A Dystopian Future? The Rise of Social Credit Systems

Edited by Liav Orgad and Wessel Reijers

European University Institute

Robert Schuman Centre for Advanced Studies

Global Citizenship Governance

A Dystopian Future? The Rise of Social Credit Systems

Edited by Liav Orgad and Wessel Reijers

Robert Schuman Centre for Advanced Studies

The Robert Schuman Centre for Advanced Studies, created in 1992 and currently directed by Professor Brigid Laffan, aims to develop inter-disciplinary and comparative research on the major issues facing the process of European integration, European societies and Europe's place in 21st century global politics.

The Centre is home to a large post-doctoral programme and hosts major research programmes, projects and data sets, in addition to a range of working groups and ad hoc initiatives. The research agenda is organised around a set of core themes and is continuously evolving, reflecting the changing agenda of European integration, the expanding membership of the European Union, developments in Europe's neighbourhood and the wider world.

For more information: http://eui.eu/rscas

Global Citizenship Governance

The project explores the international norms and structures for granting citizenship. It has six objectives: [1] to investigate the history of naturalisation and what it can teach us about 21st century challenges; [2] to identify the recent legal developments and establish the most up-to-date legal standards in the field of naturalisation law that, taken together, may form the basis for ICIL; [3] to set out the theoretical foundations and the justifications for the establishment of ICIL; [4] to analyse the normative and structural implications derived from an-ICIL approach for future citizenship policy development; [5] to examine the future of citizenship – how technology can and should remodel the way citizenship is governed; and [6] to explore the interrelationship between ICIL, global migration, and constitutional identity. In essence, the project seeks to formulate international standards by which states can admit migrants without fundamentally changing their cultural heritage and slipping into extreme nationalism. The outcome can serve as a basis for a future reform in international law, EU law, and national legal systems.

For more information: http://global-citizenship.eui.eu.

Advisory board

Abstract

Is China's Social Credit System leading the way to a dystopian future? 'Yes', claims Wessel Reijers, who identifies dangerous dictatorial tendencies that do not live up to the promise of cultivating civic virtue. 'No', argues Jens van 't Klooster, who sees it as a promising way to enhance distributive justice and an alternative for price mechanisms in market economies. In an online symposium, we have asked several eminent scholars to take sides in this timely controversy concerning an unparalleled effort of social engineering that will soon influence the lives of over a billion citizens. Is the Social Credit System a 'digital dictatorship' or a 'digital republic'? How unique it is in a comparative and historical perspective? Is a social credit system good for women? And what are the implications of the system on the institution of citizenship?

Kickoffs by Wessel Reijers and Jens van 't Klooster; **contributions** by Yongxi Chen, Jiahong Chen, John Cheney-Lippold, Jeremy Daum, Costica Dumbrava, Jelena Dzankic, Joshua Fairfield, Primavera de Filippi, Cristie Ford, Francesca Lagioia, Miriam Müller, Liav Orgad, Alberto Romele, Giovanni Sartor, Mathias Siems, Mac Sithigh.

Keywords

TABLE OF CONTENTS

How to Make 'The Perfect' Citizen?

Wessel Reijers[*]

'How to make the perfect citizen'? This question might have crossed the minds of the Chinese government officials who, in 2014, formulated the idea of a Social Credit System[1]. China is a vast country, with thousands of local units of government, innumerable businesses, and far over a billion inhabitants. Governing such a giant, while taking care of rampant corruption, public unrests, and fluctuations in the global and local economies is not an easy task. Yet, 'fortunately' for the Chinese government, the last decade has brought about a huge shift in the infrastructure, which has been transformed from a pre-digital ecosystem of cash payments and messy real-world interactions towards an ecosystem of digital payments through Baidu Pay, the online distribution of goods through Alibaba, and the sharing of millions of messages and data files through WeChat[2]. This infrastructure offers new prospects for the Chinese government's idea of the Social Credit System. It has handed it effective tools to monitor its citizens' behaviours, giving each citizen a 'score' based on behaviour, and implementing a connected system of incentives[3].

