese phenomenon. The Central Bank and Sesame Credit were learning from more advanced economies of the time, including the USA and Europe, in terms of both institutions and technological architectures.

Compared with China, however, in the US the financial and economic logic of the business of credit-reporting is more evident. In the USA, only a very limited part of an individual's non-financial (or commercial) data can be used for credit-reporting. These so-called "alternative data" include telecom, utilities, rent, and phone payments. These alternative data can be voluntarily uploaded by consumers to build positive credit profiles in the national credit bureau, Payment Reporting Builds Credit, Inc., which was established in 2002. The Big Four US credit reporting agencies (CRAs)—Experian, Equifax, TransUnion and Innovis – collect individual financial data under Metro 2 standard of the Consumer Data Industry Association, with the FICO score pervasively used in everyday scenarios (Avery et al. 2003).[72]

Besides CRAs for individual credit scores, there are also major CRAs in America for business credit reporting such as Dun & Bradstreet (introduced in Shanghai, China since 2017), Experian Business (introduced in Beijing, China since 2018)[73], Cortera, Equifax Commercial, and FICO Small Business Scoring Service (Miller 2003). These CRAs collect both financial and non-financial data of enterprises and provide scores on businesses for financial institutions to take into account when making lending decisions. There is also a not-for-profit trade association, the Small Business Financial Exchange, Inc., that is not a CRA but gathers and protects small business payment data for its Members, and then facilitates the exchange of that data to specific business CRAs.[74]

All the CRAs mentioned above in the USA are regulated under the FCRA (Fair Credit Reporting Act), Fair and Accurate Credit Transactions Act (FACTA), and Fair Credit Billing

[72] Also see FDIC report on PRBC. https://www.fdic.gov/regulations/laws/federal/2006/06c27anprfacta.pdf ; CDIA report specifications https://www.cdiaonline.org/resources/furnishers-of-data-overview/metro2-information/

[73] People's Bank of China, The 'government + market' two-wheel-driven development model of the credit reporting market has initially taken shape, 2018 (online in Chinese), at http://www.pbc.gov.cn/zhengxinguanliju/128332/128352/128402/3570533/index.html. In addition, since 2016, the Ministry of Commerce has granted equal treatment to foreign business CRAs. http://www.pbc.gov.cn/zhengxinguanliju/128332/128338/128371/3267499/index.html

[74] https://www.sbfe.org/ See also, Equifax Designated Small Business Financial Exchange Certified Vendor™ https://investor.equifax.com/news-and-events/news/2016/10-18-2016

Act (FCBA). The FTC and the Office of the Comptroller of the Currency (OCC) are two major regulatory agencies overseeing these credit-reporting agencies.[75]

In Europe, credit-reporting practices are also highly developed. Against the backdrop of the adoption of the EU's GDPR and the Equifax data breach incident of 2017, the Association of Consumer Credit Information Suppliers[76] held its 2018 annual conference in Stockholm with its 40 members and experts from the EU commission and German Central Bank, among others, taking part (Tian and Yang 2018). They arrived at a consensus that the legal basis for processing consumer credit information should be the "legitimate interest" under the GDPR. Experts from the German company SCHUFA reported that in the 2017 amendment to Germany's Federal Data Protection Act, the "legitimate interest" was explicitly adopted as the legal basis for data-processing by CRAs.[77]

As new technologies using big-data and AI are emerging, product differentiation is advancing in the CRA sector. As business lines, data-analysis and technological support are becoming more salient. They accounted for almost 30% of Experian's revenue and are on an increasing trend, while its credit-reporting now accounts for only 32.5% of the revenue. 20% of revenue is accounted for by consumer services and 20% by marketing services. It is becoming clear that these data-reporting agencies are increasingly exploiting their comparative advantage in data-analysis and software / algorithmic development to make and market various data-analytical tools, anti-fraud instruments, and credit-management decision-making algorithms (Tian and Yang, 2018).

These business line expansions require the amount of data collected to expand. As the EU PSD 2 and the UK Open Banking policy went into effect on 2018.1.13, large commercial banks have to give away consumers' bank account data to registered third parties[78]. These third parties include Account Information Service Providers (AISP), and Payment Initiation Service Providers (PISP). Experian has already registered as an AISP . Once a credit-reporting company becomes an AISP, they can, with consumers' consent, integrate consumers' payment

[75] See FCRA (Fair Credit Reporting Act), Fair and Accurate Credit Transactions Act (FACTA), and Fair Credit Billing Act (FCBA).

[76] Official website at https://accis.eu/.

[77] Bundesdatenschutzgesetz (BDSG) (geändert durch das Gesetz vom 30. Oktober 2017) [联邦数据保护法],[in Chinese] http://www.wipo.int/wipolex/zh/text.jsp?file_id=462255;[German Version] https://iapp.org/media/pdf/resource_center/Eng-trans-Germany-DPL.pdf.

[78] Revised Payment Service Directive, passed by Council of EU in November 2015 https://ec.europa.eu/info/law/payment-services-psd-2-directive-eu-2015-2366_en ; https://www.evry.com/en/news/articles/psd2-the-directive-that-will-change-banking-as-we-know-it/ By September 2019 all companies within EU and UK must comply with the revised national laws and regulations. https://www.computerworlduk.com/applications/is-2018-year-that-open-banking-becomes-reality-in-uk-3653824/ ; https://www.wired.co.uk/article/open-banking-cma-psd2-explained ; https://www.wired.co.uk/article/psd2-future-of-banking

account information, such as account balance, last 90 days' transactions, regular payments inflows and outflows, bills and subscriptions, and product holdings, not only to provide them with tailored financial products and potentially money-saving opportunities, but also to develop additional services such as external search and data-analysis. Corporate customers, especially SMEs, are also ready buyers of these services, as they are usually multi-banked and with a need for real-time account information and financial visibility. Experts from Experian claim that they have already combined these new payment account information with their readily available credit information, to provide new decision support products for lending institutions in terms of income verification and evaluating borrower's credit capacity (Tian and Yang, 2018).

In addition, Sweden is building its own credit-reporting system covering the entire society. UC AB, Sweden's largest credit-reporting agency (acquired by a Finish competitor Finland's Asiakastieto in 2018), runs the consumer credit information system for all Swedish citizens above the age 16, and all companies registered in Sweden. It collects public information and credit information for individuals and enterprises, using them to provide services ranging from consumer analysis and recommendation, credit report, credit rating, tracking and predictive warning, asset evaluation, marketing etc. It has also provided ID protection for individuals against the backdrop of increasingly common identity frauds.[79]

The Swedish system is able cover the entire society as a result of certain features of the local institutional environment: first according to the Constitution, most governmental information is published, and any individual or institutions can search other's tax information for example; and second, all such public information is integrated into the credit-reporting system, including registration for individuals and enterprises, tax, income, real-estate, court and enforcement information, economic and statistical departments' information, licence and verification, and environmental-protection sanctions and customs. Most of this public information is updated on a daily basis. Combined with the credit information collected by the company, the credit-reporting system covers nearly all citizens (aged above 16) and corporations in the country (Tian and Yang, 2018).

[79] See UC's official website, available at UC in English - UC

4.2.4 Assessment: scaling and layering in commercial scoring systems

4.2.4.1 Nature of the commercial scoring systems: data-driven socio-technical systems with financial origins from the US and Europe

Evidence shows that China has been learning from the U.S. and European financial-commercial CRAs since the 2000s. This is a process where marketisation in China led to more financial services, and the Central Bank decided to let go of its dominance in the credit-reporting business (both individual and enterprise scores). This market-financial reform is a continuation of the Central Bank's financial infrastructure reforms in the 1980s.

The financial and data-driven logic behind the trajectory of the financial-commercial SCS scoring systems closely resembles that of the US and EU. Both public and private databases and analytics services were set up to cover almost nation-wide credit risk information for the provision of a modern market economy in China.

This trajectory of market reform reveals the nature of the SCS commercial scoring systems as data-driven socio-technical systems that are essentially market information infrastructures and institutions which facilitate large-scale, cross-geographical transactions. They are algorithmic governance systems deeply rooted in how advanced economies have used massive financial datasets, metrics and algorithms to govern globalised market economies since the 1950s.

A globally convergent trend is also observed, where credit reporting agencies are moving away from traditional credit services towards providing more algorithmic data analytics services based on AI. This transition of business models relies on large tech firms' competitive advantage of their huge amount of consumer and transaction data.

The comparative analysis between FICO and Sesame credit, nevertheless, shows a number of differences in the input and output processes. On inputs, while the FICO Score only takes into account financial data collected from commercial banks, the latter three factors (per the table at the end of section 4.2.3) in the Sesame Credit algorithm consider consumption and social status information, and weigh it more heavily than the first two, financial, factors.

On outputs, FICO Score affects a person's ability to get bank loans and credit cards, rent apartments and get employed, while Sesame Credit has a much limited influence outside of finance. The most direct impact of Sesame Credit is still in consumer loans and the small and medium sized businesses doing e-commerce on Alibaba's platforms. The benefits are mostly

linked to Alibaba's services such as loyalty programs, meanwhile private bank's lending decisions are still mostly based on the Central Bank's financial database (FCIFD).

4.2.4.2 Scaling: expansion of both public and private commercial scoring

Scaling effects are observed in the development of the SCS commercial scoring systems. Both the public and private sides of the SCS commercial scoring systems provide ongoing evidence that demonstrate the scaling effects of statistical and algorithmic governance. Such scaling is not only within the national boundaries of China's national geography and economy. It is also part of the wider scaling towards global governance led by the US and European financial practices. China's national scaling of governance by transactional data and code is in competition and complementarity with the US-European governance practices, namely expanding the influence of quantified financial-commercial credits as a mode of global governance.

As evidence shows, China's transformation towards marketisation, financialisation, quantification, datafication, and ultimately, algorithmic governance with relations to data and AI, is intertwined with the development of the wider SCS. The SCS provided the major legal and policy shift since the mid-2000s, starting from the financial-commercial sectors.

This shift included the *2005 Tasks* and *2007 Instructions* entailing credit information provisions nation-wide, which involved the expansion of public and private data sets construction, measurement and more sophisticated scoring. The SCS commercial scoring systems took off in the 2010s, and culminated in the *2013 Credit Reference Regulation* and other national policies including the *2014-2020 SCS Blueprint*.

The experiments following these laws and policies were highly ambitious. The Central Bank's FCIFD is the most prominent case of making a modern and constantly updated financial public data infrastructure for citizens and enterprises of the entire country. The comparative analysis shows how the US and Europe have had similar systems of governance by numbers and code since the 1950s, such as the FICO scores and the CRAs, the Swedish UC AB and the German SCHUFA score. They scaled up with legal and socio-policy backings, for example the disclosure requirements under Swedish public law.

In the private sector the most prominent case was the expansion of Alibaba's two platforms of credit-reporting services (consumers and enterprises) towards covering additional social governance settings, most clearly in the case of Sesame Credit since 2015.

This is accompanied by the later establishment of public-private joint-ventures of Baihang Credit in 2018, Pudao Credit in 2020, and more recently Qiantang Credit in 2024 (currently awaiting for a Central Bank licence). Together these three large consumer credit-reporting corporations cover the three most economically advanced regions, namely the *Jing Jin Ji* region (Beijng - Tianji - Hebei. 京津冀), the Yangtze River Delta (Shanghai - Hangzhou - Nanjing), and the Pearl River Delta (Shenzhen - Hong Kong - Guangzhou). They showcase the ongoing scaling of governance by financial and transactional data and code.

4.2.4.3 Layering: limitations of (semi)automated scoring systems

In addition to scaling, the SCS commercial scoring systems also provide evidence on layering. The case reveals how governance by numbers and code is restricted by local geographies and economies, and evolving local social and legal governance.

(1) Central Bank databases as socio-legal foundations for algorithmic scoring

The Central Bank's FCIFD, for example, evolved over the years to cover more data and data subjects in order to provide financial information across China, but refrained from providing any credit score to measure financial risks at the national level. Even though strictly limited to financial credit reporting, the Central Bank's practices demonstrated a limitation of quantification when the scale of governance is so large.

The FCIFD as a Central Bank database that aggregates comprehensive information of citizens and organisations, can be argued as essentially a form of social and legal governance based on public institutions. The database is in essence a storage of socio-legal information related to financial and commercial transactions, compliance, fulfilment of contracts and obligations, and for citizens in particular, information including education, marriage and so on.

These state-backed institutions serve as a centralising mechanism that sustains the operations of commercial banks that generate risk scores by computers automatically and / or by banking professionals manually. This demonstrates how governance by numbers and code, some involving AI, are essentially layered, or path-dependent on social and legal governance. Law in this sense can only be quantified and computed by code to the extent that legal governance itself keeps operational and generates new data.

In fact, the Central Bank's continuous aggregation, selection and provision of qualitative information of social practices, are also the foundations of the private credit-reporting and

rating sectors. The Central Bank not only regulates the scoring industry, in the sense of providing legislation and enforcing licences such as the 2015-2018 consumer credit-reporting licence to Sesame Credit. The Central Bank also serves as a crucial recalibrator of China's local measurement metrics, namely the methods by which data are generated. When global scoring systems failed to reflect true values of bonds and securities in the 2008 global financial crisis, the Central Bank's own qualitative databases served as infrastructures and socio-legal institutions that provided the basis for recalibration.

Credit scores in the 2008 financial crisis is an example of, to some extent, how the more formal and abstracted governance modes like numbers and code have limitations when scaling up, in particular to a greater national and global scale. In face of the failure of algorithmic governance, local modes of social and legal governance like the Central Bank's FCIFD provided certain resilience in the banking industry, as consumers and commercial banks had to make their own local risk assessments instead of relying on any global / national scores. In a way, this can be seen as an example of how statistical and algorithmic governance reached their capacity limits when 'coding' global capital and assets, and how national social and legal governance remains the foundations which global algorithmic governance is layered upon.

(2) Private algorithmic scoring layered on social and legal institutions

The case of Sesame Credit further evidences the layering of algorithmic scoring systems on social and legal governance. First, despite early expansion in the 2010s, Sesame Credit using new ML / DL-driven scores ended up without a Central Bank's licence for consumer credit-reporting in 2018. Alternative data, such as citizens' social networks and consumption habits in Sesame Credit's case, is still harnessed to improve the algorithms that predict the compliance rates of consumer loan payments and so on. But these algorithmic processing have to operate with greater cautions under stricter data protection standards, and sometimes in fact grey policy areas. The production of these algorithmic governance is therefore limited to data generated based on the local socio-legal institutions.

A second evidence is the Central Bank's concern over Sesame Credit's effectiveness and independence when applied to contexts outside of Alibaba's e-commerce ecosystem. This highlights that algorithmic governance is layered on certain specific geographies, populations, and commercial-social practices. Applying the algorithms trained on data generated mostly from Alibaba consumers and businesses, does not necessarily work well across regions with different social, economic and legal realities. Local and national legal systems in this sense

provide the basic cognitive structures on which algorithmic governance is ultimately judged. The highly abstracted code systems thus can only acquire cognitive meaning and legitimacy such as 'independence' and 'fairness' by relying on social and legal concepts and the modes of governance operating on these normative concepts in specific places and settings. Sesame Credit faced functional setbacks because its algorithmic governance tried to cut across geographies and populations, and ventured into social governance settings outside of Alibaba's e-commerce governance.

Third, the Central Bank also raised antitrust concerns where each big-tech company's scoring systems were used as a way to preserve their own platform's market dominance, and / or engage in regional market foreclosure by collusion with local governments. Algorithmic governance in this sense is layered on top of local social and legal governance that segregate and / or harmonise regional markets. The three post-2018 consumer credit reporting companies, namely Baihang Credit, Pudao Credit, and Qiantang Credit, all have rather diversified shareholding structures that are required by the Central Bank or local governments, in order to prevent certain big-tech companies from gaining regional dominance. Layering thus exerts certain real spatial limits to the abstracted statistical and algorithmic governance at scales, in the form of divergent space-sensitive social and legal institutions.

4.2.5 Summary

The section provides new empirical evidence to show scaling and layering effects in China's SCS scoring systems. These systems are revealed to be data-driven socio-technical systems rooted in American and European financial credit-reporting practices, which are statistical and algorithmic modes of governance at the global scale. Evidence shows how governance was scaled up, in different ways, by the Central Bank and big-tech companies like Alibaba in making and using data and code. However, layering was also exhibited when scaling reached certain limits, and negative complementarities started to restrict the expansion of statistical and algorithmic governance and required local socio-legal institutions to recalibrate the metrics and code.

It is revealed in the next section that similar scaling and layering effects manifested in another branch of the SCS scoring systems, namely local governmental scores.

4.3 Local (semi)automated scoring systems: Urban and rural China

After examining the financial-commercial scoring systems, this section turns to another major scoring system within the SCS, that is local public scores. These public scores are based on local Public Credit Information Systems (公共信用信息系统), which operate in specific geographical / administrative regions. The experiments by local governments are nebulous, but can be broadly carved into two subsystems: (a) the urban scores and (b) the rural scores . They in general do not involve advanced AI technologies or computation, but rely on simple calculations and algorithms.

These public scores are mostly assigned to citizens and enterprises (and other organisations) by local governments in urban or rural municipalities. An exception is the scores assigned to local municipal governments themselves, usually by higher-level provincial governments, in order to formally measure and control the public performances of these rather informal, decentralised public organisations.

This section provides preliminary, new evidence on the urban and rural local scores at municipal (4.3.1) and village level (4.3.2). Scaling and layering effects were uncovered and assessed in these different practices of local public scoring. A sectional summary follows (4.3.4).

4.3.1 Scoring in urban China

By 2024 March, the author is aware of no functioning national comprehensive scores for individuals or corporations run by the central government. In general there are also no quantified scores at the provincial level.

Chinese provinces mostly have about the population of a European country, i.e. 25 to 60 million people, and some have a similar geographical span to the UK at about 0.2 million km² (e.g. Hubei or Hunan province[80]). For example, the city of Shanghai, which has a province-level status, has a population of 26 million (2019). It does not have a quantified public scoring system. Instead it has three qualitative public credit information databases under its local SCS, which do not use quantified data or metrics[81].

[80] E.g. population statistics from the National Bureau of Statistics of China, https://data.stats.gov.cn/english/easyquery.htm?cn=E0103.

[81] See e.g. Shanghai's Social Credit website. https://credit.fgw.sh.gov.cn/

Out of China's 34 provinces, only three have quantified public scoring systems: Hainan, Tianjin and Zhejiang.

(i) Hainan, an island province with a 9 million population, has a quantified scoring system named Jinye Score for individual citizens (it can be translated as 'golden coconut score', a naming which promotes Hainan island's main product).[82]

(ii) Tianjin, a city with provincial status and a population of 15 million, has had a quantified city score since 2022, the Haihe score (named after the river running across the city). The score produced can be transferred to other 'friendship cities' and used in low-stake settings.[83]

(iii) Zhejiang province is a large urban province with a high density of commercial activities. It has a population of 64.5 million (2020)[84] and is about half the geographical area of the UK (about 0.1 million km^2). [85] In principle, it has a quantified scoring system for individuals and enterprises.[86] However, there is very little evidence available on the actual implementation of Zhejiang's scoring system at the provincial level.

Quantified public credit scores, either for citizens or corporations, are mostly established at medium-sized municipal levels.

(1) The citizen scores are in general used only in low stake private and public decision-making settings for accessing small benefits, such as travel discounts in touristic places, booking privileges for gyms and museums, zero-deposits for using city bikes or lower deposits for car hires and apartment renting, higher numbers of books a citizen can borrow from public libraries, and so on.

In 2018 Hangzhou, for example, the city which is the birthplace of Alibaba and with a population of about 12 million (2023)[87], started a joint-venture with a local company to

[82] Hainan business environment department, '"Golden Coconut Points" expands the circle of friends and realises mutual recognition of personal credit across provinces', 2023, online in Chinese, available at https://www.ndrc.gov.cn/fggz/fgfg/dfxx/202304/t20230426_1354750.html

[83] Credit Tianjin, 'Tianjin's credit "circle of friends" expanded again and signed the "Strategic Cooperation Agreement on Credit Construction in Ten Cities"', 2022, online in Chinese, available at https://credit.fzgg.tj.gov.cn/detail.do?contentId=70da9420bdc04f45b13184d873e5db80&channelId=e1a5d102ea744938a58f84527fd53875

[84] National Bureau of Statistics, 'The 7th National Census Report No. 3', 2021.5.10 https://www.stats.gov.cn/zt_18555/zdtjgz/zgrkpc/dqcrkpc/ggl/202302/t20230215_1903999.html

[85] Zhejiang Government official statistics, in Chinese, available at https://www.zj.gov.cn/col/col1229631725/index.html

[86] Zhejiang Development and Reform Commission. 2019. Zhejiang Guidelines for Public Credit Evaluation of Five Types of Entities (2019) [浙江省五类主体公共信用评价指引] (2019). (in Chinese). Zhejiang Development and Reform Commission. 2020. Zhejiang Guidelines for Public Credit Evaluation of Five Types of Entities (2020) 浙江省五类主体公共信用评价指引 (2020). (in Chinese). Also see Zhejiang Development and Reform Committee, 2023 https://fzggw.zj.gov.cn/art/2023/2/10/art_1229602268_2458210.html.

[87] Hangzhou population statistics, 2023. https://www.hangzhou.gov.cn/art/2023/3/2/art_1229063404_4144634.html

implement a public scoring system for individuals, the Qianjiang Score. The score can be accessed on a mobile phone app, in the same way as Alibaba's Sesame Credit score for individuals. There are only broad indications on the five pillars of metrics, but no disclosure on specific algorithms they use, and the score's official website has not updated its news portals since April 2023.[88]

The Qianjiang score is only used for accessing small benefits in low-stake decision-making scenarios, for example, free parking and travel discounts in holidays in Hainan province, and zero-deposits for city bike hires. The score can be transferred to other city scores, including Hainan's city score, via inter-municipal agreements which cover a number of geographic regions, but again only for low-stake uses.[89]

(2) Corporate scores are usually applied in much higher-stake settings in comparison. For example, in deciding the regulatory agencies' frequency of inspection, and whether a corporation can win a bid in public procurement (Krause et al., 2023). Nevertheless, in October 2023 the National Development and Reform Committee (NDRC) issued an Instruction to ban local governments (operating at or below the provincial / prefectural level) from using credit scores as 'market barriers' in certain high-stake decision-making processes, such as public procurement. The NDRC required local governments to start abolishing some existing scoring practices or revising their procedural safeguards. This was done with the aim of preventing local governments from (ab)using their scoring power to segregate markets and create local monopolies.[90]

Since the NDRC's 2023 Instructions, there have been some revisions on local procurement scoring systems, and some retractions of scoring systems for both citizens and corporations in high-stake areas. For example, in Guangzhou, the capital city of Guangdong Province with a population of about 15 million, its 'city management' (*chengshi guanli*) department implemented a scoring system for both individual citizens and corporations in March 2021, but abolished it after February 2024. Unlike blacklist information, this city management score for citizens was in general not publicly displayed online, and access had to be formally requested by individuals or public agencies. Before being abandoned, the score was used in regulatory decision-making scenarios, such as frequencies of health and safety inspections.[91]

[88] Official websites for Qianjang Score, at http://qianjiangfen.cn/ ; http://qianjiangfen.cn/newsList.html

[89] Yang N., '34 cities have established a mutual recognition alliance, and credit points are used more and more widely', 2023. Online in Chinese, at http://qianjiangfen.cn/article.html?articleId=2202083.

[90] National Development and Reform Committee, *Notice on standardising the application of credit systems in public procurements,* [2023] No. 860. Online in Chinese, at https://www.ndrc.gov.cn/xwdt/tzgg/202311/t20231108_1361858.html.

[91] Guangzhou city management and enforcement department, [2020] No.9. Online in Chinese, available at

In public procurement settings, however, Guangzhou kept its procurement scoring system instead of abolishing it, and in fact updated the system into a 'Version 3.0' in September 2023. In the online official document the new version provided more specific rules on data categories, metrics and calculation. There is no indication of any ML / DL algorithms involved.[92]

The following part provides some field-based interview data on local public scores, focusing on Shanghai and Qingdao.

4.3.1.1 Interview data on urban scoring: Shanghai and Qingdao

A civil servant in Shanghai working on Shanghai's Public Credit Information System, suggested that the state's practices can be understood as building 'data habitats' when discussing the ways forward for Shanghai's use of public credit information:

> 'We need to build a habitat for data, like raising fish in a garden pond. It takes time for the fish to grow and then you can draw useful information from this data.'[93]

This idea of data habitat was raised when discussing the challenges in some of the risk-assessment practices in Shanghai. The civil servant said that the risk scores generated for businesses sometimes are not very useful for an experienced regulator when deciding how to allocate inspection resources.

> '[… The regulators] don't need our risk models that much because they already have a deep understanding of the places to safeguard the market order. Our public credit information is not that often useful for them'.[94]

The civil servant also mentioned that the depth of the data requires understanding of the data habitat, because data interpretation has different layers through time.

> 'You also need a habitat to understand the depth of the data. If someone has a criminal record, but because of over-defending himself or others in a violent crime, then maybe

https://www.gz.gov.cn/gfxwj/sbmgfxwj/gzscsglhzhzfj/content/post_7006788.html

[92] Guangzhou government, 2024, Public resource transactions and government procurement credit evaluation system version 3.0, https://czj.gz.gov.cn/attachment/7/7505/7505853/9206033.doc.

[93] Civil servant in Shanghai, 2021.3.

[94] *Id.*

he can be seen as brave. To understand him you need more data over a period of time and context.'[95]

To achieve deeper knowledge in the field and in its context, human expertise or tacit knowledge is also required to pass on informal knowledge alongside formalised institutional knowledge. A civil servant mentioned that:

'In market regulations, the local authorities usually are already quite familiar with the business situation because they have been there for years. They know who the major players are and how to regulate them, while smaller shops come and go (…)'[96]

Moving from Shanghai to Qingdao, a respondent who works at a state-affiliated credit-reporting company in Qingdao also described the rooms for improvement on their risk-prediction products in the longer term:

'Each commercial bank has its own way of collecting and accumulating data. They have their own risk models for lending decisions built after a long time, and often use our data analytics products as an auxiliary that often needs to be recalibrated according to their own understanding (…) They have been in the field for much longer, and know how to recalibrate polluted or distorted data. What we could do is to provide more comprehensive data models for them, and their own insights acquired through time can correct our models and predictions.'[97]

Data cycles and their path-dependence on the legal institutions was also discussed by this respondent as a state-affiliated corporate employee:

'The data on quarterly financial reports is not helpful because many small businesses don't even survive that long. The data disclosure cycle follows certain administrative practices established in the past. You need to change those then we can extract useful information.'[98]

[95] *Id.*
[96] *Id.*
[97] Corporate staff, Qingdao, 2021.2.
[98] *Id.*

4.3.1.2 Assessment: scaling and layering in urban public scoring systems

First, the documentary evidence provides that scaling attempts had been made by local governments using the SCS public scoring systems and the Public Credit Information Centres since the 2000s. However, these quantified scores have not scaled up to provincial levels. They remained at the municipality levels and mainly governed under the population of 15 million.

In addition, most of the citizen scores are applied to low-stake decision-making settings, similar to Sesame Credit's consumer credit reporting services. Some higher-stake citizen scores such as the Guangzhou city management score were abolished after 2024.

Corporate scores however, are used in rather high-stake decision-making settings such as deciding public procurement and frequencies of regulatory inspections. After the NDRC's 2023 Instruction on abandoning or revising these high-stake corporate scoring systems due to anti-competition concerns, some municipal governments have chosen to upgrade the system with more procedural and institutional safeguards. The general trend of the central government is to be more cautious and limit the assigning of quantified, standardised public credit scores to private corporations at large scale.

From the interview data, two main themes of layering effects could be further observed. They explain how quantified public scores are restrained within their data habitats and highlight the necessity of informal human expertise for recalibration of negative layering effects.

(1) Data habitats and common pool resources (CPR)

First, the fish pond analogy and 'data habitat' theme echo Ostrom's case studies on the governance of fisheries which can be understood as common-pool resources (CPR). The analogy reveals the commons or pooling aspects of data, which has sustainability and time concerns with regards to its own data cycles in its long-term institutional environments similar to CPR regeneration cycles in polycentric governance (Ostrom 2010).

This highlights the importance of legal institutions like financial disclosure rules on the generation and update of data in its habitat. The data cycles are path-dependent on how law regulates the periodical publication of, for example, a firm's financial reports. This could be extended to other situations where public data are restrained by each regulatory agency's own mandate, such as tax, social insurance paid for the employees, electricity usage etc. These different institutions and organisational operations create various time frames at local levels and thus diverse data cycles.

(2) Recalibration of negative layering by human expertise

The quote from the Shanghai local regulator also shows that layering effects requires human expertise to recalibrate numbers and algorithms. By long-term interactions with the governed space and sector, human experts can pass on informal knowledge and provide a larger picture on the data contexts. This hybrid between human expertise (regulators' informal knowledge) and formal data support (institutionalised public credit-reporting) together shape the technology and its actual use in Shanghai's public credit information centre.

The limitation of risk models also highlights that the standardisation of data meant the losing of a certain depth of understanding of that data. The empirical observation is that data and algorithms need to be interpreted in their social and legal contexts. Layering effects here mean the necessity of understanding data not just as numbers or digits, but also the embedded meaning of that data in the environments where it was generated and constructed. In this way the data applications can provide more accurate and helpful use in the long-run without being locked-in its path-dependency.

This shows that even at local levels, useful data applications are hard to achieve without human understanding of the specific needs and tacit context over time. Informal knowledge retained and exchanged by private actors like banking professionals thus are crucial for the in-depth application of standardised and large-scale data in the long run. They are epistemic communities that help retain the embedded cognition / knowledge through time.

The next part shows a similar trajectory in rural China, where local experiments expanded to provincial levels, but later lost their momentum and faced negative layering effects.

4.3.2 Rural Scoring: the rise and fall of Qingzhen's local scores

4.3.2.1 Rise of the Qingzhen experiment between 2000 and 2012

Scoring mechanisms at the village level have been documented to have proliferated since the 2000s via various local policies, rules and reports. In contrast to the more economically developed cities like Shanghai, in rural regions the state-led scoring experiments started with the "Honest Farmers" project that aimed to attract entrepreneurship and investments to poor villages.

These rural experiments started around 2006, long before ICT and unified database infrastructures were made available to rural areas. Qingzhen, as a widely reported example,

started with an 'Honest Farmers' programme and moved to the "Smart-Honest Rural-Urban SCS project".

Qingzhen, a predominantly rural county in Guizhou Province deep into the mountainous west of China, first started the 'Honest Farmers' pilot in 2006. The county has a population of just over 500,000. Part of the motivation for the pilot was a scandal over competition over a chicken-raising investment between two villages in Qingzhen, namely Xindian and Baihuahu. In 2004, when the model of 'Leading Enterprise plus Farmer Household' was first introduced in Qingzhen, a chicken-raising company invested in Xindian village, only to find the farmers secretly stealing the company's chicken feed for the hens in their own backyards. As a result, another village in Qingzhen, namely Baihuahu, took the opportunity to have the company work with its own farmers, who were claimed to be more honest and soon made the chicken-raising project prosperous. The outcome was reported as a win-win situation for both the farmers and the enterprise (Wang, 2012).[99]

Officials at the provincial level soon noticed this case and conducted case studies in Qingzhen, which led to the start of the Honest Farmer pilot in 2006.[100] Local governments then competed to promulgate high-profile policy documents prescribing the pilot's details, and soon experiments proliferated in more than 10 localities similar to Qingzhen across Guizhou Province. As for the cradle Qingzhen itself, it has been attracting investment from international companies, and selling agricultural products globally via London and Hong Kong (Wang, 2012).

The 'honest virtue' of the localities was invoked as a factor in economic, social, and judicial development, claimed to represent the enhancing spirit of contract, preserving harmonious rural neighbour relationships, and protecting food safety of products selling in China and abroad. As the experiment developed, Guizhou Province was reported to come up with a philosophy of 'Honesty Triangle' (三位一体) involving (1) local governments, (2) enterprises and (3) farmers. Since then, the Honest Farmers pilot was framed as being not only about Farmers, but including all three actors in the rural regions. In Qingzhen, for example, local governments and party committees of local villages[101] are rated on their own honesty. A

[99] Wang F., (2012). Written by the deputy secretary of Guizhou Province, published on *Qiu Shi*, The CPC's most prominent central theoretical paper.

[100] 2006 coincides with the year when the Nobel Peace Prize was awarded to Yunus and Grameen Bank 'for their efforts to create economic and social development from below.' The practice of micro-credit in Bangladesh, which helped lift people out of poverty, might have been an inspiration to Guizhou's "Honest Farmers" pilot. See e.g. the 2006 Nobel Peace Prize press release, at https://www.nobelprize.org/prizes/peace/2006/press-release/ .

[101] In China nearly every governmental institution requires a party committee embedded. These Party organisations are crucial in decision-making particularly regarding policies.

95% satisfaction rate meant that a village government could be rewarded with better business and investment opportunities from the municipal level, whereas below 85% meant that a village could lose those advantages on attracting businesses. Enterprises categorised as honest were rewarded with tax reduction and loan grants, while the "dishonest" could have their 'licence withdrawn or subsidies retracted' and other legal consequences could follow (Wang, 2012).[102]

The farmers' honesty is further regulated under 15 sections of 100 standards in 2014, a local policy document published by the village's government, namely the *One Hundred Standards of Honest Farmers*. According to this policy farmers could be granted unsecured loans with lower interest rates, subsidies and other project support. The government claimed that the local party-state led the experiment, the enterprises pushed the projects forward, and the farmers were the main actors to carry them out. The virtue of honesty is not only required and cultivated among the farmers as the Honest Farmers project implied, but also for local party-government organisations, and enterprises as connected entities (Qingzhen, 2018).

More specifically, this honesty triangle is claimed to play out in five contents with three major tasks. The five contents include cultivating farmers' honesty in relation to 1) the party and local governments, 2) the law and regulations, 3) the market rules, 4) the financial support, and 5) the spirit of contract. The three major tasks are improving education on honesty through media and propaganda; developing incentive mechanisms by constructing databases and platforms among various agencies on honesty records; and hardening sanctions for dishonest behaviours. In addition, these tasks are further integrated with four major goals: rural economic development (including finance and investment environment); innovation on governance; the cultivation of the "new farmers" with new ideas, better education, and technical and business expertise; and local Communist Party of China (CPC) development in order to better serve self-governance in rural localities (Qingzhen, 2018).

It was reported that in 2011 agricultural loans within Qingzhen reached RMB 1370 million (around $210 million), which was six times that of 2006, while the non-performing rate of rural microcredit loans dropped to 0.64% which is 8.1% less than 2006. Qingzhen's agricultural production also reached RMB 1520 million in 2010 (around $234 million), with an 85.6% increase compared with 2005 and an average of 13.2% increase each year. It was further claimed that farmers' income grew along with better agricultural business structure, farmers'

[102] Also see, Qingzhen Civilisation Office, 'Exploration and practice from "honest farmers" to "intelligent urban and rural areas"' 2018.07.18, online in Chinese, at http://gz.wenming.cn/jujiaogz/201807/t20180718_4764225.shtml [清镇市文明办, 从"诚信农民"到"智惠城乡"的探索与实践, 2018].

expertise and innovation. Better health, labour and education systems were said to be developed as well as honest norms and harmonious community reinforced; and self-governance was improved under more efficient local governments and party organisations. As of 2011, the pilots in Guizhou Province reached 206 towns, 4,484 villages and 27,000 rural groups, that are the minimum administrative unit (Qingzhen, 2018).

The official account emphasised the market logic behind the Honesty Farmers pilots between 2006 and 2012, where local party-government leadership was aiming at lifting rural regions out of poverty by leveraging loans, investments, and entrepreneurship. The following part explains the transformation of Qingzhen's SCS project after the 2010s, from a purely economic logic to embracing more high-tech integration and innovation.

4.3.2.2 Failed expansion of the rural-urban SCS: 2012 to present

The Honest Farmers experiments transformed into rural-urban experiments after 2012 against two backdrops. First the SCS was at that time being formalised at national level. The State Council in 2012 issued the first policy document on establishing the institution of 'Inter-Ministerial Joint Meeting' for SCS. Subsequently, the NDRC and China's central bank held the first national conference on implementing the SCS with a draft *2013-2020 Blueprint*.[103] In 2014 the State Council issued the top-level policy document *2014-2020 SCS Blueprint*[104].

Second, it was at the same period that information infrastructures covered more rural regions and big-data technologies became high-profile in China with increasing investments and technological progress, leading to the 2015 State Council's *Blueprint on Promoting Big Data Development* and *Instructions on Using Big Data to Serve and Supervise Market Entities* (2015 Big Data Policies).[105]

Under these two backgrounds, the Honest Farmers experiment expanded into the urban regions of Qingzhen, with a more comprehensive rural-urban scoring system. It was also claimed that more sophisticated technologies are used. Nevertheless, most of the standards, contents, and tasks of the Honest Farmers pilots were integrated and even preserved their centrality in the new scoring system.

[103] State Council, [2012] No 88. Online in Chinese, available at http://www.gov.cn/zwgk/2012-07/25/content_2191732.htm

[104] State Council, [2014] No.21. *Notice on the Blueprint of Construction of the Social Credit System (2014-2020).*

[105] State Council, [2015] No.50.

First, the *100 Standards of Qingzhen Honest Farmers* (2014), applying to a population of around 380 thousand farmers[106], later gave birth to a similar document of *100 Standards (2015)* applying to all the rural and urban population in Qingzhen, among which the urban population was around 120 thousand. [107]

Ranging from criminal records to administrative penalties such as littering on public streets, the new *100 standards* took into account nearly every aspect of people's daily life and many are similar to that of the farmers. Different rankings of stars are assigned to farmers and citizens when applying for loans. Being a citizen with a high-ranking star can mean a potential maximum loan of RMB 200 thousand ($31 thousand). Similar incentivising mechanisms on tax reduction, subsidies and project investments support were retained in the new rural-urban *100 standards* (2015).[108]

In addition, the 'honesty triangle' of government, enterprise and farmers was also preserved in the new rural-urban experiment; only slight changes were made to wordings, with the honesty of government being put at the core (as to emphasise CPC leadership), and emphasising the crux as being the need to mutually enhance reputation between enterprises and of individual '*citizens'* – previously only the farmers had been referenced. As for general categories of rating, the five contents of the Honest Farmers experiment stayed in the rural-urban one, but under the name of 'five tasks', namely the centrality of party and government, legal norms, market, finance, and spirit of contract. The key socio-legal governance backbones remain similar.[109]

For Guizhou's citizens, a unified platform was supposed to be established in 2014 under *Guizhou's 2014-2020 SCS Blueprint*.[110] It was intended to give each person a social credit ID number, used for personal recording and profiling for future sanctions and rewards. However, in the end, the credit ID number system was not adopted, and credit records are only linked to the citizen's personal ID number within the police / civil registration system under the 2022 Guizhou Social Credit Regulation (effective 2023.1.1). This Regulation also does not mention anything about a unified individual citizen score.[111]

---

[106] Qingzhen government, 2011, Qingzhen's 100 standards for honest farmers, *Modern Guizhou*, 13. [清镇市 100 条诚信农民标准,《当代贵州》2011，13 期].

[107] Bai H., 'Guizhou Qingzhen explores building a social credit system', 2015.4.21, online in Chinese, at http://zqb.cyol.com/html/2015-04/21/nw.D110000zgqnb_20150421_5-01.htm.

[108] Wu X., 'Qingzhen, Guizhou: Create "intelligent benefits" for urban and rural areas and play a concerto of integrity', thepapers.cn, 2021.4.23. https://www.thepaper.cn/newsDetail_forward_12356504.

[109] *Id.*

[110] Guizhou Provincial Government, 2014.

[111] Guizhou People's Congress, [2022] No.20. *Guizhou Province Social Credit Regulations.* Online in Chinese, available at, https://www.creditchina.gov.cn/home/dfggsy/202210/t20221018_303617.html.

In addition, another experiment in 2017 was initiated to adopt block-chain and IoT technologies to track people's behaviours on an 'Identity-chain App' in the Hongfenghu town of Qingzhen.[112] It claims to be capable of catching a panoramic record of people's behaviours with their digitised identity (a so-called 'data iron cage'), while being highly reliable regarding data security and privacy based on block-chain technology. This is part of Qingzhen's 2017 vision of 'transforming the Honest Farmers project to the Smart-Honest Rural-Urban pilot'. This latest rural-urban pilot promises to provide individuals, village groups and enterprises in Qingzhen with multiple benefits in seven major scenarios, including start-up loans, micro-business support, judicial assistance, food safety tracking, and utilising rural idle assets etc. These scenarios largely focus on benefits but mention very little on sanctions or how they can be implemented.[113]

Nevertheless, Qingzhen was left out in the twelve 'pilot cities' selected by NDRC and PBOC in 2018 (including Rongcheng, Weihai, Suqian, Hangzhou, Wenzhou, Xiamen and so on). Compared to the many reports on the early success of the Honest Farmers pilots, there is very little evidence to be found on the efficacy of the rural-urban scores after 2016. [114]

4.3.2.3 Assessment: scaling and layering in the rural scoring systems

Similar to the urban scores, the story of the rural public scores is also about its early expansions being later restricted by negative layering effects. Qingzhen's successful start and Guizhou's adoption of its Honest Farmers model demonstrate how governance by numbers can help local social norms and laws grow and achieve positive complementarities. It started from pragmatic concerns of lowering transaction costs by increasing trust and cooperation between enterprises and farmers, and promoting economic development in poor neighbourhoods to improve the living standards of rural households.

Nevertheless, later the rural-urban scores did not gain much traction compared to the early Honest Farmers scores. The plan to have a unified social credit ID system for citizens in Guizhou Province was also never implemented. This is initial evidence that the rural scoring

[112] In 2017 Qingzhen started to adopt blockchain and IoT technologies to track people's behaviours on an Identity-chain App in the Hongfenghu town of Qingzhen. Cheng X., 'ID Chain App unveiled to take Guizhou Qingzhen's Honest Rural Society Project to the Era of Big Data'，Huanqiu.com, (May 29, 2017), online in Chinese, at http://china.huanqiu.com/hot/2017-05/10762808.html. Blockchain claims to be helpful in solving problems like privacy in SCS. Xu R., et al. Constructing Trustworthy and Safe Communities on a Blockchain-Enabled Social Credits System. arXiv:1809.01031, 2018.

[113] Qingzhen Government, [2017] No 85. Online at http://gz.wenming.cn/jujiaogz/201807/t20180718_4764225.shtml

[114] NDRC, List of first batch of demonstration cities for social credit system construction announced, 2018.1.9. Online in Chinese, at http://www.ndrc.gov.cn/xwzx/xwfb/201801/t20180109_873409.html.

mechanisms were not suitable to be extended to adjacent urban populations and geographies. It was also not clear if the new technological solutions like blockchains could be implemented. Even if they were implemented, it was unclear to what extent it could solve problems such as how to decide which values, norms and behaviours should be recorded and rated across rural and urban regions.[115]

Qingzhen's attempts to transform from Honest Farmers to rural-urban SCS also demonstrate the path-dependence of the quantified scoring on local social norms. The rural score's metrics reinforced those social norms on honesty, which were based on the rural locality's close-knit, *guanxi*-based social networks and the intention to attract investments and follow the local government's leadership. The experiment to expand rural scoring systems into a rural-urban hybrid proved to be difficult when there are more diverse and less *guanxi*-based social norms at play.

In addition, from the above analysis we can see that Qingzhen's local experiments had little to do with the private / commercial digital credit platforms like Alibaba's Sesame Credit. They in fact acted as a functional substitute of the latter. Although usually mentioned under the same name of SCS, the local party-state-led pilots are mostly guided by regulations of NDRC; whereas the private / commercial pilots based on digital platforms have their own trajectory under supervision of China's Central Bank.

Compared to the government-led SCS pilots like in Qingzhen, these commercial platforms such as Sesame Credit were more akin to loyalty programs which consumers voluntarily join. Sesame Credit, for example, did not have much presence in rural regions though it once tried, partly because Alibaba's affiliated business benefits like zero-deposit car renting or hotel booking rarely reach these rural regions.[116] Other traditional commercial credit platforms focusing only on financial credits (as opposed to Sesame Credit covering more comprehensive information), also found themselves unable to set up effective financial credit systems. This deficiency also partly explains why it was the local government that initiated the Honest Farmers and later led the SCS experimentations in rural areas like Qingzhen.[117]

---

[115] *See*, e.g. Daixin's 2020 analysis raised similar question on Qingzhen.

[116] Sun J., 'Looking forward to introducing Sesame Credit to rural areas'，Fenghuang news, 2016.12.21. http://inews.ifeng.com/50448180/news.shtml.

[117] Sina Finance, 'Rural credit provision is too difficult', 2016.12.16, online in Chinese, available at http://finance.sina.com.cn/meeting/2016-12-16/doc-ifxytqav9461371.shtml. In fact, the Central Bank in 2015 required local governments to lead the SCS in rural areas and help small and medium sized companies apply for credits. People's Bank of China, [2015] No. 180, online in chinese, available at http://www.gov.cn/zhengce/2016-05/24/content_5076163.htm.

4.4 Chapter conclusion

4.4.1 Research Questions

This chapter assesses the nature of the 'scoring' systems in China's SCS. It seeks to understand how data and code work in the scoring systems, what their institutional roots are, and if there is evidence to validate the scaling and layering hypothesis, namely code expands governance scales but is limited by its dependence on legal institutions.

4.4.2 Scaling and layering in the SCS scoring systems

It has been shown in this chapter that China's (semi)automated scoring systems are essentially data-driven socio-technical systems that are rooted in the US and European credit-reporting practices and public management scores. The SCS scoring branch is therefore part of the ongoing global scaling of governance by numbers and code led by advanced economies. The SCS scoring systems are nevertheless bound by Chinese local data practices and habitats. They are institutionally rooted in the activities of the Chinese Central Bank, local governments in cities and villages, regulatory agencies, private credit-reporting industries and big-tech companies like Alibaba, which learnt from American and European practices.

The new empirical evidence provided by this chapter shows how these SCS scoring systems, based on data and algorithms, initially scaled up but eventually faced setbacks that are essentially negative layering. Data and code, as new modes of governance, are evidenced to demonstrate complex and layered interactions with law and social norms when expanding across geographies and populations. The Central Bank for example, does not provide any quantitative credit scoring system at the national level, but chose to maintain a largely qualitative database for commercial banks to access credit reference information and make their own local risk scores, some involving AI and ML. Commercial bank's risk score systems, as algorithmic governance, are thus dependent on the Central Bank's FCIFD database which is essentially socio-legal governance. These socio-legal institutions maintained by the central bank provided resilience during the 2008 financial crisis when global credit-scoring, or algorithmic governance, failed.

The Central Bank also provides rather strict regulations on China's personal credit-reporting industry. By 2024 March only three credit-reporting agencies were licensed to operate in different regions. Private scoring systems and Alibaba's private Sesame Score

initially scaled up after 2015, but was not granted a licence by the Central Bank in 2018. Sesame Credit's processing of alternative data and use of AI and ML / DL technologies, were deemed as not conducive to good social governance or scaling to a national digital economy. The reasons include conflict of interests, failure to apply across geographical regions, and antitrust concerns. Local socio-legal institutions again provided the cognitive foundations and meaning for the credit scores. A unifying score by Sesame Credit was not effective, nor fair or independent, when deployed at large scale across geographies and populations.

The public scoring systems in urban and rural China further demonstrated the scaling and layering effects in algorithmic governance. Local public scores made by governments and public-private joint-ventures are often limited by local social and legal norms and practices. It is unknown how many of these scores actually use algorithms like ML / DL, but mostly have a rather similar rating system compared to the private scoring system. In this sense the local public scoring is also a descendent of US and European credit-reporting and rating agencies, but follows the new public management approach that uses metrics and quantified scores in the everyday management of society.

Nevertheless, The NDRC after 2023 required local governments to abolish or revise the scoring systems in high-stake decision-making settings. The restricting factor includes local monopolies based on these scoring mechanisms in urban areas, and the social norms disparity between rural and urban areas. For example, when Qingzhen tried to expand its honest farmers scoring practice to urban areas, the new score did not gain much traction. The recalibration of these scores in both urban and rural regions require human expertise and professional bodies like local regulatory agencies, who have been in the local habitat for a long time and can use both formal and information knowledge embedded in that cognitive and epistemic community.

The negative layering effects evidenced in these case studies can, it is argued, be recalibrated or avoided by using human expertise in these more standardised scoring systems. The hybridisation of both formal and informal knowledge, based on local social and legal foundations, can enhance the functioning and legitimacy of statistical and algorithmic governance. Algorithmic governance, like data and legal governance, needs to grow through time from indeterminate social practices within their respective local habitats. Like an ecosystem evolving in a 'fish pond', algorithmic governance needs to be rooted in the social practices, norms and law. Effective code needs to be built layer by layer onto the socio-legal institutions of a particular society, and potentially scale to the global level.

Chapter 5 (Semi)Automated Law Enforcement: an analysis of code-driven blacklists and joint-sanctions in China's SCS

5.1 Summary of the argument

This chapter provides a case study of China's (semi)automated law enforcement systems in the forms of blacklists and joint-sanctions mechanisms (JSMs). They are the second major institutional component of China's SCS.

New empirical evidence, including online archives and field-based interviews in Shanghai, is provided and analysed to demonstrate how scaling is limited by layering in automated enforcement. The new evidence supports the overarching argument of the thesis that algorithmic governance is layered upon legal governance when applied across large geographies and populations, namely that law is computable only to the extent that computer code is sustained by law's operations within institutions.

5.1.1 Research question

The overarching question of the thesis 'is law computable', is addressed in this chapter by answering a subset of questions as follows.

- What is the nature of the 'blacklists' and 'joint-enforcement' systems in China's SCS?
- How do data and code function in these systems?
- Is there evidence to show scaling and layering of algorithmic governance upon legal governance?

It is argued in the chapter that courts, regulatory agencies and private agents, including banks and companies at local and national levels, have innovated with new ways of enhancing

law enforcement. The focused case study in Shanghai is on contract and debt, where new code-driven technologies and infrastructures were used to automate enforcement. Gains and losses perceived by local enforcers such as judges, civil servants and program managers / scientists are analysed. Challenges faced by these local enforcers from the perspective of daily operations were uncovered within their respective organisations and professional expertise. The main setbacks include (1) the rigidity of code, (2) the counter-productive behaviour changes of human enforcers, and (3) institutional overload and risks to system legitimacy.

5.1.2 Structure of the chapter

The following section (5.2) reveals the nature, roots and legal basis of the blacklists and joint-sanctions mechanisms in China's SCS. It is argued that these systems, in contrast to the scoring systems discussed in Chapter 4, are code-driven socio-technical systems that automate law enforcement procedures. It reveals the institutional roots of these systems, namely the Supreme Court's 2013 Debtors' list and joint-sanctions mechanisms (section 5.2.1), followed by empirical evidence of their national scaling/expansion before 2019 (section 5.2.2), and an analysis of the legality and legitimacy of these automated enforcement systems, as evidence of code layering on law (section 5.2.3).

Section 5.3 then focuses on the case of Shanghai's SCS debtors list and joint sanctions mechanisms. Based on new interview evidence and online legal documents, it provides a thick description of the Shanghai case and explains how code and law works as complements in civil procedures (section 5.3.1). It then provides interview evidence in detail (section 5.3.2) followed by a critical assessment of this new evidence (section 5.3.3). A sectional conclusion follows.

Section 5.4 then concludes the chapter.

The main body of the chapter is structured as follows:

5.2 Blacklists and joint-sanctions in the SCS: Roots and legal basis

- 5.2.1 Roots: Supreme People's Court's code-driven experiments
- 5.2.2 Scaling: national expansions of algorithmic governance
- 5.2.3 Layering: the legal basis of code

5.3 A case study of Shanghai's SCS

- 5.3.1 A flowchart of civil procedure and law-code interactions
- 5.3.2 Interview evidence from local enforcers
- 5.3.3 Assessment: scaling and layering effects

5.2 Blacklists and joint sanctions in the SCS: roots, expansions and legal basis

Since 2012, the Supreme Court, national ministries and local governments have established blacklists and sanctioning regimes that aim at enforcing / creating formal laws with hardened sanctioning measures. These systems usually involve low-tech databases and data sharing infrastructures, but do not involve machine-learning algorithms or AI in decision-making. Instead, AI is involved only in facial recognition for identity verification, for example, when a person is boarding a high-speed train or a flight. They are code-driven algorithmic governance in the sense that these AI and non-AI code are embedded in the physical infrastructure of the law enforcement procedures.

First, the blacklists (黑名单) serve as databases that have the function of 'push-notifications'. They aggregate and disclose the law-violation records / information of natural persons and organisations, to the general public, governmental agencies or private firms for decision-making. Then the recipient of this blacklist information can then make decisions on, for example, individual investment, public procurement, or private bank's lendings, and so on.

Second, the joint sanctions mechanism for trust-breaking behaviours (JSMs, 联合失信惩戒机制) facilitate automatic enforcement of punishments to the natural persons and organisations on the blacklists. This specific enforcement agent, for example, a boarding gate with AI facial recognition at a train station or an airport, can automatically verify if a person is on a blacklist and prevent them from travelling by high-speed train or by air.

The blacklists and JSMs together serve as the SCS experiments that uses code to enforce obligations under the law, following the SCS's instructions of creating a trustworthy environment for governments, business, society and the justice systems (entailed in the *2014-2020 Blueprint*)[118]. They work hand-in-hand to facilitate inter-agency regulations in crucial industries[119]

The blacklists and JSMs are mostly established at the national level, but the provincial and local governments can also specify their own implementation rules based on the national legal frameworks.[120] They are all state-enforced compulsory SCS regimes, which are similar

[118] State Council, [2014] No.21.
[119] See, e.g. NDRC, [2017] No.1798, at https://www.ndrc.gov.cn/xxgk/zcfb/ghxwj/201711/t20171103_960925.html.
[120] See, e.g. one of the earliest JSMs was actually established in 2013 in Hei Longjiang Province for enterprises. http://www.hlj.gov.cn/wjfg/system/2014/03/21/010640613.shtml.

to the local public scoring systems analysed above in section 4.3. In contrast, most private financial-commercial scoring systems discussed in section 4.2 are mostly voluntary where users can opt in and opt out (though in order to access certain financial services from commercial banks it is almost impossible to opt out).

By June 2018, it was documented that 37 blacklists and JSMs had provided more than 12.9 million pieces of information, restricting 11.60 million flight bookings and 4.41 million high-speed train bookings; and that (notionally) 2.5 million debtors have actively fulfilled their obligations by complying with judgments or other mechanisms (the actual number of debtors will be less than this, as one person or entity can be blacklisted for multiple times).[121]

By November 2018, the NDRC had facilitated altogether 40 inter-departmental JSMs based on MOUs.[122] Most of these 40 experiments started taking off from the 2013 Supreme Court's debtors list and its JSM, and proliferated since 2016 with the State Council and NDRC pushing for progress.[123]

5.2.1 Roots: the Supreme People's Court's (SPC) experiments

(1) The SPC debtors list (2013-present)

The first national blacklist-style databases and joint punishment programs were established by the China Securities Regulatory Commission in 2012. They were aimed at improving regulatory enforcement of the securities and futures market (2012.9.1, Interim Measures), by establishing industry-specific integrity *Dang'an* (archives) similar to Blacklists[124]. The second national blacklist was established on food and drug safety by the NDRC (2012.9.4, Instructions)[125].

[121] Credit China website, 'The National Public Credit Information Center will regularly publish monthly reports on blacklists', 2018.6.11, online in Chinese, available at http://www.creditchina.gov.cn/hangyexinyong_824/zonghedongtai/zhengfubumen/201806/t20180611_117781.html

[122] NDRC press release, 2018.11, online at http://www.ndrc.gov.cn/xwzx/xwfb/201811/t20181115_919731.html. Also see http://www.gov.cn/fuwu/2018-11/02/content_5336733.htm 2018.11.2. With two more in November on academic research, http://www.gov.cn/xinwen/2018-11/09/content_5338654.htm; and notary. http://www.creditchina.gov.cn/xinyongdongtai/buwei/201811/t20181108_130634.html.

[123] NDRC, *Notice of the NDRC, the Supreme People's Court, the People's Bank of China, and Other Departments on Issuing the Memorandum of Understanding on Taking Joint Disciplinary Actions against Dishonest Persons Subject to Enforcement,* [2016] No.141.

[124] China Securities Regulatory Commission, [2012] No.80, *Interim Measures for the Supervision and Administration of Integrity in the Securities and Futures Markets*, effected on 2012.9.1, Amended on 2014.10.15 (Adding the 3 year and 5 year limit of the effects of the bad records, and establishing public search platform to open the records). Online at https://www.gov.cn/gongbao/content/2012/content_2266931.htm. Later this was replaced by the China Securities Regulatory Commission, [2018] No.139, *Measures for the Supervision and Administration of Integrity in Securities and Futures Markets*, online at https://www.gov.cn/gongbao/content/2018/content_5294430.htm.

[125] NDRC, [2012] No. 2829. Online at http://images.mofcom.gov.cn/sczxs/accessory/201211/1352271472709.pdf.

At the local level nevertheless, there were already JSMs in April 2013 established in Hei Long Jiang Province for enterprises (2013 Hei Long Jiang Regulation) based on the province's own corporate credit information database.[126]

However, a consistent language of the Blacklists and JSMs did not take form until July 2013 when the Supreme People's Court's judgment debtors list program (the 'SPC debtors list') was formally launched.[127] For example, the 2012 Interim Measures for securities and futures market did not use the term JSM or blacklist, but instead stated an intention to build the 'integrity *Dang'an* (archive) database' (诚信档案数据库).[128] This Interim Measure did not adopt the formal JSM language until it turned into a non-provisional Measure in 2018 with its Article 43 explicating the JSM terms.[129]

The July 2013 Judicial Interpretation of the SPC debtors list sets the prototype for other SCS blacklists. The 2013 Judicial Interpretation defined six elements of the debtors list:

(1) Legal basis: namely, under China's Civil Procedural Law;

(2) Non-compliance conduct: six types of severe, law-violating are defined, each involving 'being able to comply with civil judgement but refuse to do so'[130];

(3) Data processor: identified as the Supreme People's Court (SPC);

(4) Data categories: including names, ID numbers, obligations in the judgments, and the (non-)compliance situation and so on[131];

(5) Data subjects: defined as 'trust-broken judgment defaulters', including natural persons, legal persons and other organisations;

(6) Procedures (these are briefly defined, in contrast to the 2017 amendment)

By that time, the punitive measures were very limited as there were not yet expansive JSMs. The main measures were from the *2010 SPC Judicial Interpretation of Limitation on High-end Spending*, which provided the basis for restrictions on defaulters' golf plays, airline tickets, or luxury hotel bookings.[132] Nevertheless, Article 4 of the *2010 Judicial Interpretation* also stipulated that usually the restrictions are triggered by enforcement applicants' written

[126] Established under a *Hei Long Jiang Provincial Regulation for Corporate Credit Information* 2007, after the PBOC started establishing the new FCIFD in 2006. Available at http://www.hlj.gov.cn/wjfg/system/2014/03/21/010640613.shtml.

[127] Supreme People's Court, Judicial Interpretation [2013] No.17. *Judicial Interpretations on the Publication of Information of Dishonest Judgement Defaulters*, online in Chinese, available at https://www.chinacourt.org/law/detail/2013/07/id/146217.shtml.

[128] China Securities Regulatory Commission, [2012] No.80.

[129] China Securities Regulatory Commission, [2018] No.139.

[130] Supreme People's Court, Judicial Interpretation [2013] No.17. Article 1.

[131] *Id.* Article 4.

[132] Supreme People's Court of China, *Judicial Interpretation on High-end Consumption Restrictions for Judgement Defaulters,* [2010] No.8.

request, and only granted by the court after reviewing it. The court can only apply the restrictions 'when deemed necessary by its duty'.[133]

(2) JSMs for the SPC debtors list (2014-present)

In 2014, the SPC established a JSM linked to its debtors' list. The goal was to restrict debtors' luxury travel by trains (since June 18th 2014) and planes (since July 1st 2014). This JSM entered into effect by a Memorandum of Understanding (MOU), signed by the SPC with the China Railway Corporation and the China Civil Aviation Information Network Co., Ltd.. These two corporations are both state-owned and oversee the railway and airline booking services.[134] Early in August 2015, the restriction extended to first class seats on high-speed trains, and by August 2016 this JSM already blocked 1.55 million train bookings and 4.705 million flight bookings.[135]

On January 20th 2016, the original JSM on the SPC debtors list expanded. A new MOU was signed by 44 ministerial departments, with the NDRC, SPC and the Central Bank as the most important three overseers and enforcers. Since then, the JSM has started to restrict debtors nation-wide with 32 measures. It is no longer just focusing on high-spending restrictions like travelling and sending children to luxury private schools, but also on multiple financial restrictions, subsidies, and being banned from becoming civil servants or high-level executives in state-owned enterprises and so on.[136]

Looking at the intent of the JSM, the MOU quoted three legal documents to state its three major goals: solving the courts' enforcement difficulties (per the Instructions by the Central Political and Legal Committee), promoting fair market competition and normal market order (per the Instructions by the State Council), and promoting honesty and trust in society (per the 2014-2020 Blueprint by the State Council). It manifests a hybrid of the concerns of the justice system and the market, dealing with these under the umbrella of the SCS.[137]

---

[133] *Id.* Article 4.

[134] Li W., '44 departments signed a memorandum to jointly punish dishonest judgement defaulters', China News, 2016.1.21, online in Chinese, at http://www.chinanews.com/gn/2016/01-21/7726810.shtml. [44 部门签署备忘录 联合惩戒失信被执行人].

[135] Tang Y., 'Supreme Court: It has restricted 4.7 million flights for those subject to execution for dishonesty', China News, 2016.9.22, online in Chinese, at http://www.chinanews.com/gn/2016/09-22/8011149.shtml.

[136] See, NDRC, [2016] No.141. *Notice of the NDRC, the Supreme People's Court, the People's Bank of China, and Other Departments on Issuing the Memorandum of Understanding on Taking Joint Disciplinary Actions against Dishonest Persons Subject to Enforcement.* Online at http://www1.lawinfochina.com/display.aspx?id=22323&lib=law; [国家发展改革委、最高人民法院、中国人民银行等关于印发对失信被执行人实施联合惩戒的合作备忘录的通知]

[137] NDRC, [2016] No.141. Preamble.

In addition, on September 25, 2016, a top-level policy document co-issued by the Central CPC Office and State Council Office, prompted the construction of a more comprehensive JSM with a more detailed repairing / delisting system. It provides certain procedural safeguards. These include an advanced notice procedure (legal notice must be given before putting debtors onto the SPC debtors list), a facility for debtors to apply for correction of the list in 10 days upon receipt of the listing notice following (which the court must decide in three days whether to correct), and a requirement that the debtor must be removed within three days of compliance confirmation.[138]

The statistics on blacklisting and removal can help one to understand the impact of the SPC debtors list and JSM in their scale. In June 2018 the SPC debtors list added 266,447 pieces of new information, including 191,872 debtors (160,994 natural persons and 30,878 entities). Meanwhile, 132,035 debtors were delisted from the debtors list that month (116,047 natural persons, and 15,988 entities).[139] Later in October 2018, the numbers all decreased. 200,596 pieces of new information were added, including the listing of 171,943 debtors (145,858 natural persons and 26,085 entities); while 88,542 debtors were delisted (76,268 natural persons and 12,274 entities).[140]

Since June, 2018, there have been monthly statistical reports from the National Public Credit Information Centre, which provide an important window to look into the operation and effects of the national compulsory SCS pilots. The Centre was established by the NDRC in September 2017 and has been staffed and functioning since June 2018[141].

5.2.2 Scaling: national expansions of algorithmic governance (2016-)

Three waves of Blacklists and JSMs can be identified by three policy interventions since 2016. The first one of these interventions was the *June 2016 State Council JSM Instructions*,

[138] Central CPC Office and State Council Office, [2016] No.64. *Opinions on accelerating the construction of credit supervision, warning and punishment mechanisms for dishonest judgement defaulters.*[中共中央办公厅 国务院办公厅印发《关于加快推进失信被执行人信用监督、警示和惩戒机制建设的意见》].

[139] National Public Credit Information Center, 'Announcement and explanation on newly listed dishonesty judgement defaulters', 2018.7.2. Online in Chinese, at http://www.zhengxinbao.com/6669.html.

[140] Credit China, 'Announcement and explanation on newly listed dishonesty judgement defaulters in 2018 October' 2018.11.2. . Online in Chinese, at http://www.gov.cn/fuwu/2018-11/02/content_5336733.htm.

[141] Credit China, 'The National Public Credit Information Center will regularly publish monthly reports on the blacklist' http://www.creditchina.gov.cn/hangyexinyong_824/zonghedongtai/zhengfubumen/201806/t20180611_117781.html 2018.6.11.

leading to the first wave of 23 JSMs (with corresponding Blacklists).[142] Major examples include the regulations of the energy industry.

The second wave was led by the more detailed *October 2017 NDRC and PBOC JSM Instruction,* which resulted in 11 JSMs being set in one year from October 2017 to October 2018.[143] An important example is on labour protection and the regulations of the employers of domestic migrant workers.

The third wave was preceded by the *October 2018 State Council JSM Instruction,* entailing the NDRC to 'accelerate the SCS by cooperating with other departments' and 'facilitate more than 50 MOUs for JSMs in crucial areas by the end of 2018'.[144]

Regulations of the energy industry

China's energy industries include coal, oil, gas, electricity, new energy, and renewable energy. The Blacklists and JSMs for this industry emerged in the first wave of national expansion of the SCS. They mostly target enterprises and their accountable representatives. The response of the relevant regulatory agencies could exemplify not only how a typical set of post-2016 SCS pilots of Blacklists and JSMs were established, but also how these pilots innovated their own paths during the experiment.

First, instead of a simple Blacklist, the National Energy Administration (NEA) created a comprehensive ranking, listing and inspection system for the entire energy industry. This was required on 2016.12.21 when the NEA issued the *Implementation Instruction on the Construction of Credit Systems in the Energy Industry (2016-2020)* to mandate credit systems, including blacklists and JSMs, being set up in the energy industry.[145] On 2017.8.1 the NEA set up the specific details of the Energy Credit Evaluating System which is more comprehensive than a blacklist and more like a rating system. In the NEA's *Provisional Regulations on the Credit Evaluating of Market Entities in Energy Industries,* this new evaluation system categories the energy enterprises into five levels: AAA, AA, A, B, C. Some of these five main ranks can be further specified into six sub-categories: BBB, BB, B, CCC, CC, C (Article 8).

[142] State Council [2016] No.33. [国发〔2016〕33 号].

[143] NDRC [2017] No.1798. [发改财金规〔2017〕1798 号]

[144] State Council [2018] No.104. [国办发〔2018〕104 号]

[145] National Energy Administration (NEA) [2016] No.350, online in Chineses, available at http://zfxxgk.nea.gov.cn/auto79/201612/t20161220_2369.htm.

These ranks link to the ranked entity's public procurement, regulatory inspection, project subsidies, and licensing opportunities.[146]

Second, two JSMs were set particularly for the Electricity Industry (2017.5)[147], and the Oil and Gas Industries (2017.8.2) by the NDRC[148]. In May 2017 the JSM for the Electricity industry was set. 13 measures were set to punish serious violations of related laws in the Electricity industry.[149] Following this JSM, later on 2018.2.3 the NDRC further issued a joint 2018 Implementation Rules on JSM for the Electricity Industry and added another Particular (with more attentions being paid to) Observation List （重点关注名单) that increase the inspection rates and regulatory attention for listed entities.[150] The specific evaluation and rule-making was delegated to the China's Alliance of Electricity Enterprises (中国电力企业联合会), a NDRC & NEA-led self-organised business alliance.[151] The JSM for Oil and Gas Industries was established on 2017.8.2, which is similar to the Electricity JSM with different particular industrial rules and focuses[152].

Regulations of employers of domestic migrant workers

After the Oct 2017 NDRC & PBOC Instructions, the second wave of Blacklists and JSMs were set up. A key example in this wave is the ones for domestic rural-urban migrant worker protection. The Ministry of Human Resources and Social Security (MOHRSS) established a Blacklist (2017.9.25) and a JSM (2017.11.29) to help protect the social security and welfare of

[146] NEA [2017] No.37. Article 8. Online at http://www.gov.cn/xinwen/2017-08/25/content_5220483.htm.. Also Credit China, 'The national energy sector is accelerating into the era of 'credit' based regulation, 2018.10.18. http://www.creditchina.gov.cn/lianhejiangcheng/lianhejiangchenganliguiji/lianhejiangchengbuwei/201810/t20181017_128320.html.

[147] NDRC [2017] No.946. *Memorandum of Understanding on Joint Sanctions of Serious Illegal and Untrustworthy Market Entities and Related Personnel in the Electric Power Industry.* Online at http://www.chinatax.gov.cn/n810341/n810755/c2665761/content.html [发改运行〔2017〕946 号]. Later in 2018 the NDRC made a further regulation on the electricity industry. http://zizhi.nea.gov.cn/html/2018/tzgg_0226/1620.html

[148] NDRC [2017] No.1455. Online at http://www.ndrc.gov.cn/zcfb/zcfbtz/201708/t20170828_858891.html [发改运行〔2017〕1455 号]

[149] NDRC [2017] No.946.

[150] NDRC [2018] No.233. Online at https://www.ndrc.gov.cn/xxgk/zcfb/ghxwj/201802/t20180213_960940.html [发改运行规〔2018〕233 号] 2018.2.3.

[151] *Id.* See the NDRC's attachment, https://www.ndrc.gov.cn/xxgk/zcfb/ghxwj/201802/W020190905495706791868.pdf.

[152] NDRC [2017] No.1455. 'Memorandum of understanding on joint sanctions of serious illegal and dishonest entities in the oil and gas industry', at https://www.gov.cn/xinwen/2017-08/28/content_5221065.htm [发改运行〔2017〕1455 号]

domestic rural-urban migrant workers, by blacklisting their employers who do not pay their salary on time153.

This JSM is part of the effort to achieve the 2020 goal of the MOHRSS 3-year-plan on eliminating payments in arrears for migrant workers (issued on 2017.7.7, in response to 2016.1.7 State Council Instruction on the goal of 2020 – worth noting it was the first policy issued by the State Council in 2016).[154]

Compared to the same time period of 2017, in the first three seasons of 2018 the JSM allegedly has helped reduce the numbers of migrant workers with payments in arrears by 43%, and the amount owed reduced by 32.5%—nearly 1.313 million workers with a RMB 12.98 billion payment (around $2 billion).[155] By way of comparison, the reduction in the rate of migrant workers being paid in arrears had been around 14% during 2015-2016, much lower than the 2017-2018 reduction rate after the JSM was established in late-2017 (again, 43%).[156] If the data published are reliable, such reduction of migrant workers with payments in arrears could be partly attributed to the JSM.

In addition, in September, 2018, the MOHRSS further established detailed *Rules for the 2018 Task of Guaranteeing the Payment of Migrant Worker Salary*. Rule 4 focuses on the employers' payment creditability / trustworthiness.[157] These Rules further reached to provincial level, and some provinces have made their own specific rules based on the national one.[158]

---

[153] Ministry of Human Resources and Social Security (MOHRSS), [2017] No.16; and National Development and Reform Commission (NDRC), [2017] No.2058.

[154] MOHRSS, 'Three-year action plan to combat arrears and guarantee payments, striving to achieve basically no arrears in wages for migrant workers by 2020', 2017.7.21, online in Chinese, at http://www.mohrss.gov.cn/ldjcj/LDJCJgongzuodongtai/201707/t20170725_274594.html; Also see MOHRSS [2017] No.80. State Council [2016] No.1, online in Chinese, at http://www.gov.cn/zhengce/content/2016-01/19/content_5034320.htm

[155] Credit China, 'Joint punishment enables millions of workers to recover over 10 billion in unpaid wages', 2018.11.2 online in Chinese, at http://www.creditchina.gov.cn/xinyongdongtai/buwei/201811/t20181101_129920.html [联合惩戒使百万劳动者追偿拖欠工资超百亿, 信用中国].

[156] See, National Bureau of Statistics, '2017 Migrant Workers Monitoring Survey Report', 2018.4.28, online in Chinese at https://app.www.gov.cn/govdata/gov/201804/28/423387/article.html [2017 年农民工监测调查报告]; Liang Z., '2.369 million migrant workers suffered from wage arrears last year. Ministry of Human Resources and Social Security: Focus on solving wage arrears in the engineering field', Jiemian News, 2017.3.14. Online in Chinese, at https://www.jiemian.com/article/1172388.html [去年 236.9 万农民工遭遇拖欠工资 人社部：重点解决工程领域欠薪] ; According to the National Bureau of Statistics, in 2016 there are around 281.71 million migrant workers, and according to the MOHRSS in 2016 there were only 2.369 million of them with payments in arrear (around 0.84%), see MOHRSS, 'We are confident in managing wage arrears', 2018.10.19, online in Chinese, available at http://www.mohrss.gov.cn/SYrlzyhshbzb/dongtaixinwen/buneiyaowen/201810/t20181019_303164.html.

[157] Han Bing., 'Detailed assessment rules for ensuring the payment of wages to migrant workers are announced. Strengthen credit punishment for untrustworthy companies with unpaid wages', www.gov.cn, 2018.9.11. Online in Chinese, at http://www.gov.cn/xinwen/2018-09/11/content_5320898.htm.

[158] MOHRSS, 'Qinghai issued the "Detailed Assessment Rules for Guaranteeing the Wage Payment of Migrant Workers in 2018"', 2018.11.09. http://www.mohrss.gov.cn/SYrlzyhshbzb/dongtaixinwen/dfdt/gzdt/201811/t20181109_304494.html; Shandong Department of Human Resources and Social Securities [2018] No.48, , http://www.lyrs.gov.cn/info/1208/15218.htm [鲁人社发〔2018〕48 号] 2018.10.11.

These provinces and localities were not only interested in establishing rules, but have also shown enthusiasm in specifying their own blacklists. To better enforce the national blacklist according to local needs, Hebei Province in 2018 established its own *Provisional Rules for Blacklist on Migrant Workers Salary Payment in Arrears*. Hebei's capital city Shi Jiazhuang also in 2018 published a municipal Blacklist with four enterprises and two natural persons, which together owed almost RMB 7 million ($ 1 million) salary payment to 529 migrant workers at that time.[159] It might be worth further looking into the local specificity and variations of the national Blacklist and JSMs to see how implementations of these institutions can practically assist migrant workers in different urban areas.

According to the National Bureau of Statistics, in 2016 there are around 281.71 million migrant workers and the number is still increasing every year, approaching the U.S. population (325 million 2017)[160]. According to the MOHRSS in 2016, 2.369 million of these migrant workers were owed payments in arrear (around 0.84%, albeit this may be an underestimate given the likelihood of underreporting of the problem).[161] If the Blacklist and JSM can really help in solving this perennial problem, this would enhance social justice in China with significant socio-political impacts. Despite it still being at an early stage with lots of political risks, the promise of eliminating the problem of payment in arrears for migrant workers is one reason why the SCS pilots could win popular support. This is particularly, so considering that half of the Chinese population (around 582 million, 2017 World Bank)[162] are rural with an average monthly wage of around 150$, and half of this rural population are migrant workers (the 286 million) who work in urban areas to support their families living in rural areas, with a monthly income of 400$.[163] A migrant worker's salary not being paid could mean not only increasing inequality and social injustice in urban areas, but also devastation for families in rural areas.

[159] Li and Zhang, 'The arrears of wages are nearly 7 million. These people and companies in Shijiazhuang are on the "blacklist"!', hebei.com.cn, 2018.11.13. Online at http://report.hebei.com.cn/system/2018/11/13/019250714.shtml. http://www.erdsrsrc.com/jgzn/rszyzc_jzb_/11372.html. Hebei Department of Human Resources and Social Securities [2018] No.9. [冀人社规〔2018〕9号] 2018.4.26.

[160] US population, World Bank Statistics. 2017

[161] Liang Z., Jiemian News, 2017.3.14. Online at https://www.jiemian.com/article/1172388.html

[162] World Bank Statistics 2017, China Rural Population https://data.worldbank.org/indicator/SP.RUR.TOTL?locations=CN.

[163] National Administration of Agriculture Statistics 2017, quoted from Sina News, 2018.1.2. available at http://www.sohu.com/a/214090338_603780.

Further expansions

There were altogether 38 JSMS established by MOUs in the first two waves. The third wave of Blacklists and JSMs pilots emerged intensely since October 2018. It was largely due to a State Council Instruction on Oct. 29th 2018 mandating the NDRC to 'accelerate the SCS by cooperating with other departments' and 'establish more than 50 MOUs for JSMs in crucial areas by the end of 2018' (*State Council 2018 JSM Instruction Focusing on Corporate Concerns to Further Promote the Implementation of the Business Environment Improvement Policy*).[164]

This *State Council 2018 Instruction* first highlights the urgent problem of financial difficulties for small and medium, and micro-sized enterprises (SMMEs), which must be relieved by the Central Bank (PBOC) and other departments. It demands the 'China's Banking and Insurance Regulatory Commissions and the State Administration of Taxation should actively promote the interaction between banking and tax agencies, encourage commercial banks to rely on tax credit information to innovate credit products, promote tax and bank information interconnection, and alleviate the financing problems of small and micro-sized enterprises (SMiEs)'.

It then mandates the State Administration of Taxation, the General Administration of Customs and other departments to 'strengthen cooperation, further accelerate the progress of export tax rebates, simplify the procedures for enterprises with high credit ratings and good tax records, and shorten the time for tax refunds.' It further requires that all departments should 'actively promote new types of supervision methods such as … credit supervision, big data supervision… not only to improve regulatory efficiency, but also effectively reduce interference to enterprises'.[165]

5.2.3 Layering: legal basis for code in the SCS

In terms of legal basis, each Blacklist is established in a single field or industry by a particular state agent with specific disclosure purposes, based on certain higher-level laws. For example, in 2013 the Supreme People's Court (SPC) established the first debtors list to disclose persons and organisations that have refused to comply with civil judgments. It mainly relies on

[164] State Council [2018] No.104, at http://www.creditchina.gov.cn/toutiaoxinwen/201811/t20181109_130729.html [国办发〔2018〕104号]

[165] *Ibid.*

China's Civil Procedural Law and the SPC's own judicial interpretations as its legal basis.[166] However, the ministries of energy, social security, and tourism respectively established their own Blacklists for the purposes of energy industry disclosure, rural-to-urban migrant workers protection, and tourism regulations. The formal legality of these Blacklists rests not only on each agency's own delegated authority but also on two Regulations enacted by the State Council: the 2008 Open Government Information Regulation (2008 OGI Regulation), and the 2014-2020 Blueprint. Accordingly, the legitimacy of these Blacklists lies in the crucial values of both transparency / disclosure, as enshrined in the 2008 OGI Regulation, and trust, as the core of the 2014-2020 Blueprint.[167]

In contrast, the JSMs are stipulated among various agents to facilitate their data-sharing and inter-departmental sanction mechanism. The details of how each mechanism works are usually stipulated in the Memorandum of Understandings (MOU，合作备忘录) signed by these cooperating agents (mostly MOUs are signed by more than 20 agents). For example, in 2016 the NDRC and SPC led together 44 agencies to sign the MOU on establishing the JSM for the SPC debtors lists, in order to penalise those on this particular Blacklist in order to promote civil judgement enforcement.[168]

Therefore, the formal legality of these MOUs (that create the JSMs) is generally traced back to each law and policy that grants each agency's mandate and enforcement authority. These laws and policies, such as contract law, civil procedure, energy regulations, environmental and migrant worker protection regulations, link the respective joint enforcement measures to the law-violating behaviours which are recorded in their respective Blacklists. As some linkages cannot find a field-specific legal basis, they are solely connected through the more general *State Council 2014-2020 Blueprint*. This situation raises a point of legal ambiguity since under the current Chinese Legislation Law only Laws (*Falv,* 法律) enacted by the National People's Congress (NPC, China's highest legislature) can impose extra obligations upon people and organisations; it can, however, be argued that many of these MOU-based JSMs only involve restricting benefits rather than imposing 'obligations'.

[166] See, e.g. Supreme People's Court, Judicial Interpretation [2013] No.17. Later amended in 2017, SPC Judicial Interpretation, [2017] No. 7.

[167] See State Council, OGI 2008; and State Council, Blueprint 2014-2020.

[168] NDRC, Financial Department, [2017] No 1798. Online in Chinese, available at https://www.ndrc.gov.cn/xxgk/zcfb/ghxwj/201711/t20171103_960925.html [发改财金规〔2017〕1798 号].

5.2.3.1 Blacklists: Judicial interpretations and ministerial regulations

The Blacklists are in essence field-specific databases of law-violating behaviours. They are set up by state regulatory agencies with formally enacted ministerial regulations (*Guizhang,* 规章). Each of these regulations, alongside setting up a particular Blacklist, will also state the detailed mechanisms for the inputting of information on people and organisations into this designated database. More specifically, it will stipulate the legal basis of the Blacklist (namely, under which currently existing laws), the data subjects (whose information would be collected), the specific type of serious law-violating behaviours that should be aggregated (usually severe violations selected under own mandates), the data categories that are recorded (namely, what concrete information are recorded, such as personal information like names; case information like the facts, legal basis and dates of violation; and repairing or correcting information), and the procedures for inputting or erasing that data (including the ministry's own procedural obligations, and the correction and removal mechanisms). These detailed mechanisms are important for the public and the ministry itself to understand how to operate the Blacklist. Therefore, the ministerial regulations are of great value in the following analysis.