Scoring Citizens

The Chinese Social Credit System gets easily likened to dystopian science fiction scenarios in the West, which at least in part seems to be related to the authoritarian character of the Chinese state. However, it seems wrong to assess the system merely in relation to the potentially nefarious tendencies of Xi Jinping's administration, exemplified by rampant human right abuses, oppression of minorities, and political persecutions[4]. Notwithstanding these terrifying practices, we should ask: is the implementation of *a* Social Credit System leading to a dystopian political system? More specifically, I ask: is such a system more akin to a digital republic or a digital dictatorship? The answer I give is that the 'making' of citizens implied in Social Credit Systems limits the political freedom of citizens to act virtuously and promotes dictatorial rule.

In terms of the technical properties, I define a Social Credit System as any system of governance that uses surveillance capacities of the state and/or private entities to *trace* behaviours of citizens, assigns *scores* to citizens, and incentivises behaviours by means of *awards* or *sanctions* based on the scores. I depart from what I believe is the crucial difference between such a system and other means for either keeping public order or scoring individuals based on their behaviour. This essential difference lies in (1) tying the score based on multiple behaviours to *who* someone is, as an entity that persists through time (i.e., to a person's moral character), (2) tying the score to a person as a *citizen,* and not merely for instance as a taxi passenger (e.g., in getting an Uber score) or a car driver (e.g., in getting traffic penalty points).

With regard to the first issue, I would like to point at considerations for tying offences such as major crimes to a person's *permanent criminal record* and for excluding minor offences from such a record. Tying offences to a permanent record entails something about the person in question: they become part

---

[*]   Max Weber Fellow, Robert Schuman Centre, European University Institute.

[1]   Legal Reform Agenda of the Fourth Plenum (2014), Translated version accessed at https://chinacopyrightandmedia.wordpress.com/2014/10/28/ccp-central-committee-decision-concerning-some-major-questions-in-comprehensively-moving-governing-the-country-according-to-the-law-forward/ on 16-11-2018.

[2]   Diab, R. S. (2017), 'Becoming-Infrastructure: Datafication, Deactivation, and the Social Credit System', *Journal of Critical Library and Information Studies*, *1*(1), p.p. 1–23.

[3]   Creemers, R. (2018), 'China's Social Credit System: An Evolving Practice of Control', *SSRN*. University of Leiden.

[4]   Human Rights Watch (2017), 'China, Events of 2017'. Accessed at https://www.hrw.org/world-report/2018/country-chapters/china-and-tibet on 18-01-2019.

of someone's history as a good or bad citizen in the eyes of the public. A Social Credit System extends the notion of a permanent record tied to a person to many kinds of everyday behaviours. The second point can be illustrated by referring to the notion of civic virtue invoked by the Chinese policy document on the Social Credit System, which does not merely relate to be a good civil servant or a good consumer, but to be a good *citizen*. In other words, it is not just the score but also its *integral* character (i.e., linking heterogeneous activities such as jaywalking to having access to train tickets) that sets it apart from other scoring systems.

Promoting Civic Virtue?

Essentially, a Social Credit System aims to make good citizens, *qua* citizens, through promoting civic virtues such as 'trustworthiness'. This aim seems to resonate with republican thought, which emphasises the vital importance of the cultivation of civic virtue. However, the dispositions cultivated by a Social Credit System cannot be rightfully designated as civic virtues for two reasons: (1) because a score constitutes an aim external to any 'virtuous action', and (2) because the resulting activity tends to conformity rather than to distinction in the public sphere.

First, as Aristotle[5] indicated, we can consider three 'ultimate goods' of action – pleasure, honour and eudaimonia. 'Pleasure' and 'honour' are not self-contained, meaning that pleasure depends on some external stimulus (e.g., consuming food) and honour depends on the external approval of others (e.g., having a reputation in making a work of art). Only actions aimed at 'eudaimonia' are self-contained, meaning that their aim is achieved in acting (e.g., courage is attained in acting courageously). Virtue, according to this reading, is the disposition that aims at eudaimonia and does not entail aiming at anything external, such as money or reputation. A Social Credit System seems to contradict this notion of virtue. On the one hand, it promotes activities such as labouring for the community because it offers a standard of reputation that can be a source of honour – for instance, in seeking to receive the highest credit score in the neighbourhood. On the other hand, it promotes activities based on expected pleasure (e.g., receiving discounts on certain consumption items) or pain (e.g., being put on a blacklist).