Before going into those details, it is also important to understand the formal legality of these regulations in the normative hierarchy of Chinese laws. According to China's Legislation Law, the legal authority of ministerial regulations situates under that of the State Council-enacted Regulations (*Fagui*, 法规), which means they can be reviewed and revoked by the State Council if they are found in violation of the Regulations or higher laws (that granted them with mandates). In the Blacklists context, besides the laws setting up each agencies' mandate, two Regulations are the most relevant, the *2008 Open Government Information Regulation* ('2008 OGI Regulation') and the *2014-2020 Blueprint*, both enacted by the State Council.[169] Above the Regulations and under the Constitution, at the top of the normative hierarchy is the Law (*Falv*, 法律) enacted by the National People's Congress (NPC). By March 2024 there is only a *Social Credit Systems Construction Law* (Draft, 2022) proposed by the NDRC and the Central Bank to the NPC. If passed by the NPC this will be China's first comprehensive SCS legislation at the Law level.[170] Therefore, the current highest legal authorities that the Blacklists regulations can be based on are the *2008 OGI Regulation* and *2014-2020 Blueprint*.

[169] State Council, 2008 OGI Regulation. 2014-2020 Blueprint

[170] NDRC, *Social Credit Systems Construction Law (Draft for public consultation, 2022),* 2022.11.14, online at https://yyglxxbs.ndrc.gov.cn/file-submission/20221116185938435168.pdf.

(Nevertheless, the NPC-enacted Legislation Law could be indirectly related to the legality of these Blacklists, as noted in later discussions.[171])

The only exception is the SPC debtors list. Instead of a ministerial regulation, it was enacted by an SPC Judicial Interpretation (*Sifa Jieshi*, 司法解释) under the SPC's own mandate, which is outside the normative hierarchy of Chinese laws. Unlike the U.S. Supreme Court, the Chinese SPC does not have the mandate to review legislations or regulations, but it could issue Judicial Interpretations as quasi-legislative documents.[172]

Apart from the formal legality in the normative hierarchy, the more substantive legitimacy or values of the Blacklists can also trace back to their higher laws, namely the 2008 OGI Regulation and 2014-2020 Blueprint. The legitimacy or value they provide is disclosure (evident in the name 2008 'Open Government Information'). The 2014-2020 Blueprint as a normative policy document, however, besides disclosure, also uses the fuzzy ideas of honesty and trust to provide legitimacy.[173]

This is particularly interesting when comparing the foundational values of the Blacklists and JSMs, namely disclosure and trust, to that of the Central Bank-led financial-commercial pilots, that are consent of the consumers and the legitimate interests of third parties such as the data controller (for example, the Central Bank or the credit-reporting companies' legitimate interest in building a robust financial infrastructure for banks). Their distinctive logics also help us understand the different types of data, databases, and the purposes of these two categories of SCS pilots, that are the NDRC-led Blacklists and JSMs discussed in this chapter, and the Central Bank-led scoring experiments in the previous chapter (Ch 4).

5.2.3.2 JSMs: a weaker legality and legitimacy

Based on an individual Blacklist database that contains law-violating information, each JSM connects to a particular Blacklist and shares its data to a central database, that is the National Credit Information Sharing Platform established in 2015 (NCISP, or the Credit China Website)[174]. From this public platform, the set of regulatory agencies involved in the relevant JSM as parties to the effectuating MOU are required to together carry out sanctions and rewards

[171] *Legislation Law of the People's Republic of China*. 2000, with latest 2023 Amended.
[172] Supreme People's Court, Judicial Interpretation [2013] No.17.
[173] See the preamble of the State Council, 2014-2020 Blueprint.
[174] Liu K., 'National Credit Information Sharing: Let Debtors Have No Place to Hide', Guangming Daily, 2018.8.10，online in Chinese, at http://www.gov.cn/xinwen/2018-08/11/content_5313109.htm?_zbs_baidu_bk.

on people and organisations as outputs of the system. This joint enforcement of existing laws by multiple agencies is usually not directly related to the original law-violating behaviours.

For example, according to the Measure 17 stipulated under the JSM for the SPC debtors list (2016), the Central Organizing Bureau (中组部), Ministry of Human Resources and Social Security (MOHRSS), and the Civil Servant Bureaus (among other agencies), are required to restrict those on the SPC debtors list from being hired as civil servants. Here the sanction in the form of professional restrictions is not directly related to the behaviours that send these people onto the SPC debtors list (which are non-compliance to law enforcement on debt). This, on the face of it, imposes new obligations on citizens and organisations without a Law's mandate, and thus creates a legitimacy gap and violates the Chinese Legislation Law.

However, this legitimacy gap was somehow bridged by the *Civil Servants Law* after its *2018 Amendment*, which in Article 26 added the clear statement that those who are blacklisted on the JSMs cannot be hired as civil servants. The legal mandate or legitimacy of professional restriction therefore does not come from the SPC's mandate, but directly from the new *Civil Servants Law*.[175]

In each MOU, there are specific stipulations on who (or which agencies, e.g. MOHRSS), are responsible to jointly enforce which laws (e.g. the *Civil Servants Law*), and with what regulatory sanction or reward measures (e.g. Measure 17 of the JSM for the SPC debtors list) for the people and organisations on the particular Blacklist.[176]

Nevertheless, the MOUs that establish the JSMs have a more ambiguous legal status in the normative hierarchy. They could be seen as jointly enacted ministerial regulations, which are similar to the Blacklist regulations but with more agencies recognising them. Then their formal legality would be similarly based on State Council's regulations like the *2008 OGI Regulation* and *2014-2020 Blueprint*, as long as the relevant agencies all act within their own mandates. Nevertheless, as different agencies have different mandates, such linkage must be established by multiple pieces of legislation (such as the *Civil Servants Laws* in the Measure 17 case). Many measures without a clear and particular legislative basis had to resort to the ambiguous languages in the *2014-2020 Blueprint*. For example, Measure 18 of the JSM for the SPC debtors list stated that blacklisted individuals and organisations cannot be awarded by governments as 'moral models' (which often bring certain material benefits).[177]

[175] See Article 17, NDRC [2016] No.141. Also see *Civil Servants Law, 2018.*
[176] *Id.*
[177] *Id*. Measure 8.

The problem for legality here is that, as mentioned above, the Chinese Legislation Law requires that when the state is imposing extra obligations on citizens and organisations by way of administrative actions (such as regulatory sanctions), this imposition must be stipulated under NPC-enacted Laws. The *2014-2020 Blueprint* is not a NPC-enacted Law but a State Council regulation, and therefore is, on the face of it, unable to provide legitimacy to these JSMs. Nevertheless, there is ambiguity here in the interpretations of 'extra obligations'. As these measures are mostly restrictions on benefit-granting state behaviours or increase inspection frequencies, it can be said that they are not legally, as a matter of Chinese law, exerting extra obligations.

For example, the only three other measures in the JSM for the SPC debtors list that are solely based on the *2014-2020 Blueprint* are (1) requiring the foreign exchange agency to 'carefully consider' the annual quota granted to Qualified Domestic Institutional Investors (QDII) and Qualified Foreign Institutional Investors (QFII) (Measure 8); (2) requesting the NDRC, MOHRSS, and the Treasury, among others agencies, to restrict subsidies and social security support (Measure 10)[178]; and (3) asking all market inspecting and industry regulatory departments to heighten routine inspections on those enterprises on the debtors list (Measure 12)[179]. Since these agencies are only restricting extra benefits to the debtors or increasing the chance of finding their law-violating behaviours, these measures can be interpreted as not necessarily imposing 'extra obligations' (with comparable negative impacts to a direct administrative punishment). It thus could be argued that in terms of legality these measures are not in clear violation of the *Legislation Law*.

However in terms of legitimacy, the lack of a national Law's backing brings more problems. The substantive legitimacy of some measures in the JSMs are undermined particularly considering that the value of disclosure provided by the 2008 OGI Regulation is not very effective in granting legitimacy to the joint enforcement on parties whose behaviours are not punishable under an agency's original legal mandate. This also explains why the NDRC drafted and proposed a national Social Credit Systems Construction Law (Draft 2022) for the NPC addressing the legitimacy concern of the JSMs.[180]

In addition, more efforts were made by the NDRC besides its proposal of a Law on the SCS. These efforts include the NDRC and the Central Bank together establishing and updating

[178] *Id*, Measure 10.
[179] *Id*, Measure 12.
[180] NDRC, Social Credit System Construction Law (Draft for public consultation, 2022).

an annual 'Foundational List' (基础清单) for all JSMs in China since 2021, with the current version being the *National JSM Foundational List 2024*. These 'Lists' state the sanctioning measures and their legal basis clearly in a chart, and in 2024 there are only three categories of the sanctioning measures with a total of 14 measures. These 14 measures have remained consistent throughout 2021-2024.[181]

These efforts of the NDRC and the Central Bank were required by the State Council and the Central Committee of the Communist Party of China (CPC) in 2020 and 2022 with two top-level policy instructions on providing more central supervision, procedural safeguards, and substantive legitimacy to the JSMs.[182]

5.2.4 Section conclusions

This section has revealed the nature, roots and legal basis of the blacklists and JSMs. in contrast to the scoring systems discussed in Chapter 4, it has been that these SCS systems are code-driven automation of mainly civil procedures and state regulations. It has been revealed that the institutional roots of these systems lie in the Supreme Court's 2013 debtors list and its JSM.

Empirical evidence was used to show how automated law enforcement scaled across the nation after 2016 in three waves. Two examples were provided, respectively, of the regulations of the energy sector and of the blacklisting of employers who seriously delay salary payments to domestic migrant workers. Both are reported to have achieved scaling effects at the national level.

It is followed by an analysis of the legality and legitimacy of these automated enforcement systems, as evidence of code layering on law. The blacklists doctrinally have a better legal foundation compared to the JSMs, in terms of both legality and legitimacy. The JSMs however were provided with more procedure safeguards since the 2017, and in particular 2021, which aimed to solidify its legal foundations. This evidence shows that legislative feedback can help recalibrate or counter the negative layering effects that happened during the scaling of algorithmic governance.

The following section provides a more focused case study on contract and debt enforcement in Shanghai's SCS, which is a local branch of the Supreme People's Court's

[181] NDRC, 2024, https://www.ndrc.gov.cn/xxgk/zcfb/ghxwj/202402/P020240228351343988147.pdf

[182] State Council 2020 https://www.gov.cn/zhengce/content/2020-12/18/content_5570954.htm, Central Committee of the CPC and State council, 2022 https://www.gov.cn/zhengce/2022-03/29/content_5682283.htm

debtors list. This case further explains the workings of code-law complementarity and how scaling and layering are evidenced by interviews data collected from local enforcers working with these systems.

5.3 Case study: an assessment of code-driven SCS using interview evidence from Shanghai

The section provides fieldwork-based evidence to examine the use of code-driven automation in daily law enforcement practices, within the wider framework of the SCS. It focuses on the case of contract and debts, and their enforcement in Shanghai, one of the largest cities in China with a population of 25 million in 2021 (about three times the population of London in 2021)[183]. The interview evidence, combined with online archival research, uncovers a more complete understanding than previously available of the detailed workings of these systems and of their perceived impacts, both positive and negative, in the field.

It is argued in this section that in order to prevent and / or recalibrate 'concept drift' created by automation, human agency and expertise from civil services and judges are needed. These human interventions serve to continuously realign and re-fit code-based systems to text-driven laws and social norms, operating in specific spatio-temporal environments. The section starts with an introduction, followed by a detailed flowchart description of the SCS as automated law enforcement, a discussion of interview evidence, an analysis of the evidence, and a section conclusion. The details of the interview method, as well as the reason for choosing Shanghai as a case study, have been previously explained in the methodology chapter.

This case study starts by contributing to a more informed understanding of the SCS based on evidence. It shows that blacklists and joint sanctions together constitute the two automated actions of law enforcement required by Chinese laws: (1) automated sharing of enforcement information and (2) automated penalisation or rejection of access / privileges.

Existing literature on automated law enforcement has discussed its negative consequences, such as de-skilling the police (Joh, 2019), the potential of restricting freedom afforded by law (Yeung, 2017), reducing contextualised, ethical judgements, removing useful indeterminacy (Hartzog et al., 2015; Shay et al., 2016), and having chilling effects online (in the case of copyright enforcement)[184]. However, there are very few detailed empirical studies

[183] Shanghai population statistics from the National Bureau of Statistics of China, https://data.stats.gov.cn/english/easyquery.htm?cn=E0103. (In Chinese).

[184] J. Nathan Matias, et al., 'Do Automated Legal Threats Reduce Freedom of Expression Online? Preliminary Results from a Natural Experiment' (OSF 2020) <https://osf.io/nc7e2/> accessed 7 June 2023. J. Nathan Matias et al., 'Comments

of these effects, and those with case studies often largely focus on the police settings in developed countries, such as criminal or traffic law enforcement (with one exception looking at traffic enforcement in developing cities where population is dense and with poor infrastructures, Saha 2020). Moreover, despite the logics of automation being similar across these cases, in the sense that they rely on machines to replace human action, this line of literature has not so far discussed any case like China's SCS, which mainly focuses on automating information-sharing and civil and administrative sanctions。

5.3.1 A flowchart of civil procedure in the SCS and law-code interactions

5.3.1.1 Civil procedure in two parts: human judges in adjudication and enforcement

This section first draws an original flowchart (*figure. 5*) as below to explain how law and code work together in enforcing civil procedural laws in China under the SPC debtors list and its JSM.

responding to US Copyright Office Notice of Inquiry Docket No. 2021-10', (OSF 2022), <https://osf.io/498ja/> accessed 7 June 2023.

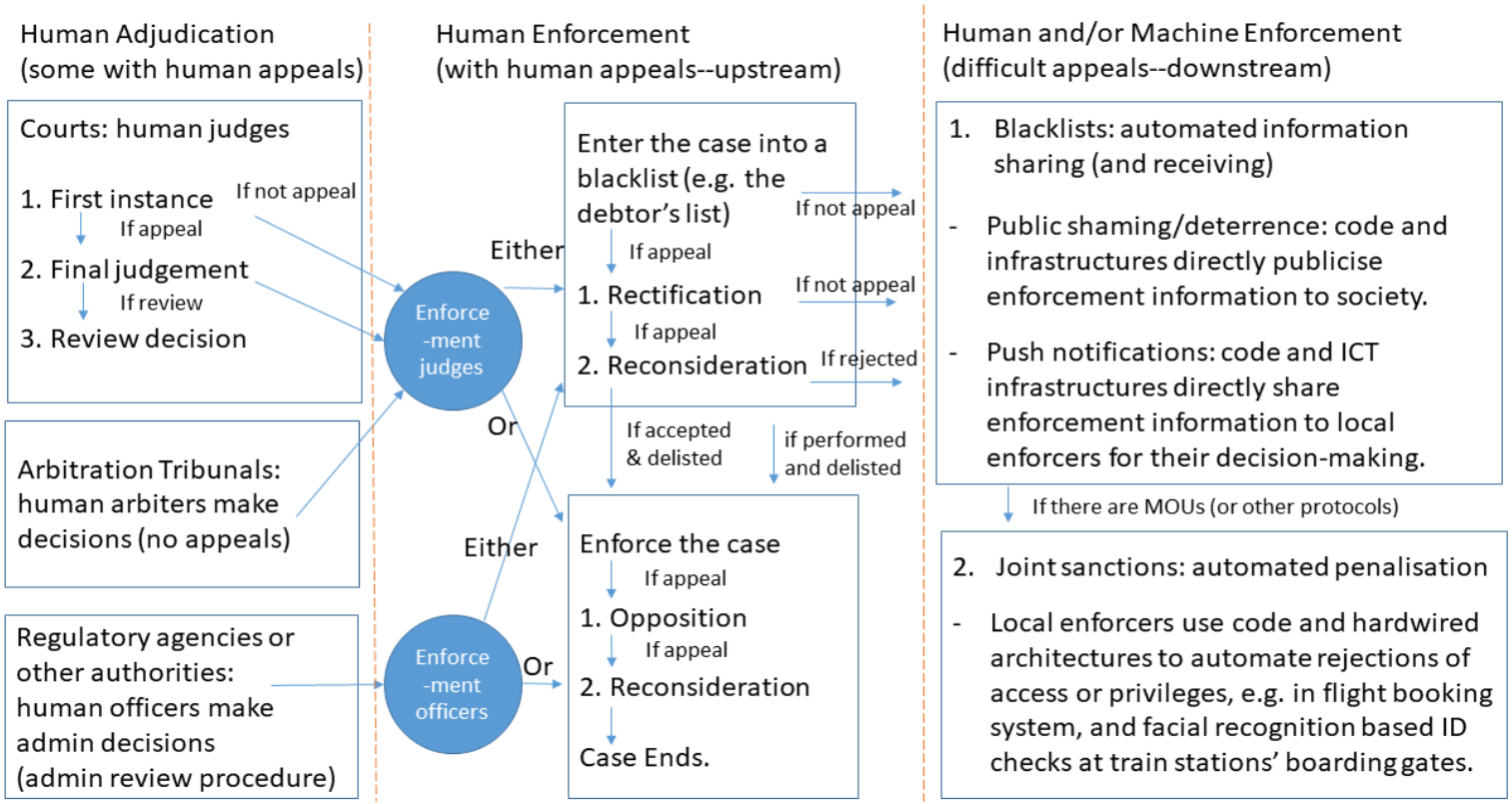

Figure 6. A flowchart of how human and machine interact in civil procedures in the SCS

According to China's Civil Procedure Law, legal enforcement has independent procedures after the adjudication stage. It also requires specialised staff, namely the enforcement judges who form a department separate from the adjudication bodies of courts in China. These enforcement judges oversee and operate the daily tasks of reviewing the written 'application to enforce' (*shenqing zhixing shu*, 申请执行书) sent by litigating parties to claim for, e.g. debt payments or compensations[185].

These enforcement judges decide which and when cases can be moved from the adjudication procedure to the enforcement procedure. They are also responsible for tracing and examining, for example, a debtor's enforceable financial assets or property, and finally ensuring the legally required transfer of money and / or property is made from one party to another on time (sometimes involving freezing, seizing and auctioning these assets). At the end of the day, their jobs only finish when debts or compensations are paid in the physical world and confirmed by the litigating parties, and the enforcement cases can thus be labelled as 'closed' on the court's digital files system for the court administrators (i.e. the secretary, IT and human resource department) to archive and review [186].

The enforcement judges can only start an enforcement procedure usually after a final verdict is made by an adjudication body, such as a court or an arbitration tribunal. In China the court system follows a single appeal procedure (plus review), which means the second instance judgment will be the final verdict (*ershen zhongshen*, 二审终审). The arbitration tribunals, however, follow a no appeal procedure (*yishen zhongshen*, 一审终审)[187].

To take a closer look at the operation of the SCS, the author conducted an online search on Shanghai's debtors list on a date randomly chose (June 26th, 2022), and picked the first two names on the list[188].

(1) The first one is a named individual 'W' (whose full name in Chinese is published online, but pseudonymised here for privacy reasons). W's case (2020) was based on an arbitration verdict which was the final verdict upon which the enforcement procedure

[185] For a primer see the official online SPC litigation guidelines on how to apply to enforce a court judgment. SPC, 'Necessary Notes on the Execution Procedure' [执行程序须知] (SPC 26, Jan 2016) <https://www.court.gov.cn/zixun-xiangqing-78732.html> accessed 7 June 2023.

[186] *Id.* Also see China's Civil Procedure Law (2017) Article 225 mainly and Article 227 (Article 232 and 234 after the 2021 Amendment).

[187] *Id.*

[188] Shanghai Municipal List of Dishonest Judgement Debtors [上海市失信被执行人名单] (Shanghai High People's Court, 26 June 2022)

was initiated, or more accurately, 'recovered' (*huifu zhixing*, 恢复执行) after certain enforcement-suspending circumstances were removed in 2022[189].

(2) The second one is a manufacturing company Shanghai Fengyi, which in contrast, involved a first instance court's judgment (2021). This was the final verdict as the parties decided not to appeal, and hence the basis for the initiation of enforcement procedures by enforcement judges[190].

These enforcement procedures also have a two-appeals safeguard under the Civil Procedural Law and its 2015 SPC judicial interpretation. According to these laws, courts (same and higher level) can review a decision if there have been mistakes or unfair practices during enforcement. Enforcement applicants thus can use these procedures to explain their specific situations to and / or check the work of enforcement judges.[191]

5.3.1.2 Blacklists: invoke the dormant information-sharing procedures and automate them

The first type of machine involvement in the enforcement loop is by automating laws related to blacklists. China's highest court, the Supreme People's Court (SPC), first issued a Judicial Interpretation 2013 that established the national Debtors List, which is a proto-typical text-driven blacklist law (and later amended in 2017).[192]

---

[189] See 2022 W's case (2020, 2022) Case No. 0677 of Shanghai Arbitration [沪仲案字第 0677 号]. Shanghai 02 Execution Reinstatement Case No.72 [沪 02 执恢 72 号] (Shanghai High People's Court, 2022).

[190] See 2022 Shanghai Fengyi case (2021, 2022) Shanghai 0120 Civil First Instance Case No.11522 [沪 0120 民初 11522 号]; Shanghai 0120 Execution Case No. 3708 [沪 0120 执 3708 号] (Shanghai High People's Court, 2022)

[191] For enforcement appeal procedures, see China's Civil Procedure Law (since 1982, with 2013, 2017 and 2021 amendments) and the SPC 2015 interpretation 'Regulations on Several Issues Concerning the Handling of Execution Objections and Review Cases by the People's Court' [(关于人民法院办理执行异议和复议案件若干问题的规定)]. Also see the official primer for case execution applicants by the SPC (n43). More details see the later parts on appeal. Recent Smart Court developments also allow for an online real-time update of (enabled by Wechat) and even the live-streaming of the enforcement judges' daily work. The enforcement applicants can thus watch, for example, enforcement judges go on trips to look for remaining assets of a judgement debtor and provide real-time feedback on the judges' work. Interview with judges M and N. Automation, or data-driven facilitation, of *adjudication* procedures falls into the Smart Court project rather than the SCS project. This includes, for example, the use of legal-predictive analytics, case recommendation systems and algorithmic case allocation systems. Also see Chinese literature on Smart Court in general, e.g. Zuo W (2018).

[192] Supreme People's Court's Several Provisions on the Disclosure of the Dishonest Judgment Debtors List Information, (SPC Judicial Interpretation No. 17, 16 July 2013) [最高人民法院关于公布失信被执行人名单信息的若干规定 法释〔2013〕17 号, 2013 年 7 月 16 日] <https://www.chinacourt.org/law/detail/2013/07/id/146217.shtml> accessed 7 June 2023; Later amended in 2017, Supreme People's Court's Decision on Amending the 'Supreme People's Court's Several Provisions on the Disclosure of the Dishonest Judgment Debtors List Information', (SPC Judicial Interpretation, No. 7 2017).[最高人民法院关于修改《最高人民法院关于公布失信被执行人名单信息的若干规定》的决定，法释〔2017〕7 号. 2017.2.8] <https://www.chinacourt.org/law/detail/2017/02/id/149233.shtml> accessed 7 June 2023.

The *2013 Interpretation* essentially takes the quality of national law and regulates the processing and sharing of debtors' information. It interprets two articles of China's *Civil Procedure Law* (2013, later amended in 2017, 2021 and 2023), that are Article 13 and Article 255 (now Article 266 after the 2023 amendment).

Article 13 of the *Civil Procedure Law* (2013) sets out the principle of honesty and credibility in civil litigation, and Article 255 is a broadly worded mandate for the publication and sharing of non-enforced judgement information. More specifically, Article 255 requires (1) the full recording and publicising of non-enforced obligation information (*bu lvxing yiwu xinxi*, 不履行义务信息) of the listed subject to credit-reporting agencies and media, and (2) limiting the listed subject from travelling abroad (*xianzhi chujing*, 限制出境).[193]

Article 255 on information sharing and border control was a difficult procedure to enforce before the *2013 Interpretation* established the details for building a national blacklist system. The *2013 Interpretatio*n, by initiating the blacklist system, essentially created two concrete actions under this Article 255 *ad hoc* procedure, which the chapter names respectively as 'push notifications' and 'public shaming / deterrence'. These actions then can be (semi)automated by machines.

(1) First, push notification refers to the direct information-sharing between agencies in order to coordinate economic penalisations by agencies under laws or social norms, which save the agencies' information-searching and decision-making costs. (2) Second, public shaming or 'deterrence' means the socio-financial risk-labelling of the debtors as both 'dishonesty' and 'has the capacity but refuses to perform' legal verdicts.

According to the interviews, these two easily automated procedures are used by enforcement judges to reduce their workload.

(1) First, the enforcement judges have large discretion in deciding whether a case should be entered into the debtors list.[194] In fact, in most cases the reason for listing falls into the vague 'other situations' with limited evidence[195]. According to a respondent who is a public official in Shanghai, once a case is listed, the enforcement judges only need to wait for further evidence of a debtor's enforceable assets to emerge in order to decide whether to restart actively pursuing

[193] *Civil Procedure Law 2013*. Article 13, Article 255. Notice of the Supreme People's Court on Conscientiously Implementing the 'Several Provisions on the Disclosure of the Dishonest Judgment Debtors List' (SPC, 2013)

[194] This is because of an ambiguity in the legal text that allowed the enforcement judges to enter a case into the debtors list with only preliminary evidence. A vague article in the Judicial interpretation (2013, 2017) on 'other situations (*qita,* 其他) where one has the capacity but refuses to perform'.

[195] See e.g. cases of W 2022 and Shanghai Fengyi 2022. Any day's random search will return similar results on the reason for listing.

those assets. This makes the debtors list in essence a database for judges to 'buy time' in their work, and clear their backlogs of written enforcement applications that have yet to be reviewed[196].

(2) Second, once a case is put into the debtors list, the push notification and public shaming / deterrence procedures are automatically triggered. This then does not require more efforts from the enforcement judges to follow up or actively close these enforcement cases. Enforcement judges no longer need to, for example, proactively search for hidden financial assets or properties of the listed debtors. The responsibilities are then shifted to the courts and other agencies to negotiate on information-sharing, and to the IT department in courts to write the code and build the hardwire infrastructures, according to two respondents in Shanghai working with these systems. Once built, these machines run enforcement procedures automatically at scale. They only require a human enforcer's label inputs and periodical IT maintenance[197].

According to two respondents working in one of Shanghai's intermediary courts, these code infrastructures are deemed as cost-efficient investments by the Supreme Court in saving human enforcers the time and effort. They said that local courts are often short of enforcement judges facing the large number of backlogs of contract and debt enforcement cases. The debtors list to some extent externalises the enforcement judges' work to other public and private agents, like government procurers and commercial banks, who will use this blacklist information to reject lending or procurement bids from the debtors. These negative consequences in turn push the debtor to comply with the court rulings and save local enforcers time and effort in investigating each case.[198]

---

[196] Government official (supervisor rank) Shanghai, 2021. It is worth noting that the SPC debtors list (since 2013) oversees Shanghai's local debtors list. They are the first widely used national blacklist and specifically aimed to resolve the 'difficult problem of judgment enforcement' (*zhixing nan*, 执行难) in China's rule-of-law and judicial reform (*sifa gaige,* 司法改革), at a time when high non-enforcement rates was a problem that endangers litigants' respect and confidence in the court system. See *Notice of the Supreme People's Court on Conscientiously Implementing the Several Provisions on Announcement of the list of Dishonest Persons Subject to Enforcement (SPC, 2013).* These blacklist laws on enforcement are generally overseen by the National Development and Reform Committee (NDRC), and from national to local levels, have a unifying goal to enhance enforcement of laws and regulations at a time when human public enforcement resources are limited 'Establish a system of lists for joint incentives for trustworthiness and joint sanctions for untrustworthiness.' Guidance Opinion on Strengthening and Regulating the Management of Lists for Joint Incentives for Trustworthiness and Joint Sanctions for Untrustworthiness by the National Development and Reform Commission and the People's Bank of China. (NDRC, [2017] No.1798) [国家发展改革委 人民银行关于加强和规范守信联合激励和失信联合惩戒对象名单管理工作的指导意见.]

[197] See interview with a manager, Shanghai 2021.

[198] These code infrastructures are also welcomed by cooperating agencies whose enforcers also reduced daily work of searching for non-enforcement information and deciding whether to afford service or privileges, such as public funds, to certain applicants. See interviews in Shanghai 2021. This Supreme Court-made 2013 Judicial Interpretation later inspired multiple blacklist laws and regulations promulgated by regulatory agencies. These agencies invoked similarly 'dormant' clauses in other legislations, e.g. energy sector regulations, environmental protection regulations, and labour protections. See

5.2.1.3 Joint sanctions: extended automatic penalisation and upstream appeal procedures

The second type of machine involvement in enforcement is the automation of joint-sanctions, where different public and private regulators restrict access of services to debtors (through the JSMs as discussed above). For example, boarding gates at airports and train stations, with AI facial recognition cameras embedded, can automatically reject debtors' access to services like flights and high speed trains. These access restrictions, such as limitations on travel methods deemed as high consumption, are aimed to push the debtors themselves to comply with court rulings, and thus save local enforcers the time and effort of investigation.

The (semi)automated joint sanctions, in essence, extend the debtors list's push-notification and shaming / deterrence functions by imposing automatic penalisations of debtors. The penalisations are effectuated through code-drive infrastructures, like the boarding gates embedded with AI facial recognition technologies. With the joint-sanctions carried out by human and / or machine enforcers at each end-point of the debtor's spending or travelling, the enforcement judges have less incentives to investigate the undiscovered assets and properties of that debtor. Instead, the enforcement judges can wait for the debtors to reveal their assets (for example, when those debtors are paying their children to go to expensive private schools), or to pay their debts with their own initiatives when the negative restriction become too inconvenient (for example, when the debtors are restricted from booking flights, or winning a public procurement bid).

Similar to most blacklists, these JSMs are also made by negotiations and agreements between public and private agents. The Supreme People's Court (SPC) as mentioned above, was not content with just invoking the dormant public information-sharing procedures with the debtors list in 2013. It wanted to minimise the discretion of public and private local enforcers in their imposition of sanctions once they received the debtor's information. With this aim, the

Notice on Issuing the Memorandum of Understanding on Joint Punishment of Serious Violations and Dishonest Market Entities and Related Persons in the Electric Power Industry (NDRC and others, No.946, 2017) [印发《关于对电力行业严重违法失信市场主体及其有关人员实施联合惩戒的合作备忘录》的通知 2017]; Notice on Issuing the Memorandum of Understanding on Joint Punishment of Serious Violations and Dishonest Entities in the Oil and Natural Gas Industry (NDRC and others, No.1455, 2017) [印 发《关于对石油天然气行业严重违法失信主体实施联合惩戒的合作备忘录》的通知 2017]; Notice on Issuing the Memorandum of Understanding on Joint Punishment of Dishonest Production and Operation Units and Their Related Personnel in the Field of Environmental Protection (NDRC and others, No.1580, 2016) [印发《关于对环境保护领域失信生产经营单位及其有关人员开展联合惩戒的合作备忘录》的通知]; Notice on Issuing the Memorandum of Understanding on Joint Punishment of Units and Their Related Personnel Who Severely Arrears Wages of Migrant Workers (NDRC and others, No.2058, 2017) [印发《关于对严重拖欠农民工工资用人单位及其有关人员开展联合惩戒的合作备忘录》的通知]. For.

SPC negotiated two Memorandum of Understandings (MOUs) respectively in 2014 and 2015 with the China Railway, and the China Aviation Network. Both MOUs are facilitated by the National Development and Reform Committee (NDRC).

These two MOUs are text-driven legal documents that concretised the required actions of China Railway and China Aviation Network. These two state-owned enterprises then use code and ICT infrastructures to translate these laws into the physical actions of their boarding gates and / or employees on-site. These human and machine actors who control the access to service then can refuse the booking transactions and / or boarding of the debtors. Other MOUs established between the SPC and other public or private agents follow a similar implementation logic.

In the perspective of code-driven laws, the computational code directly enforces these MOUs, blacklists, and other laws, regulations and individual legal verdicts. The text-driven laws were coded and built into the hardwired infrastructures of the front-line enforcers, that are machines supervised by humans. For example, these machine enforcers include the online booking systems of high-speed trains and flights, and the boarding gates at train stations and airports with ID checkers and facial recognition scanners. The joint sanctions are thus the automatic executions of penalisations that merely extend the 'push notification' procedure required by blacklists. These joint sanctions reduce the involvement of human enforcers at local levels, or make it very costly for the local enforcers to deviate from their centrally instructed sanctions, such as rejection of booking or entry, usually required by blacklist laws.

Even though (semi)automation reduces the discretion of enforcement, there is an in-built appeal procedure of 'rectify and reconsider'. This appeal procedure, however, is a civil procedure placed at the start of the blacklist procedure, i.e. when the enforcement judges / regulators make the decision on whether an enforcement case should enter a particular blacklist. In the flowchart of automated enforcement, this 'rectify and reconsider' procedure is situated upstream, which cannot be used at the downstream users' ends.

This 'upstream-blacklists, downstream-joint sanctions' structure entailed by the civil procedure laws on enforcement made the appeal procedure centralised at the upstream control of the enforcement judges. According to the SPC's latest 2017 Amended Judicial Interpretation, which added Article 12 to the 2013 Judicial Interpretation, the debtor can first 'apply to rectify' (*shenqing jiuzheng*, 申请纠正) a decision of the enforcement judge which enters a person's case into the debtors list database. This application should be reviewed within

15 days, and if successfully accepted, the rectification should be made within three working days; otherwise the rectification application will be rejected (Article 12).[199]

If the applicant (debtor) is unsatisfied with the rejection of their appeal made by an enforcement judge, they can apply for a 'reconsideration' (*fuyi,* 复议) at a higher level court within ten days of the rejection decision. The higher level court will have to make a decision within 15 days. The enforcement process will not be paused during the period of reconsideration (A12)[200].

5.3.2 Interview evidence from local enforcers: Perceived gains and loss of automated enforcement

'Social Credit governance is a direction, because governance only by coercive force cannot work. There needs to be people's recognition from their heart, so that we don't need to give sanctions every day.'

– Government official, Shanghai, March 2021

5.3.2.1 Perceived gains: efficient scaling, effectiveness and informed decision-making

Blacklists and joint sanctions have gained wide recognition in achieving enforcement efficiency and effectiveness, and better informed decision-making from the interviewees. The degrees and focuses of gains, however, are perceived within a variety of ways.

A respondent who is a government official (supervisor rank) agreed that there had been positive scaling of the SCS where judgement compliance rates increased. However, this official also worried about the potential *over-effectiveness* of the system. This could undermine its own legitimacy. The danger they stated was that courts and governments overused the SCS, creating new forms of path-dependencies.

Another public official (middle rank) respondent claimed that the debtors list is good in particular for raising the expectations of corporations of the need to abide by laws and contracts,

[199] Article 12 of the *SPC 2017 Amendment*.

[200] *Id.* This is different from a litigation appeal under China's Civil Procedure Law, but almost the same to the enforcement case appeal procedures, i.e. the 'opposition and reconsideration' procedure under A225 of the Civil Procedure Law (2017, and Article 232 after the 2021 Amendment). See the 15 days of review, and if rejected within ten days apply for reconsiderations at a higher court. Provisions of the Supreme People's Court on Several Issues concerning the Handling of Enforcement Opposition and Reconsideration Cases by People's Courts (SPC Judicial Interpretation, [2015] No.10). This proximity indicates that the listing decisions by enforcement judges is also deemed as a standard enforcement action in civil procedure which can be appealed.

which ‘is good for Shanghai’s general business environment and the examinations of such environment (by the IMF’s Doing Business Report).’ This respondent also stated that other blacklists and joint sanctions are very useful for custom high-level accreditations, tax inspections, public finance projects approvals and so on. In addition, the respondent mentioned that in the state of emergency during Covid-19, the temporary blacklists of those who severely violated health regulations were also useful for public health governance. Another respondent, a public data centre manager (middle-rank), claimed similar effectiveness and efficiency achieved by the blacklist systems, and emphasised the potential economic values of the enforcement data being processed, shared and maybe even transacted on the data market.

Four other respondents, as front-line government officers, commented that the SCS automatic enforcement is particularly effective in cities like Shanghai because people here are less afraid of direct sanctions, but care more about their business reputation. Compared to centralising the SCS blacklists and joint sanctions, a public official respondent in particular believes that local governments and legislators should have more space to make their own automated enforcement systems.

Another respondent, who is a state-affiliated corporate employee (front-line manager) of a credit-reporting agency in Shanghai, also argued that the blacklist information the agency pushes to different agencies, namely their customers, were considered as useful in their decisions-making or automated enforcements. The respondent claimed that Shanghai's SCS information-sharing project started in collaboration with Shanghai’s High Court; negotiations with different government departments on sharing information were later coordinated by the local NDRC and courts. Even though the negotiation, maintenance and standardisation of code and ICT infrastructures took time and effort in the past years, the respondent believed those obstacles were worth overcoming for enforcement and informed decision-making. A few other respondents who are employees of a state-owned enterprise in Hefei also share this opinion on the positive scaling effects. They manage the data of SCS and smart city projects for the Hefei government, which learnt largely from Shanghai.

Two respondents, who are court administrators / judges at an intermediary court in Shanghai, did not directly comment on the blacklists’ effectiveness. They however briefly discussed the huge amounts of enforcement work accumulated at the court due to the Person-to-Person (P2P) financial fraud cases in the past few years. They explained that the enforcement judges in the court were focusing on these big caseloads using new technologies like block-chain and ‘enforcement live stream’ on WeChat, but did not discuss if the debtors list was

useful. Instead, the court administrators suggested some alternatives to the automated enforcement measures of blacklists and joint sanctions in e.g. the Hangzhou's internet court.

5.3.2.2 Perceived loss: rigidity, human perverse behaviours, and institutional overload

Rigidity of enforcement

The two respondents who are judges pointed out some limitations of the Shanghai debtors list. They also referred to alternative arrangements other courts were making in order to let the debtors have certain flexibility in terms of automated sanctions. They claimed that:

> 'The automatic enforcement by the SPC is very rigid (*gangxing*, 刚性). Once listed the limitations on high-expenditure activities are automatically controlled by the SPC (in Beijing). Due to technological coerciveness (*jishu qiangpo xing*, 技术强迫性), an enforcement case once listed cannot be closed unless it goes through the SPC. We wished to have the local enforcement judges operate the details, but it proves to be infeasible.'

> 'This might not be the best for a company's development. It is like giving water for fish to grow–we need to let them slowly pay back their loans. In some Hangzhou courts there are new buffering attempts of 'instituting credits' (*shuxin jizhi*, 树信机制), where the courts underwrite for the debtors so that they can operate normally to repay their debts. As in if you suddenly restricted someone's business it is for sure they won't be able to repay their debts. Some buffering measures can be more effective.'

Their views indicate that there are certain counter-productive aspects embedded in the rigidity of the automatic law enforcement systems. The SPC debtors list and joint sanctions, which have a very centralised code and ICT architecture, are perceived as sometimes inefficient when individual cases require more flexible interpretations and deviations from clear-cut penalisations in local environments.

Human perverse behaviours and dis-incentivisation

A government official respondent (supervisor rank) stated that the redundancy of human talents and disincentivisation are key problems.

> ‘Currently our enforcers are not publicly well-rated. When a case cannot be enforced, the enforcement judges just throw it (into the debtors list). The state feeds these batch of enforcement judges in the team of judges to do things. You should fulfil your obligations. Go nudge people to repay their debts when you should nudge them, so that they can be delisted from blacklists. Otherwise this database grows larger and larger.’
>
> ‘If the enforcement judges just throw cases into the blacklists, and leave them there, this social credit institution will lose its functions. This cannot become an excuse for lazy government (*lanzheng*, 懒政).’

The respondent went on to discuss the problem of wasting public official’s human expertise, which is a public resource accumulated by training before and in the organisation. It also led to oversimplification of enforcement tasks at a place where problems should be dealt with more granularity.

> ‘In fact, the personnel department has very important responsibilities. We cannot just simplify the question (with automation). There are so many people every year taking civil service exams which consume large amounts of public resources. These people want to work properly, but you do not train them anymore…
>
> The court is the same. When you have this blacklist at the back end to resolve all things, enforcement becomes over-simplified. But this is exactly a situation that requires the granularization of institutions (*zhidu jingxihua*, 制度精细化).’