Second, as Hannah Arendt[6] argued, virtuous action requires a citizen to be able to distinguish her/himself in the public sphere. Crucially, this means that the outcomes of a citizen's actions can always be unexpected because they constitute what might be an exceptional, hence 'excellent' activity. In other words, acting virtuously is not merely acting in an agreeable manner or following known social conventions, but it means standing out by distinguishing oneself. This notion of the distinction entails that for virtuous actions, there can be no fixed measurements or standards based on past, generic behaviours. A Social Credit System contradicts this notion of virtuous action. In order to score citizens, a certain list of categories or a catalogue need to be established, based on observable, expected behaviours that lead to either positive or negative evaluations. For instance, donating money to charity is a measurable, observable type of behaviour that could be positively evaluated. However, it in no way constitutes an activity by which a citizen distinguishes herself in the public sphere. By subjecting 'good' and 'bad' actions to upfront categorisations based on what can be expected and measured, a Social Credit Systems promotes conformity and therefore demotes civic virtue.

Hence, it is highly questionable that a Social Credit System would lead to the emergence of a 'digital republic', neither in China nor elsewhere. Would it contain a tendency towards a digital dictatorship? To answer this question, we would first need to clarify the meaning of a dictatorial governance. This type of governance derives from the master-slave relation of the ruler and the ruled, characterised by the absence of political freedom. I argue that Social Credit Systems contain a *tendency* towards a dictatorial

---

[5]   Aristotle (1999), *Nicomachean Ethics*, Irwin, T. (ed.). Indianapolis: Hackett Publishing Company.

[6]   Arendt, H. (1958), *The Human Condition*. Chicago: University of Chicago Press.

rule, which I explicate according to its (1) output (i.e., the resulting behaviour of the ruled), and (2) input (i.e., the method of governance applied by the ruling).

Virtuous Governance?

Firstly, I claim that the Social Credit System promotes 'slavish' behaviour. Being a slave means essentially to be unfree, which can be explicated as (1) being subject to necessity (i.e., having to 'make a living' through labour), *and* (2) being excluded from the public realm in which one can act and speak amongst other citizens. By linking civic behaviour, such as having online discussions, with production and consumption, such as having access to train tickets or cleaning the neighbourhood, the incentives mechanisms of a Social Credit System can become part of life's necessity. Certainly, many of the behaviours linked to the categories of the Social Credit System remain voluntary (e.g., one can choose to donate to a designated charity) but *reckoning* with the social credit score itself becomes a necessity. In Arendt's terms, civic behaviour in the public realm gets progressively transformed into labour, the human activity aimed at continual production and consumption, which is essentially an unfree activity. In other words: the more people's lives are dominated by a-political or anti-political concerns 'in public', such as by labouring for the sake of the social credit score, the less they are free to participate in political activities. Hence, the more citizens will have to reckon with the effects of their everyday actions on their credit scores, the less they will be free to engage in political action in the public sphere.

Secondly, I argue that the type of governance involved in a Social Credit System promotes master-like rule. The prototype of such rule is the mastery of a craftsman over his craft, because it implies full control over both means (e.g., materials, instruments) and ends (e.g., a function of a technical object). For this reason, Plato likened his ideal statesman to a craftsman, who would use his social engineering skills to control his populous. It is not hard to see how the Social Credit System implies an effort of social engineering, but what sets it apart from other forms of social engineering is its *instrumentalisation* of the citizen. While the factory would instrumentalise the worker, the worker would be free from its control outside of its walls and outside of the bounds of the working day. Citizenship, however, permeates public life, which led Aristotle to ask whether the good citizen is in any way different from the good human being. The underlying notion of a Social Credit System is therefore that the ruler, be it one man or a collective, instrumentalises his subjects, and treats them as means towards a definite end – which in the Chinese case is social order and the centralised power of the Communist Party.

Conclusion

We ought to be suspicious of Social Credit Systems like the one that is implemented in China, because of their tendency to demote civic virtue and promote dictatorial rule. However, we need to make some reservations concerning this conclusion. First, a perspective of western political philosophy is likely to be incompatible in some respects with the Chinese Confucian political philosophy. Second, matters of scale and scope of any Social Credit System should not be neglected. Although we did not discuss those, it becomes immediately clear that a scoring system is less problematic when applied to a limited set of activities (i.e., limited scale) or not to a person *qua* citizen but only, for instance, *qua* responsible traffic participant (i.e., limited scope).

Rewarding Virtuous Citizens

Jens van 't Klooster[*]

A Social Credit System rewards those citizens who are seen as virtuous and punishes those who are not. Consider someone with a low social credit rating resulting from traffic violations, a few missed loan repayments, a divorce, and clipping nails in a public space. Would it be fair to ban this person from flying? Few would think that is particularly fair. Indeed, the Chinese Social Credit System, as presented by the western media, is widely seen as the height of technological dystopia.

But is that intuition well-founded? Wessel Reijers has sought to identify features that he takes to justify a rejection of the Chinese Social Credit System but forgoes an equally critical consideration of the alternatives. Relying on the market, the default solution of western societies is not more just. Indeed, as I argue, while Social Credit Systems may have their flaws, (1) we tend to be very reluctant to apply similarly stringent moral criteria to western capitalist institutions and, in fact, (2) a Social Credit System may actually serve to make western societies more just. Nonetheless, while I am more sanguine than Reijers about shaping the behaviour of citizens through Social Credit Systems, I agree that it is potentially an incredibly powerful and dangerous social technology. We should not focus on the technology itself, but rather on the political institutions that govern it.

What is a Social Credit System?

In its 2014 *Planning Outline for the Construction of a Social Credit System*, China formulated the objectives of the Social Credit System as that of 'raising the honest mentality and credit levels of the entire society[1]'. To this end, it 'uses encouragement to keep trust and constraints against breaking trust as incentive mechanisms'. The Outline promises that the system will mark a major step in 'comprehensively implementing the scientific development view and building a harmonious Socialist society'.

The 2014 Outline goes on to list a wide range of pathologies that the Social Credit System is meant to solve. Most fundamentally, there are not enough social institutions to incentivise trustworthy behaviour. For the Chinese state, being trustworthy means primarily complying with the law and adhering to contracts. This has become a problem in China because the rapid economic development has eroded traditional social structures and given rise to a semi-autonomous private sphere that is unstable and plagued by a range of social pathologies. Amongst those that the Outline identifies are 'grave production safety accidents, food and drug security incidents […] commercial swindles, production and sales of counterfeit products, tax evasion, fraudulent financial claims, academic impropriety and other such phenomena', which 'cannot be stopped in spite of repeated bans'. In addition to flaws in the legal system, financial markets are underdeveloped, services provided are insufficiently regulated, and 'the mechanisms to protect the rights and interests of credit information subjects are flawed'. There is widespread discontent about these matters and 'a difference between the extent of sincerity in government affairs and judicial credibility, and the expectations of the popular masses'.

Because these problems are so diverse, it is not surprising that the 2014 Outline has prompted an incredible range of initiatives at the levels of the central government, local government, and the private sector. There is, accordingly, not one Chinese Social Credit System, but a range of initiatives aimed at documenting the behaviour of citizens to punish breaches of social trust and reward trustworthiness.

---

[*] Max Weber Fellow, Department of Social and Political Science, European University Institute.

[1] Made available in an English translation by Rogier Creemer. Accessed at https://chinacopyrightandmedia.wordpress.com/2014/06/14/planning-outline-for-the-construction-of-a-social-credit-system-2014-2020/

Defending the Social Credit System

The 2014 Outline lists severe problems facing China today. If Social Credit Systems indeed contribute to the trustworthiness and responsible use of autonomy, that would be a strong case in its favour. If, instead, all it does is to create fear and oppression, inhibiting individuals from using the freedom that the private sphere provides them, that would be a strong case against the Social Credit System.

The truth is somewhere in the middle. Unlike Reijers, I do not take the justification of the Chinese Social Credit System to depend on big questions such as our ideas about citizenship and utopic visions of the just society. It also strikes me as unfair to cast the Social Credit System as a dystopian technology by invoking moral standards, which we forgo applying to our western institutions. Indeed, as I show, political philosophers tend to apply quite modest moral standards to western capitalist institutions.

The reason why philosophers have been less demanding in theorising the moral demands that apply to western institutions, in particular those specific to a capitalist society, is that they do not do very well in giving individuals what they morally deserve. Consider some well-established failings of a market economy. First, in response to the specific concern raised by Reijers, markets, too, 'make' citizens by incentivising a set of behaviours, which are unlikely to correspond to standards of virtuous action. The price mechanism that underlies market exchanges is mostly sensitive to the skills needed to anticipate market prices successfully. Second, even where the market does reward virtuous actions, it rarely does so in a way that is proportionate to individual virtue. People who go to work all their lives diligently are rarely those who are most rewarded, while many markets provide high rewards to only a small group of participants. Finally, the price mechanism is past-focused. It tends to look more at what parents of beneficiaries have done than what they have done. This is, in part, due to the dramatic failure of the existing system of private credit assessment, one of the issues that the Chinese system is meant to address, but also reflects a more general backward-looking tendency of the price mechanism.