The respondent also critically stated that the normalisation of the *ad hoc* enforcement procedures is not sustainable, and the need for an accountability mechanism to re-incentivise human enforcers to delist people from the blacklists.

> 'You have to make sure that every measure has been exhausted at this step of the task (of enforcement), then consider going there (the blacklists) ... Even though every measure has been exhausted, you (the enforcer) still need to provide me with a promise on when I will be delisted…
>
> We need to clearly allocate the responsibilities in the step of social credit repair. The governments, courts or credit-reporting agencies who enter people into lists need to also be responsible for delisting. It has to become a legal obligation too. In this case the next time you enter someone into the list there will be expectations (to delist) ... This (blacklist) in essence is administration-backed hard regulations, which thus should have the same redress mechanisms with other kinds of administrative powers... This should be common sense.'

The problem of human redundancy and disincentivisation was considered by an official as leading to the overstretched institutional capacity (*zhidu rongliang*, 制度容量) of the SCS, an apprehension the dissertation calls 'institutional overload' (*zhidu chaozai,* 制度超载).

Institutional overload

A respondent who is a public official at supervisor ranks, explained the problem of over-stretched institutional capacity as follows:

> 'For example, there cannot be more than 10% of the population being blacklisted, even 3-5% of the population is too much. The proportion should have certain limits. Imagine among the 1.4 billion Chinese people, there are more than 3% people under social credit related sanctions, then this will exceed the institutional capacity (of the SCS). This is related to the quantity of the total population.'

> ‘This is a problem of institutional capacity… the database may be unable to run with such large numbers. This is not a problem of limited technological conditions, but an institutional problem. If too many people are put into the (blacklist) database, they will get used to it and stop being deterred by this (blacklist).’

This concern echoed some news reports since 2019 on the worrying number of entries in the debtors list (excluding other national and local blacklists). The reported accumulated entries between 2013-2019 almost amounted to 1% of the national population (more than 13 million entries by 2019, but may have large numbers of repetitions and without deducing those delisted)[201].

This impression of the national picture is corroborated by localised data; the accumulated numbers of people in the debtors list from 2013-2021 has exceeded over 1% of the population in some provinces. For example, the accumulated number of debtors in Zhejiang in 2021 amounts to 1.2% of the province’s population (784,100 people), while Shanghai reached 0.92% (220, 000 people), and Jiangsu hit 0.83% (704,500 people), according to a news report that scraped online data from QiChaCha (a widely used private database on corporate enforcement information in China)[202]. These three regions of the Yangzi Delta are historically highly active in commercial and financial activities while consistently constitute large proportions of China’s GDP (e.g. Li and Van Zanden, 2012). In addition, though the existing numbers of debtors should be much lower than these accumulated numbers, there are other blacklists adding to the total numbers of people who are currently under the automated enforcement effects of the SCS in these provinces[203].

In this sense, there is evidence that the above-quoted official’s worries are justified. Even the apprehension about the total number of blacklisted people hitting a 3% national population (as in the accumulated number of people who were ever listed on any kind of blacklist) is not completely unwarranted. The respondent further claimed the essence of the SCS institutions in Shanghai should be to promote social integrity, and only penalise a very limited number of people in the society.

[201] Miao Q., ‘The blacklist has nearly 13 million entries, what's the difference between Jia Yueting and Dai Wei?’ (Yicai.com, 16 Jan 2019, in Chinese) <https://www.yicai.com/news/100100754.html> accessed 7 June 2023.

[202] See e.g. Xu J., ‘There were an additional 1.34 million 'Deadbeats' in China in 2021, with a cumulative total of 704,500 in Jiangsu Province’, (Yangtse.com, 23 March 2022, in Chinese) <https://www.yangtse.com/zncontent/2052424.html> accessed 7 June 2023.

[203] See e.g. interviews with government official, Shanghai 2021. The eight blacklists in Shanghai that generate eight types of data in addition to the SPC debtors list.

‘Governance does not rely on coercive force, even though at first it might borrow some of the state's coercive power as administrative commands. But this is premised on costing the state’s credit (people’s confidence in the state). The state pays for all the resisting moods of the regulated parties…

We must look at why we built this institution. At first there is an institutional arrangement to adjust people’s ideals, such as having integrity (*jiang yiqi*, 讲义气). Being an honest (*shouxin*, 守信) person is an honour (*guang rong*, 光荣). The institution wants to avoid disintegrity (*shixin*, 失信), and adjust people’s behavioural habits, in order to make people abide by laws and social norms…

Shanghai’s social credit law has a promotional task. It should only set penalising arrangements for very few people in the society, and enforce these arrangements. It is absolutely not about regulating everyone’s behaviours using this law.’

To solve the problem of institutional overload, the respondent made several points: (1) the need to assign balanced accountability mechanisms for those in charge of the automated enforcement SCS, e.g. the enforcement judges, (2) the desirability of making these enforcers responsible for raising the public awareness of the SCS laws; (3) the need to provide alternative arrangements for debtors to perform their legal obligations before imposing automatic sanctions, and (4) the need for long-term quantified evaluations of institutional capacities of the blacklists and joint sanctions to make sure people still acknowledge the overall legitimacy and effectiveness of these automated systems.

5.3.3 Assessment: layering effects, concept drift and human agency in automated enforcement

The debtors list and its joint sanctions were developed by the SPC to increase the scale of law enforcement under the pressures of limited resources available for human enforcement; the number of judges is too small in proportion to the large numbers of non-enforcement cases.

The SPC supplied a concrete interpretation of the *ad hoc* rules (Article 255) that were practically 'dormant' in existing civil procedural texts before 2013.

Once the legal interpretation was made clear by the SPC, the 'hermeneutic gap' in Diver's words, is diminished. 'Law's affordance of delay' is also removed once this interpretative ambiguity is limited (Diver, 2022). This paved the way for the SPC and NDRC to complete the negotiation of MOUs with different public and private agencies, and then the building of code and ICT infrastructures at scale. These code and hardwired infrastructures automated both (1) the public information-sharing legal procedure (blacklist), and (2) the extended penalisation imposed by local enforcers (joint sanctions laws). This further reduced the human interpretative spaces at local enforcers' level, and thus allowed for more fidelity and immediacy in the local enforcement of these extended sanctions at large scales. The perceived gains of code-driven legal enforcement in the SCS also corroborates this scaling effect. The more formalised mode of code-driven governance, some of it embedded in ICT infrastructures, saved on human labour and enabled large-scale enforcement in Shanghai, with its high density of commercial and financial activities. This code-driven enforcement also allowed for better informed decision-making across different public and private agencies and their enforcers, which contributed to an overall improved efficiency and effectiveness of governance at scale.

However, as explained earlier, the appeal procedures of the joint sanctions are situated in the upstream blacklists end of the blacklist process. A debtor cannot explain their situation, such as the rejection of flight booking or boarding on high-speed trains, to a local enforcement officer, but has to file a rectification application to the enforcement judge if the listing decision is a mistake. Moreover, as described by Shanghai's court administrators / judges, the rigidity of the code and ICT infrastructure have led to high costs for localising the code and laws. The local enforcement judges have to go through the central SPC system to delist a debtor, without much discretion at each individual case. This rigidity of the automated enforcement system provides very little space for accommodations of specific needs at the level of granularity needed. As a respondent argued, the result can be counterproductive oversimplifications of many enforcement tasks.

This mismatch between the code-driven enforcement on the one hand and granular social reality on the other is evidence of the 'layering effect' at the spatial dimension. The mismatch requires human agency to realign and re-fit the code with laws and social norms. When the use of code-driven automatic enforcement becomes so rigid and imprinted that it can no longer represent or satisfy the real need of a niche space, bottom-up human interventions are required. The Hangzhou court's innovative attempts to underwrite protection for debtors is a good

example of human-created alternative arrangements, as explained by the two respondents who are court administrators / judges. This echoes the public official's argument to create local alternative institutions before imposing immediate and standardised sanctions. The conscious human choice at local levels to deviate from the code's path-dependence can restore law's 'affordance of delay' and the 'hermeneutic gap'. Relatedly, the human effort to realign code to its local socio-legal order can potentially prevent the negative effects of layering and maintain the positive effects of scaling. The 'layering effect' can also help us see how the SCS code is path-dependent on the highly standardised legal rules it translates, and thus guide the human efforts to change both the law and code to allow for a mitigation space before penalisations (Diver 2022).

Moreover, the 'institutional overload' problem in the SCS reveals deeper costs of the layering effects in the long-run. As described by one respondent, with human enforcement judges and officers disincentivised over time, the code-driven enforcement systems of blacklists will keep accumulating more people in their databases, reflecting code's path-dependency. These code-driven blacklists and joint sanctions can freeze people's futures by not allowing people to change through time, for example, not giving them loans to restart a business to pay back past debts. It is particularly so when the code follows a path-dependence to incentivise human enforcers to input data into the blacklists, while not giving them obligations to clear people out of the database.

More severely, the foundation of the automated enforcement element of SCS can be eroded as code-driven governance of this kind reaches its capacity limits. As explained in the apprehension of 'institutional overload', the entire governance system of automated enforcement would fail if people's foundational cognitions of laws of contract, or social norms of concepts like 'honesty' and 'justice', start to drift. When too many people are trapped by the blacklists without much hope of leaving the lists soon, they will perceive a 'new normal' that defies the current law and code. This can mean that all previous modes of governance, from social norms to text-driven laws, to the code that enforces these laws and norms, by losing their legitimacy, cease to be 'normative'. Without continuous human judgements, it is likely some code-driven institutions would have already reached their capacity and collapsed.

For example, after the 2017 Amended Judicial Interpretation added the 'rectify and reconsider' procedures (Article 12), enforcement judges in Shanghai could handle appeals and address grievances of those being mistakenly listed. These judges could also consciously limit the data size of the debtors list by setting redlines of, for example, not reaching 0.5% of the local population. It was important, according to interviewees, that local enforcers kept adjusting

the fit between code and its foundational social norms and laws to maintain their positive complementarity, and prevent problems like concept drifts. In the debtors list case, judges should make sure only 'very few people are listed', as argued by one respondent who is a government supervisor. As one respondent argued, after all the SCS is designed to promote honesty and legal compliance, rather than making people stop believing in the courts and the values that uphold the laws.

The three problems exhibited here—code's rigidity, disincentivised behaviours, and institutional overload—all constitute parts of the negative layering effects. This is a negative complementarity which happens when states and courts over-rely on code and statistics and fail to adjust the code to laws and social norms over time and space. Code-driven automation created the social 'drift' at the interior layers, which erodes the code's own functioning and legitimacy at the exterior layer.

5.3.4 Section conclusions

This section has used new empirical evidence to uncover the workings of blacklists and joint sanctions systems from the viewpoints of those who supervise, manage and / or operate them. Methodologically, the section combines new interview evidence with doctrinal and online archival research to tell a more detailed story about the SCS in a particular context (Shanghai). It provides a more complete understanding and empirical grounding which complement existing narratives or theorisations of the SCS[204].

The case study provides a thick description of the workings of these systems, and their perceived positive and negative impacts. It argues that the scaling effect of code-driven 'laws' cannot be achieved without humans consciously managing the problems of 'layering'. The perceived evidence of rigidity shows that layering happens in niche spaces, where local contingencies require deviations from the code's central control of enforcement. Fears of institutional overload demonstrate inter-temporal layering in governance. Code-driven enforcement proves to be path-dependent and complementary to the less formalised modes of governance, namely the social norms and text-driven laws. If individual perceptions of, for example, good faith or laws of contract, change due to the overgrowth of automated enforcement, code would be deprived of its own meaning within human beings' lived time-space. Like legal devices, code also 'wears out'.

[204] See Chapter 1, Section 1.4 a literature review on China's SCS.

Human agency and expertise are therefore required to continuously realign and re-fit these 'layered' modes of governance, from social norms to laws, to data and code. Only in this way, code can maintain sustainable scaling, and achieve positive complementarities between modes of governance in specific spatio-temporal environments.

5.4 Chapter conclusion

This chapter makes three major contributions. First, it provides new empirical evidence to show a more realistic picture of the blacklists and joint-sanctions mechanisms in the SCS. It uses the evidence to uncover their nature as semi-automated law enforcement systems driven by data and code.

Second, it reveals the interactions between law and code as complements instead of substitutes. Specifically Shanghai's case study provides new evidence on how code enhances law enforcement rather than replacing it in the SCS, and how dormant civil procedures on debt enforcement is activated by code at large scales.

Third, the chapter shows how the scale of governance increased when code was introduced to automate enforcement and verification processes, but with the cost of negative layering. Layering effects like rigidity of code (spatial layering), and institutional overload (inter-temporal layering) have been revealed by interviews evidence from human enforcers in their daily operations of the SCS.

More specifically, the chapter started by providing an overview and the institutional roots of the SCS blacklists and joint sanctions. It shows that the 2013 Supreme People's Court's (SPC) debtors list is an institutional prototype for blacklists. It started as a concretised judicial interpretation of an existing dormant *ad hoc* rule in China's Civil Procedure Law (Article 255).

This new SPC juridical interpretation was then enforced with code and ICT infrastructures for public information sharing, which was constituted of both (1) 'push notification' to agencies, and (2) intergovernmental enforcement of sanctions. Empirical evidence from Shanghai shows that the ambiguity of the judicial interpretation allowed enforcement judges to mass-label cases into the debtors list. In this way human judges used the new semi-automated debtors list to save time and work, in the face of mounting case backlogs. The code-driven blacklists also reduced the information-searching efforts of human enforcers in other agencies, who need this court information for their own decision-making processes.

In addition, the 'push notification' part of SCS requires extensive negotiations and coordination in unifying the code and hardwired infrastructures to scale up automation. This

led to the NDRC's involvement in establishing MOUs between agencies, which extended the 'push notification' to the joint sanctions mechanisms. The MOUs required immediate and full enforcement from local enforcers, which incentivised building of new code and infrastructures at local levels for joint enforcement. This extended automation while reducing the workloads of front-line enforcers, further squeezing their decision-making spaces. These code-driven joint sanctions, nevertheless, also increased the fidelity of local enforcement required by central interpretations of laws and regulations. The joint sanctions, as a result, became the direct code-translation of the MOUs and related substantive laws that entail penalisation. To this point, the blacklists and joint sanctions have become legal technologies that use code to directly represent laws. They are embodied in computational code that, for example, automatically reject a debtor's flight bookings or public funds applications.

Meanwhile, problems of (1) code's rigidity, (2) counter-productive incentive change, and (3) institutional overload were raised. First, human appeals from the enforcement of these code-driven 'laws' are difficult to make at local levels. Local enforcers have to refer the case back to the enforcement judges / officers who are 'upstream' in the chain of automation. These front-line officers, including some of the judges involved, do not necessarily have the authority and / or system permission to alter the code's decisions. The hardwired code structure of this long stream of enforcement procedures, as seen in the flowchart, can become rigid and costly in niche environments. This has led to the Hangzhou court's reported innovative attempts (as revealed in the interviews) to provide public underwriting for debtors and help them operate through difficult periods of times.

Second, the automation of enforcement made local enforcers like judges / civil servants less proactive in fulfilling their duties. They in turn mass labelled cases into blacklists, leading to an overgrowth of data entries. As noted by an interviewee at the supervisor level, this ever-growing number of people being blacklisted can lead to a general change of perceptions over the system's overall legitimacy and purpose. If the blacklisted individuals start to realise that so many others around them are also blacklisted, they would perceive it as 'normalised' and thus stop believing the system as 'just' or 'promoting honesty and law-abiding values'. This potential problem of 'institutional overload' requires re-incentivising human judges / officers by accountability systems, as suggested by the interviewee. Human agency and judgement are required to continuously limit blacklist data entries and monitor the self-execution of code infrastructures at local levels. Laws should change and require human enforcers to be accountable for new jobs, such as clearing people out of the blacklists. Local enforcers should

be required to diligently root out enforcement errors, and ensure the self-execution of code do not overburden individuals and the system as a whole.

Finally, these thick descriptions of automated enforcement systems in the SCS evidence the governance framework of 'scaling and layering' proposed by this thesis. It is argued that such automated legal technologies represent a highly formalised mode of governance, namely 'code'. This high formalisation of code provides scale effects in governance, as compared to other less formalised modes of governance such as social norms and text-driven laws. The high levels of standardisation and fewer ambiguities in interpretations of code are shown to be further entrenched in physical infrastructures. This increased level of formalisation guarantees fidelity of the human and machine enforcers' actions to the laws. Governance by code, however, does not allow for the 'hermeneutic gap' or 'delay by law' embedded in human interpretations of texts, and showcased negative layering effects. These negative effects, as mentioned, can be prevented by institutional actors within their organisations by human feedback, training and institutional knowledge.

Chapter 6 Data Experiments: scaling and layering towards a national digital economy

6.1 Summary of argument

This chapter is concerned with data and algorithmic governance in the co-production of a large-scale, national digital economy by the Chinese state and market. It is a case study on China's various data experiments which aims to answer the overall research question of the thesis: to what extent can text-based law be computed systematically, namely replaced / substituted by computer code? It proceeds by addressing the following sub-questions:

- What are the emerging approaches or 'strategies' that the Chinese state adopts to the governance of massive amounts of public and private data and code in the contexts of wider local and national data experiments?
- How do data and code, as modes of governance, work in practice in these experiments?
- Is there evidence of scaling and layering in these experiments?

The chapter empirically investigates these questions by analysing China's data experiments in practice. The chapter conceptualises China's strategies in creating its national digital economy as 'institutionalisation', 'activation' and 'layering', by drawing from theories of institutional law and economics, especially layering from the institutionalist framework proposed in Chapter 3. It argues that these three types of strategies deal with China's three real data predicaments: firstly, security and compliance concerns under concentrated corporate (and state) data power; secondly, the fragmentation of data access and use; and thirdly, the lack of valuable data use and long-term, sustainable scaling.

More specifically, (1) first, the chapter uses 'institutionalisation' to describe the state's adoption of formal institutions to co-produce the economy by inducing informal knowledge

exchange with private actors. This strategy addresses data harms / risks and stabilises common knowledge in the epistemic communities of data governance, which can potentially lead to a soft 'Beijing Effect' similar to a Brussels Effect[205]. Evidence appeared intensively in China in summer 2021, including the cyber / data security review after DiDi's initial public offering (IPO) at the New York Stock Exchange, the completion of China's data protection regime by passing the Personal Information Protection Law (PIPL) featuring 'gatekeeper' and managerial liabilities, and antitrust practices that restrain data power and facilitate interoperability.

(2) Second, the state creates flexible legal and policy spaces for the practical experimentations of 'activating' data transaction, sharing, and use. Evidence of the activation strategy is drawn from the struggle between Ant Group and the Central Bank over personal credit information, the competing approaches to data property between Shenzhen and Beijing, and the case of 'Party manages data' in Tianjin.

(3) Third, in order to achieve valuable public data applications, the state adopts the strategy of 'layering' that is to build long-term data habitat / ecosystems with in-depth understanding of the institutional contexts. Evidence includes practices in Smart Court and social / financial credit systems, complemented by other secondary sources from Huawei, Alibaba and research institutes.

In conclusion, it is categorised by the chapter that three data strategies, namely institutionalisation, activation and layering, have emerged that help the Chinese state co-produce the digital economy, contest data property, and co-evolve with emerging technologies under path-dependency. The story shows that the Chinese government is using data and code as more abstract modes of governance to scale up a national digital economy, but these algorithmic modes of governance are still restricted by layering effects and sustained by the socio-legal institutions at their foundation. It provides new empirical evidence to the institutionalist framework and the scaling and layering hypothesis.

In other words, on top of the two case studies of China's SCS (Chapter 4 and 5), this final case study of China's other data experiments provides further support for the central argument of the thesis, namely law is computable *only to the extent* that computer code is (empirically) sustained and conditioned by the operations of law in its meaning and institutional contexts.

6.2 Introduction: China's data experiments and strategies

[205] Cf, a hard Beijing Effect discussed in Erie and Streinz (2021).

On paper the Communist Party of China (CPC) and state actors (the party-state[206]) have stipulated many data policies and laws with strategic goals. Yet this chapter uses doctrinal and qualitative evidence to argue that in practice, the Chinese party-state adopts data strategies that can be conceptualised as institutionalisation, activation and layering, which help co-produce a knowledge-based regulatory state and a national digital economy. This conceptualisation draws from the institutionalist framework detailed in Chapter 3, and emphasises law's role in co-producing the economy (Deakin et al., 2017; Ostrom 2010), coding / contesting property (Pistor, 2019; Deakin and Meng, 2020), and the law's coevolution with technology in a path-dependent, non-ergodic and historically layered fashion (Peters, 2019; Deakin and Markou 2021).

6.2.1 China's data strategies in paper / policies

China's data strategies on paper officially appeared in 2015 under the rise of big data, when the CPC for the first time raised the concept of 'national big data strategy' in its top-level policy report.[207] Soon the State Council (SC)—China's top executive branch—stipulated *The Action Plan for Promoting Big Data Development* (2015), aimed at competing with Obama's 2012 *Plan on Big data*.[208]

In March 2016 the National People's Congress (NPC), namely China's highest legislator, passed *China's 13th Five-Year Plan (2016-2020)*, with Chapter 27 named as 'Implementing the National Big-data Strategy'. It specifically required the state to 'accelerate the opening and sharing of government data' (section 1), and 'promote healthy development of the big data industry' (section 2).[209]

The State Council followed up accordingly that year by issuing the *Provisional Measures on Government Information Resources Sharing* (2016) with the aim to regulate and facilitate

[206] This chapter uses the phrase 'party-state' to simply mean both the CPC and state actors, with their meshed relationship in the policy-making and strategy-implementation processes as evidenced below. Party-state here is used as a more descriptive term, without addressing the 'party-state' question in Chinese studies, for example, on China's unique model of party-state capitalism or the cadre responsibility system. See e.g. Pearson, Rithmire, and Tsai, 2021. Edin, 2017.

[207] Report of the Fifth Plenary Session of the 18th Communist Party of China (CPC) Central Committee, 2015. [授权发布:中国共产党第十八届中央委员会第五次全体会议公报]

[208] State Council's instructions on the action plans for promoting big data development, 2015. SC [2015] No.50. [国务院关于印发促进大数据发展行动纲要的通知（国发〔2015〕50号）](last accessed 2023.2.8); Cf. Obama administration unveils "big data" initiative: announces $200 million in new r&d investments (Executive Office of the President, 2012)

[209] 中华人民共和国国民经济和社会发展第十三个五年规划纲要 [The 13th Five-Year Plan of the People's Republic of China on economic and social development, 2016] 2016. available at http://www.gov.cn/xinwen/2016-03/17/content_5054992.htm (last accessed 2023.2.8).

the data flows between governmental agencies[210]. In December 2016 *The Big Data Industry Development Plan (2016-2020)* was promulgated as a leading industrial policy for big-data and the digital economy by the Ministry of Industry and Information Technology (MIIT). It opens with the claim that 'data is a national and foundational strategic resource; and it is the 'diamond mine' in the 21st century.'[211]

Moreover, China's data strategy can be seen as a foundation for, and against the background of, China's other national development strategies (at least on paper). In 2017 the Central Committee of the CPC emphasised in its top-level policy document (the *19th Report)* the need to 'integrate the Internet, *big data*, and artificial intelligence (AI) into the real economy (*shiti jingji*)', in order to 'build a Strong Manufacturing Nation'. [212]

The 'Strong Manufacturing Nation' goal is listed under Chapter Five of the *Report*-- 'Implementing New Development Visions', which also includes Sec. 2 'building an Innovative Nation'. The Innovation Nation project envisages an accelerated application of China's fundamental research and critical technologies in order to support projects like 'Strength in Cyberspace', 'Digital China', and 'Smart Society'.[213] It was around the same time frame (2015-2017) that national projects like 'Social Credit' (2014), 'Smart Cities' (2016), and 'Smart Courts' (2017) were promoted by high-level policy documents and action plans. These projects all require data-sharing, building data centres, and innovating data applications[214].

In 2019, China's data strategies on paper expanded to agriculture and geographically the rural regions with its 'Digital Villages'[215]. This project is supposed to cover more than 540 million people living in rural areas (2020), i.e. around 38.5% of China's entire population

[210] State Council on the provisional measures to government information resources sharing and management. 2016. SC [2016] No. 51. [国务院关于印发政务信息资源共享管理暂行办法的通知（国发〔2016〕51号）]

[211] MIIT on issuing the Big-data Industry Development Plan. MIIT [2016] No. 412. [工业和信息化部关于印发大数据产业发展规划（2016- 2020年）的通知]

[212] Xi Jinping: Report Delivered at the 19th National Congress of the Communist Party of China --Secure a Decisive Victory in Building a Moderately Prosperous Society in All Respects and Strive for the Great Success of Socialism with Chinese Characteristics for a New Era, 2017.10.18 [习近平：决胜全面建成小康社会 夺取新时代中国特色社会主义伟大胜利——在中国共产党第十九次全国代表大会上的报告]

[213] *Ibid*.

[214] E.g. State Council 2014. SCS Blueprint 2014-2020; SPC, [2017] No. 12 Opinions of the Supreme People's Court on Accelerating the Construction of Smart Courts [最高人民法院关于加快建设智慧法院的意见] ; The General Office of the Central Committee of the Communist Party of China and the General Office of the State Council issued the "Outline of the National Informatization Development Strategy" 2016 [中共中央办公厅 国务院办公厅印发《国家信息化发展战略纲要》].

[215] The General Office of the Central Committee of the Communist Party of China and the General Office of the State Council issued the "Outline of the Digital Rural Development Strategy" 2019 [中共中央办公厅 国务院办公厅印发《数字乡村发展战略纲要》].

(2020) and more than 1.5 times of the U.S. population (2020)[216]. It started with the central Party-state's joint policy document *Strategic Plan for Digital Villages Development (2019)*[217]. This *Strategic Plan* was followed by the *2019-2025 Development Plan (2020)*[218], *Key Working Points (2020)* [219] and most recently the *Directives 1.0* (2021) for details of implementations[220]. These implementation policies and directives are all stipulated by the Cyberspace Administration of China (CAC) together with other top-level ministries, and they intend to shape China's rural data strategies and digital economy, for example by building public data platforms, big-data centres and introducing 'smart agriculture' applications, which are similar to their urban counterparts. Some case studies were raised in the *Directives 1.0* (2021) [221], and practices like the *Blockchain Chicken Farm* also emerged where blockchain technologies were used in rural China facilitating new supply chains of local chicken farming (Wang, 2020).

While expanding agriculturally and geographically, China's data strategies also became more clearly market-oriented on paper. In 2020 the central Party-state for the first time jointly defined data as the fifth 'essential factor of production' following land, labour, capital and technology, in the top-level policy document *Instruction on Improving the Market Mechanism for Allocating Essential Factors (2020).* The document requires the state to 'foster a data factor market' by first, 'advancing open government data and sharing' (section 20), second, 'increasing the value of data resources' (section 21), and third, 'strengthening the integration and security protection of data resources' (section 22). Some emphases include 'guid[ing] and cultivat[ing] the big-data trading market, and start data-trading according to laws and regulations' (section 26), and 'establish[ing] and improv[ing] mechanisms on data property transaction (*chanquan jiaoyi,* 产权交易) and self-governance by industrial associations' (section 27).[222]

[216] Rural population - China | Data (worldbank.org); Rural population (% of total population) - China | Data (worldbank.org); Population, total - United States | Data (worldbank.org) US population: 329,484,123.

[217] The General Office of the Central Committee of the Communist Party of China and the General Office of the State Council issued the "Outline of the Digital Rural Development Strategy" 2019 [中共中央办公厅 国务院办公厅印发《数字乡村发展战略纲要》].

[218] Notice of the Ministry of Agriculture and Rural Affairs and the Office of the Central Cyber Security and Information Commission on Printing and Distributing the Digital Agriculture and Rural Development Plan (2019-2025) [2019] No.33.[农业农村部 中央网络安全和信息化委员会办公室关于印发《数字农业农村发展规划（2019-2025年）》的通知）》]

[219] The Cyberspace Administration of China and other four departments jointly issued the "Key Points of Digital Rural Development in 2020" 2020 [中央网信办等四部门联合印发《2020年数字乡村发展工作要点》]

[220] Ministry of Agriculture and Rural Affairs, 2021 Digital Rural Construction Directives 1.0.

[221] *Id*. See, e.g. chapter 3 and 4 of the Directive 1.0, 2021.

[222] CPC central committee and State Council: Instruction on Improving the Market Mechanism for Allocating Essential Factors, 2020.3.30.[中共中央 国务院关于构建更加完善的要素市场化配置体制机制的意见]

6.2.2 Practical questions and gap of understanding

While China's data strategies have been laid out extensively on paper, this chapter seeks to understand the practical questions which include: is data 'secured' and 'integrated' as the Party-state required in 2020? To whom and what extent do governments share data? Is data trading or data 'property transaction' working well? Has the 'value of data' been realised?

For example, in contrast to the policy vision of data 'property transaction', as of March 2024 Chinese national laws in practice (deliberately) leave a grey zone for data trading and ownership, which encourages local and industrial experimentations (see below the Activation strategy in section 6.5). In fact, China's new Civil Code (2021) reaffirms that there is no clear data 'property' or 'ownership' at the national level (yet), which means doctrinally Chinese courts still cannot enforce traditional contractual clauses that cover data transaction[223]. Shenzhen also failed to write 'data property rights' into its municipal data legislation (2021)[224]. Many Chinese scholars in fact have argued against general data property rights (Kong, 2022; Peng 2022). After the State Council's 2022 *Opinions on Building the Data Basic Regime to Better Exploit the Value of Data Factors* (the 'Basic Regime'), some Chinese scholars argued that a bundle-of-rights approach was taken by the state to commercialise data in a more flexible and emergent way. The debate on data trading and ownership, nevertheless, is still ongoing (Xiong et al., 2023).

A potential alternative to 'property transaction' is to use licensing, sharing or 'contact' (*jiechu*) agreements under intellectual property (IP) laws to approximate data sets as IP objects, but legal professionals have pointed out the risk is that current Chinese IP laws do not clarify data as IP objects[225]. Legal scholars recently raised possibilities for 'data compulsory licensing' (*qiangzhi xuke*) based on 'essential facilities' provisions in antitrust and IP laws (Zeng and Zhu, 2021; Chen YW, 2021; Li SJ, 2021). However, *the 2021 Antitrust Guideline on Platform Economy* in its final version scrapped the part on 'considering data as an essential facility' (cf. Draft 2020)[226]. Kong (2022) recently argued that instead of a general data property, China

[223] China's Civil Code Article 127 also stated that 'abide by laws where data and internet virtual assets are protected'. This is understood more as a protection for virtual assets like coins in online game purchases rather than data sets.

[224] See Section 6.5.2. Also see legal scholar's discussions in news: 'Does the local government have no right to legislate on the "right to data"? The draft of the Shenzhen Data Regulations has aroused heated discussions among experts' 2020 [地方无权对"数据权"立法？深圳数据条例意见稿引专家热议.]

[225] See practising lawyers' online opinions, e.g. Zhai M., 2017 'Legal issues of data transaction contracts'. [数据交易合同的法律问题 | China Law Insight]; Liu Y., 2021 'Data transaction architecture and data licensing' [数据交易架构和数据许可]

[226] Compare Antitrust Guideline 2021 and 2020 Draft.

should start by stipulating 'commercial data rights' under IP law which delineate industrial data from more sensitive data types like personal / public data. This is an attempt to specify data governance regimes and use trade / commercial secret law as an inspiration to create new 'industrial IP rights' categories (Kong, 2022).

Moreover, Feng has argued for an unfair competition law approach to protect those data interests that do not qualify for IP protections (Feng, 2021). For example, the Hangzhou Internet Court ruled (and Hangzhou Intermediary Court affirmed) in the 'Wechat / Tencent unfair competition case' (2019, 2020) that the two plaintiffs (corporations of Tencent) have 'competition interests (*jingzheng quanyi*)' with regards to their 'data resource totality (*shuju ziyuan zhengti*)'. It also explicitly ruled that these corporations are only limited to the use rights to their 'individual data units (*danyi shuju geti*)' rather than the whole bundle of property rights[227].

Besides property, IP and compulsory licensing and unfair competition protection, legal professionals have also discussed the potential of data service agreements (as a service, AAS) to help data trading and exchange, especially concerning increasing data localisation rules around the world[228]. There are also five local Data Exchanges as emerging data brokers who experiment new modes of data transactions as of January 2022 (Guizhou 2015, Shaanxi 2015, Beijing 2021, Shanghai 2021, Changsha 2022), where Beijing raised the idea of 'use without retaining' (similar to AAS)[229]. But in short, China's data trading and 'property transaction' (*chanquan jiaoyi*, 产权交易) is still in a legal grey zone, and there is a gap of understanding on these rapidly evolving practices.

In addition, very few works in the existing literature have comprehensively examined China's data strategies and experiments in practice. Ding briefly discusses China's global data strategy based on existing legal and policy documents (Ding, 2020). Dai uses a law and economics approach to conceptualise government's strategies to use reputation. This conceptualisation touches on China's data strategies in the case of the Social Credit System (SCS), where for example public data is shared and used by authorities in public administration.

---

[227] （2019） 浙 8601 民初 1987 号 [2019, Zhejiang, 8601 Civil first instance case no. 1987] ；（2020）浙 01 民终 5889 号 [2020, Zhejiang, 01 Civil appeal case no. 5889]

[228] Zhai M., 2017 'Legal issues of data transaction contracts'. [数据交易合同的法律问题 | China Law Insight]; Liu Y., 2021 'Data transaction architecture and data licensing' [数据交易架构和数据许可] .

[229] See, China News. 2022 'The only new data exchange in the central region was put into trial operation in Changsha' 中部地区唯一新型数据交易所在长沙试运营|长沙市_新浪财经_新浪网 (sina.com.cn).

But Dai's paper does not intend to systematically examine the data practices or strategies involved. He also suggests a deficit of empirical evidence in his research (Dai, 2020).

This chapter fills a gap of understanding on China's data strategies in practice. It builds on doctrinal and empirical evidence and adopts an institutionalist approach in order to provide better conceptual models. The chapter argues that China's data strategies can be understood in practice as:

(1) Institutionalisation[230], or the co-production of formal institutions by public and private actors to address data harms and stabilise the expectations of these actors;

(2) Activation, or the party-state's adoptions of flexible policies and practices in order to mobilise data sharing and use, and

(3) Layering, or the construction of long-term data habitats / ecosystems in order to achieve high-value data use and interpretation.

These practical strategies emerged to deal with three real predicaments in China's data governance:

(1) security and privacy concerns under concentrated corporate (and state) power;

(2) fragmented data access and use by private and public actors; and

(3) the lack of valuable data use, interpretation, and long-term scaling.

6.2.3 Scope and structure of the chapter

The chapter limits the scope of the discussion to China's data strategies practised within its borders. China does not have explicit laws or policies on its extraterritorial data strategy, except for a 'Global Data Security Initiative' published by the Chinese Foreign Ministry (2020)[231]. However, recent Chinese laws and regulations have in practice formed a territorially 'defensive' data strategy that prevents certain data from crossing China's geographical border. These practices include the stipulation of data localisation rules[232] and security reviews for

[230] This is broader than Dai's definition of the state's 'institutionalization strategy' in (Dai, 2020), which means the government's use of its own public data for reputational regulations (G2G), such as the Shanghai public credit information system that provides public risk-scores for corporations. In this chapter however, institutionalision means setting up formal legal frameworks as focal points in order to allow organisations and individuals to compete and coordinate, mostly for private actors but not for governments. The government's data sharing would be more relevant in the second strategy of *activation*. Institutionalisation in this chapter nevertheless includes Dai's other three state data strategies in terms of reputational governance, i.e. regulation (of private reputation data market like Uber and Airbnb, B2B), searchlight (public information used by private individuals and corporations for decision-making, G2B), and incorporation (private data used for public decision-making, B2G).

[231] Foreign Ministry: global data security initiative (2020) [外交部全球数据安全倡议]

[232] See Civil Code 2021, Personal Information Protection Law 2021, and Data Security Law 2021.

cross-border data transfer, which culminated in the 2021 CAC review of DiDi's IPO in New York.

In contrast to this formally instituted defensive strategy of retaining data within Chinese borders, it can be argued that China in practice has informally carried out an extraterritorial data strategy under the name of the 'digital silk road' (DSR). The DSR idea was first raised in a speech given by President Xi Jinping to corporate and government partners around the world at the '2017 Belt-and-Road Initiative (BRI) Summit'[233]. Since then, many public and private practices were identified as the DSR, including the early 'China-Myanmar land cable networks' (2014) built by China Unicom[234], the 2017 DSR summits respectively held between China and Czech, Hungary and Serbia, the 'Digital Hub' jointly set up by China's Alibaba and a Malaysian company in Kuala Lumpur (2018), the 'China-Philippine Data Harbour' (2018) and so on[235]. These practices, together with the recent expansion of Chinese big-tech companies abroad, contributed to the apprehensions on whether China's DSR would in practice 'control the global internet' and shape the order of global digital governance. Even without explicit state policies, China's investments, businesses, and ICT infrastructure constructions are viewed as aspects of China's expansive data strategy in African, Asian, European and Latin American countries via the BRI[236]. Some argued that these expansive practices constitute a hard 'Beijing effect' compared to the soft 'Brussels effect' (Erie and Streinz, 2021).

This chapter focuses on China's domestic data strategy in practice, and all qualitative evidence is collected in cities rather than villages, due to practical limits of the fieldwork and the scope of this chapter.

The chapter proceeds by first introducing its institutionalist framing and the conceptual models of institutionalisation, activation, and layering (section 6.3). It then discusses these three data strategies respectively by analysing doctrinal and qualitative evidence (sections 6.4; 6.5; 6.6), and in the end concludes (section 6.7).

---

[233] Xi J. (2017) Work together to promote the construction of the "Belt and Road". [携手推进"一带一路"建设--时政--人民网 (people.com.cn)]

[234] Shi Y. (2016) Big Data and the Belt and Road Initiative[大数据与"一带一路"--理论-人民网 (people.com.cn)]; Guangxi Daily, (2014) The China-Myanmar International Land Optical Cable Project has been completed in its entirety [中国-缅甸国际陆地光缆工程全线贯通_地方报道_新闻_中国政府网 (www.gov.cn)]

[235] Zheng QH. (2018) China's big data companies embark on the "Belt and Road" [中国大数据企业走上"一带一路" (xinhuanet.com)]

[236] See e.g. Will China Control the Global Internet Via its Digital Silk Road? - Carnegie Endowment for International Peace (2020); China's Digital Silk Road and the Global Digital Order – The Diplomat (2021).

6.3 Institutionalist theory and conceptual model of state strategies

This section proposes three conceptual categories of institutionalisation, activation, and layering, as distinct state strategies in China's data experiments. This conceptualisation is a recapitulation and deepening of the institutionalist framework (or 'layered' model) proposed in Chapter 3. It relies on the earlier institutionalist analysis, and demonstrates its usefulness in examining empirical case studies.

Institutionalisation draws from the state's instituting capacity that co-produces the economy. The activation strategy derives from law's capacity to code / contest property rights and facilitate access, transaction, and pooling practices. The layering strategy is inspired by theories of long-term law-tech coevolution, where law and technologies are institutionally 'layered' and imprinted through time. These groups of theories together hold that institutions are path-dependent, non-ergodic (time irreversible) (Peters, 2011, 2019)[237], and embedded in history (Deakin and Markou, 2021; Deakin 2003, 2011; Deakin and Wilkinson, 2005). In such evolutionary processes institutions retain shared cognition and knowledge of governance practices, as the 'sediments of time' (Aoki, 2010; Koselleck, 2018; Deakin and Meng, 2020). This part explains the conceptual models of institutionalisation, activation, and layering, and their groundings in institutionalist theories.

6.3.1 Institutionalisation: State-market Co-production

First, 'institutionalisation' is used to describe the state's use of formal institutions and practices to initiate knowledge exchange with private actors and co-produce the economy. This conceptual category is inspired by institutionalist theories of state-market co-production. In early literatures Ostrom and co-authors used 'co-production' to explain public services being carried out by both paid (public) and unpaid (private) labours (Parks et al., 1981). For example, 'safety' as a public good achieved by both the police walking on beat and the communities' anti-crime efforts. Her case studies in Indianapolis (Ostrom and Whitaker, 1973) and neighbourhoods of African Americans (Ostrom and Whitaker, 1974) show that instead of large-scale patrolling in their cars, the police walking on the streets (as formal institutions) exchange informal / tacit knowledge with people in the communities, who despite unpaid, voluntarily contribute to local crime prevention and reduction. Her later case studies shows similar

[237] The non-ergodic concept here is drawn from an emerging school of ergodicity economics developed by Ole Peters.

synergetic mechanisms of co-production in U.S. education policy (Davis and Ostrom, 1991), urban infrastructure policy in Brazil, and education in Nigeria (Ostrom, 1996).