Faced with these well-established facts, those who favour markets have tended to give up on the idea that societal institutions should be evaluating in terms of how they reward virtue. Indeed, this is a striking point of convergence between authors on the (moderate) left, such as John Rawls and Elizabeth Anderson, as well as right-wing thinkers such as Friedrich Hayek. While these authors assign an important role to the individual autonomy that the market enables, they believe that markets should not be expected to reward virtue. Rather, markets should be evaluated by asking whether their outcomes are justifiable. Unequal outcomes are just as long as they are to the greatest benefit of the least advantaged. Hayek and Anderson have their conceptions of justice, but they agree that, if markets are justifiable, there is no further question of whether virtue is adequately rewarded. By this standard, the Social Credit System seems to do just fine.

A Social Credit System for the West

Let me go further and suggest that a well-designed Social Credit System might make western capitalist societies more just (although not all systems currently promoted by the Chinese state are Social Credit Systems in that sense).

Similar to Reijers's definition, I take a Social Credit Systems to be an institution that has two features: (1) a centralised way of rating the virtuousness of citizens based on observable traits, linked to (2) societal rules that reward individuals with high ratings and punish those with low ratings. In light of these criteria, some systems introduced under the Social Credit System-heading are better understood as traditional coercive punishment systems. They focus on specific behaviours, such as entering the cockpit during a flight, and involve proportionate punishments, such as losing access to credit or being banned from air flight.

A Social Credit System is different than a punishment system in that a social credit rating mediates between behaviour and individual freedom. The Chinese Social Credit Systems seek to incentivise a

range of behaviours associated with adherence to promises and rules as well as other social virtues. To this end, a Social Credit System can rate individuals for behaviours that are neither legally prohibited nor morally wrong – e.g., parking cars, divorcing, and clipping nails on public transport. Based on such a broad set of socially desirable behaviours and personality traits, the Social Credit System rates the virtuousness of individual citizens, which in turn affects the opportunities and resources that are available to them.

A Social Credit System of this targeted kind can ameliorate all three objections that I listed. First, although a Social Credit System will be coercive on the individual level, its design can result from collective, democratic deliberation and thus it is potentially sensitive to a much wider range of behaviours that are seen to merit reward. Basically, any behaviour that can be reliably measured is a potential input for a social credit rating. Second, ratings can be designed to ensure a reward in proportion to a social virtue. Finally, a social credit rating is future-oriented in that it seeks to predict individual behaviour, rather than merely rewarding virtuous behaviour that occurred in the past.

The Politics of Social Credit Systems

There are many moral objections that critics might wield against Social Credit Systems, and I do not want to suggest that there are no real worries. In particular, there is a range of concerns that, while they do not apply to Social Credit Systems per se, illustrate the need for an appropriate set of background institutions. For one, there are procedural concerns over the design and operation of the Social Credit System. For the Chinese Social Credit System, it is currently not possible to appeal both individual ratings and the consequences of these ratings for individuals. Likewise, there is a serious concern regarding privacy stemming from the fact that its aggregative ratings are based on widespread surveillance of individual behaviour. Finally, the rules that govern the assignment of ratings are often opaque. Without wanting to downplay any of these objections, I think it is fair to say that they are not about Social Credit Systems per se, but rather about the set of political institutions in which they are embedded. All of this is not surprising since Social Credit Systems are potentially an incredibly powerful social technology, which comes with potentially large benefits and yet equally with risks.

Treating China as a 'Normal' Country

Mathias Siems[*] and Mac Sithigh[**]

Introduction

In comparative law, it is often said that the law of countries of the East Asian legal family is markedly different from western conceptions of law. For example, while western countries prefer a 'struggle for law' with winners and losers, in East Asia personal relations and networks are often seen more important. However, there have been some challenges to this position. For example, Bruce Aronson – an expert on Japanese law – suggests that '[o]ur perceptions of Japan, related to (…) a preoccupation with cultural explanations, may hinder, rather than aid, careful analysis of Japanese law and its impact on society'; thus, he proposes that we should 'treat Japan as a "normal" country that has both similarities to and differences with other advanced societies[1]'.

An analogous dialectic can be applied to the Chinese Social Credit System. Some observers emphasise its uniqueness[2], also calling it a 'great innovation'[3], or – from a critical perspective – equating it with an 'Orwellian nightmare'[4]. Taking the position of Chinese exceptionalism, it may be argued that it would be inappropriate to apply western notions of virtue or justice (e.g., by Aristotle or Hannah Arendt, see the contribution by Reijers; or by Rawls, see the contribution by van't Klooster).

This blog post – and the corresponding paper on which it is based[5] – suggests that it is preferable to regard it as a specific instance of a wider phenomenon. In this respect, China may be considered as a 'normal country' that starts experimenting with rating-based forms of governance. This position does not imply an uncritical perspective, given that any rating system operates in a regulatory context; reflecting on the question also addressed in the other blog posts of this debate (Reijers and van't Klooster), this post will suggest that whether or not these systems present a 'dystopian future' depends on their regulatory environments.

Variants of Social Credit Systems in China and the West

Generally, three versions of the Chinese Social Credit System can be distinguished: China-wide blacklists, compliance scores by pilot cities, and social credit scores by financial institutions. It should also be noted that many of those elements are a moving target. For example, the local schemes are deliberately used to experiment with different forms of ratings, in particular in the input data that may be considered (e.g., whether these mainly include violations of the law or also other behavioural data). As regards to the scores by financial institutions, western observers focus on the Sesame Credit scores

---

[*]   Professor of Private Law and Market Regulation, Department of Law, European University Institute.

[**] Professor of Law and Innovation, Queen's University Belfast.

[1]   Bruce, E. Aronson, B.E. (2014), 'My Key Phrase for Understanding Japanese Law: Japan as a Normal Country…With Context'. *Michigan State International Law Review,* 22, p. 815.

[2]   Dai, X. (2018). 'Toward a Reputation State: The Social Credit System Project of China' *Working Paper*, available at https://ssrn.com/abstract=3193577.

[3]   Backer, L.C. (2017). 'Measurement, Assessment and Reward: The Challenges of Building Institutionalized Social Credit and Rating Systems in China and in the West', In: Proceedings of the Chinese Social Credit System, Shanghai Jaiotong University, 23 September 2017, p. 15.

[4]   See, e.g., Asian Institute for Policy Studies (2017), 'Orwell's Nightmare: China's Social Credit System'. Accessed at http://en.asaninst.org/contents/orwells-nightmare-chinas-social-credit-system/.

[5]   Sithigh, D.N. and Siems, M. (2018). 'The Chinese Social Credit System: A Model for Other Countries?', *EUI Department of Law Research Paper* No. 2019/01, available at https://ssrn.com/abstract=3310085.

that consider not only financial credit records but also, for example, personal information and social relationships; yet, these scores are about to be incorporated in a scoring system by the People's Bank of China with the precise scope not yet decided.

In the West, many countries know financial consumer credit scores or ratings, and recent years have seen a proliferation of rating systems in relation to online platforms and the 'sharing economy', such as eBay, Uber, and Airbnb. Recent innovations reflect the availability of novel sources of data. A subject's use of social media or of sharing economy sites can, for instance, be used as a proxy for creditworthiness. Thus, an Australian scheme for alternative (and significantly cheaper) deposits on rented properties, Trustbond, makes use of both types of data[6]; a Singapore-based lender, Lenddo, claims to use 'non-traditional data…to economically empower the emerging middle class', which includes various social media sources[7]; and a U.S. firm, Tala, operates in developing nations in Africa and Asia, claims to utilise up to 10,000 'data points', such as social media and smartphone use, in order to create a new type of credit score to the advantage of lower-income customers[8]. There have also been forms of aggregation: for example, Traity (a former partner of Trustbond) explains its role as assisting users to 'gather … reputation from different data sources so that [they] can control it, own it, and leverage [it][9]'.

Comparing China and the West, some specific parallels can be identified. For example, as part of the Social Credit System, China uses blacklists of persons who have violated the law in order to exclude them from many aspects of social life with the help of data sharing. This may be related to the use of background checks and 'no-fly lists' in the U.S.[10]; or that in Israel, codes (based on various data) are assigned to passengers as part of a distinctive 'risk-based' security system[11]. In many countries, criminal records and details of insolvent debtors are also collected and may be shared more or less widely in a number of European states. And even in respect of private sector services (e.g., in the sharing economy), a poor reputation can exclude individuals from these emerging markets, with far-reaching consequences for their financial position and participation in some aspects of urban life.