Her collaborative empirical studies on common-pool resources further demonstrated how the state and private communities need each other to co-produce both formal and informal institutions. Law in her framework is thus not merely about legal rules, but also the constitutive processes and practices that facilitate public-private knowledge exchange and collective actions (Ostrom 2010).

In more recent literature, Aoki uses the concept of 'embedded cognition' to describe institutions' capacity to store and transmit knowledge through time. He argued that states and corporations are not 'bundles of contracts' but epistemic communities that maintain focal points and cognitive structures embedded in society (Aoki, 2010). Deakin et al. further raised 'legal institutionalism' as a law and economics paradigm to explain how the rule-of-law state co-produces the modern market economy at its current global scales. They argue that industrialisation and modern economic development requires a clear boundary between the state and market, where co-production of public service becomes possible between public and private actors working with complementarity (Deakin et al., 2017).

Deakin, Chen and collaborators further used qualitative evidence to show that China's large scale industrialisation has been accompanied with the emergence of a rule-of-law state, which enforces contracts, property rights, corporate governance structures and so on. They argue that this recent rising rule-of-law state complements and constitutes China's traditional *guanxi*-based economy. Such state-market complementarity helps the Chinese economy to break from its previous limits, which come from traditional modes of cooperation and enforcement that heavily rely on informal social norms within close-knit groups and geographical regions. A Chinese rule-of-law state, at least by enforcing private and commercial laws, thus co-produces the Chinese modern market economy (Chen and Deakin, 2015; Chen et al., 2017). Deakin and Meng (2020) also use the case of COVID-19 governance in Wuhan to further illustrate how legal institutions at national and international levels helped retain public health knowledge and coordinated large-scale collective actions.

In the context of data strategy, evidence shows that in 2021 China had intensive legislation and enforcement practices on cyber / data security, data protection, and antitrust. They are examples where the state uses data institutionalisation to initiate state-market co-production of common knowledge to address data harms / risks, stabilise corporate expectations, and scale up the digital economy.

6.3.2 Activation: Coding Property or Pooling Commons

Second, the term 'activation' is used to describe the experimental processes of state-market co-production before new forms of practices and knowledge are stabilised and coded into laws. In this dynamic period, the state strategically encourages or acquiesces to the flexible legal and policy spaces which allow for competing practices to emerge, and which 'activate' new forms of market transactions, pooling and / or the use of undefined things or assets, including data. It can be viewed as a strategy that often chronologically precedes the institutionalisation strategy, or is adopted in between stabilised periods of 'punctuated equilibrium' to trigger change in legal evolution (Deakin, 2003, 2005, 2011).

'Activation' is inspired by institutionalist theories that reveal the processes of legal 'coding' of things into property and capital, and the polycentric governance of common-pool resources (CPR) co-produced by the public and private actors. This was discussed more in detail in Chapter 3 (institutionalist framework), and in particular on Ostrom's theory of polycentric governance of CPR in 3.4.

In the context of data, there has been no consensus on how law can code 'data', as a new 'thing', into an existing or novel legal category. This leaves an elastic legal and political space for major jurisdictions including the U.S., EU, and China to compete in innovating data ownership and access practices (Litman, 2000; Evans, 2011; Victor, 2013; Wiebe 2017). Evidence will show that the Chinese state has used the activation strategy to allow for legal innovations on data trading, sharing, and use against data fragmentation. These legal ingenuities include the Central Bank's grant of a provisional licence to Ant Group on handling personal credit information, techno-legal arrangements in the Beijing International Data Exchange (BIDE) which will, below, be contrasted with Shenzhen's local legislative attempts on establishing data property rights, and Tianjin's emerging practice of 'Party manages data'.

6.3.3 Layering: Law-Tech Coevolution

The third major strategy being conceptualised is 'layering', which describes the state's efforts to construct sustainable ecosystems of institutions for long-term public governance in changing environments. This aligns with the argument of the thesis that code is layered on law, and is path-dependent on social norms and law's modes of existence and operations.

The state, in this sense, uses the layering strategy to make timely adjustments to the often fragile fits between formal institutions and their environments layered in history with path-dependency. In particular, data layering means the state builds useful data cycles and

sustainable data habitats / ecosystems for valuable data applications. It is concerned with better designs of law and technologies in order to achieve valuable and long-term data use in public governance. It is a strategy of sustainability which can complement the stabilising strategy of institutionalisation, and the contestation strategy of activation. This layering strategy is inspired by institutionalist and design theories that understand law as co-evolved with technology through time. In early literature, Deakin argued that juridical categories like 'employment' evolved in history with path-dependency (Deakin and Wilkinson, 2005). Similar to the evolution of technologies such as the computer keyboard arrangement (QWERTY), these juridical concepts can have 'frozen accidents' and exhibit 'punctuated equilibrium' through time (Deakin, 2003, 2011, 2015).

Applying design and affordance theories to law, Hildebrandt argues that modern positive law is made possible by printing technologies, and the emerging pervasive smart technologies are likely to afford the design of data-driven laws (Hildebrandt, 2016). Deakin and Markou combine this observation with legal evolution theory. They use the case of recent machine-learning applications in law to reveal the co-evolution between law and technology as institutionally path-dependent and non-ergodic, i.e. embedded in irreversible time. They argue that machine-learning and other data-driven technologies can lock-in the past, just like law's path-dependency in evolution, but with even more closed ends. They present the risk that while law is afforded by natural language that allows for flexible interpretations and breaking from the past, the numbers and code presented in data can be more rigid and thus trap society into its past once applied to law and governance (Markou and Deakin, 2020). Hildebrandt uses the term 'freezing' to describe shared apprehensions that pervasive data and code-driven technologies will exacerbate law's negative imprinting of the past (Hildebrandt, 2020b, 2021).

For example, when the UK courts decided to define drivers working on ride-hailing platforms as 'workers' (Uber BV v Aslam [2021] UKSC 5), new populations will be added to existing data categories or require the creation of new data categories in the mathematical models. This requires timely updates of both data sets and the algorithms to prevent model decay caused by such 'concept drifts'. Otherwise, as time passes by the shift in social reality will impede valuable long-term data interpretation and / or prediction, and the data use will not be sustainable or scalable.

The notion of 'layering' as drawn from the law-tech evolutionary literature discussed above, is useful to illustrate the need for long-term interpretation of data in a sustainable 'habitat / ecosystem'. Evidence for this claim is drawn on the practices in Smart Court

experiments, and is complemented by secondary sources from Huawei, Alibaba and research institutes.

6.4 Institutionalisation: Data Security, Protection, and Antitrust

Heldt and Hennemann argued that the 'Roaring Twenties' of EU technology regulations are dawning. Drafts of the EU *Digital Services Act* (2020), *Digital Markets Act* (2020), *Data Governance Act* (2020), and the *AI Regulation* (2021) are at the forefronts, which they described as a strategy of 'upward convergence' in regulatory standards similar to the Brussels Effect[238]. This part evidences a similar legislative and regulatory trend in China that started with the 'Roaring Summer of 2021'. The Chinese strategy is however conceptualised as 'institutionalisation'. In the data context it means the state's initiation of a market-state co-production of public services and common knowledge within epistemic communities of data governance, in order to stabilise and scale up a national digital economy. Three major institutionalisation efforts respectively addressed emerging forms of harms / risks in data security, privacy, and competition are: (1) the cybersecurity review of the Chinese ride-hailing giant DiDi Chuxing after its IPO at the New York Stock Exchange (NYSE: DiDi Global Inc); (2) the completion of China's data protection legal framework featuring 'gatekeeper' and managerial liabilities; (3) antitrust actions that restrain data power and facilitate interoperability.

6.4.1 DiDi's 2021 Cyber / Data Security Review

6.4.1.1 Security Concerns over DiDi before its IPO in New York

In 2016 China's *Cybersecurity Law* [239] was passed as the first high-level law that regulates data security[240]. It created a new legal category, the 'critical information infrastructure (CII) Operators', with high-level obligations on cyber / data security (Article 31) and localisation requirements (Article 37). These high-level compliance requirements on CII Operators are stipulated because they are, by name, 'critical' in handling / managing data and its

[238] Heldt and Hennemann, The Roaring Twenties of Technology Regulation - Tagesspiegel Background 2021.7.15.
[239] Effective since June 1st 2017
[240] Cybersecurity Law of the People's Republic of China 2017 [中华人民共和国网络安全法]

infrastructures. The Cyberspace Administration of China (CAC) is the top-level party-state organ that oversees such high-level security compliance. [241]

Soon after DiDi Chuxing got listed at the NYSE in the U.S. June 2021, the CAC started a security review / investigation under the *Cybersecurity Law* (2017), which effectively deemed DiDi as a 'CII Operator' [242]. Such security review after an outbound IPO, is arguably functionally similar (as it does not stop the IPO) to the U.S. national security review by The Committee on Foreign Investment in the United States (CFIUS) which can pause and stop cross-border mergers and acquisitions (e.g. Griffin, 2016)[243].

DiDi is a ride-hailing tech giant born in China, and a major competitor with Uber in the global ride-hailing market.[244] DiDi claims in its SEC filing that between March 2020 and March 2021 it had 377 million annual active users (which is larger than the U.S. population) and 156 million monthly active users (MAU) in China[245], compared to Uber's 93 million MAU in the world (last quarter of 2020)[246].

Evidence shows that the regulatory agency had difficulty in addressing DiDi's cyber / data security issues long before DiDi's IPO in NYSE (June 2021). First, DiDi Shanghai was already investigated and raided in 2018 by various Shanghai regulators based on security concerns. These include 'fake cars', illegal drivers and incomplete data filing to the ride-hailing regulatory agencies. But nothing was found when the enforcement team arrived. DiDi was reported to have even prepared several boxes of chaotic and useless files in advance for the enforcement team's third inspection.[247] One respondent, a governmental official and legal scholar in Shanghai, complained in a closed group discussion that DiDi is not concerned with their everyday fines from violations of regulations: 'they see such administrative fines as the cost in their financial reports'[248].

Second, another respondent raised in the same closed group discussion that 'the CAC is afraid of reviewing data security in cross-border transfers because they don't have the technical expertise yet.' This was especially acute with regards to the massive amount of geographical data generated each day on ride-hailing platforms like DiDi. The respondent claimed that one

[241] *Id.* Chapter 3 section 2 on CII operators. Including article 37 on storing important data in China.
[242] 2021 CAC announcement on DiDi.
[243] Also see a classic case of Shuanghui's acquisition of Smithfield. China Inc. and the CFIUS National Security Review – The Diplomat
[244] The New Global Order in Ride Hailing: Didi vs. Uber — The Information;
[245] SEC Filing, 2021, Didi Global Inc Prospectus under 424(B)(4) (sec.report)
[246] Uber's users of ride-sharing services worldwide 2017-2020 | Statista
[247] He, YH. (2018). 'The fourth on-site inspection of Didi by the Shanghai inspection team has been accessed and will be reviewed', Thepaper.cn. [上海检查组第四次上门检查滴滴 数据已接入将复查|滴滴|检查组|上海_新浪科技_新浪网 (sina.com.cn)]
[248] Government official, Shanghai, 2021.3

concern of the CAC was that such geographical big-data can be used to infer and reveal the locations of military bases or other sensitive places. However, the CAC did not have the technical team to understand how to evaluate such risks (March 2021, Shanghai).

Third, the sensory and audio-visual big-data collected by the cameras on self-driving cars were of particular concerns with the rapid roll-out of autonomous vehicles by companies Tesla, and DiDi's robo-taxi in cities like Shanghai[249]. Two respondents, who both are government officials and legal scholars, discussed the data risks of Tesla's self-driving cars in Shanghai. They claimed that a Tesla's cameras and sensors collect so much real-time visual data of its surroundings including people's faces, buildings, other vehicles and so on, and can transfer them abroad. Sensitive information can be inferred from such big-data where data security and privacy stakes are high (March 2021, Shanghai).

These security concerns over sensitive geographic data was later confirmed in May 2021 with the CAC's stipulation of *Several Provisions on the Management of Automobile Data Security (Draft for Comment)*[250]. Article 3 listed six types of 'important data' which include for example,

> A3 (1) 'Data on the flow of people and traffic in military administrative areas, national defence science and industrial units or other units that involve state secrets, or sensitive, important areas of Party and government administrative units above the county level, etc.';
>
> A3 (3) 'Data on the operation of automobile charging networks'; and
>
> A3 (5) 'Audio-visual data of individuals' faces, voices, and licence plates, etc., outside the vehicle'[251].

6.4.1.2 Institutionalisation in the DiDi Security Review 2021

DiDi's 2021 security review is not that surprising against the background of CAC and local regulators' security concerns. On July 2nd 2021, two days after DiDi's IPO at NYSE, the CAC announced its review of DiDi's cybersecurity compliance, and required that no new users be registered on DiDi's ride-hailing platform immediately. The announcement cited the *Cybersecurity Review Measure (2020)* as the applicable law, and used *National Security Law*

[249] AutoX launches its RoboTaxi service in Shanghai, competing with Didi's pilot program | TechCrunch 2020

[250] Translation: Several Provisions on the Management of Automobile Data Security (Draft for Comment) | DigiChina (stanford.edu)

[251] *Id.* '(2) Survey and map data that is more precise than maps publicly issued by the state (4) Data on types and traffic volume, etc., of vehicles on the road (…); (6) Other data that might affect national security and the public interest, as specified by the state cybersecurity and informatization department and relevant departments of the State Council.'

(2015) and the *Cybersecurity Law (2017)* as the higher-level legal sources.[252] In this way CAC effectively deemed DiDi as a 'CII Operator' which is the only subject regulated under the scope of Article 1 and 2 of the current *Measure (2020)* ('national security risks were brought about by the CII Operators' procurement of network products and services'). Data security is listed in Article 9 (1) as one of the risk factors for review ('theft, leak or damage of important data').[253]

On July 4th, CAC confirmed DiDi's 'severe violation of data regulations' and required all App Stores in China to delist DiDi's online ride-hailing app, despite allowing existing riders and drivers to keep using the app. It demanded DiDi to make changes to comply with the *Measure (2020)* and other security standards. [254] Two months after this sanction it was reported that DiDi's daily active users dropped 30% (comparing August to June)[255].

Besides DiDi, on July 5th 2021 CAC initiated similar investigations and disciplines on three more Chinese tech giants that recently got listed in the U.S. These big-techs are all deemed as CII operators, which have cyber and data security concerns in their public listing abroad. Two other Chinese tech giants later cancelled their listing plans after the CAC's four reviews[256].

Within a week's time, on July 10th the CAC also released a new *Cybersecurity Review Measures (2021, Draft for Public Comments)* that has provisions targeting big-tech outbound IPOs. It requires any 'CII Operator with more than 1 million user information that wish to be publicly listed abroad, must file a cybersecurity review to the Cybersecurity Review Office' (Article 6).

Seven 'major factors' were listed concerning risks to national security (Article 10), including the risk that 'after an IPO abroad is completed, the CII, core data, important data or massive amounts of personal information are influenced, controlled or maliciously used by foreign governments'. It also expanded the CAC's general review mandate to all 'data handlers' instead of just the 'CII Operators' (Article 2)[257]. More clarity would be provided

[252] CAC 2021 announcement on DiDi.
[253] CAC 2020 Review Measure.
[254] CAC 2021 announcement on DiDi
[255] Didi loses 30% of daily users after Beijing crackdown following IPO | Financial Times 2021 Sep
[256] See various analysis articles written by lawyers in China. Song H. et al. (2021) 'From the implementation of cybersecurity review to the removal of APP, data security lawyers sorted out eight core issues in the Didi incident-Mobile Payment Network' [从实施网络安全审查到 APP 下架，数据安全律师对滴滴事件梳理了八个核心问题-移动支付网]. And in the U.S., After 5 Years, China's Cybersecurity Rules for Critical Infrastructure Come Into Focus | DigiChina (stanford.edu)].
[257] Notice of the CAC's on Soliciting Public Opinions on the 'Cybersecurity Review Measures (Draft for Comments)' [国家互联网信息办公室关于《网络安全审查办法（修订草案征求意见稿）》公开征求意见的通知-中共中央网络安全和信息化委员会办公室 (cac.gov.cn)]

regarding security concerns for both data handlers and CII operators when the revised *Measure 2021* is finalised and taken into effect.

A more detailed *CII Security Protection Regulation (2021)* was also later passed in August 2021. This new *CII Regulation* (2021) adds granularity to the existing *Cybersecurity Law (2017)* and *Review Measure (2020*), and clarifies important obligations for CII Operators to, first perform data security accountability and establish protection institutions (Article 15); second, to prevent data leakage (Articles 2, 9 ,18), and third, to maintain data completeness, confidentiality and usability (Article 6).

In addition, it is worth noting that China passed its long-due *Data Security Law (2021)* in June 2021 about three weeks before DiDi's IPO. It is a high-level legislation as important as the *Cybersecurity Law (2017)*, and came into effect on the same date with the *CII Security Protection Regulation (2021)* on September 1st 2021[258]. Its Chapter IV detailed the 'Data Security Protection Obligations' for all 'handlers of important data' (a broader concept than CII Providers), where two articles are crucial.

(1) Article 27 requires 'The conduct of data handling activities using the Internet or other such information networks shall perform the data security protection obligations described above on the basis of the cybersecurity Multi-Level Protection System (MLPS)'.[259] The MLPS 1.0 was a ministerial standard for cybersecurity that functioned as early as 2007 before the big-data and algorithmic technologies took off in China[260]. Since 2019, the MLPS 2.0 is the revised standard in effect, which was updated by legal and technical communities entering the big-data era, and shall be used for more granular reviews on DiDi's data security issues[261].

(2) Article 31 stipulates the data security checks for 'outbound security management' for CII Operators and other data handlers working in China pursuant to the *Cybersecurity Law (2017)*.[262] This article raises data security concerns when data handlers change their corporate governance structures, e.g. by listing outside of China.

In DiDi's case the CAC cybersecurity review might be complicated if the U.S. regulatory authority Security Exchange Commission (SEC) requires data disclosure or access to third-

[258] Data security law of the People's Republic of China 2021. [中华人民共和国数据安全法]

[259] Translation: Data Security Law of the People's Republic of China | DigiChina (stanford.edu)

[260] MLPS 2007

[261] MLPS 2.0, 2019

[262] 'The provisions of the Cybersecurity Law of the PRC apply to the outbound security management of important data collected or produced by critical information infrastructure operators operating within the mainland territory of the PRC; outbound security management measures for other data handlers collecting or producing important data within the mainland territory of the PRC are to be jointly formulated by the national cybersecurity and informatization department and relevant departments of the State Council.' Data security law of the People's Republic of China 2021. [中华人民共和国数据安全法]

party auditing due to DiDi's IPO at NYSE. For example, under the U.S. *Holding Foreign Companies Accountable Act 2020 (HFCAA),* the SEC and the Public Company Accounting Oversight Board (PCAOB) can access third-party audit papers of non-US firms listed in the U.S. required by the Sarbanes-Oxley Act 2002.[263] Besides, outbound data transfer concerns can also be triggered when the voting structure within the board changes and new decisions can be made on data transfer. Article 31 aligns with the latest proposed rules on security reviews of public listing abroad under the new *Cybersecurity Review Measures (2021, Draft for Public Comments).*

There are also rapid legislative moves on the security of cross-border data transfer. Most recently on October 29th 2021, the CAC also issued a new *Security Assessment Measures on the Cross-border Data Transfer (Draft for comments)*. It specified the conditions, procedures, materials necessary for filing, and redress.[264] Shanghai's local legislature also passed a *Shanghai Data Regulation (2021)* which requires the making of a 'low-risk catalogue' for certain types of cross-border data transfers in order to facilitate data flow in its *Lin Gang* free trade zone pilot (Articles 66, 67).[265]

6.4.1.3 Co-Production of Knowledge in Cyber / Data Security

The above legislative and regulatory practices did not only establish rules, but more importantly shaped new common knowledge within emerging cyber and data securities communities. As the evidence shows, the Shanghai local regulators had problems enforcing regulations against DiDi's massive real-time operations, and CAC started with little expertise facing new data security risks on such large-scale ride-hailing platforms and the roll-out of autonomous vehicles.

DiDi's security review thus became a point of institutionalisation where existing concerns are addressed by the state's legislative and enforcement practices. These CAC practices also initiate informal knowledge exchange with private actors to make more detailed regulations.

[263] 15 U.S.C. § 7211, Holding Foreign Companies Accountable Act 2020 (HFCAA); Congress Passes the "Holding Foreign Companies Accountable Act" (harvard.edu) 2021; Senators call on US securities regulator to investigate Didi IPO | Financial Times (ft.com) 2021.

[264] Notice of the CAC on Soliciting Public Opinions on the Draft 'Measures for Security Assessment of Data Cross-Border Transmission'[国家互联网信息办公室关于《数据出境安全评估办法（征求意见稿）》公开征求意见的通知]

[265] Shanghai Data Regulation (2021) [上海市数据条例] ; Zhang YT. (2021) Interpretation of the Shanghai Data Regulations (Draft) 丨 The first proposal in China to formulate a low-risk cross-border data flow catalogue. [《上海市数据条例（草案）》解读一丨全国首提制定低风险跨境流动数据目录]

DiDi's case sets an example for future security compliance by other private actors such as Tesla, whose autonomous vehicles collect massive amounts of data in China.

It is debatable whether DiDi's review was political in nature, but the process at least helped stabilise market expectations on**:** (1) what is a 'CII operator', (2) what are the review procedures and requirements; (3) what knowledge and expertise the CAC needs to retains in order to review automobile data (including self-driving cars), foreign listings and cross-border data transfer.

The process of the review also enabled the wider legal and technical community of cyber / data security to update common knowledge, for example, by revising the MLPS 1.0 (2007) standards to MLPS 2.0 in 2019 which aligns with high-level legislation and emerging technologies. The informal knowledge exchange within these epistemic communities engenders self-governance of data risks. Such private efforts evolved around but also shaped the state's centralised instituting efforts to keep updating common knowledge and practices. In this sense the CAC and regulators have used an institutionalisation strategy to initiate a market-state co-production of knowledge on data security governance that stabilises corporate expectations and scales the digital economy.

6.4.2 Data Protection, Gatekeeper, and Managerial Liabilities

In addition to security, data protection is the second area that shows extensive evidence of China's institutionalisation strategy. China's Personal Information Protection Law (PIPL) was recently passed in August 2021 and came into effect on Nov 1st.[266] It is considered as China's equivalent of the EU's General Data Protection Regulation (GDPR, 2018), but also features rather unique provisions such as gatekeeper and managerial liabilities inspired by / in competition with both the U.S. and the EU.

6.4.2.2 'China's GDPR': Personal Information Protection Law (PIPL) 2021

To emphasise the importance of this legislation, the *National People's Congress (NPC) Legislative Note on PIPL (2020)* claims at the beginning that 'by March 2020 China's internet users reached 900 million.'[267] The *Note* goes on to observe, however, that data violations and harms on people's daily life are rampant and require change: 'an objective demand to enhance

[266] Personal Information Protection Law of the People's Republic of China, 2021 [中华人民共和国个人信息保护法]

[267] This is more than the total population of the EU, U.K. and U.S. combined, but only about 65% of the Chinese population.

PI protection' and 'a real demand for maintaining a healthy ecosystem of cyberspace'. To highlight this real predicament the *Note* cites a few NPC reports and initiatives which showcase the society's demands as well as the central Party's concerns over societal complaints. [268]

In addition to addressing the data privacy concerns, the PIPL is also considered as an 'important measure to promote healthy development of the digital economy'. The *Note (2021)* adopted the SC's understanding of 'data as an essential factor (of production)' and claimed that 'data competition becomes important in global competition, and PI is the core and foundation of big-data.'[269] It further raised that 'in coping with COVID-19', many big-data applications are useful and become more pervasive, but a comprehensive data protection framework needs to be built and applied.[270]

Doctrinally, the PIPL 2021 completes China's data protection framework. Its ultimate legal authority of data protection laws comes from the Chinese Constitution: Article 33 on human rights, Article 38 on human dignity, and Article 40 on freedom and privacy in communication. In 2021 the *State Council's Human Rights Action Plan (2021-2025)* also highlighted in the protection of 'privacy rights' and 'personal information interests (PII , *geren xinxi quanyi*)'[271].

However, the earliest high-level data protection law in effect is the *Ninth Amendment of the Criminal Code (2015)*. It stipulates a maximum of seven years of prison for the most severe offences, namely those of stealing, selling or abusing other people's personal information. This was targeted at the privacy concerns at that time with the rise of big-data technology. The *Ninth Amendment (2015)* was later followed by the *Judicial Interpretations on Personal Information Crime (2017)* issued by China's Supreme People's Court (SPC) and the Supreme People's Procuratorate (SPP). Some high-profile criminal cases have been reported for example, in Shanghai and Hangzhou where big-data companies illegally collect and store personal information. They were heavily fined and their legal representative and CTO were sentenced to three years of prison (but with reprieve periods).[272]

[268] Liu JC. 2020.10.13. Explanation on the "Personal Information Protection Law of the People's Republic of China (Draft)"[关于《中华人民共和国个人信息保护法（草案）》的说明_中国人大网 (npc.gov.cn)]

[269] *Ibid*.

[270] *Ibid.*

[271] National Human Rights Plan, chapter 2, section 3. (2021). [国家人权行动计划（2021-2025年）(scio.gov.cn)].

[272] See, e.g.（2020）浙 0106 刑初 437 号 [2020, Zhejiang, 0106, Criminal Court, first instance case no. 437]. Also see 淘宝 11.8 亿条用户隐私被泄露 两男子被判刑 [Taobao 1.18 billion pieces of user information leaked, two men sentenced. Sina news, 2021.6.17]. 2021. Available at https://www.sohu.com/na/472780318_120957976 (last accessed 2023.2.8)

Besides the criminal code, China's new *Civil Code 2021* sets two layers of protection for personal information. First, Article 1032 stipulates the 'privacy right' that is similar to the Article 8 'right to private life' in the European Convention of Human Rights (ECHR). The Article 1032 right protects private spaces, activities and communication at the right-level with human dignity as the foundation. Although privacy is afforded with higher-level protection, the scope of privacy is narrower compared to 'personal information interests' (PII).

Art. 1034 *Civil Code 2021* stipulates the protection of PII, which is not protected at the level of 'rights', but includes a larger scope of interests related to both private and non-private personal information. After the *PIPL 2021* came into effect, PII is protected in detail by this field-specific legislation. This completes the two-layer civil protection of personal information, and it is worth noting that as private information is part of personal information, it is doctrinally also protected under specific rules of the PIPL 2021[273].

Other important pieces of legislation on data protection include the *Cybersecurity Law* 2017 and the *Data Security Law 2021,* together with other specific regulations concerning security issues. These security laws function as major complements to the PIPL on data protection.[274]

The *Consumer Protection Law 2014* and *E-Commerce Law 2019* broadly touch upon personal information with articles that link the jurisdictions to PIPL 2021[275]. There are also many low-level personal information protections standards (e.g. *Guidance for personal information security impact assessment 2020*)[276] and local items of legislation (e.g. *Shenzhen Data Regulation 2021*, *Shanghai Data Regulation 2021*)[277].

The Chinese courts are also active in the data protection regime, with various courts having made high-profile decisions. Even before the new *Civil Code 2021*, the Beijing Internet Court for example in 2020 had already made decisions that WeRead (of Tencent, linked to Wechat) and TikTok (of Bytedance) respectively violated personal information protections laws. The court found WeRead had misled users to consent on displaying their favourite / liked books to Wechat friends and contacts, and required WeRead to stop accessing user's Wechat

[273] Civil Code of the People's Republic of China 2021. Articles 1032, 1034 [中华人民共和国民法典]

[274] *Cybersecurity Law of the People's Republic of China 2017; Data security law of the People's Republic of China 2021.*

[275] *Consumer Protection Law 2014*; *E-Commerce Law 2019.*

[276] Information security technology -- Guidance for personal information security impact assessment (GB/T 39335-2020) [信息安全技术—个人信息安全影响评估指南(GB/T 39335-2020)]

[277] Shanghai Data Regulation (2021) Effective 2022.1.1. [上海市数据条例] ; Shenzhen 7th People's Congress Plenary Committee Report No. 10. On passing the Shenzhen Economic Special Zone's Data Regulation, 2021. Effective 2022.1.1 [深圳市第七届人民代表大会常务委员会公告（第十号）] (last accessed 2023.2.8)

friendship list and delete all relevant information. It also required WeRead to stop showing each user's data (such as favourite books, what is being read and so on) to the user's friends. The Court also found violations of 'personal information interests' where TikTok's Beijing operator accessed the name and phone number of the plaintiff Ling via Ling's friend's contact book, but without consents from either Ling or Ling's contact. It ruled the deletion of related data including the name, phone number and the geographical data of Ling, and an apology with compensation of 1,000 RMB to Ling.[278] After the new *Civil Code 2021*, the first civil collective action case appeared where the Hangzhou Internet Court ruled both criminal punishment and civil liabilities to Sun who bought and then publicly sold more than 45,000 pieces of personal information online for profit.[279]

Recently the Supreme People's Court also made the *Judicial Interpretations on Facial Recognitions (2021)*[280]. The latest facial recognition related case also was accepted in Changsha on data protection basis, where a legal intern sued the building management company of his law firm for imposing compulsory facial recognition entrance registrations (pending after the PIPL)[281].

The *PIPL 2021* builds on these previous laws and practices, and completes China's data protection framework. Detailed practices and the empirical effects of PIPL await to be observed after it came into effect on Nov 1st 2021. Nevertheless, below-discussed PIPL legislative process, featuring gatekeepers and managerial liabilities to a certain degree manifests an institutionalisation strategy that deliberately clarifies the state and market divide in order to co-produce common knowledge in China's data protection practices.

6.4.2.3 Gatekeeper and Managerial Liabilities in the PIPL 2021

PIPL 2021 features 'gatekeeper' and managerial liabilities that require big-techs in China to take up more public enforcement obligations with regards to their third-party suppliers. The

[278] Xiao and Yang, 2020.7.31. People.com.cn, Beijing Internet Court: WeChat Reading and TikTok infringe personal information rights[北京互联网法院：微信读书、抖音侵犯个人信息权--法治--人民网 (people.com.cn)]

[279] Xiao et al. 2021.1.9. People.com.cn, After the implementation of the Civil Code, the first civil public interest litigation case for personal information protection in China was pronounced in court[民法典实施后全国首例个人信息保护民事公益诉讼案当庭宣判--法治--人民网 (people.com.cn)]

[280] SPC [2021] on facial recognition.

[281] Fan YW. 2021.9. www.nandu.com, The first case of Hunan face recognition: the building was forced to "brush face", and the plaintiff hoped to promote the implementation of the personal protection law [湖南人脸识别第一案：大楼强制"刷脸"，原告望促个保法落地_] (accessed 16 December 2022).

legislative process followed an international trend in regulatory practices of mandating big-techs to self-regulate their third-party suppliers or App developers.

First, after the Cambridge Analytica incident, the U.S. Federal Trade Commission (FTC) in 2020 required Facebook (U.S. v Facebook 2020) to take more public obligations of regulating its third-party developers / suppliers in terms of data protection. It not only deemed Facebook liable for third-party breach of consumer data protection, but also stated that Facebook's Mark Zuckerberg can be personally liable in the future with potential civil, regulatory, and even criminal punishment[282].

Van Loo theorised this practice by the FTC as a move towards 'gatekeeper liability', whereby regulators like the FTC can require big-tech platforms to be law-enforcers that supervise and regulate their third-party suppliers. He argued that this liability originates from vicarious / third-party liability in torts that can penetrate corporate protection of individual officers or board of directors. The Facebook case manifests what he calls the 'revival of the *respondeat superior'*: where gatekeeper liability re-emerged alongside the big-tech's rising power. Big-tech's control over third-party suppliers' data access means vicarious liability (Van Loo 2020a, 2020b).

The FTC's gatekeeper and managerial liabilities imposed on Facebook is a sanction-backed legal arrangement to clarify big-tech's obligations to self-govern and prevent systematic harms / risks. This approach aligns with Shanghai's 'cooperation and deterrence' attitude towards big-tech companies, as expressed by a respondent who is a government official in Shanghai.

> 'The Chinese state has historically adopted this cooperation and deterrence attitude towards large businesses… Render unto market what is the market's, and unto state what is the state's. Then we won't disturb the foundational ecosystems of the citizens.' (Government official, March 2021).

However, gatekeeper liability was not in the first draft of China's PIPL (October 2020). Only personal liability for executives of data-handlers was stipulated in this draft with moderate fines and bans from the profession various periods of years. These managerial liabilities do not require the data-handlers to be large platforms or 'gatekeepers'[283]. Comparatively, while the GDPR does not stipulate personal / managerial liabilities, the UK Data Protection Act (2018) has provisions on the 'Liability of directors etc.' that penetrate the corporate protection.[284]

[282] United States v. Facebook, Inc., No. 19-2184 (TJK), 2020 WL 1975785 at *5 (D.D.C. Apr. 23, 2020).
[283] See PIPL first legislative draft for comments.
[284] Data Protection Act 2018 c.12. Section 198

Many other Chinese laws that focus on risk-related security compliance also have managerial liabilities, including the *Food Safety Law 2019, Cybersecurity Law 2017,* and *Data Security Law 2021*[285].

Nevertheless, Article 58 featuring gatekeeper liabilities appeared in the second draft of China's PIPL (2021 April). It stipulated a group of obligations for big-techs which are similar to those in EU's *Digital Services Act (DSA, 2022)* on 'very large online platforms' (VLOPs).[286] Article 58 and the DSA share similar requirements on transparency, compliance structures and audits for big-tech companies designed to mitigate systemic risks and better regulate third-party suppliers or App developers[287]. The MIIT also accompanied the PIPL second draft with an *APP Personal Information Protection Regulations (Draft for comments)* in 2021 April, which proposed more details on how to regulate third-party APPs on gatekeepers (Zhang, 2021).

It is worth noting that the Chinese legal academia and legislators became more aware of the term 'gatekeeper' after the EU's *Digital Markets Act (DMA, 2022)* explicitly proposed 'gatekeeper obligations' for big-tech platforms in December 2020 (draft). The focus of the *DMA* (2022)*,* however, is on fair competition for third-party developers / suppliers and interoperability, which is different from the focus of the obligations on VLOPs stipulated by the EU's *DSA* (2022) and the FTC's mandated gatekeeper's liabilities concerning data protection[288].

The third and final version of the PIPL (2021 August) solidified Article 58 as imposing a set of obligations for 'personal information handlers that provide important internet services, with very large amounts of users and complex business categories'. It added two more sections which further specified gatekeeper liabilities: (i) 'platforms should set internal compliance institutions', and (ii) 'platforms should obey transparent, just and fair principles, and should set platform rules that specify third-party obligations on handling data.'[289] These provisions to some extent establish the state-market boundary and the division of labours and in regulating third-party providers.

Legal scholars' efforts are also considerable in the PIPL's drafting of gatekeeper liabilities. For example, a comparative law article by Zhang (2021) introduced the concept of gatekeeper obligations in 2021 after the second draft of PIPL and the EU 2020 DMA and DSA

[285] See China's *Food Safety Law 2019, Cybersecurity Law 2017,* and *Data Security Law 2021*

[286] Second draft PIPL 2021[《个人信息保护法（草案二次审议稿）》]

[287] Regulation (EU) 2022/1925 (DMA).

[288] Regulation (EU) 2022/1925 (DMA), Regulation (EU) 2022/2065 (DSA).

[289] See PIPL 2021.

proposals. Wang (2021) advocated a 'state protection obligation' which requires the state to take up its protection role on personal information and coordinate both public and private efforts in data protection. Professor Zhou Hanhua at the Chinese Academy of Social Science has also been deeply involved in the 15 years legislative process and the three versions of PIPL between 2020-2021[290].

6.4.2.4 State-market Co-Production of Data Protection and a soft 'Beijing Effect'

The above evidence describes China's legislative and regulatory processes on data protection, in particular featuring gatekeeper and managerial liabilities. Together they show the state has an institutionalisation strategy that deliberately provides a division of labour between the state and market actors and which clarifies separate spaces for the co-production of common knowledge in data governance communities.

Data protection practices before the PIPL 2021 already started addressing data violations using criminal punishments which established redlines for data-handling activities. They did not however provide stable private expectations on what constitutes compliant data handling practices in daily operations. PIPL 2021 is the institutionalisation effort that tries to stabilise these accumulated practices and knowledge. The PIPL's gatekeeper provisions in particular provided the state with new tools to enforce data protection through big-techs like Alibaba and DiDi. This aligns with the state's 'cooperation and deterrence' attitude that involves big-techs in the daily data governance of their third-party supplies and users.

PIPL also competes with the EU's digital governance framework - GDPR (2018), DMA and DSA (2022) - for high standards of privacy and data protection. A respondent claimed that 'by implementing the law well at home, we can set good examples and reach our laws beyond China as best practices (Government official, Shanghai, 2021.3). This raises the potential of a 'Beijing Effect' in data protection practices in competition with the 'Brussels Effect' (Bradford, 2020). More accurately, this can be a soft 'Beijing Effect' (which is more similar to the 'Brussels Effect') that does not necessarily requires expansive foreign investments or infrastructure constructions, as compared to the 'hard' Beijing Effect as recently argued through the case of the Digital Silk Road (cf. Erie and Streinz, 2021).

[290] See, e.g. Top Scholar Zhou Hanhua Illuminates 15+ Years of History Behind China's Personal Information Protection Law | DigiChina (stanford.edu).

6.4.3 Antitrust: Data Power and Interoperability

The Chinese state also started an antitrust movement against big-tech companies during 2020-2021 alongside its institutionalisation on security and data protection. Major cases include first, the Central Bank's 2020 investigation of Ant Group (affiliate of Alibaba Group). Ant is China's Fintech giant for e-payment and consumer loans, and claims to have more than 1 billion annual active users within China[291]. Second, the $2.75 billion fine (comparable to $2.8 billion EU antitrust fine on Google[292]) issued by China's antitrust agency State Administration for Market Regulation (SAMR) on Alibaba Group in April 2021. Alibaba is China's e-commerce tech-giant and equivalent of Amazon[293]. Third, the SAMR $530 million fine of Meituan, which is China's food delivery tech-giant and a rough equivalent of Deliveroo / UberEats (food delivery) plus Yelp (consumer rating and booking). It serves 340 million annual transacting users by 2018 as claimed in its 2018 IPO at the Hong Kong Stock Exchange (HKEX)[294].

As to the first of these cases, on November 3rd 2020 the Shanghai Stock Exchange paused the public listing plan of Ant Group after China's Central Bank and other financial regulators launched a joint investigation[295]. The Central Bank later in its press conference listed 'using market advantage status to exclude competitors' as one of the problems of Ant Group. It also raised 'breaking monopolies … and maintaining a fair and competitive market order' as the first principle of the future regulations of big-tech platforms.[296]

Soon in November 2020, data became an important factor in the *Antitrust Guideline on Platform Economy (Draft for comments, 2020)* issued by the State Council Antitrust Committee at the SAMR. This top-level *Antitrust Guideline (2021)* was passed in February 2021, which listed 'the capacity to hold and process relevant data' as a determining factor in

[291] Beijing to break up Ant's Alipay and force creation of separate loans app, at https://www.ft.com/content/01b7c7ca-71ad-4baa-bddf-a4d5e65c5d79; China plans to break up Ant's Alipay and force creation of separate loans app - FT | Reuters

[292] Google loses challenge against EU antitrust ruling, $2.8-bln fine | Reuters 2021.11.10

[293] China fines Alibaba record $2.75 bln for anti-monopoly violations | Reuters ; SAMR, 2021 April 10th. Administrative penalties on Alibaba Group Holding Limited for its "choose one of two" monopoly in the online retail platform service market in China. [市场监管总局依法对阿里巴巴集团控股有限公司在中国境内网络零售平台服务市场实施"二选一"垄断行为作出行政处罚]

[294] China Fines Meituan $530 Million in Second Tech Antitrust Case - The New York Times (nytimes.com) ; SAMR, 2021 October 8th. Administrative penalties on Meituan for its "choose one of two" monopoly in the online food delivery platform service market in China [市场监管总局依法对美团在中国境内网络餐饮外卖平台服务市场实施"二选一"垄断行为作出行政处罚]; Meituan's Global Offering (IPO on HKEX).

[295] Shanghai Stock Exchange, 2020. Decision on Suspending the Listing of Ant Group Co., Ltd. on the Science and Technology Innovation Board [关于暂缓蚂蚁科技集团股份有限公司科创板上市的决定]

[296] Sohu.com, Central Bank: Joint interview with Ant Group announced! Pointed out existing problems and 5 major rectification requirements [央行：联合约谈蚂蚁集团情况公布！指出存在问题与 5 大整改要求！_监管 (sohu.com)]

the evaluation of a platform's market dominance, under the category of the platform's 'financial and technical conditions' (Article 11 section 3) [297].