It can also be shown that the design and regulation of today's rating systems are already clear examples of policy diffusion. The Chinese system partly derives from its western counterparts: the financial credit scoring systems of western countries have been an explicit source of influence[12], and tools such as the Sesame Credit scores may be seen as an example of the 'gamification' of rating systems in recent time. In return, the growing global influence of China can mean that the Social Credit System may well be a regulatory tool that could inspire the West: it may show that China now 'appears to have ascended to the position of principal global driving force in political theory and action', with 'the potential to change law and government as we know them in China and beyond[13]'.

---

[6] http://www.trustbond.com.

[7] http://www.lenddo.com.

[8] https://www.devex.com/news/how-alternative-credit-scoring-is-transforming-lending-in-the-developing-world-88487.

[9] http://www.traity.com.

[10] Chorzempa, M., Triolo, P., and Sacks, S. (2018). 'China's Social Credit System: A Mark of Progress or a Threat to Privacy?' *Peterson Institute for International Economics*, Policy Brief 18-14, p.p. 4 and 7.

[11] See e.g., Jonathan-Zamir, T., Hasisi, B., & Margolioth, Y. (2016). 'Is It the What or the How? The Roles of High-Policing Tactics and Procedural Justice in Predicting Perceptions of Hostile Treatment: The Case of Security Checks at Ben-Gurion Airport, Israel'. *Law & Society Review, 50*(608), p.p. 616-7.

[12] See, e.g., Knight, A. (2018). 'Credit: The God of China's Big Data Era', *ECFR China Analysis,* p. 7.

[13] Backer, above note 3, at p 2; Dai, above note 2, at p 1.

Conclusion

Will China's Social Credit System lead the way to a dystopian future? Our answer is that this very much depends on the way the Chinese models as well its western counterparts evolve. Then, a number of related questions emerge. For example, could there be mixtures between the Chinese and western models? If reputation and rating systems consolidate in western markets in a similar fashion, what opportunities, features, controversies, and pitfalls will arise? And how could lawmakers intervene if this happens?

Crucially, any assessment also depends on the legal and regulatory tools that states will adopt. Thus, by contrast to the binary distinction suggested in the blog post by Reijers, we take the position that it is these tools which determine whether it may be rather seen as a form of 'digital republic' or 'digital dictatorship'. Technically speaking, any such regulatory tools either be implemented through the introduction of a new regulatory regime or the application of general requirements to a particular context. It may also be unlikely that there will be a single 'law on ratings' given the relevance of many overlapping policy considerations and corresponding fields of law, such as e-commerce law, privacy and data protection law, anti-discrimination law, tort law, competition law, sector-specific regulation on financial services etc. However, the growing interest in the regulation of powerful 'platforms', as seen in the recent work of the European Commission, may indicate the type of approach likely to be favoured in the coming years.

Social Reproduction and Social Credit Apparatuses

John Cheney-Lippold[*]

Dystopia is a peculiar concept. Like dark matter or 'the market', dystopias are generally difficult to find in the wild even when there are always 'future dystopias' looming on the horizon. While many places on earth may already deserve the title – Congolese cobalt mines, the Gaza Strip, secret Chicago prisons – our interpretation of dystopia is chiefly one based on temporality. Things are changing, and we fear the worst will happen. Or, what's happening is bad, but don't worry – it gets worse.

The types of futuristic societies envisioned by Reijers and van 't Klooster provide outwardly opposing views. But whether dystopian or not, the 'social' of the Social Credit System under debate remains largely untroubled. For Reijers, it's simply an instrumentalised 'social order' within China; for van 't Klooster, it's an ideal market with potential for just consequences. The social serves as a placeholder for something more, an unstated premise that hides in plain sight.

In the case of China, the dystopian algorithmic future of this placeholder is seasoned by Orientalist appreciation. The widely cited Bonaparte quotation 'China is a sleeping lion. Let her sleep, for when she wakes, she will shake the world' locates this futurity on a terrain of awe-inspired fear. In a 2014 speech in Paris, even Chinese President Xi Jinping recited the passage, marshalling the implied danger that establishes China as an imminent, not-to-be-messed-with seismic force to an apprehensive French audience.