'Platform's data occupation situation (平台占有数据情况)' also became a 'considering factor' in defining whether a platform can be deemed as an 'essential facility' (*bixu sheshi*) (Article 14). Whether data itself can be deemed as an 'essential facility', however, is not clear under the *Antitrust Guideline 2021*. In fact, the 2021 Guideline explicitly deleted the part on 'deeming data as essential facility' in Art. 14 of the 2020 Draft. Data was instead added as a new factor for deeming whether platforms are essential facilities (cf 2020 Draft). [298]

Nevertheless, data is listed as one of the 'essential resources and facilities' （必要资源和必需设施）among 'technology, IP', 'channel (*qu dao*), and user' in the consideration of platform-related merger controls (Article 20).These mixed signals imply the regulator's ambiguous attitudes towards considering data as an essential facility, and have triggered debates among Chinese antitrust scholars (Wang and Wu, 2021; Chen YW, 2021; Li SJ, 2021).

'Data and algorithms' were also listed as technical tools to 'restrict transactions' (Art. 15), and added as ways to form horizontal (Article 6), vertical (Article 7) or 'Hub-and-spoke' agreements that are monopolistic (Article 8). 'Depriving of data' (Article 21 section 1) and 'opening data infrastructure' (Article 21 section 2) were added as antitrust enforcement tools[299].

After the passing of this *Guideline (2021),* two massive antitrust fines were respectively issued to Alibaba and Meituan. Among others, the use of 'technical tools of data and algorithms' was deemed by SAMR as conducive to Alibaba and Meituan's abuses of market dominance that excluded competitors by forcing suppliers to 'choose only one (platform) between two (options) (*er xuan yi*)'. Besides these two fines, SAMR also issued two similarly drafted 'administrative guidance' that respectively prohibit Alibaba and Meituan's future anticompetitive behaviours including those that involve 'the use of data and algorithms'[300].

It is worth noting that SAMR, the antitrust agency, did *not* claim that data or algorithmic power were considered as defining factors for Alibaba or Meituan's market dominance. Therefore, no antitrust measures on firm structures were taken[301].

---

[297] State Council Antitrust Committee [2021] No.1. Antitrust Guideline on Platform Economy (2021.2.7) [国务院反垄断委员会关于平台经济领域的反垄断指南]

[298] SAMR, 2020, Anti-monopoly Guidelines for the Platform Economy (Draft for Comments). State Council Antitrust Committee [2021] No.1. Antitrust Guideline on Platform Economy.

[299] *Id.*

[300] SAMR, 2021, Administrative penalties on Alibaba; SAMR, 2021 Administrative penalties on Meituan..

[301] *Id.*

Another high-profile antitrust practice focused on data and platform interoperability. In September 2021 the MIIT (the industrial regulator but not the antitrust agency) held an administrative guidance joint-meeting with big-tech companies including Douyin (the Chinese TikTok), Taobao (Alibaba), Wechat (Tencent) and others. The MIIT required these tech-giants to stop blocking each other's links and data flow on their platforms or affiliated companies within a week in order to improve interoperability[302]. This is similar to the EU's *DMA (2022)* which requires gatekeepers to ensure fair competition for their third-party suppliers[303]. While Wechat is still reported to be partly resisting Taobao (Alibaba) links[304], major platforms including Wechat, Alibaba and Douyin changed their technical architectures to comply with this administrative guidance[305].

These antitrust practices combine new legislative efforts with state enforcement and administrative guidance, which together require big-techs to develop self-governance rules and help maintain a competitive digital market order. These practices institute a new antitrust regime in the digital economy and co-produce common knowledge in the new epistemic communities of antitrust agencies and private actors.

6.4.4 Summary: State-Market Co-Production via Institutionalisation

The above evidence shows that the Chinese state has in recent years intensively adopted formal institutions in the fields of data security, data protection and antitrust to deal with data harms / risks in the rising digital 'economy'. This institutionalisation strategy co-produces and stabilises common knowledge in the epistemic communities of data governance in order to scale a national digital economy. Co-production means the efforts from both the state's formal practices and the private actors' informal knowledge exchange and self-governance at scales. The data security, privacy, and competitive market order in China's digital economy are provided by both formal and informal labour from public and private actors.

[302] The Ministry of Industry and Information Technology held an "Administrative Guidance Meeting on Blocking Website Links" 2021.9.11 [工信部召开"屏蔽网址链接问题行政指导会"]

[303] DMA 2022.

[304] Fan J. 2021.10.1. thepaper.cn. How can the Internet industry "tear down the wall"? "Security" becomes the focus, and the phenomenon of garbled characters still exists [互联网行业如何"拆墙"？"安全"成焦点 乱码现象仍存在]

[305] Ren XN. 2021.9.17. eeo.com.cn. WeChat can now redirect to Taobao, and the "unspoken rules" of the Internet wall will become a thing of the past [微信能跳转淘宝了，互联网围墙"潜规则"将成往事]

6.5 Activation: Data Property, Commons, or 'Party manages data'

Evidence in China manifests data fragmentation at all levels. Public and private data handlers keep data to themselves and lack the incentives to share data as an 'essential factor' in the market. To deal with data fragmentation the Chinese state adopted what this chapter calls the 'activation' strategy since 2015.

In the context of China's data governance, activation is used to contest data access and ownership, where the state innovates ways to 'activate' data flow and usage, by experimenting with data licensing, ownership, 'use without retaining', and Party management. These different legal, technical or political arrangements 'code' or contest the socio-legal nature of data, before stabilised institutions and common knowledge can be co-produced.

Evidence of 'activation' is drawn from three cases. First the push and pull between Ant Group and Central Bank over personal credit information, second the contrasting approaches to data property between Shenzhen and Beijing, and third Tianjin's case of 'Party manages data'.

6.5.1 Ant Group v the Central Bank—Licensing or Commons?

6.5.1.1 Central Bank's licence to Ant Financial (2015-2018)

Between 2015-2018 the Chinese Central Bank used private licensing to activate the market use of personal financial credit information (which are different from personal public credit information or 'social credit'). Before 2015 the consumer credit-reporting (*geren zheng xin*) business was monopolised by the Central Bank, which meant that only public reporting of personal credit information was legal. This meant first, only the Central Bank's database can be analysed to provide risk-assessment services for commercial banks' loan decisions (though commercial banks can always build their own risk models based on internal data for internal use). Second, the data types are also strictly limited, for example, to personal credit cards payment, mortgages payment, and other financial information.

However, in 2015 the Central Bank for the first time granted eight provisional licences of consumer credit-reporting to Big-Techs including Ant Financial (before changing its name to Ant Group in 2020) and Tencent (whose consumer credit product was never effectively

online)[306]. This gave birth to Sesame Credit (of Ant Financial), which uses people's behavioural data to produce financial / social risk-scores.[307]

Central Bank's licences in 2015 brought two main changes in terms of the use of personal information. First, it allowed private actors to handle personal credit information that they privately collected or acquired, giving up the Central Bank's monopoly on consumer credit-reporting. Second, the types of data being handled expanded from the narrow consumer credit data to the use of 'alternative data' including social networks, consumer habits, and other information collected by the FinTech giants.[308]

It is worth noting that in 2015 there were very few data protection laws in China except for the *Ninth Amendment of the Criminal Code* which served as a redline of deterrence towards novel data practices.[309]Meanwhile the 2013 *Regulation on Credit-reporting Industry* (similar to the U.S. *Fair Credit Reporting Act 1970*) only provides the legal framework for handling traditional consumer credit data in the credit-reporting industry[310]. The big-techs' handling of 'alternative data' for multiple analytical purposes was thus left in a legal grey zone, before the Central Bank's 2015 licences granted them with more degrees of regulatory certainty and protections.

A respondent who is a law firm partner in Beijing, argued that these provisional licences were taken not so much as an opening of the private consumer credit-reporting market, but as a regulatory haven for big-techs to collect data and innovate new technologies.

> 'The tech giants were desperate to get the Central Bank's provisional licences not because of the consumer credit-reporting business itself. They wanted to have regulatory protections in collecting and analysing new data. That was a regulatory haven.' [311]

In practice, Central Bank's 2015 provisional licences provided a window period for data-driven and AI technological innovation. Sesame Credit (of Ant Financial) for example was able to experiment 'smart' projects on Alibaba's massive e-commerce ecosystem as well as

[306] People's Bank of China, 2015. Notice on Preparing for Personal Credit Reporting Business.

[307] See, explanation by sesame credit's former ITO [助力共享经济，芝麻信用背后的技术 - 蚂蚁集团金融科技 (antfin.com)].

[308] *Id.* See Sesame Credit's ITO explains how different types of data are used besides traditional financial credit-reporting data.

[309] *Ninth Amendment of the Criminal Code 2015.*

[310] State Council, *Regulation on Credit-reporting Industry* [2013] No.631. [征信业管理条例]

[311] Law firm partner. Beijing, 2021.4

collaborating with local governments in early social credit and smart city projects, though the real effects and normative aspects of these projects are debatable (Dai, 2020).[312]

6.5.1.2 Central Bank's Data Pooling Attempts after 2018

In 2018 however, the Central Bank did not renew the eight consumer credit-reporting licences after realising the problem of data fragmentation and data silos created by these companies. It criticised the tech-giants like Ant and Tencent who kept their data strictly to themselves as 'data islands'[313].

The Central Bank instead established a fully licensed joint-venture, Baihang Credit in Shenzhen, with a shareholding structure that gave Ant Financial, Wechat (Tencent) and other seven companies each 8% shareholdings. The goal was to incentivise these private tech giants to pool their personal credit data into this joint venture. However, such pooling effort is reported to have stagnated as of 2019[314]. This may have contributed to the motivation of the state to use antitrust measures against Ant Group as discussed above.

In December 2020, the Central Bank fully licensed another consumer credit-reporting company, Pudao Credit, which situates in Beijing and is also a joint venture between state and corporate holders, including big-techs in China like JD.com (25%), Xiaomi (17.5%), and Kuangshi Technology (17.5%)[315]. It is reported to be trialling new ways of big-data innovation for micro-finance, but it is unclear if data pooling is arranged between these shareholders as of 2021[316].

This history of the central bank v Ant Group since 2015 shows the push and pull between the state and corporations trying to deal with data fragmentation by activation. Moreover, the following section presents a local competition between Beijing's 'data pooling' approach, and Shenzhen's attempted data property right approach.

---

[312] Also see news, e.g. 'Use Sesame Credit to borrow and return books at Hangzhou Library' 2017 [刷芝麻信用"进"杭图借还图书]

[313] Chen Y. 2017.4.26. Full text of Wan Cunzhi's speech: None of the eight pilot personal credit reporting institutions are qualified, and qualified institutions should be like this. [万存知演讲全文：八家个人征信试点机构无一合格，合格机构应该是这样]

[314] Hao X. 2019.9.30. The deadlock in the credit reporting industry among hundreds of banks, a battle between Tencent and Alibaba [百行征信僵局，腾讯阿里困兽之斗]

[315] People's Bank of China's licence to Pudao Credit 2020.

[316] Pudao Credit: Dislocated development, collecting credit data other than credit [朴道征信：错位发展，采集信贷之外的信用数据] 2021

6.5.2 Beijing v Shenzhen: 'Use without Retaining' or Property Right?

6.5.2.1 Data Fragmentation due to Ownership and Compliance Uncertainties

A respondent, who is a Shanghai official, claimed that it is very difficult for local governmental agencies to share data with each other because there are risks that internal mistakes would be made known to another public department. There are not many incentives to share data, while sharing means potential exposure of one's weaknesses.[317]

Another respondent complained about the lack of data access, as a data scientist and product manager in a state-affiliated credit-reporting company in Qingdao. The Central Bank's credit payment data is the most important data type to train and test the prediction accuracy of the algorithms. Machine learning (ML) techniques like back-propagation cannot work without this crucial data set.

> 'My team cannot train useful machine-learning models because we don't have the data on loan paybacks from the Central Bank's Qingdao branch. Without that data we cannot label our algorithmic predictions as good or bad.'[318]

The access to credit payment data was expected to be done by negotiation between the Qingdao's municipal (horizontal state) and the local central bank (vertical state). However this respondent was uncertain about the progress during the time of the interview.

The respondent also mentioned that other important public data like financial reports, tax, and payment for employees' social insurance, all needed negotiation with powerful local governmental agencies like the SAMR, tax agency, and the social security agency. Meanwhile accessing electricity usage data was expensive because large state owned enterprises (SOEs) like the State Grid Corporation of China already charge fees based on data access entries.[319]

There were also greater uncertainties of data ownership as well as data protection by the time of the interviews before PIPL 2021 was passed in August 2021 (which still left the data ownership problem unanswered).

> 'Without clear ownership arrangements and data protection rules how do we access public data? Can we keep the data we used in our hard drives?'[320]

In spite of these difficulties, this state-affiliated credit-reporting company took an innovative method called 'use without retaining the data'. The technical team basically goes to

[317] Government official, Shanghai, 2021.3
[318] Manager, Qingdao, 2021.2
[319] *Id.*
[320] *Id.*

different local branches of the ministries, accesses their public data and finishes the calculation in their office, and takes away the software product / algorithms with some predictive capacities without taking the data away. But the problem was that they needed to go to the offices and update the algorithms quite often.[321]

The above interview evidence only reveals the data fragmentation experienced by this particular firm in relation to public information access in Qingdao. However, the respondent claimed that they have extensively consulted and learnt from practices in other cities like Hangzhou and Shanghai. There has been active informal knowledge exchange between local governments and firms on how to access and trade data.[322]

Respondents at an internal discussion in Shanghai also reported their concerns on how to define data property and transaction under uncertain compliance regimes, especially concerning Shanghai's big-data centre and the pilot of *Lingang* free trade zone with the aim to be a data harbour (Shanghai, 2021.3) [323]. Interviews with officials at Hefei's big-data centre also show similar questions regarding data ownership and data protection (Hefei, 2021.3). The reported rise of China's big-data centres and companies also evidence the local attempts to 'activate' data use when data is actually fragmented without clear legal coding on data ownership and protection[324].

Beijing and Shenzhen are two poles within these myriads of local activation experiments. Beijing adopts a common's approach to data access, while Shenzhen tries to formally establish data property rights and interest in its legislation (yet without success).

6.5.2.2 Beijing: 'Use without Retaining'—a Data Pooling Approach

The Beijing International Data Exchange (BIDE) was established in March 2021 under the Beijing Municipality[325]. It partly functions as a traditional trading centre for big-data products like analytics instruments, their future financial gains, i.e. securitised data assets, and cross border data transactions (Art. 5 sec. 2)[326].

[321] *Id.*
[322] *Id.*
[323] Also see Shanghai Data Regulation (Draft for comments) 2021.
[324] NDRC [2020] No.1922. [关于加快构建全国一体化大数据中心协同创新体系的指导意见 (发改高技〔2020〕1922 号)]. Also see, Crn.cn, Big Data Centre: The "Granary" of the Internet Age [大数据中心：互联网时代的"粮仓"]
[325] Beijing International Data Exchange website. 北京国际大数据交易所 (bjidex.com)
[326] Notice of Beijing Municipal Bureau of Local Financial Supervision and Administration and Beijing Municipal Bureau of Economy and Information Technology on Issuing the Implementation Plan for the Establishment of Beijing International Big Data Exchange. [北京市地方金融监督管理局 北京市经济和信息化局关于印发北京国际大数据交易所设立工作实施方案的通知] 2020.9.29

But more importantly, BIDE features the 'use without retaining' approach to data access (Art. 5 sec. 2) that is similar to the informally adopted method in Qingdao[327]. The difference is that in Beijing the BIDE is expected to function as a 'pool' that helps gather data from different data handlers, while in Qingdao's case the data scientist had to go to each different branches of the government to 'use without retaining' the data[328].

Experts have commented that BIDE is better than previous data trading centres (e.g. in Guizhou) because the 'use without retaining' method circumvents the ownership question and replaces it with the 'right to use' question. It also creates a safe environment for public and private actors to access data and run their computer programs based on specific needs. The fees charged are based on the entries of data access and quantity. It is safe both in terms of data security and data protection, which adopts a 'regulatory sandbox' approach inspired by the UK's financial regulations. The sandbox environment creates a virtual space to experiment the release of data or financial instruments before taking these practices into the real market environment.[329]

This chapter calls these BIDE practices as attempts towards 'data pooling', where data are being collectively governed by data communities who regularly contribute and use such data commons, with the state backing its security and compliance regimes. This categorisation is inspired by Ostrom's polycentric governance of common-pool resources (CPR), which defies a dichotomy of state v privatisation on the governance of commons like forests or fisheries (Ostrom, 2010).

The effects of Beijing's data pooling practices, however, still need to be observed in the long-run. Like Qingdao, the 'use without retaining' practices can also face the problems of regular data updates and data fragmentation when critical data providers like the Central Bank refuse to share data.

6.5.2.3 Shenzhen: Data Property Right and Interest?

Beijing's 'use without retention' is in contrast to Shenzhen's (unsuccessful) attempt to code 'data rights' into its local legislation, which leaned towards a property rights approach. In 2020 October, the central Party-state issued an implementation plan (2020-2025) which

[327] Manager, Qingdao, 2021.2.
[328] *Id*.
[329] Expert opinions in 2021 news article; 2021.4.1 Beijing International Data Exchange established. [北京国际大数据交易所成立].

instructed Shenzhen to start piloting data property and transaction institutions, and data protection mechanisms[330].

As a result, in July 2020 Shenzhen proposed its first draft of *Shenzhen Data Regulations (2020)*. It was the first time in China that a formal legislation draft tried to stipulate 'data right' articles. They include Art. 4 on a personal data right to 'self-determination, control, processing / disposal, benefit and redress', and Art. 52 protects 'data rights' that are exclusive to its collector or generators[331]. This triggered constitutional debates on whether the local laws can establish such data rights even before higher national laws allowed, and on whether this 'data right' would prevent the flow of data[332]. It was also argued that in the 'Wechat unfair competition case' (2020) the Hangzhou Internet Court decided that personal data is protected as a 'right of control' at the individual level, but the corporations only have competition rights with regards to the 'data pool'[333]. This was used by legal scholars as evidence against writing 'data rights' into the latter two versions of the *Shenzhen Data Regulation (2021)*[334].

In July 2021 Shenzhen passed its final version of the *Data Regulation* (effective January 2022) after the second draft (May 2021). It deleted all articles regarding data rights, and explained in a companion legislative note that there has been 'no consensus on data rights', and it is difficult to create this 'new right category' in local legislation. However, in Article 4 it does recognise the property right of data products and services made by natural and legal persons, which is not controversial in the existing legal framework.[335]

Shenzhen's data legislation saga is to some extent similar to the Central Bank's 2015-2018 private licences, though Shenzhen's early efforts were rejected in later drafts. The state started by proposing privatised legal arrangements that incentivised corporations to innovate

[330]The General Office of the CPC Central Committee and the General Office of the State Council issued the "Implementation Plan for the Comprehensive Reform Pilot of Shenzhen to Build a Pilot Demonstration Zone of Socialism with Chinese Characteristics (2020-2025)" 2020. [中共中央办公厅 国务院办公厅印发《深圳建设中国特色社会主义先行示范区综合改革试点实施方案（2020- 2025 年）》]

[331] Notice of the Shenzhen Municipal Bureau of Justice on soliciting public opinions on the "Shenzhen Special Economic Zone Data Regulations (Draft for Comments)". 2020[深圳市司法局关于公开征求《深圳经济特区数据条例（征求意见稿）》意见的通告] (accessed 16 December 2022).

[332] Yuan and Zhang, 2020.7.20. Nandu.com. 'Does the local government have no right to legislate on the "right to data"? The draft of the Shenzhen Data Regulations has aroused heated discussions among experts' [地方无权对"数据权"立法？深圳数据条例意见稿引专家热议.]

[333] （2019） 浙 8601 民初 1987 号 [2019, Zhejiang, 8601 Civil first instance case no. 1987] ；（2020）浙 01 民终 5889 号 [2020, Zhejiang, 01 Civil appeal case no. 5889]

[334] Yuan and Zhang, 2020.7.20. Nandu.com.

[335] Shenzhen 7th People's Congress Plenary Committee Report No. 10. On passing the Shenzhen Economic Special Zone's Data Regulation, 2021. Effective 2022.1.1 [深圳市第七届人民代表大会常务委员会公告（第十号）]

new ways of using their data, but later challenged by the needs of bettering data access and flow. In contrast, the data pooling idea also appeared in both the Central Bank's post-2018 approach on consumer credit information and the Beijing 'use without retaining' cases.

6.5.3 Tianjin: 'Party manages data'

A final form of the activation strategy is by resorting to the Party's leadership. It is more politically flexible than the Central Bank's licensing and the experiments on data access or ownership (Beijing v Shenzhen).

In Tianjin Municipality, the party-vanguard raised the idea of 'Party manages data' (*dangguan shuju*) since 2019, which so far only focuses on facilitating the use of public information collected by local government agencies[336]. News reported that this strategy is crucial for data security and improving governance performances. For example, it was reported to be very useful in the COVID-19 pandemic governance. Public big data was used in conjunction with the 'grid management' in streets and neighbourhoods. The 'red, orange and green QR code' system was claimed to be effective where civil servants and community workers use these systems in respective 'grids' (streets) to help identify risks[337].

Wang Yun, the chief officer of the Tianjin Big Data Centre (a public department in Tianjin government) argued in a news report that: **'**Party members should have some spaces and immunity to trial and error in their innovative use of data for the public good'[338]. In 2020 the Tianjin Big Data Centre passed an *Implementation Measure* (2020) to list certain immunity measures to promote the local party cadres to take more responsibilities and adopt new technologies in governance. This is a follow-up policy of the Tianjin Municipality's 2018 *Action Plan (2018-2020)* which aimed to tackle the problem of party members' 'inaction and not taking responsibilities'[339].

---

[336] Insist on the Party manages data method to build digital Tianjin, www.cac.gov.cn, 2019.4.15 [坚持党管数据建设数字天津]

[337] Zheng G. 2019.9.8. "Party management of data" is the only way to ensure data security ["党管数据"是保证数据安全的必由之路]

[338] Chen RX. 2021.1.5. "Party manages data" Tianjin sample - Interview with Wang Yun, Deputy Director of the Publicity Department of Tianjin Municipal Committee and Director of the Cyberspace Administration of Tianjin Municipal Committee ["党管数据"天津样本——专访天津市委宣传部副部长、市委网信办主任王芸]

[339] Tianjin Daily, 208.4.3. Muncipal, Regarding the in-depth implementation of the three-year action plan for special governance of inaction and irresponsibility (2018-2020) [关于深入开展不作为不担当问题专项治理三年行动方案（2018—2020 年）]

Wang also claimed that the Party's leadership is important to break 'information islands' and 'data chimneys', and big data is crucial to the Party's governance capacity and performance. She categorised the major achievements of this strategy since 2019 as enhancing top-level designs of governmental data', 'promoting the development of digital economy' and 'services to the people'. Wang provided some examples. (1) Governmental data sharing with national and ministerial agencies to coordinate financial and market regulations and building risk-models and emergency management systems. (2) Using data to help with economic recovery after the COVID-19 pandemic in 2020 by analysing economic data (e.g. on tax) to help private companies decide when to get back to work and alleviate their pressures on cash-flow and loan payback. The Tianjin Big Data Centre also designed a technology called 'code for frozen food'. It traces the supply of frozen food from sea ports to the serving tables, as COVID-19 virus proved to survive in frozen environments via the supply chain. (3) Providing public information to the local communities on daily life, including municipal construction plans, environment concerns, fire safety, noise pollution, police alerts, community disputes and so on.[340]

Outside of Tianjin, this 'Party manages data' approach has also been advocated, for example by large state-owned enterprises in their data security compliance work in general.[341]

Despite its emphasis on a political and flexible handling of data, Tianjin's 'Party manages data' strategy still needs to comply with data laws including the PIPL 2021 that regulates both public and private 'data handlers'. The party-state most likely, can only legally access and use public information collected by state agencies themselves under legal procedures. Against private actors, Tianjin is actually one of the first few cities that passed a local legislation (effective January 2021) that bans the private collection of biometric data including facial data (Art. 16)[342]. Government news also claim that the city is at the forefront of personal data protection practices against abusive private information collection and processing.[343] Nevertheless, there is rather limited information on the protection against public abuse. The

[340] Guangming Daily, 2021.8.18. Tianjin's "smart" management of data by the Party, [党管数据的津门"智"理, 光明日报] 2021.8.18.

[341] Ou, Li. 2021. *Qizhi Magazine,* Adhere to the Party's management of data and ensure data security [坚持党管数据保障数据安全]

[342] Tianjin Social Credit Regulations, Tianjin 17th People's Congress Plenary Committee Report, No 63. Effective on 2021.1.1 [天津市社会信用条例（天津市第十七届人民代表大会常务委员会公告，第 63 号）]

[343] Tianjin steadily develops work on data security and personal information protection, www.tjcac.gov.cn, 2020.8.20 [天津扎实开展数据安全和个人信息保护工作]; Tianjin works on regulating APP illegal collections of personal information, www.gov.cn, 2019.11.23 [天津市开展 APP 违法采集个人信息集中整治]

empirical and normative aspects of Tianjin's 'Party manages data' practices still require further fieldwork and long-term observations.

6.5.4 Section conclusion: Legal coding / pooling of data assets via activation

Attempts have been made to 'code' data with legal arrangement in China, yet without conclusive agreements among stakeholders. Data fragmentation is revealed at different ministerial (vertical), regional (horizontal) and corporate levels. To cope with data fragmentation, the Chinese party-state adopted different forms of activation strategies as shown in case studies. These include the cases of Central Bank v Ant Group, Beijing v Shenzhen and Tianjin's 'Party manages data'.

In these three case studies the state provided legal and policy spaces for the coding and contestation of data's ownership and access. If we borrow Roman law's concept, data so far, in its individual form, seems to be *res nullius*—physical things that have not or have never had an owner. This legal category includes the individual 'wild animal', or *ferae naturae*, which will not become property until being captured / killed.

Following this metaphor, the activation strategy allows individual units of data, i.e. the new 'wild animal', to be captured in a window period by public and private 'hunters'. The strategy allows them to try ways that best use the captured data in privatised or collective ways. The competition between ministerial (vertical), regional (horizontal) and corporate interests jointly contest the technical-legal arrangements of data access and use, with informal knowledge exchange happening in data governance communities including scholars, lawyers, and social groups across geographic or administrative regions.

While the activation strategy provides dynamic periods of state-market co-production of new practices and knowledge, this paper also observes an emerging 'data pooling' approach from these practices. Instead of hunting 'wild animals', China's data governance in practice may lead towards a common-pool resources (CPR) governance approach, similar to the governance of forestry or fisheries studied by Ostrom. If so, activation may transit into institutionalisation which could solidify a CPR-based practices of data governance.

6.6 Layering: Data Cycles, Local Habitats, and Sustainable Design

As required in the party-state's top-level data policies, data security, data markets, access and transaction are all foundations for the ultimate goal to 'increase the values of data resources'[344]. Evidence however finds a lack of valuable and sustainable use of data in governance settings, which the local public actors try to address.

This section 'Layering' describes the local state actors' attempts to build sustainable data cycles and habitats with in-depth understanding of data's local institutional contexts. This can be seen as a data strategy that is concerned with better designs of laws and technologies in order to achieve valuable and sustainable data use in governance. The strategy provides an insight that data cycles are generated locally and data scaling is limited by institutional layering. Evidence is mainly drawn from empirical observations of China's smart court and the social / financial credit risk-assessment projects.

6.6.1 Smart Court: National v Local Practices

The 'Smart Court' project in China has experienced difficulties when applying data-driven technologies to the court's daily operations. Not only did such difficulties happen at the Supreme People's Court's (SPC) national case recommendation system, but also many provincial High Courts' databases and models. Evidence shows limitations in these efforts that use China's judicial big-data with the goal to help judges make consistent judgements and accelerate the proceedings.

More specifically, the national (SPC) and provincial (High Court) case recommendation systems are not effective enough to gain popularity among judges because the recommended cases were not always very relevant. This low relevance in matches can be partly due to the insufficiency of digitised court judgments before the 1990s / 2000s to form long-term time series, and sometimes the concentrated caseloads in certain years or regions that were weighted over-heavily in the algorithmic models (Zuo W., 2018). A respondent, who is a judge and court administrator in one of Shanghai's intermediary courts, said:

> 'While technology can improve the efficiency of the court, the quality of judgments cannot be rushed in the Smart Court progress.'[345]

[344] CPC central committee and State Council: Instruction on Improving the Market Mechanism for Allocating Essential Factors, 2020.3.30.

[345] Judge and Court Administrator, Shanghai, 2021.3.

The judges in this Shanghai court use another locally designed 'typical case' system, instead of the national system, the respondent also mentioned. This local system is built by more traditional ways of human-selection, i.e. senior judges come together to discuss what new cases can be seen as important and typical, and circulate these cases to the entire court and lower courts.[346]

These 'typical cases', without any high-tech involved, are claimed to be more helpful in the long-term. It alludes to the temporal aspects of layering—data and its knowledge accumulated through time within institutions with path-dependency. Local courts can have their own concerns and ways of generating judicial data which can be difficult when scaling to the national level. This means data-driven technologies in governance contexts might be more effective at local level with longer time-scales.

Another example is the 'algorithmic case allocation' system used in this Shanghai intermediary court to automate the process of case assignment. It claims to help prevent judges from arbitrarily picking cases they wish to decide, and also makes more efficient division of labour by sending simple cases to junior judges while the complex ones to senior judges.[347] This local court practice gained much support by judges at the court, which is largely due to the active participation of judges in the design and updates of this algorithmic system. At the initial design of this system, data scientists and judges sit together to discuss the data categories and the labelling of the cases' cause of actions. Periodically the senior judges come back to evaluate the system's performance, following the operational cycles of the court and the time when legal institutions change. For example, when new types of cases emerge due to legal amendments or new judicial interpretations, the senior judges would make decisions on how to change data categories in terms of the cause of actions, and ask the court's information team (with computer scientists) to update the system. This also follows the court's own judgement cycles which determine the data lifecycle of generation, processing and use.[348]

The smart court cases show that legal institutions operated at the local level are coded into the data cycle construction and modelling. These legally and historically embedded technologies retain shared knowledge and practices, like the 'sediments of time', shaped by interactions between lawyers, judges, administrators and the wider legal epistemic community.[349] These technologies, like laws, thus require timely updates and alignments to the changes of external environments, including their legal institutional context and the court's

[346] *Id.*
[347] *Id.*
[348] *Id.*
[349] A concept borrowed from Koselleck, 2018.

own operational environments. Both legal and tech epistemic communities are crucial in this law-tech coevolution process.

6.6.2 Law-tech Coevolution and Layering in Design

The above qualitative evidence reveals aspects of the practical hurdles that need to be surpassed before achieving sustainable data governance for valuable data applications. Compared to industrial big-data, the massive amount of public information (including court judgements) and / or personal information may not be as easy to use, for example in judicial and social / financial governance settings.

As a result, data assets are hard to evaluate, especially when the legal contestation on data property and commons is still ongoing. Alibaba and Deloitte's joint report (2019) also pointed out that data are not yet deemed as assets in accounting, and thus cannot be represented in financial reports[350].

Some secondary grey sources also reveal emerging trends to make use of the public 'data sediments'. Huawei for example claimed that 'we help the government have its data assets sedimented.[351] It also tries to provide services on big data ecosystems[352]. Various businesses and reports also actively use terms like 'awakening the sleeping data sediments'[353]. Although these slogans can be corporate marketing strategies, they share common themes of releasing the value of long-term accumulated public information in governance, which requires building a better data habitat.

Moreover, Shi's research on COVID prediction puts 'emphasis on heterogeneous, unstructured data (…) and let unfiltered data become sediments.' He explained that medical and socio-economic data need to be accumulated over time to achieve a good prediction of COVID-19 cases. 'Only high quality data can have good results'[354]. The big-data predictions

[350] Deloitte and Alibaba Research Institute, Path towards data assetisation, 2019, Research Report. [德勤携手阿里研究院联合发布《数据资产化之路——数据资产的估值与行业实践》报告 (deloitte.com)]

[351] Tian, L. Sediments of government data assets, realising data sharing upon demands. Huawe.com. [沉淀政府数据资产，实现数据随需共享 - 华为期刊 (huawei.com)] last accessed 2021.11.6.

[352] Theme on Big data ecosystem, See Huawei Cloud website] last accessed 2021.11.6 [大数据生态系统-华为云 (huaweicloud.com)].

[353] See e.g. Awake the sleeping data, wakecloud.com [WakeData 惟客数据-唤醒沉睡的数据]. Rapid development in big-data, awakening the 'sleeping' data to empower all industries [我国大数据快速发展，唤醒"沉睡"的数据赋能各行各业！]; Wu, S. 'Sleeping' data restricts big-data industrial development, jjckb.cn, 2017.3.23 [数据"沉睡"制约大数据产业发展 _ 经济参考网 (jjckb.cn)] last accessed 2021.11.6.

[354] Ren, Fangyan: a person that lets data precipitate. 2021.1.5 [任芳言：让数据沉淀的人—新闻—科学网 (sciencenet.cn)]. Reporting on Shi Yong's work in predicting COVID case growth in Wuhan. (last accessed 2021.11.6)

in turn shape laws and policies on COVID governance. This can also be seen as data layering practices that utilise long-term institutional data for governance and shape new data cycles of generation and use.

Attempts have been made to layer data institutions and build data habitats in Smart Court, social and financial credit systems, and some corporate or research use of public data. These emerging cases demonstrate a law-tech coevolution process where layering is a key strategy to the timely adjustment between institutions and the technological design and prevent pathological path-dependencies. For example, data cycles can be co-designed and updated by judges, regulators and data scientists in areas like market regulations and disclosure cycles. In specific industries like banking for instance, market regulators and commercial banks can be involved in the design of public data infrastructures that helps update data cycles and bank's lending decisions and financial regulations.

Public and private epistemic communities in this sense can generate new knowledge and practices for valuable data use. They can help provide a deeper understanding of how past data is locally layered through time shaped by law-tech coevolution. With emerging evidence of layering, this chapter conceptualises it as a strategy that helps better design institutional data habitats / ecosystems.

6.7 Chapter conclusion

The chapter conceptualises three data strategies of the Chinese party-state in practices: institutionalisation, activation and layering. It concludes that the Chinese party-state uses these strategies to co-produce a data-driven regulatory state and a national digital market. It contributes to the more general theoretical debate on state-market relationship, and argues that a rising rule-of-law state in China co-produces the digital economy, shapes data 'property' and commons, and co-evolves with new technologies. More specifically, the three state strategies are conceptualised based on institutionalist theories of law and economics, and are examined by doctrinal and qualitative evidence drawn from China's recent developments since the 2010s.

Institutionalisation is shown to work as a stabilising strategy. It facilitates state-market co-production of common knowledge by stabilising formal institutions that deal with data security, data protection and antitrust concerns. Cases are drawn from DiDi's IPO security review, the legislative process of PIPL featuring gatekeeper and managerial liabilities, and the antitrust actions on Ant Group, Alibaba, Meituan and other big-techs. These cases together show the rise of stabilising legal institutions and a knowledge-based regulatory state co-produced with a

national digital market. They also reinforce a public-private divide that facilitates division of labours between state and market actors in co-producing security, data marketing and market order facing a rapidly evolving digital society.

Activation is observed to serve as a contestation strategy. It triggers co-production of epistemic communities which experiment with data commons, ownership, or political leadership approaches to tackle data fragmentation. Activation allows for dynamic and provisional ways of activating the data market and use, while also exhibiting tendencies to be later institutionalised into stabilised common knowledge and coded into formal laws. Cases of Central Bank v Ant Group, Beijing v Shenzhen, and Tianjin's 'Party manages data' were provided. They demonstrate that a dynamic data commons / pooling approach is emerging, while data property is highly contested. These cases also show that data property or commons don't emerge by themselves, but require legal-political contestations.

Layering is evidenced to function as a sustainable development strategy. It only recently emerged as a strategy to build data cycles and sustainable habitats that facilitate the in-depth understanding of data and the valuable use of data. It reveals that while big data is useful in large-scale governance projects, new data technologies are layered, i.e. path-dependent, on existing socio-legal institutions. Cases of China's Smart Court, the social and financial credit systems, and other grey resources provide initial evidence on the state's layering approach towards co-design of data cycles and habitats for more valuable applications and deployments of data technologies.

The conceptual model of institutionalisation, activation and layering contributes to a more realistic understanding of how China's party-state is trying to scale up a national digital economy. It also provides evidence to validate the scaling and layering framing of the thesis, where in a variety of cases algorithmic governance is demonstrated to have helped scale up legal governance, but is still path-dependent on social norms and law. This is most salient in the layering strategy, where data and algorithms are shown to 'grow' through time in their socio-legal habitats.

This model of the emerging strategies and its theoretical foregrounding pave the way for more comparative studies and empirical research which is beyond the scope of this chapter. The normative implications of these strategies and data governance practices also requires further examination, ideally based on more empirical evidence.

Chapter 7 Conclusions

7.1 Research questions

To recapitulate, the thesis asks the overall research question: to what extent can text-based law be computed systematically, namely replaced / substituted by computer code.

After a chapter on introduction and a chapter on methodology, Chapter 3-6 each answers a specific set of sub-questions:

Chapter 1.

- Introduction

Chapter 2.

- Methodology

Chapter 3.

- What is the nature of algorithmic governance, compared to more established ways of text-based legal governance?
- What is the relationship between legal and algorithmic modes of governance, in relations to other established modes of governance like social norms and numbers?
- Is there initial evidence to support the hypothesis that algorithmic governance is 'layered' upon legal governance, namely that code scales up governance but is limited by its path-dependencies on laws?

Chapter 4.

- What is the nature of the 'scoring' systems in China's SCS?

- How do data and code work in the scoring systems, and what are their institutional roots?
- Is there evidence of scaling and layering effects, where code expands governance scales but is limited by its path-dependency on legal institutions?

Chapter 5.

- What is the nature of the 'blacklists' and 'joint-enforcement' systems in China's SCS?
- How do data and code function in these systems?
- Is there evidence to show scaling and layering of algorithmic governance upon legal governance?

Chapter 6.

- What are the emerging approaches, namely 'strategies', that the Chinese state adopts, to the governance of massive amounts of public and private data and code in the contexts of wider local and national data experiments?
- How do data and code, as modes of governance, work in practice in these experiments?
- Is there evidence of scaling and layering in these experiments?

7.2 Synopsis of the argument

The aim of the thesis is to answer the central question: to what extent can text-based law be computed systematically, namely replaced / substituted by computer code? The thesis argues that law is computable *only to the extent* that computer code is (empirically) sustained and conditioned by the operations of law in its meaning and institutional contexts.

This is similar to the way in which social norms (for example, customary law) can be 'codified' into laws, *only to the extent* that these laws are cognitively sustained by the operations of the accepted norms of that economy and community. This complementary relationship between social norms and laws is well evidenced within the law and economics literature, and the thesis argues that a similar 'layered' relationship exists between law and code, where code can scale up governance but is path-dependent on law.

Chapter 1 introduced the emerging field of computational law, and the limiting contours of the debate on 'code v law'. It provided a literature review on the question of law's computability, where some scholars proposed a 'legal singularity' hypothesis that law is functionally complete and thus code can fully automate / replace law, while some were more

sceptical of this singularity hypothesis. The main gaps of understanding were identified in the debate, which were the lacking of a comprehensive theoretical framework, and the lacking of empirical testing and evidence.

This chapter then explained the contributions and motivations of the thesis as to filling both gaps in theory and empirical evidence. For theoretical contribution, the thesis proposed an institutionalist framework to ground the 'code v law' debate, which connected the emerging field of computational law with more established theories in institutional law and economics. For empirical contribution, the thesis tested the validity of this new framework and the 'scaling and layering' hypothesis, by providing three case studies on the site of China's Social Credit Systems (SCS) and other data experiments.

In order to pave ways for the empirical chapters of the thesis, chapter 1 then provided a literature review of the rapidly growing scholarship on China's SCS and data experiments. It analysed three strands of literature on the SCS and explained the gaps of understanding. It showed the existing SCS literature lacked more flexible and realistic theoretical framings as some of them rely on more rigid matrices of orthodox law and economics. In addition there was a lack of detailed understanding of how code and law work in the SCS with relations to both traditional ICT infrastructures and new AI technologies. Empirical evidence from the field was also generally lacking. While identifying these gaps, the literature review laid out the basic frameworks of the SCS and also the structures of the empirical chapters, respectively on (1) the scoring systems of the SCS; (2) the blacklist and joint-sanctions systems of the SCS, and (3) other data experiments.

Chapter 2 explained the three methodological choices of the thesis. Firstly, it explained theoretical synthesis as a method used to draw together various threads of theories in law and social sciences. It explicated the institutionalist foundations of the thesis and the ontological claims on governance and institutions (e.g. Deakin 2024; Xu, 2019; Aoki, 2010; Supiot 2017). This paved the way for Chapter 3 that proceeded to propose an institutionalist framework. Secondly, the chapter carved out the bounded time and scope of the thesis, and uncovered the advantages and limitations of empirical methods. It was clarified that the case studies cover a timeframe between the 1990s and April 2023, after which the thesis mostly stopped incorporating new empirical evidence and only updated on a few important legal changes. The thesis thus did not cover the impacts of generative models like GPT-4 deployed after March 2023. Interview data were collected between December 2020 and May 2021 mainly in Beijing and Shanghai. Thirdly, interview data and methods were explained in detail. This included how the interviewees were recruited through snowballing methods, how data was collected

confidentially and transcribed later, and so on. By explaining the methodological choices and details on empirical data collections, this chapter started laying out the theoretical foundations of the thesis which is essentially (neo-)realist and institutional.

Chapter 3 proposed the theoretical framework of the thesis which explicates the nature of algorithmic governance as the most abstract form of governance layered on top of the less abstracted modes, which are social norms, laws, and numbers / statistics. The framework drew on institutionalist theories in three parts, which respectively explained how each pair of modes of governance is layered, that is between social norms and laws, between laws and numbers, and between numbers and code.

This institutionalist framework thus provided a 'layered model' of algorithmic governance. Starting from the bottom up, the indeterminate social practices, or 'magma', cool down and form the social norms as the first governance layer. Social norms were then codified into legal governance, and gradually moved to governance by numbers, and to governance by code.

Through this three-fold analysis, the chapter postulates the 'scaling and layering' hypothesis, that the more abstract mode of governance has the potential to rule larger geographies and populations (scaling), but is inherently restricted and path-dependent on the less abstract modes of governance (layering). This is because the more abstract modes of governance such as code and statistics, decay or wear out through time if they are not sustained by the cognitive framing effects of the less abstract forms of governance, like law and social norms.

More specifically, (1) first, scaling and layering between social norms and law was drawn from theories of institutional law and economics (e.g. North 1990, Deakin et al 2017, Hodgson 2023). This tradition of literature shows how formal legal systems functioned as an institutional force to scale up modern, capitalist market economies by overcoming the limitations of local or tribal trade networks based on social norms and / or *guanxi.* Empirical cases include for example, the emergence of modern capitalism in 18th century England (Hodgson, 2023), and the more recent developments of private law, courts and a national economy in China (Chen et al, 2017). Their theories and empirical fieldwork show that state laws co-evolve with informal institutions like social norms and *guanxi* in a complementary way. Legal governance is imprinted and path-dependent on the social practices and norms of that economy and community, often with negative layering effects that restrict the scale of governance.

(2) Second, between law and numbers there is also theoretical and empirical evidence for scaling and layering, mostly drawn from literature on critiques of governance by numbers

(Supiot 2017), mathematical modelling in decision-making and quantification of law (Deakin and Markou, 2020; Cobbe, 2020; Xu 2021, 2023), and the legal 'coding' of assets and wealth into quantifiable global capital (Pistor 2019). Here statistical and quantitative modes of governance were shown to expand governance scales to a global economy and financial market. But these numerical modes of global governance first require the legal 'coding' of assets from various jurisdictions into standardised numbers. Governance by numbers thus are limited and imprinted by their path-dependencies on legal governance. The global 2008 financial crisis, Enron accounting scandals, the limitations of global economic indicators and structural adjustment plans by the IMF and World Bank, and the restraints of mathematical modelling in governing public health risks in the Covid-19 global pandemic, were used as examples to show how governance by numbers has to be sustained and recalibrated by national and international legal systems and public institutions (e.g. Supiot 2017; Pistor 2019; Deakin and Meng 2021).

(3) Third, scaling and layering was preliminarily shown between numbers and code by drawing from theories of computation law (e.g. Cobbe 2020, Hildebrandt 2020a, Diver 2022, Deakin and Markou 2020), and surveillance and information capitalism (Zuboff 2018, Cohen 2019). Here ML and DL algorithms were revealed to be essentially mathematical models that are imprinted and path-dependent on data, which is in essence constructed by non-computational concepts. Here, data do not, in themselves, produce more data. The effective functioning of ML and DL algorithms and other AI systems, for example, requires timely updates on their data inputs. Problems like concept drift and cold start showed the layering effects which can potentially lock-in pathological institutions from social norms and laws, which frames the data and statistics and in turn, ML and DL algorithms.

It followed that societies must constantly (re)align these four modes of governance and recalibrate their 'fit' to social practices, if negative layering effects were to be avoided or mitigated. Data is in essence framed and 'grown' in a local socio-legal habitat. In the same way, governance by code requires reference to laws and social norms. As hypothesised by the 'layered model', this is so no matter how formally complete 'learning' algorithms might be, or become, in terms of replicating human decision-making and / or behaviours.

Chapter 4, 5 and 6 then used three empirical case studies to test the validity of this institutionalist framing and the 'scaling and layering' hypothesis.

Chapter 4 was a case study of the (semi)automated 'scoring' systems in China's SCS. It is shown these systems are in nature, data-driven algorithmic governance that are rooted in the US and European credit-reporting practices and public management scores. New empirical

evidence was provided on the SCS experimentations in order to show scaling and layering effects where code scaled up governance, but was limited by its path-dependency on laws. This supported the overarching argument of the thesis that law is computable *only to the extent* that computer code is sustained and conditioned by the operations of law in its meaning and institutional contexts.

The first half of the chapter focused on the financial-commercial scoring systems, with case studies on the Central Bank v Alibaba's Sesame Credit. (1) First, the Chinese Central Bank has maintained a national database, FCIFD, since the 1990s which mainly provides individuals and commercial banks with qualitative credit reference information. The individual or corporate users of the FCIFD then make their own local risk scores, some using ML / DL algorithms. The Central Bank's database essentially provided socio-legal information and resilience to a diversified scoring ecosystem that are modes of data-driven algorithmic governance.

(2) Private scoring systems like Alibaba's Sesame Score initially scaled up in the 2010s, but lost their momentum after 2018. Sesame Credit's processing of alternative data and use of AI and ML / DL technologies, were deemed as not conducive to cross-regional and / or cross-populational governance. The reasons listed by the Central Bank include conflict of interests, failure to be independent, neutral and just, and antitrust concerns over regional and commercial monopolies. They manifest that local socio-legal institutions again provided the cognitive foundations and meaning for these private credit scores.

The second half of the chapter zoomed in on local public score, with cases on urban v rural China. (1) In municipal cities, many local governments have created public scoring systems for both citizens and corporations. They are usually run by public-private joint-ventures, and some have attempted to use algorithms like ML / DL to automatically generate scores. Most citizen scores are found to be similar to private scoring systems like Sesame Credit, and usually applied in low-stake settings like zero-deposits for renting city-bikes or tourism discounts. High-stake settings are more often found in corporate scores which usually use simple quantified metrics. Some have been observed to help scaling local governance in specific settings such as public procurement and city management (Guangzhou's case), but were required to be scaled back, revised or abandoned after 2023 by the NDRC with concerns on regional monopolies. Interviews from Shanghai and Qingdao also showed that local corporate scores were often not effective. Local experts such as regulators and bankers with more field-based experience and informal knowledge were needed to recalibrate these standardised scores and make effective use of them in local contexts.

(2) In rural experiments, cases like Qingzhen in Guizhou showed that disparity of social norms between rural and urban areas have limited the expansion of the Honest Farmers experiments into its adjacent urban regions. The new scores which tried to apply across rural and urban regions did not gain traction and showed how layering happens when scores are applied to larger and more complex geographies and populations.

Together these four stories showcased the rise (scaling) and fall (layering) of the SCS's scoring systems. The negative layering effects were argued as can be avoided and recalibrated by using human expertise and informal local knowledge accrued in time. Evidence showed that this hybridisation of both formal and informal knowledge, based on local social and legal foundations, was better at effectuating statistical and algorithmic governance with functional and legitimate operations.

Data-driven algorithmic governance, like statistical and legal governance, was shown to grow through time from indeterminate social practices within their respective local habitats. Like a data ecosystem evolving in a 'fish pond', algorithmic governance was revealed to be rooted in the social practices, norms and laws of a specific habitat. Rule-makers thus need to be aware that quantified scores and their algorithms should be constantly maintained and recalibrated by local knowledge and human agents based on socio-legal institutions. When using (semi)automated scoring mechanisms, layering is always an inherent restraint to scaling of governance to complex and diverse geographies and populations.

Chapter 5 provided a second case study on the SCS that focused on the blacklist and joint sanctions mechanisms (JSMs) at national and local levels. They are in nature very different from the scoring branch of the SCS (Chapter 4) and have different roots. The first half of the chapter provided documentary evidence to show a more realistic picture of the national blacklists and JSMs, and explained their nature as code-driven automated law enforcement systems. Most of the technologies were revealed as not involving AI (except those facial recognitions embedded in the boarding gate cameras), but the simple yet large-scale code infrastructures. It was shown that the roots of these national algorithmic governance systems are China's Supreme People's Court's (SPC) experimentations on enforcing contracts and debt payments. These included (1) the *2013 Judicial Interpretation* which established the SPC debtors list as a pro-typical national blacklist, and (2) the two MOUs signed by the SPC with China's railway and airline networks which established the first large-scale automated JSMs, using AI facial recognition at boarding gates that can reject listed debtors from boarding high-speed trains and flights.

It was then evidenced how these judicial experiments were learnt by the central government (State Council and the NDRC) after 2014, and used to scale up the law enforcement of industrial regulations, such as those in the energy sector and in the regulations of employers of domestic migrant workers. A doctrinal analysis was followed to showcase the layering effects, where the legality and legitimacy of code-driven automated law enforcement systems were revealed. It was demonstrated how the code which runs the blacklists and JSMs, becomes infrastructures that are path-dependent on laws, and their functioning is sustained by the systematic updates and maintenance of the legal system, including its legal concepts, interpretations and daily operations.

The second half of the chapter zoomed in on a field-based study of local SCS in Shanghai, in particular its debtors list and JSMs maintained by the Shanghai local courts and civil service. It provided new interview evidence and triangulated it with doctrinal and documentary evidence. This new set of evidence was assessed to uncover the detailed workings of code and (civil procedural) laws in the enforcement of contracts and debts, through the views of local enforcers like judges, civil servants and corporate employees. It revealed gains and losses perceived by these local enforcers and uncovered three main layering effects that include (1) the rigidity of code, (2) the counter-productive behaviour changes of human enforcers, and (3) institutional overload and risks to system legitimacy.

It was again evidenced that human expertise and agency are crucial to allow for recalibrations of code by acting upon local socio-legal norms. This is particularly important in preventing potential problems like institutional overload, where the foundational social norms of 'honesty', contractual obligations, and good faith are undermined by code's blind self-reinforcements. The supervisors of these code systems were shown to worry in the interviews that if the numbers of listed debtors reach a critical mass, these debtors will stop believing in the legitimacy of the socio-legal institutions of contract, property and courts, and never comply as a result. Layering therefore happens not only at the granular level of code's operations at each enforcement decision (e.g. whether to reject a loan or project bid by a listed debtor), but also at the systemic and organisational level of cognition and legitimacy. It highlighted the argument that code is cognitively layered and sustained by social and legal governance.

Above all, it was shown in this chapter that the blacklists and JSMs are code-driven algorithmic governance that automates law enforcement, and are in essence sustained by the operations and maintenance of the legal system. More specifically, the case studies showed that the tasks of enforcement judges and officials at both national and local levels were being partially automated, from debt enforcement to the regulation of industries and businesses.

Scaling was observed as code's higher level of abstraction helps standardise enforcement practices and expand governance in various judicial or regulatory settings. However this high-level abstraction of code was still cognitively based on social and legal meaning and contexts, and these (semi)automated systems required legality and legitimacy. Algorithmic governance thus exhibited inherent limitations when applied across geographies and populations, and needed to be consciously recalibrated by human expertise and feedback systems based on the interpretations and maintenance of law and social norms.

Chapter 6 was the final case study that addressed the central research question using empirical evidence. It concerned the wider practices of China's most recent data experiments. New empirical evidence was provided to support the thesis's central argument that law is computable *only to the extent* that computer code is sustained and conditioned by the operations of law in its meaning and institutional contexts.

Based on the proposed institutionalist model of chapter 3, chapter 6 conceptualised three 'strategies' that the Chinese party-state has adopted in data and algorithmic governance, namely institutionalisation, activation and layering. (1) First, institutionalisation is the stabilising strategy that facilitates state-market co-production of common knowledge. In China's various data experiments the institutionalisation strategy was used by the state to establish formal institutions that deal with data security, data protection and antitrust concerns. Specific cases studies were drawn intensely from the period between 2020-2021, which include (i) the central government's data security review of DiDi's public listing on the New York Stock Exchange, (ii) the legislative process of China's Personal Information Protection Law (PIPL), and (iii) the central government's antitrust enforcement actions on Ant Group, Alibaba, Meituan and other big-tech companies. These cases first revealed how the national digital economy was scaled up by centralising public institutions, which facilitated cross-regional information and knowledge exchange, formation and governance. They also presented the evolving legal foundations of China's data experiments, and showed how code and data can only extend across geographies and populations by layering on legal governance.

(2) Second, activation is a contestation strategy used by the state to trigger co-production of epistemic communities. In data experiments, activation was evidenced by China's local innovations of competing approaches on data commons, ownership, or political leadership. Cases included (i) first, the struggle between Ant Group and the Central Bank. The Central Bank first attempted an industry-based 'data licensing' approach by giving Ant Group and other large corporations *de facto* legal safe havens for processing alternative data, but later found this to be sub-optimal and led to data fragmentation. It eventually adopted a 'data

commons' approach to facilitate the regional and national flow of consumer credit data, by facilitating pooling practices between tech companies and establishing joint-ventures instead. (ii) The second case was the opposing approaches Beijing and Shenzhen each took on data ownership. Beijing experimented with a 'use without retention' method that is similar to the Central Bank's data pooling practices. The chapter argued that it follows a 'Common Pool Resources' (CPR) logic of Ostrom's polycentric governance framework (e.g. Ostrom, 2010). Meanwhile, Shenzhen tried to establish a 'data property rights' in its local legislation, but failed due to constitutional restrictions. (iii) Third, the case of Tianjin's 'Party manages data' approach revealed the most flexible approach in facilitating data flow. This approach was reported to rely on the local Party's authority to break departmental data silos and make public data sharing possible for improving local services. These cases together demonstrated that a dynamic data commons / pooling approach is emerging, while data property is highly contested. It also became apparent that rules and institutions on data property or commons do not emerge by themselves, but require legal-political contestations. The cases revealed that the regional and national flow of data was heavily path-dependent, namely, layered on the social and legal institutions which defined data access, ownership and processing.

(3) Finally, layering is a sustainable development strategy that constructs data cycles and habitats which facilitate the contextual and effective use of data. This layering strategy was mainly observed in China's 'smart court' experiments and tech-companies' various data practices. In the case of smart courts, data and code, some involving AI and blockchain, were used to help judges make adjudication decisions, automate case allocations to judges which can prevent conflicts of interest, and secure evidence in enforcement. However it was revealed that judicial data of previous rulings so far is only useful when judges and computer scientists co-design and maintain the data cycles, categories, and the recommendation algorithms together. Otherwise the accuracy of case recommendations or predictions decay easily. This aligned with previous case studies on the negative layering effects that code is path-dependent on law and social norms, which need to be interpreted and made sense of by local human expertise. The layering strategy was a conscious circumvention of this negative potential, and deliberately paid attention to the construction of data cycles and cultivation of local habitats that involved expertise in both legal and computational sciences.

The conceptual model of institutionalisation, activation and layering was aimed to provide a more realistic understanding of how China's party-state was trying to scale up a national digital economy, when faced with negative layering restrictions. It also provided evidence to validate the scaling and layering hypothesis, and showcased in a variety of data experiments

that algorithmic governance was path-dependent on law. Data and algorithms were shown to 'grow' through time and be sustained in their socio-legal habitats, where human expertise remained the key to the contextual understanding and interpretation of social and legal norms that generate data and code for governance.

Above all, the broader motivation of the thesis was introduced in Chapter 1 as to seek the 'spirit' of human law in our time of emerging AI and automata. It was provoked by German legal theorist von Jhering in his writing on the 'Spirit of Roman law' in the 1800s. Von Jhering claimed that Rome unified and governed Europe first through violence (the state), then through religion (the church), and finally by law. In a broader sense, the thesis continues the query and asks whether after law, 'code' is becoming the next mode of unification and governance of the globe, with the help of both advanced AI technologies and large-scale infrastructures.

In von Jhering's query, Rome was the maker of the traditional modes of governance in Europe, with soldiers, statesmen, clergies and lawyers as the protagonists. It was shown in this thesis how the global north and China were making statistical and algorithmic modes of governance, with the protagonists being judges, technocrats, bankers, economists, and computer scientists in multinational big-tech corporations. By studying China's SCS and data experiments, the thesis demonstrated how a hybrid of state and corporate actors, learned from advanced economies like the US and European countries, co-produced algorithmic governance from the local to the national, and potentially the global level. More importantly, these empirical case studies revealed how code was path-dependent, namely layered, on law's operation, in a similar way laws always had to maintain a precarious 'fit' to the indeterminate social norms and practices (including those of the state, religion, culture / *guanxi* and so on).

7.3 Future research agenda and concluding remarks

For future research, this thesis opens the way to the pursuit of a series of long-term empirical and collaborative endeavours to better understand the 'law v code' debate in different jurisdictions. Researchers across the globe are invited to apply the proposed institutionalist frameworks to their own case studies and qualitative fieldwork. More evidence-based studies hopefully can further test the validity of the hypotheses of complementarity (that code complements law) and 'scaling and layering' (that code is layered on law). This institutionalist framing also pushes for future theoretical developments in the field of computational law, drawing on theories in law, economics, and computer science.

This thesis also illustrated the challenges and benefits of tackling the complex subject of China's SCS and data experiments using fieldwork-based methods and interview evidence. The approach could be used in future to study (1) not just the SCS but also wider socio-technical experimentations, and (2) not just in China but in other jurisdictions as well. Practical and policy-oriented studies, based in public and private institutions like local courts, civil services, tech companies, need to be conducted. These organisational agents could seek to understand how far negative layering effects such as 'institutional overload' are occurring due to automation, and how far negative layering can be successfully averted by human feedback in various inter-temporal and spatial settings. Further studies may also provide more details on how agents reinvent and operationalise reflexive laws and local practices in mitigating the risks associated with the negative path-dependencies of 'code'.

In addition, the thesis poses some questions that transcend jurisdictional boundaries. For example, on the relevance and applicability of computational law theories originating from common law and / or continental law jurisdictions in other regions, and the influence of Chinese legal and technological systems on the development of common law and / or continental legal theories in computational law. The thesis serves as a start to these questions.

Ultimately, it is hoped that the thesis can contribute to better legal and algorithmic governance across the world. As generative AI models like GPT-4 keep being upgraded and deployed in society since 2023, the thesis emphasises the truly 'generative' core of human society lies in the indeterminate social practices (or 'magma') stemming from human connections and agency. Text-based law has been protecting this human generativity by precariously 'fitting' law's abstracted, sometimes mathematical logic, to the dynamic, heated and fluid patterns of social practices. But it is unclear whether statistics and algorithms, including the most recent generative AI models, can to the same extent provide or protect this bottom-up generativity.

It may be that algorithmic governance in our time is helping modern law to scale-up across geographies and populations with ever complex governance needs. But it is argued by the thesis that we should not forget that the 'spirit' of the law lies in its foundations in generative social practices. The generative AI models, no matter how advanced in the future, are not generative in themselves, but always need to 'grow' just like data and law grow, layer by layer, from the local 'fish ponds' of human interactions. Code thus should not be expected to magically bring generativity, but be harnessed in order to foster and protect social connections. The proposed layered model of the thesis hopes to reveal, and provoke future research on, this intricate complementarity between the social, legal, statistical and computational.

Bibliography

Books

Abdi Aidid and Alarie, B. (2023). *The Legal Singularity: How Artificial Intelligence Can Make Law Radically Better*. Toronto: University of Toronto Press.

Alain Desrosières (2002). *The politics of large numbers : a history of statistical reasoning*. Cambridge, Mass.: Harvard University Press.

Alpaydin, E. (2020). *Introduction to Machine Learning, fourth edition*. Cambridge, MIT Press.

Amoore, L. (2020). *Cloud Ethics: Algorithms and the Attributes of Ourselves and Others*. Durham, NC: Duke University Press.

Aoki, M. (2010). *Corporations in evolving diversity : cognition, governance, and institutions.* Oxford ; New York: Oxford University Press.

Bevir, M. (2013). *A theory of governance*. Berkeley: University Of California Press.

Bradford, A. (2020). *The Brussels effect: How the European Union rules the world*. USA, Oxford University Press.

Buchanan J, Chai D.H and Deakin S (2012). *Hedge fund activism in Japan: The limits of shareholder primacy*. Cambridge University Press.

Bucher, T. (2018*). If...Then: Algorithmic Power and Politics*. Oxford: Oxford University Press.

Calo M., Froomkin, M. and Kerr, I. (2016). *Robot law*. Cheltenham, Uk: Edward Elgar Publishing.

Cohen, J.E. (2012). *Configuring the networked self : law, code, and the play of everyday practice*. New Haven Conn.: Yale University Press, Cop.

Commons, J.R. (1924). *Legal foundations of capitalism*. New York:Macmillan.

Commons, J.R. (1950). *The economics of collective action*. Madison: University of Wisconsin Press.

Cooter, R.D. (2013). *Solomon's knot : how law can end the poverty of nations*. Princeton: Princeton University Press.

Cormen, T., Leiserson, C., Rivest, R. and Stein, C. (2009). *Introduction to Algorithms. Third ed*. Cambridge, Mass.: Mit Press.

Deakin, S. and Markou, C. (2020). *Is Law Computable*? Oxford: Bloomsbury Publishing.

Deakin, S. and Wilkinson, F. (2005). *The law of the labour market : industrialization, employment and legal evolution*. Oxford: Oxford University Press.

Deng, L. and Liu, Y. (2018). *Deep learning in natural language processing*. Singapore: Springer.

Diver, L. (2022). *Digisprudence: Code As Law Rebooted*. Edinburgh University Press.

Dupont, A. (2002). *The visible hand : the managerial revolution in American business.* Cambridge Mass. Etc.: The Belknap Press Of Harvard University Press.

Ellickson, R.C. (1994). *Order without law.* Cambridge, Mass. Harvard University Press.

Ferguson N., (2012). *Civilization : the West and the rest*. London: Penguin Books.

Foucault, M., (2008). *The Birth of Biopolitics: Lectures at the Collège de France*, 1978–1979.

London: Picador.

Foucault, M. (2011). *The Government of Self and Others: Lectures at the Collège de France, 1982–1983*. New York: Picador/Palgrave Macmillan.

Foucault, M. (2012). *The Courage of Truth (The Government of Self and Others II): Lectures at the Collège de France, 1983–1984*. London: Picador.

Gareth Michael James, Witten, D., Hastie, T.J. and Tibshirani, R. (2013). *An introduction to statistical learning : with applications in R*. New York: Springer.

Hayek, F.A. (1973). *Law, Legislation and Liberty; Volume 1: Rules and Order.* London: Routledge and Kegan Paul.

Hildebrandt, M. (2020a). *Law for computer scientists and other folk*. Oxford, United Kingdom: Oxford University Press.

Hodgson, G.M. (2015). *Conceptualizing Capitalism Institutions, Evolution, Future*. University Of Chicago Press.

Hodgson, G.M. (2023). *The wealth of a nation: Institutional foundations of English capitalism*. Princeton University Press.

Howarth, D. (2013). *Law as Engineering: Thinking about What Lawyers Do*. Edward Elgar

Kooiman, J. (2003). *Governing as governance*. Sage, New York.

Koselleck, R. (2018). *Sediments of time: on possible histories*. Stanford University Press.

Latour, B. (2013). *An inquiry into modes of existence*. Cambridge, Massachusetts Harvard University Press.

Lessig, L. (2006). *Code and Other Laws of Cyberspace, Version 2.0.* New York: Basic Books.

Llewellyn, K. (1960). *The bramble bush.* Doobs Ferry, N.Y.: Oceana Publications.

Lin J. (2003). *Theories of the Social Credit System*. Fangzheng Publishing, Beijing. (In Chinese).

Liu X. (2016), *Credit Information Service and Big Data*, CITIC Press Group. (in Chinese)

Luhmann, N. (1995). *Social systems.* Stanford: Stanford University Press.

Mackay, D.J.C. (2003). *Information theory, inference, and learning algorithms*. Cambridge: Cambridge University Press.

Miller, M. J. (Ed.). (2003). *Credit reporting systems and the international economy*. MIT Press.

North, D. (1990). *Institutions, Institutional Change, and Economic Performance*. Cambridge: Cambridge University Press.

North, D.C. and Thomas, R.P. (1973). *The Rise of the Western World : a New Economic History*. Cambridge: Cambridge University Press.

Ostrom, E. (2005). *Understanding institutional diversity*. Princeton, Nj: Princeton University Press.

Parsons, T. (1951). *The social system*. Glencoe, IL: Free Press.

Pasquale F. (2015). *The Black Box Society.* Harvard University Press.

Pedhazur, E.J. and Schmelkin, L.P., (1991). *Measurement, Design, and Analysis: An Integrated Approach (1st ed.)*. Lawrence Erlbaum Associates Hillsdale.

Pistor, K. (2019). *The Code of Capital: How The Law Creates Wealth And Inequality*. Princeton, Nj: Princeton University Press.

Polanyi, K. (1960). *Personal knowledge: Towards a post critical philosophy*. Routledge, London.

Poteete, A.R., Janssen, M.A. and Ostrom, E. (2010). *Working together: collective action, the commons, and multiple methods in practice*. Princeton University Press.

Pound, R. (1923). *Interpretations of Legal History*. Cambridge: Cambridge University Press.

Pound, R. (1954) . *An introduction to the philosophy of law*. New Haven: Yale Univ. Press.

Primavera De Filippi and Wright, A. (2018). *Blockchain and the law: the rule of code*. Cambridge, Massachusetts: Harvard University Press.

Reder, M.W. (1999). *Economics: The culture of a controversial science*. Chicago: University of Chicago Press.

Rhodes, R.A. (1997). *Understanding governance: Policy networks, governance, reflexivity and accountability*. London: Open University Press .

Smith, A. (1776). *An Inquiry into the Nature and Causes of the Wealth of Nations*. Oxford: Clarendon Press.

Spiegelhalter, D.J. (2019). *The art of statistics : how to learn from data*. New York: Basic Books.

Strahilevitz, L. J. (2011). *Information and exclusion*. New Haven Conn.: Yale University Press.

Supiot, A. (2007). *Homo Juridicus: On the Anthropological Function of the Law*. Verso 2007. (Translated by Saskia Brown from the French edition of 2005).

Supiot, A. (2017). *Governance by numbers: The making of a legal model of allegiance*. Oxford and Portland, Oregon: Hart Publishing. (Translated by Saskia Brown from the French edition of 2015).

Susskind, R.E. (2010). *The end of lawyers? : rethinking the nature of legal services*. Oxford ; New York: Oxford University Press.

Susskind, R.E. (2017). *Tomorrow's lawyers : an introduction to your future*. Oxford: Oxford University Press.

Upton, G., Cook, I. (2008). *Oxford dictionary of statistics*, Oxford: Oxford University Press.

Wang, X.. Blockchain chicken farm: and other stories of tech in China's countryside. FSG Originals, 2020.

Zuboff, S. (2019). *The Age of Surveillance Capitalism: The Fight for a Human Future at the New Frontier of Power.* New York: Public Affairs.

Articles

Alarie, B. (2016). The path of the law: Towards legal singularity. *University of Toronto Law Journal*, 66(4), pp.443–455.

Alarie, B., Niblett, A. and Yoon, A. (2016). Using Machine Learning to Predict Outcomes in Tax Law. *SSRN Electronic Journal*. 58, pp.231-254.

Aletras, N., Tsarapatsanis, D., Preoţiuc-Pietro, D. and Lampos, V. (2016). Predicting judicial decisions of the European Court of Human Rights: a Natural Language Processing perspective. *PeerJ Computer Science*, [online] 2(1), p.e93.

Avery, R. B., Calem, P. S., Canner, G. B., & Bostic, R. W. (2003). An overview of consumer data and credit reporting. *Fed. Res. Bull.*, 89, 47.

Backer LC, (2018). Next Generation Law: Data-driven Governance and Accountability-based Regulatory Systems in the West, and Social Credit Regimes in China. 28 *S Cal Interdisc* LJ 123.

Benvenisti, E. (1998). Margin of appreciation, consensus, and universal standards. *NYUJ Int'l L. & Pol.*, *31*.

Bobadilla, J., Ortega, F., Hernando, A. and Bernal, J. (2012). A collaborative filtering approach to mitigate the new user cold start problem. *Knowledge-Based Systems*, 26, pp.225–238.

Boyer, W.W. (1990). Political Science and the 21st Century: From Government to Governance. *PS: Political Science and Politics*, 23(1), pp.50-54.

Breiman, L. (2001). Statistical modeling: The two cultures. *Statistical science*, 16(3),pp.199-231.

Brownsword, R. (2015). In the year 2061: from law to technological management. *Law, Innovation and Technology*, 7(1), pp.1–51.

Buckner, C. (2019). Deep learning: A philosophical introduction. *Philosophy Compass*, 14(10).pp.12625.

Campbell-Verduyn, M., Goguen, M. and Porter, T. (2016). Big Data and algorithmic governance: the case of financial practices. *New Political Economy,* 22(2), pp.219–236.

Casey, A.J. and Niblett, A. (2015). The Death of Rules and Standards. *Ind. LJ*.92, pp.1401-1448.

Castoriadis, C. (1994). The logic of magmas and the question of autonomy. *Philosophy & Social Criticism*, 20(1-2), pp.123–154.

Chen Y, Lin C, and Liu H. (2018) Rule of trust: The Power and Perils of China's Social Credit Megaproject 32 *Colum J Asian L* 1.

Chen, D. and Deakin, S. 2015. On Heaven's Lathe: State, Rule of Law, and Economic Development. *Law And Development Review* 8(1), pp.123-145.

Chen, D., Deakin, S., Siems, M. and Wang, B. (2017). Law, trust and institutional change in China: evidence from qualitative fieldwork. *Journal of Corporate Law Studies*, 17(2), pp.257–290.

Chen YW., Should the principle of necessary facilities apply to the data? — Based on the analysis of "two kinds of errors", *Competition Policy Research* 2021:4. [陈永伟: 数据是否应适用必需设施原则？—基于"两种错误"的分析，《竞争政策研究》2021:4.]

Cheung, A.S.Y. and Chen, Y. (2022). From Datafication to Data State: Making Sense of China's Social Credit System and Its Implications. *Law & Social Inquiry*, 47(4), p.1137.

Citron DK, and Pasquale F. (2014) 'The Scored Society: Due Process for Automated Predictions' 89 *Wash L Rev* 1.

Coase, R., (1960). The Problem of Social Cost. *Journal of Law and Economics*, 3,pp. 1–44.

Cobbe, J., (2020). Legal Singularity and the Reflexivity of Law. In Deakin S. and Markou C. (eds), *Is Law Computable? Critical Perspectives on Law and Artificial Intelligence*. Hart Publishing.

Cobbe, J., Veale, M., & Singh, J. (2023, June). Understanding accountability in algorithmic supply chains. In *Proceedings of the 2023 ACM Conference on Fairness, Accountability, and Transparency* (pp. 1186-1197).

Coglianese, C. and Lehr D., 2019. Transparency and algorithmic governance. *Admin. L. Rev.* 71, pp.1-56.

Commons, J.R., (1932). The problem of correlating law economics and ethics. *Wis. L. Rev*., 8, pp.3-26.

Creemers, R. (2018). China's Social Credit System: An Evolving Practice of Control. *SSRN Electronic Journal* .

Dai, X. (2020). Toward A Reputation State: A Comprehensive View of China's Social Credit System Project. *Social Credit Rating: Reputation und Vertrauen beurteilen*, pp.139-163.

Danaher, et al. (2017). Algorithmic governance: Developing a research agenda through the power of collective intelligence. *Big Data & Society*, 4(2).

Davis, G. and Ostrom, E. (1991). A Public Economy Approach to Education: Choice and Co-Production. *International Political Science Review*, 12(4), 313–335.

Davis, K. E., Kingsbury, B., & Merry, S. E. (2012). Indicators as a technology of global governance. *Law & Society Review*, *46*(1), 71-104.

Deakin, S. (2001). The Contract of Employment: A Study in Legal Evolution. *Historical Studies in Industrial Relations*, (11), pp.1–36.

Deakin, S. (2003). Evolution for Our Time: A Theory of Legal Memetics. *Current Legal Problems* 55(1), pp. 1–42.

Deakin, S. (2011) Legal Evolution: Integrating Economic and Systemic Approaches. *Review of Law & Economics*, Vol. 7 (Issue 3), pp. 659-683.

Deakin, S. (2015). Juridical ontology: the evolution of legal form. 40(1),*Historical Social Research/Historische Sozialforschung* pp.170–184.

Deakin, S. (2023). The legal construction of management: a neo-realist framing and genealogical case study. *Journal of Corporate Law Studies*, pp.1-22.

Deakin, S. and Markou, C. (2021). Evolutionary law and economics: theory and method. *Northern Ireland Legal Quarterly*, 72(4), pp.682–712.

Deakin, S. and Markou, C. (2022). Evolutionary Interpretation: Law and Machine Learning . *Journal of Cross-Disciplinary Research in Computational Law*, 1(2).

Deakin, S. and Meng, G. (2020). The governance of Covid-19: anthropogenic risk, evolutionary learning, and the future of the social state. *Industrial Law Journal*, *49*(4), pp.539-594.

Deakin, S. and Meng, G.,(2022). Resolving Douglass C. North's ‘puzzle’ concerning China's household responsibility system. *Journal of Institutional Economics*, *18*(4), pp.521-535.

Deakin, S., Gindis, D., Hodgson, G.M., Huang, K. and Pistor, K. (2017). Legal institutionalism: Capitalism and the constitutive role of law. *Journal of Comparative Economics,* 45(1), pp.188–200.

Diab, R.S. (2017). Becoming-Infrastructure: Datafication, Deactivation and the Social Credit System. *Journal of Critical Library and Information Studies*, 1(1).

Diver, L. (2021). Computational legalism and the affordance of delay in law. *Journal of Cross-disciplinary Research in Computational Law,* 1(1).

Edin, M. (2017). Remaking the communist party-state. In *Critical Readings on the Communist Party of China (4 Vols. Set)* .

Engelmann S, Chen M, Dang L and Grossklags J. (2021) Blacklists and Redlists in the Chinese Social Credit System: Diversity, Flexibility, and Comprehensiveness. *Proceedings of the 2021 AAAI/ACM Conference on AI, Ethics, and Society* 78.

Engstrom, D. F., & Ho, D. E. (2020a). Algorithmic accountability in the administrative state. *Yale J. on Reg*., 37, 800.

Engstrom, D.F. and Ho, D.E., (2020b). Artificially Intelligent Government: A Review and Agenda. In: Roland Vogl, (ed.), *Big Data Law (2020).*

Erie, M. S., & Streinz, T. (2021) The Beijing Effect: China's Digital Silk Road as *Transnational Data Governance*, 54 N.Y.U. J. Int'l L. & Pol. 1 .

Evans, B.J. (2011). Much Ado about Data Ownership. *Harv. JL & Tech.* 25(1), pp.69.

Feng X. (2021). Theoretical interpretation and construction of data property management and its legal regulation. *On Politics and Law* (04) [冯晓青, 2021.数据财产化及其法律规制的理论阐释与构建. 《政法论丛》(04)] .

Feng, L. (2017). Research on regulatory issues of China's credit rating industry. *Southwest Finance.*(8).pp.5.(In Chinese).

Fukuyama, F. (2013). What Is Governance? *Governance: An International Journal of Policy, Administration and Institutions*, 26(3), pp.347–368.

Gama, J., Žliobaitė, I., Bifet, A., Pechenizkiy, M. and Bouchachia, A. (2014). A survey on concept drift adaptation. *ACM computing surveys (CSUR)*, *46*(4), pp.1-37.

Greif, A. (1993). Contract enforceability and economic institutions in early trade: *The Maghribi Traders' Coalition 3 American Economic Review*. 83(3), pp.525–548.

Grimmelmann, J. (2004). Regulation by software. *Yale LJ* 114:1719.

Hansen K.H. and Mühlen-schulte, A. (2012) The Power of Numbers in Global Governance, *Journal of International Relations and Development* 15(4): 455–465.

Hartzog,W. Conti,G. Nelson, J. and Lisa A. Shay. (2015) 'I*nefficiently Automated Law Enforcement*' Mich St L Rev 1763.

Hildebrandt, M. (2016). Law as Information in the Era of Data-Driven Agency. *The Modern Law Review,* 79(1), pp.1–30.

Hildebrandt, M. (2018). Law as computation in the era of artificial legal intelligence: Speaking law to the power of statistics. *University of Toronto Law Journal*, 68(supplement 1), pp.12–35.

Hildebrandt, M. (2020b). 'Code Driven Law: Scaling the Past and Freezing the Future' in Simon Deakin and Christopher Markou (eds), *Is Law Computable? Critical Perspectives on Law and Artificial Intelligence* (Hart 2020)

Hildebrandt M. (2021). The adaptive nature of text-driven law. *Journal of Cross-disciplinary Research in Computational Law*. (1).pp.1.

Hildebrandt, M. (2023). Grounding computational 'law' in legal education and professional legal training. In *Research Handbook on Law and Technology* (pp. 99-127). Edward Elgar Publishing.

Ho, D.E. (2012). Fudging the Nudge: Information Disclosure and Restaurant Grading. *Yale LJ*, 122, pp.574-689.

Horsley, J. (2018). China's Orwellian Social Credit Score isn't Real. *Foreign Policy* 16.

Hurley, M. and Adebayo, J. (2016). Credit scoring in the era of big data. *Yale JL & Tech.*, 18, pp: 148-216.

Huttner, L. and Merigoux, D. (2022). Catala: Moving towards the future of legal expert systems. *Artificial Intelligence and Law* 1.

Jackson, P.M. (2011). Governance by numbers: what have we learned over the past 30 years?. *Public Money & Management* 31:1. 13-26.

Joh E. (2019). The Consequences of Automating and Deskilling the Police, 67 *UCLA L. Rev. Discourse* 133.

Katzenbach, C. and Ulbricht L. (2019). Algorithmic governance. *Internet Policy Review* 8.4: 1-18.

Kehl, D., Guo P., and Kessler, S. (2017). Algorithms in the Criminal Justice System: Assessing the Use of Risk Assessments in Sentencing. *Responsive Communities Initiative, Berkman Klein Center for Internet & Society*, Harvard Law School.

Kong, X. (2022). Commercial data rights: New industrial property rights in the digital era, *Research on Comparative Law*, (01) pp 2-3. [孔祥俊, 2022.商业数据权:数字时代的新型工业产权,《比较法研究》,（01）pp 2-3;]

Krause, T., Chen, M., Wassermann, L., Fischer, D. and Grossklags, J. (2022). China's corporate credit reporting system: A comparison with the United States and Germany. *Regulation & Governance* 17(3), pp. 755–771.

Lam, X.N., Vu, T., Le, T.D. and Duong, A.D. (2008). Addressing cold-start problem in recommendation systems. *Proceedings of the 2nd international conference on ubiquitous information management and communication*. pp.208-211.

Lewkowicz G. et al. (2020). Unpacking smart law: how mathematics and algorithms are reshaping the legal code in the financial sector. *Lex Electronica* 25: 171.

Lewicki, K., Lee, M. S. A., Cobbe, J., & Singh, J. (2023). Out of Context: Investigating the Bias and Fairness Concerns of "Artificial Intelligence as a Service". In *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems*. pp. 1-17.

Li, B. and van Zanden, J.L. (2012). Before the Great Divergence? Comparing the Yangzi Delta and the Netherlands at the Beginning of the Nineteenth Century. *The Journal of Economic History*, 72(4), pp.956–989.

Li SJ. (2021) On the standard of data composition of necessary facilities —— Comment on the revision of Article 14 of the Anti-monopoly Guide on the Field of Platform Economy, *Journal of Henan University of Economics and Law*. 36(05) [李世佳: 论数据构成必需设施的标准——兼评《关于平台经济领域的反垄断指南》第十四条之修改, 河南财经政法大学学报. 2021,36(05)]

Liang, F., Das, V., Kostyuk, N. and Hussain, M.M. (2018). Constructing a Data-Driven Society: China's Social Credit System as a State Surveillance Infrastructure. *Policy & Internet*, 10(4), pp.415–453.

Liebman, B. L., Roberts, M. E., Stern, R. E., & Wang, A. Z. (2020). Mass digitization of Chinese court decisions: How to use text as data in the field of Chinese law. *Journal of Law and Courts*, *8*(2), 177-201.

Lin, J. (2012). The Theoretical Inheritance and Innovation of the Social Credit System, *Credit Reference*, 2012(1), pp.1-12.

Lin, J. (2019), 'Why do we claim that the construction of the Social Credit Systems started in 1999?', (in Chinese).

Litman, J. (2000). Information Privacy/Information Property. *Stanford Law Review*, 52(5), pp.1283-1313.

Liu P. and Shi L., (2006). Legal Discussions on Administrative 'Blacklists', *Fazhi Luncong*. 21(2), pp.68-73.

Llewellyn, K. (1942). The Crafts of Law Re-valued. *Rocky Mountain Law Review*, 15, pp.1-7.

Loubere, N. and Brehm, S. (2018). The Global Age of Algorithm: Social Credit and the Financialisation of Governance in China. *Made in China Journal*, 3(1).

Luhmann, Niklas (1976). Generalized media and the problem of contingency. *In: Explorations in general theory in social science, essays in honor of Talcott Parsons. Bd. 2, Jan J. Loubser , New York, Free Press of Glencoe,* 1976, S. 507-532.

Mac Síthigh, D. and Siems, M. (2019). The Chinese Social Credit System: A Model for Other Countries? *The Modern Law Review*, 82(6), pp.1034–1071.

Macaulay, S. (1936). Non-contractual relations in business: A preliminary study. *American Sociological Review* 28(1), pp. 55-67.

Markou, C. and Deakin, S.F. 2019. Ex Machina Lex: Exploring the Limits of Legal Computability. SSRN Electronic Journal.

McQuillan, D. (2017). Data Science as Machinic Neoplatonism. *Philosophy & Technology*, 31(2), pp.253–272.

Medvedeva, M., Vols, M., & Wieling, M. (2020). Using machine learning to predict decisions of the European Court of Human Rights. *Artificial Intelligence and Law*, *28*(2), 237-266.

Medvedeva, M., Wieling, M., & Vols, M. (2023). Rethinking the field of automatic prediction of court decisions. *Artificial Intelligence and Law*, *31*(1), 195-212.

Moore, D., (1992). Teaching Statistics as a Respectable Subject. *F.* Gordon; S. Gordon (eds.). Statistics for the Twenty-First Century. Washington, DC: *The Mathematical Association of America*. pp.14–25.

Offe, C., (2009). Governance: An 'Empty Signifier'?. *Constellations*, *16*(4), pp. 550-562.

Ostrom, E. (1996). Crossing the great divide: Coproduction, synergy, and development. *World Development*, 24(6), pp.1073–1087.

Ostrom, E. (2010). Beyond Markets and States: Polycentric Governance of Complex Economic Systems. *American Economic Review*, 100(3), pp.641–672.

Ostrom, E. and Whitaker, G. (1973). Does Local Community Control of Police make a Difference? Some Preliminary Findings. *American Journal of Political Science*, 17(1), pp.48-76.

Ostrom, E. and Whitaker, G. (1974). Community control and governmental responsiveness: The case of police in black neighborhoods. *Improving the quality of urban management*, 10(4), pp.303-334.

Parks, R.B., Baker, P.C., Kiser, L., Oakerson, R., Ostrom, E., Ostrom, V., Percy, S.L., Vandivort, M.B., Whitaker, G.P. and Wilson, R. (1981). Consumers as coproducers of public services: Some economic and institutional considerations. *Policy Studies Journal*, 9(7), pp.1001–1011.

Pasquale, F.A. (2019). A Rule of Persons, Not Machines: The Limits of Legal Automation. *Social Science Research Network,* 87(1), p.332-336.

Pearson, M., Rithmire, M., & Tsai, K. S. (2021). Party-state capitalism in China. *Current History*, *120*(827), 207-213.

Peng C. (2021). The Legal Dilemma and Solution for Joint Sanctions Mechanisms in China. *Studies In Law and Business* 5 (in Chinese).

Peng H., 2022. Logical structure and empowerment boundary of data ownership, *Research on Comparative Law*, (01). [彭辉,2022. 数据权属的逻辑结构与赋权边界，《比较法研究》,（01）].

Peters, O. (2011). Optimal leverage from non-ergodicity. *Quantitative Finance*, 11(11), pp.1593–1602.

Peters, O. (2019). The ergodicity problem in economics. *Nature Physics*, 15(12), pp.1216–1221.

Pistor, K. (2013). A Legal Theory of Finance. *Journal of Comparative Economics,* 41(2), pp.315-330.

Rhodes, R.A.W. (1996). The new governance: governing without government. *Political studies*, 44(4), pp.652-667.

Rhodes, R.A.W. (2007). Understanding Governance: Ten Years On. *Organization Studies*, 28(8), pp.1243–1264.

Rodrik, D (2008). Thinking about governance. *Governance, growth, and development decision-making*, Bloomsbury Academic, London and New York, pp.17-24.

Saha, S. (2020). Automated Traffic Law Enforcement System: A Feasibility Study for the Congested Cities of Developing Countries, *International Journal of Innovative Technology and Interdisciplinary Sciences 346*.3(1).

Schmidhuber, J. (2015). Deep learning in neural networks: An overview. *Neural Networks*, 61(61), pp.85–117.

Schrepel T and Teodora G, The Adoption of Computational Antitrust by Agencies: 2021 Report. (2022) 2 *Stanford Computational Antitrust* 78.

Shen, Kui [沈岿], The Approach Consistent with the Rule of Law to Constructing the Social Credit System (2019) 5 *China Legal Science* 25.

Siems, M. and Nelken D. (2017) Global social indicators and the concept of legitimacy. *International journal of law in context* 13.4 436-449.

Stern, R. E., Liebman, B. L., Roberts, M. E., & Wang, A. Z. (2020). Automating fairness? Artificial intelligence in the Chinese courts. *Colum. J. Transnat'l L.*, *59*, 515.

Sun, G., 2017. From FinTech to RegTech, *Internet Finance,* 2017(5). pp.94-96.

Sun, G., 2018. Developing RegTech to Construct a New Financial Ecosystem, *Tsinghua Financial Review*, 2018(3), pp.16-19.

Surden H, (2019). Artificial Intelligence and Law: An Overview. *35 Georgia State University Law Review* 19.

Surden, H. (2021). Machine learning and law: An overview. *Research Handbook on Big Data Law*, pp.171-184.

Surden, H. (2024). ChatGPT, Artificial Intelligence (AI) Large Language Models, and Law. *Fordham Law Review*, (92).

Tian and Yang. (2018). 'The Development and Transformation of the European Credit Reporting Industry in the Digital Economy Era', *China Credit Reference*, 5. (in Chinese).

Tsymbal A. (2004). The problem of concept drift: definitions and related work. *Computer Science Department, Trinity College Dublin*, *106*(2), p.58

Twining, W. et al. (2015). A symposium on global law, legal pluralism and legal indicators. *The Journal of Legal Pluralism and Unofficial Law*, 47(1), 1-8.

Van Loo, R. (2020a). The revival of respondeat superior and evolution of gatekeeper liability. *Geo. lJ*, *109*, 141.

Van Loo, R. (2020b). The New Gatekeepers. *Virginia Law Review*, *106*(2), 467-522.

Vanderstraeten, R. (2002). Parsons, Luhmann and the theorem of double contingency. *Journal of Classical Sociology*, *2*(1), 77-92.

Varian, H.R., (2010). Computer mediated transactions. *American Economic Review* 100.(2), pp.1-10.

Varian, H.R., (2014). Beyond big data. *Business Economics* 49.(1), pp.27-31.

Victor, Jacob M. (2013). The EU general data protection regulation: Toward a property regime for protecting data privacy. *Yale LJ*. 123, pp.513.

Waddington, M. (2020). Rules as code. *Law in Context*, *37*(1), pp.179-186.

Wang F., (2012), Integrity attracts the golden phoenix—Practice and reflections on the construction of honest farmers in Guizhou Province, *Qiushi*, 4 (in Chinese).

Wang X. and Huang Z. (2021). On the Rule-of-Law Constraints on the Restrictions on Trust-Breaking" 1 *China Law Review* 96. (in Chinese).

Wang X. (2021). Personal information's state protection obligations and its implications, *China Legal Science*, (1).(in Chinese).

Wang and Wu. (2021). Research on Data as an Essential Facility under the Antitrust Law, *Research on Rule of Law*, (2) pp.102-114. [王健 吴宗泽：论数据作为反垄断法中的必要设施]

Wei, J., He, J., Chen, K., Zhou, Y. and Tang, Z. (2017). Collaborative filtering and deep learning based recommendation system for cold start items. *Expert Systems with Applications*, 69, pp.29–39.

Widmer, G. and Kubat, M. (1996). Learning in the presence of concept drift and hidden contexts. *Machine Learning*, 23(1), pp.69–101.

Wiebe, A. (2017). Protection of industrial data – a new property right for the digital economy? *Journal of Intellectual Property Law & Practice* 12(1), pp. 62–71.

Williamson, B. (2014). Knowing public services: Cross-sector intermediaries and algorithmic governance in public sector reform. *Public Policy and Administration*, 29(4), pp.292–312.

Williamson, O.E. (1979). Transaction-Cost Economics: the Governance of Contractual Relations. *The Journal of Law and Economics*, 22(2), pp.233–261.

Williamson, O.E. (1998). Transaction Cost Economics: How it Works; Where it is Headed. *De economist*, 146(1), pp.23-58.

Williamson, O.E. (2002). The Lens of Contract: Private Ordering. *American Economic Review*, 92(2), pp.438–443.

Williamson, O.E. (2005). The economics of governance. *American Economic Review*, *95*(2), pp. 1-18.

Xiong, B., Ge, J. and Chen, L. (2023). Unpacking data: China's 'bundle of rights' approach to the commercialization of data. *International Data Privacy Law*, *13*(2), pp.93-106.

Xu, T. and Gong, W. (2020). Community-based individual property rights: developing the 'bundle of rights' perspective in the Chinese context. *Asia Pacific Law Review*, *28*(1), pp.138-158.

Xu, T. (2017). Towards an Evolutionary Theory of Property: A Longitudinal Analysis of Property Regime Transformation in China. *J. Comp. L.*, *12*, p.496.

Xu, T. (2019). A law-and-community approach to compensation for takings of property under the European Convention on Human Rights. *Legal Studies*, *39*(3), pp.398-414.

Xu, T. (2021). Uncertainty, Ignorance and Decision-Making: Looking Through the Lens of Modelling the Covid-19 Pandemic. *Series 2 Vol. 3 Amicus Curiae*, p.10.

Xu, T. (2023). 'Mathematical modelling, rule-making, and the COVID-19 pandemic', in Andrea Saltelli, and Monica Di Fiore (eds), *The Politics of Modelling: Numbers Between Science and Policy* (Oxford, 2023)

Xue H., (2018), 'The model boundaries and missions of Baihang Credit', *Financial Economics,* (4):3. [薛洪言, 2018, 百行征信的模式边界与使命,《金融经济》]

Yeung, K. (2017). Blockchain, Transactional Security and the Promise of Automated Law Enforcement: The Withering of Freedom Under Law? *SSRN Electronic Journal.*

Yeung, K. (2018). Algorithmic regulation: A critical interrogation. *Regulation & Governance*, 12(4), pp.505–523.

Young, T., Hazarika, D., Poria, S. and Cambria, E. (2018). Recent Trends in Deep Learning Based Natural Language Processing [Review Article]. *IEEE Computational Intelligence Magazine*, 13(3), pp.55–75.

Yu, K. (1999). Introduction to Governance and Good Governance. *Marxism and Reality*, 5, pp.37-41. [俞可平：治理和善治引论，《马克思主义与现实》，1999 (05):37-41.]

Yu, K. (2002). Introduction to Global Governance. *Marxism and Reality*, 1, pp.20-32. [俞可平: 全球治理引论，《马克思主义与现实》，2002 (01):20-32]

Yu, K. (2018). Governance Reform in China (1978-2018). *Wuhan University Journal,* 71(3), pp.48-59. [俞可平:中国的治理改革(1978-2018), 《武汉大学学报(哲学社会科学版)》2018, 71(03) : 48-59].

Zeng C., Zhu X. (2021). The rationality, basic principles and regulatory ideas of the construction of mandatory data licensing system under the background of digital economy. *E-government*,. [曾彩霞，朱雪忠：数字经济背景下构建数据强制许可制度的合理性，基本原则和监管思路. 《电子政务》2021.]

Zhang, X. (2021). Research on Internet ecosystem's Gatekeepers special protection obligations to personal information. *Comparative Law Journal* (3). [张新宝：互联网生态"守门人"个人信息保护特别义务设置研究，《比较法研究》，2021 年第 3 期].

Zhong and Peng, (2014), 'The inspiration of Ali's credit reporting model on the construction of my country's financial credit reporting system', *Credit Reference,* 2.

Zimmermann, R. (2015) 'Roman Law in the Modern World', in D. Johnston (ed.) *The Cambridge Companion to Roman Law*. Cambridge University Press (Cambridge Companions to the Ancient World), pp. 452–480.

Žliobaitė, I., Pechenizkiy, M. and Gama, J. (2016). An overview of concept drift applications. In: *Big data analysis: new algorithms for a new society*, pp.91-114.

Zuboff, S. (2015). Big Other: Surveillance Capitalism and the Prospects of an Information Civilization. *Journal of Information Technology*, 30(1), pp.75–89.

Zuo, Weimin. (2018). 'How to achieve similar case with similar judgements by AI'. *Chinese Legal Commentaries,* 2018 (2).[左卫民，如何通过人工智能实现类案类判，《中国法律评论》，2018 年第 2 期。]

Major laws and regulations (including drafts)

15 U.S.C. § 7211, Holding Foreign Companies Accountable Act 2020 (HFCAA)

15 U.S.C. § 1681 Fair Credit Reporting Act

Bank of England (2023). PS17/23 – Implementation of the Basel 3.1 standards near-final part 1.

Central CPC Office and State Council Office, [2016] No.64. *Opinions on accelerating the construction of credit supervision, warning and punishment mechanisms for dishonest judgement defaulters*. [中共中央办公厅 国务院办公厅印发《关于加快推进失信被执行人信用监督、警示和惩戒机制建设的意见》].

China Securities Regulatory Commission, [2012] No.80, *Interim Measures for the Supervision and Administration of Integrity in the Securities and Futures Markets*, effected on 2012.9.1, Amended on 2014.10.15.

China Securities Regulatory Commission, [2018] No.139, *Measures for the Supervision and Administration of Integrity in Securities and Futures Markets*.

Civil Code of the People's Republic of China, 2021, 2023

Civil Servants Law of the People's Republic of China*,* 2018, 2023.

CPC central committee and State Council: Instruction on Improving the Market Mechanism for Allocating Essential Factors, 2020.3.30., available at http://www.gov.cn/zhengce/2020-04/09/content_5500622.htm (accessed 16 December 2022) [中共中央国务院关于构建更加完善的要素市场化配置体制机制的意见].

Cybersecurity Law 2017. Chapter 3 section 2 on CII operators. Including article 37 on storing important data in China. 中华人民共和国网络安全法_中国人大网 (npc.gov.cn)

Cybersecurity Law of the People's Republic of China, 2017

Data security law of the People's Republic of China, 2021

Guangzhou city management and enforcement department, [2020] No.9.

Guizhou People's Congress, [2022] No.20. *Guizhou Province Social Credit Regulations.*

Hebei Department of Human Resources and Social Securities [2018] No.9. [冀人社规〔2018〕9 号] 2018.4.26.

Information security technology – Guidance for personal information security impact assessment (GB/T 39335-2020) [信息安全技术 - 个人信息安全影响评估指南(GB/T 39335-2020)]

International Organisation of Securities Commissions (IOSC), (2017) *Objectives and*

Principles of Securities Regulation.

Legislation Law of the People's Republic of China. 2000, 2023 Amended.

MIIT, Mobile Internet Application Personal Information Protection Provisional Regulations (Draft for Comments), 2021.4.26

MIIT on issuing the Big data Industry Development Plan. MIIT [2016] No.412., available at http://www.cac.gov.cn/2017-01/17/c_1120330820.htm (accessed 16 December 2022) [工业和信息化部关于印发大数据产业发展规划 (2016-2020 年) 的通知]

National Energy Administration (NEA) [2016] No.350,

National Development and Reform Committee (NDRC), [2016] No.141. *Notice of the NDRC, the Supreme People's Court, the People's Bank of China, and Other Departments on Issuing the Memorandum of Understanding on Taking Joint Disciplinary Actions against Dishonest Persons Subject to Enforcement,*

NDRC [2017] No.1455. *Memorandum of understanding on joint sanctions of serious illegal and dishonest entities in the oil and gas industry*. [发改运行〔2017〕1455 号]

NDRC [2017] No.1798. [发改财金规〔2017〕1798 号]

NDRC [2017] No.946. *Memorandum of Understanding on Joint Sanctions of Serious Illegal and Untrustworthy Market Entities and Related Personnel in the Electric Power Industry.* [发改运行〔2017〕946 号].

NDRC [2018] No.233. [发改运行规〔2018〕233 号] 2018.2.3.

NDRC, [2016] No.141. *Notice of the NDRC, the Supreme People's Court, the People's Bank of China, and Other Departments on Issuing the Memorandum of Understanding on Taking Joint Disciplinary Actions against Dishonest Persons Subject to Enforcement.*

NDRC, *Social Credit Systems Construction Law (Draft for public consultation, 2022),* 2022.11.14.

NDRC, [2023] No. 860. *Notice on standardising the application of credit systems in public procurements.*

Personal Information Protection Law of the People's Republic of China, 2021

People's Bank of China, *Interim Measures for the Administration of the Basic Data of Individual Credit Information*, [2005] No.3.

People's Bank of China, [2006] No.117. Notice of the General Office of the People's Bank of China on Issuing the 'Switching Plan for the Corporate Credit Information Foundational Database' [中国人民银行办公厅关于印发《企业信用信息基础数据库切换方案》的通知]（银办发[2006]171 号）

People's Bank of China, *Specification for User Management of Financial Credit Information Basic Database* [2014], JR/T 0115-2014 [金融信用信息基础数据库用户管理规范]

People's Bank of China, [2015] *Instruction on Individual Credit Information Reporting,*.

People's Bank of China, National Development and Reform Commission, Ministry of Finance, and China Securities Regulatory Commission, *Interim Measures for the Administration of credit rating industry.* [2019] No.5.

Regulations of Shanghai Municipality on Social Credit (Shanghai Municipal People's Congress, No. 54, 2017).

Regulation (EU) 2022/1925 (DMA).

Regulation (EU) 2022/2065 (DSA).

Shanghai Data Regulation (2021) [上海市数据条例]

State Council, Several opinions of the General Office of the State Council on the construction of the social credit system, [2007] No.17. [国务院办公厅关于社会信用体系建 设 的 若干 意 见(2007)].

State Council, Regulation on Credit-reporting Industry [2013] No.631, [国务院征信业管理条例国务院令 631 号]

State Council's instructions on the action plans for promoting big-data development, 2015. SC [2015] 50. [国务院关于印发促进大数据发展行动纲要的通知国发〔2015〕50]

State Council on the provisional measures to government information resources sharing and management.2016. SC [2016] 51.], available at http://www.gov.cn/zhengce/content/2016-09/19/content_5109486.htm (accessed 16 December 2022) [国务院关于印发政务信息资源共享管理暂行办法的通知国发〔2016〕51 号]

State Council Antitrust Committee [2021] No.1, Antitrust Guideline on Platform Economy (2021.2.7)], [国务院反垄断委员会关于平台经济领域的反垄断指南]

Supreme People's Court of China (SPC), [2010] No.8, *Judicial Interpretation on High-end Consumption Restrictions for Judgement Defaulters,*.

SPC, Judicial Interpretation [2013] No.17, *Judicial Interpretations on the Publication of Information of Dishonest Judgement Defaulters.*

SPC, 'Necessary Notes on the Execution Procedure' [执行程序须知] (SPC 26, Jan 2016)

SPC, Judicial Interpretation, [2017] No. 7. *Decision on Amending the 'Supreme People's Court's Several Provisions on the Disclosure of the Dishonest Judgment Debtors List Information'*

SPC, [2017] No. 12. Opinions of the Supreme People's Court on Accelerating the Construction of Smart Courts [最高人民法院关于加快建设智慧法院的意见]

Tianjin Social Credit Regulations, Tianjin 17th People's Congress Plenary Committee Report, No 63. Effective on 2021.1.1 [天津市社会信用条例（天津市第十七届人民代表大会常务委员会公告，第 63 号）]

Zhejiang Development and Reform Commission. 2019. Zhejiang Guidelines for Public Credit Evaluation of Five Types of Entities (2019) [浙江省五类主体公共信用评价指引] (2019). (in Chinese).

Zhejiang Development and Reform Commission. 2020. Zhejiang Guidelines for Public Credit Evaluation of Five Types of Entities (2020) 浙江省五类主体公共信用评价指引 (2020). (in Chinese).

Cases and administrative actions

(2019）浙 8601 民初 1987 号[2019, Zhejiang, 8601 Civil Court, First instance case no 1987]

(2020）浙 01 民终 5889 号[2020, Zhejiang, 01 Civil appeal case no. 5889]

(2020）浙 0106 刑初 437 号[2020, Zhejiang, 0106, Criminal Court, first instance case no. 437]

2022 Shanghai Fengyi case (2021, 2022) Shanghai 0120 Civil First Instance Case No.11522 [沪 0120 民初 11522 号]

2022 W (anonymised) (2020, 2022) Case No. 0677 of Shanghai Arbitration [沪仲案字第0677号].

Harber v Commissioners for His Majesty's Revenue and Customs [2023] UKFTT 1007 (TC)

Uber BV v Aslam [2021] UKSC 5

United States v. Facebook, Inc., No. 19-2184 (TJK), 2020 WL 1975785 (D.D.C. April 23, 2020)

SEC Filing, 2021, Didi Global Inc Prospectus under 424(B)(4) (sec.report)

Shanghai 0120 Enforcement Case No. 3708 [沪 0120 执 3708 号] (Shanghai High People's Court, 2022)

Shanghai 02 Enforcement Reinstatement Case No.72 [沪 02 执恢 72 号] (Shanghai High People's Court, 2022)

State Administration for Market Regulation (SAMR), 2021 April 10th, Administrative penalties on Alibaba Group Holding Limited for its "choose one of two" monopoly in the online retail platform service market in China. [市场监管总局依法对阿里巴巴集团控股有限公司在中国境内网络零售平台服务市场实施"二选一"垄断行为作出行政处罚]

State Administration for Market Regulation (SAMR), 2021 October 8th. Administrative penalties on Meituan for its "choose one of two" monopoly in the online food delivery platform service market in China [市场监管总局依法对美团在中国境内网络餐饮外卖平台服务市场实施“二选一”垄断行为作出行政处罚]

Others

Artificial Intelligence (AI) Guidance for Judicial Office Holders. (2023). Available at: https://www.judiciary.uk/wp-content/uploads/2023/12/AI-Judicial-Guidance.pdf

Bai H., 'Guizhou Qingzhen explores building a social credit system', 2015.4.21, online in Chinese, at http://zqb.cyol.com/html/2015-04/21/nw.D110000zgqnb_20150421_5-01.htm.

Bank of England, 2023. *PS17/23 – Implementation of the Basel 3.1 standards near-final part 1. 2024*. Available at:https://www.bankofengland.co.uk/prudential-regulation/publication/2023/december/implementation-of-the-basel-3-1-standards-near-final-policy-statement-part-1

Bian, W., 'Interview with Zhao Yi Han Chairman of Pudao Credit: Complementing each other and developing in conjunction with the Central Bank Credit Information and Baihang Credit, finance.sina.com, 2021.6.7. Online in Chinese, at https://finance.sina.cn/2021-06-07/detail-ikqciyzi8197100.d.html.

Bundesdatenschutzgesetz (BDSG) (geändert durch das Gesetz vom 30. Oktober 2017) [联邦数据保护法] [in Chinese]; http://www.wipo.int/wipolex/zh/text.jsp?file_id=462255 [German Version];

Business Credit Scores & Reports; "15 Business Credit Reporting Agencies You Should Know About", https://businesscreditblogger.com/2018/05/03/business-credit-reporting-agencies/

Cambridge Dictionary (2022). *computational.* [online] Available at: https://dictionary.cambridge.org/dictionary/english/computational

Chen, G., 'None of the 8 personal credit reporting agencies has qualified so far. The development of credit reporting agencies must strengthen the protection of personal information', www.ce.cn, 2017.4.27 [online] Available at: http://www.ce.cn/xwzx/gnsz/gdxw/201704/27/t20170427_22415553.shtml.

Chengxin Tong official website home page. [online] Available at: https://cxt.1688.com/

Corporate Sesame Credit scores website, http://www.xin.xin/#/detail/4-1-0

Credit China, 'Announcement and explanation on newly listed dishonesty judgement defaulters in 2018 October' 2018.11.2. . Online in Chinese, at http://www.gov.cn/fuwu/2018-11/02/content_5336733.htm2018.6.11.

Credit China, 'Joint punishment enables millions of workers to recover over 10 billion in unpaid wages', 2018.11.2 online in Chinese, at http://www.creditchina.gov.cn/xinyongdongtai/buwei/201811/t20181101_129920.html [联合惩戒使百万劳动者追偿拖欠工资超百亿, 信用中国].

Daum J. 2019. 'Untrustworthy: Social Credit isn't What You Think it is' (Verfassungsblog, 2019)

European Central Bank (2023). https://www.ecb.europa.eu/press/financial-stability-publications/macroprudential-bulletin/focus/2023/html/ecb.mpbu202312_focus01.en.html

Experian official website, https://www.experianplc.com/

Federal Reserve (2023) *Agencies request comment on proposed rules to strengthen capital requirements for large banks*.

FICO official website, https://www.fico.com/

Financial Crisis Inquiry Commission, The Financial Crisis Inquiry Report, 2011. https://www.govinfo.gov/content/pkg/GPO-FCIC/pdf/GPO-FCIC.pdf

Foreign Ministry: global data security initiative (2020), available at http://www.xinhuanet.com/world/2020-09/08/c_1126466972.htm (accessed 16 December 2022)[外交部全球数据安全倡议]

Guangzhou government, 2024, Public resource transactions and government procurement credit evaluation system version 3.0, https://czj.gz.gov.cn/attachment/7/7505/7505853/9206033.doc.

Hainan business environment department, '"Golden Coconut Points" expands the circle of friends and realises mutual recognition of personal credit across provinces', 2023, online in Chinese, available at https://www.ndrc.gov.cn/fggz/fgfg/dfxx/202304/t20230426_1354750.html

Han Bing., 'Detailed assessment rules for ensuring the payment of wages to migrant workers are announced. Strengthen credit punishment for untrustworthy companies with unpaid wages', www.gov.cn, 2018.9.11. Online in Chinese, at http://www.gov.cn/xinwen/2018-09/11/content_5320898.htm

Hangzhou population statistics, 2023. https://www.hangzhou.gov.cn/art/2023/3/2/art_1229063404_4144634.html

He, YH. (2018). 'The fourth on-site inspection of Didi by the Shanghai inspection team has been accessed and will be reviewed', Thepaper.cn. [上海检查组第四次上门检查滴滴数据已接入将复查]

Hua M., 'Open Sesame: Revealing the secrets of big data credit reporting system', 2015.1.30. Online in Chinese.

International Bureau of Weights and Measures, 2008. International Vocabulary of Metrology – Basic and General Concepts and Associated Terms (VIM) (3rd ed.) ECD. 2008. OECD Glossary of Statistical Terms.

Jiang, Zeming. *Report of the 16th National Congress of the Communist Party of China (CPC),* 2002. Available at: http://en.people.cn/200211/18/eng20021118_106983.shtml;

Li, J., 'Four Major Changes in the 2017 Spring Festival Travel Rush: 'Face Scanning' Entry Available at Some Train Stations' [2017 年春运四大变化：部分火车站可"刷脸"进

站] (People's Daily Online, 15 Dec 2016)

<http://politics.people.com.cn/n1/2016/1215/c1001-28951537.html> accessed 7 June 2023

Li and Zhang, ‘The arrears of wages are nearly 7 million. These people and companies in Shijiazhuang are on the "blacklist"!’, hebei.com.cn, 2018.11.13. Online at http://report.hebei.com.cn/system/2018/11/13/019250714.shtml. http://www.erdsrsrc.com/jgzn/rszyzc_jzb_/11372.html

Li, W., ‘44 departments signed a memorandum to jointly punish dishonest judgement defaulters’, China News, 2016.1.21, online in Chinese, at http://www.chinanews.com/gn/2016/01-21/7726810.shtml. [44 部门签署备忘录 联合惩戒失信被执行人].

Liang, Z., ‘2.369 million migrant workers suffered from wage arrears last year. Ministry of Human Resources and Social Security: Focus on solving wage arrears in the engineering field’, Jiemian News, 2017.3.14. Online in Chinese, at https://www.jiemian.com/article/1172388.html [去年 236.9 万农民工遭遇拖欠工资 人社部：重点解决工程领域欠薪]

Liang, Z., Jiemian News, 2017.3.14. Online at https://www.jiemian.com/article/1172388.html

Liu, JC. 2020.10.13. Explanation on the "Personal Information Protection Law of the People's Republic of China (Draft)"[关于《中华人民共和国个人信息保护法（草案）》的说明_中国人大网 (npc.gov.cn)]

Liu, K., ‘National Credit Information Sharing: Let Debtors Have No Place to Hide’, Guangming Daily, 2018.8.10，online in Chinese, at http://www.gov.cn/xinwen/2018-08/11/content_5313109.htm?_zbs_baidu_bk.

Matias, J.N., Mou, M.E., Penney, J., Klein, M. and Wright, L. 2020. *Do Automated Legal Threats Reduce Freedom of Expression Online? Preliminary Results from a Natural Experiment.* Available at:https://osf.io/nc7e2/ accessed 7 June 2023.

Matias, J.N., Penney, J., Klein, M. and Wright, L. 2020. *Comments responding to US Copyright Office Notice of Inquiry Docket No. 2021-10. OSF.2022.* Available at:https://osf.io/498ja/ accessed 7 June 2023.

Merriam-Webster (2019). *Definition of ALGORITHM*. [online] Merriam-webster.com. Available at:https://www.merriam-webster.com/dictionary/algorithm

MOHRSS, 'Qinghai issued the "Detailed Assessment Rules for Guaranteeing the Wage Payment of Migrant Workers in 2018"', 2018.11.09. http://www.mohrss.gov.cn/SYrlzyhshbzb/dongtaixinwen/dfdt/gzdt/201811/t20181109_304494.html

MOHRSS, 'Three-year action plan to combat arrears and guarantee payments, striving to achieve basically no arrears in wages for migrant workers by 2020', 2017.7.21, online in Chinese, at http://www.mohrss.gov.cn/ldjcj/LDJCJgongzuodongtai/201707/t20170725_274594.html

MOHRSS, 'We are confident in managing wage arrears', 2018.10.19, online in Chinese, at http://www.mohrss.gov.cn/SYrlzyhshbzb/dongtaixinwen/buneiyaowen/201810/t20181019_303164.html

National Administration of Agriculture Statistics 2017, quoted from Sina News, 2018.1.2. available at http://www.sohu.com/a/214090338_603780.

National Bureau of Statistics, '2017 Migrant Workers Monitoring Survey Report', 2018.4.28, online in Chinese at https://app.www.gov.cn/govdata/gov/201804/28/423387/article.html [2017年农民工监测调查报告];

National Bureau of Statistics, 'The 7th National Census Report No. 3', 2021.5.10 https://www.stats.gov.cn/zt_18555/zdtjgz/zgrkpc/dqcrkpc/ggl/202302/t20230215_1903999.html

National Public Credit Information Center, 'Announcement and explanation on newly listed dishonesty judgement defaulters', 2018.7.2. Online in Chinese, at http://www.zhengxinbao.com/6669.html

NDRC, List of first batch of demonstration cities for social credit system construction announced, 2018.1.9. Online in Chinese, at http://www.ndrc.gov.cn/xwzx/xwfb/201801/t20180109_873409.html.

Obama administration unveils "big data" initiative: announces $200 million in new r&d investments (Executive Office of the President, 2012), available at https://obamawhitehouse.archives.gov/sites/default/files/microsites/ostp/big_data_press_release_final_2.pdf (accessed 16 December 2022)

OECD, 2011. Policy framework for investment user's Toolkit, Public Governance. http://www.oecd.org/investment/toolkit/policyareas/publicgovernance/41890394.pdf

Official websites for Qianjang Score, at http://qianjiangfen.cn/ ; http://qianjiangfen.cn/newsList.html

PBCCRC official website, online at http://www.pbccrc.org.cn/zxzx/zxgk/gywm.shtml.

People's Bank of China, Report on China's Credit Reporting Industry, 2013 (online in Chinese) [中国人民银行，中国征信业发展报告]

People's Bank of China, The 'government + market' two-wheel-driven development model of the credit reporting market has initially taken shape, 2018 (online in Chinese), at http://www.pbc.gov.cn/zhengxinguanliju/128332/128352/128402/3570533/index.html [征信市场"政府+市场"双轮驱动的发展模式初步形成]

People's Bank of China, The 'government + market' two-wheel-driven development model of the credit reporting market has initially taken shape, 2018 (online in Chinese), at http://www.pbc.gov.cn/zhengxinguanliju/128332/128352/128402/3570533/index.html.

People's Bank of China, Press Release, 2020.12.13, Online in Chinese, at http://www.pbc.gov.cn/goutongjiaoliu/113456/113469/4142965/index.html

People's Bank of China, The approval of the personal credit reporting business licence of Pudao Credit Information Co., Ltd., 2020. 12.25 http://www.pbc.gov.cn/goutongjiaoliu/113456/113469/4153264/index.html ;

People's Bank of China, The number of corporate credit reporting agencies across the country, 2022. Online in Chinese at http://www.pbc.gov.cn/zhengxinguanliju/128332/128352/2875623/index.html

Qingzhen Civilisation Office, 'Exploration and practice from "honest farmers" to "intelligent urban and rural areas"' 2018.07.18, online in Chinese, at http://gz.wenming.cn/jujiaogz/201807/t20180718_4764225.shtml [清镇市文明办, 从"诚信农民"到"智惠城乡"的探索与实践, 2018]

Qingzhen government, 2011, Qingzhen's 100 standards for honest farmers, *Modern Guizhou*, 13. [清镇市 100 条诚信农民标准].

Report of the Fifth Plenary Session of the 18th Communist Party of China (CPC) Central Committee, 2015. [中国共产党第十八届中央委员会第五次全体会议公报]

Sesame Credit website: https://www.zmxy.com.cn/

Shanghai population statistics from the National Bureau of Statistics of China, https://data.stats.gov.cn/english/easyquery.htm?cn=E0103. (In Chinese).

Shanghai Social Credit website. https://credit.fgw.sh.gov.cn/

Sina Finance, ‘Rural credit provision is too difficult’, 2016.12.16, online in Chinese, available at http://finance.sina.com.cn/meeting/2016-12-16/doc-ifxytqav9461371.shtml. In fact, the Central Bank in 2015 required local governments to lead the SCS in rural areas and help small and medium sized companies apply for credits. People’s Bank of China, [2015] No. 180, online in chinese, available at http://www.gov.cn/zhengce/2016-05/24/content_5076163.htm

State Council, Report on the Five Year Economic Plan (2006-2010), [2016] No.12, approved by the National People's Congress on 2006.3.14. Online in Chinese, at http://www.gov.cn/ztzl/2006-03/16/content_228841.htm

Sun J., ‘Looking forward to introducing Sesame Credit to rural areas’，Fenghuang news, 2016.12.21. http://inews.ifeng.com/50448180/news.shtml.

Tang Y., ‘Supreme Court: It has restricted 4.7 million flights for those subject to execution for dishonesty’, China News, 2016.9.22, online in Chinese, at http://www.chinanews.com/gn/2016/09-22/8011149.shtml.

Taobao 1.18 billion pieces of user information leaked, two men sentenced. Sina news, 2021.6.17 [淘宝 11.8 亿条用户隐私被泄露两男子被判刑]

The 13th Five-Year Plan of the People’s Republic of China on economic and social development, 2016. [中华人民共和国国民经济和社会发展第十三个五年规划纲要]

Tianjin steadily develops work on data security and personal information protection, www.tjcac.gov.cn, 2020.8.20 [天津扎实开展数据安全和个人信息保护工作]

Tianjin works on regulating APP illegal collections of personal information, www.gov.cn, 2019.11.23[天津市开展 APP 违法采集个人信息集中整治]

Torstensson, P. 2023. *Basel III finalisation in the EU: the key elements and how they make the EU banking system more resilient*. Available

at:https://www.ecb.europa.eu/press/financial-stability-publications/macroprudential-bulletin/focus/2023/html/ecb.mpbu202312_focus01.en.html

UC's official website, available at UC in English - UC

US population, World Bank Statistics. 2017 https://data.worldbank.org/indicator/SP.POP.TOTL?locations=US

Wan Cunzhi, 2017, Accelerating the construction and improvement of a Credit Referencing system that covers the whole society, (in Chinese) [万存知：加快推进建立健全覆盖全社会的征信系统] https://huzcredit.huzhou.gov.cn/art/2017/11/5/art_1229635293_4583.html.

Wang Wei, 'Beijing Station Pilots 'Face Scanning' for Entry Using Facial Recognition Technology'[北京站试运行"刷脸"进站 使用人脸识别技术] (Souhu.com, 5 Feb 2016) <https://www.sohu.com/a/58100500_119703> accessed 7 June 2023;

World Bank Statistics 2017, China Rural Population https://data.worldbank.org/indicator/SP.RUR.TOTL?locations=CN

Wu X., 'Qingzhen, Guizhou: Create "intelligent benefits" for urban and rural areas and play a concerto of integrity', thepapers.cn, 2021.4.23. https://www.thepaper.cn/newsDetail_forward_12356504.

Xi Jinping: Report Delivered at the 19th National Congress of the Communist Party of China – Secure a Decisive Victory in Building a Moderately Prosperous Society in All Respects and Strive for the Great Success of Socialism with Chinese Characteristics for a New Era, 2017.10.18], available at http://www.gov.cn/zhuanti/2017-10/27/content_5234876.htm (accessed 16 December 2022) [习近平决胜全面建成小康社会夺取新时代中国特色社会主义伟大胜利——在中国共产党第十九次全国代表大会上的报告]

Zhang Y., ‘Wan Cunzhi: The implications of eight market institutions divesting their personal credit reporting business’, Caixin.com, 2018.5.27. Online in Chinese, available at https://finance.caixin.com/2018-05-27/101257470.html

Zhejiang Government official statistics, in Chinese, available at https://www.zj.gov.cn/col/col1229631725/index.html

List of Major Abbreviations and Acronyms

General

Artificial intelligence (AI)

Common-pool resources (CPR)

Deep learning (DL)

Initial Public Offering (IPO)

Machine learning (ML)

State owned enterprises (SOEs)

European Union laws and concepts

Digital Services Act 2022 (DSA)

Digital Markets Act 2022 (DMA)

General Data Protection Regulation 2018 (GDPR)

Very large online platforms (VLOPs)

U.S. laws and entities

Federal Trade Commission (FTC)

Holding Foreign Companies Accountable Act 2020 (HFCAA)

New York Stock Exchange (NYSE)

Public Company Accounting Oversight Board (PCAOB)

Security Exchange Commission (SEC)

Chinese laws, concepts and entities

Beijing International Data Exchange (BIDE)

Belt-and-Road Initiative (BRI)

Communist Party of China (CPC)

Critical information infrastructure (CII)

Cyberspace Administration of China (CAC)

Digital Silk Road (DSR)

Financial Credit Information Foundational Database (FCIFD)

Hong Kong Stock Exchange (HKEX)

Joint-sanctions Mechanisms (JSMs)

Ministry of Human Resources and Social Securities (MOHRSS)

Ministry of Industry and Information Technology (MIIT)

Multi-Level Protection System (MLPS)

National Development and Reform Commission (NDRC)

National Energy Administration (NEA)

National People's Congress (NPC)

People's Bank of China (PBOC)

Personal Information Protection Law 2021 (PIPL)

Personal information interests (PII)

State Administration for Market Regulation (SAMR)

State Council (SC)

Supreme People's Court (SPC)

Supreme People's Procuratorate (SPP)