And for the West, with the narrativised help of Black Mirror's 'Nosedive', China's Social Credit System supplies the 'ur text' of the intersection between technology and dystopia for the contemporary era. But while Social Credit Systems of any form are appalling for many reasons, I want to resist a relying on a good vs. bad framing without taking a breath to assess a presupposition. What do all Social Credit Systems purportedly want; what is the 'social' in social credit? I answer with conventional wisdom from digital studies, science and technology studies, and critical theory: to reproduce the social, just as asymmetric, biased, and discriminating as it was when it was born.

Social, So What?

I am no expert of China. And while I am a professor who writes and teaches on digital technology and cultural studies, my interest in the issue of China's Social Credit System is one that does not necessarily fit the typical Orwellian shade of authoritarianism. I propose we briefly remove China from our analysis and think about 'Social Credit System' in the abstract, alleviating some of the more problematic points of cross-cultural pressure - like racial othering and cultural fetishization - normally assigned to China by a western audience. In doing so, I structure my perspective in an Althusserian form, concerned not only with the power to make 'good citizens', but also the process by which that power is constructed and maintained.

Abstracting a Social Credit System allows us to ask more general questions. Most important: what in the world is 'social'? Any social system, credit or not, needs to reproduce itself. For Althusser, the social conditions of life are reproduced - just like labour power - through repression and/or ideological interpellation. While Althusser's perspective has been long-troubled by scholars, it remains a useful analytic of the social that refuses a vague, neutral origin story. Rather, the social is a constantly reorganised definition of community that always exploits some for the sake of others. Far from Hobbesian or Lockean mythologies of harmonious nation-building, Althusser's formulation apprehends empowered difference, i.e., those with power vs. those without, as its structural a priori.

---

[*] Professor, Department of American Studies, University of Michigan.

But Social Credit Systems, in general, are more intimate than some rudimentary ideological relation. Reijers: such a system 'extends the notion of a permanent record tied to a person to a great many kinds of everyday behaviours', a tying that combines singular institutions of power (the police, education, the media) into one agglomerated ideological social factory. This definition of society routes through the calculations of biopolitics and statistics, scarred and disfigured by the founding limitations of datafication.

Limits of Datafication

One limitation of datafication: as much as dystopian imaginations tend to universalise their anxieties onto all aspects of life, there can be no completeness of surveillance and control. From Chinese Social Credit to Facebook's 'Trustworthiness' scale and to even Pinochet's 'Directory of Commercial Information', these systems rely on complex, but still incomplete, methods of deducing an individual's value to society. Subjects are not interpreted but forced to prefigure their actions based on how they are datafied: smiling at a stranger won't raise your score while buying vegetables will.

In this case, civic obligations are configured by the recognised perimeter of datafication, a streetlight effect of virtue where civic virtue that is untranslatable into data gets demoted or ignored. While virtue itself can never be separated from existing reward structures in society (in the U.S., 'I Voted' stickers, not votes, operate as the real designation of virtue), an ideal of virtue configured according to data and its algorithmic outcome pursue a value structure according to what is interpreted to be measured as virtuous, not what may be socially deemed as virtuous. The social is reproduced in intermediary, data-based, form.

Another limit of datafication: that Social Credit Systems rely exclusively on data means; subjects are not just seen by power but seen by power through data. The production and reproduction of the social is how data produces and reproduces its datafied subjects. Yet instances of datafication are exceedingly social – and I do not mean this in a good way. Just like the social reproduces the social, readers interested in questions of digital technology may be familiar with the fact that algorithmic processing of data also inevitably reproduces existing biases and oppressions. Facial recognition systems are notoriously racist and transphobic. Algorithmic hiring systems redouble sexism. Quantified self-tracking systems further alienate and dispossess users with disabilities.

Datafication is thus an intensely political process, and that 'any behaviour that can be reliably measured is a potential input for a social credit rating', citing van 't Klooster, should worry us. What it means to 'reliably measure' is not merely a question of scientific method as measurement itself can only exist within the realm of politics. We see this in China. While the persecuted Uyghur ethnic group in the country's western Xinjiang region are certainly repressed by security forces, the Uyghurs – their faces and names – are also studied in ways that may technologically improve that repression. Just like how the post-9/11 U.S. push to hire more Arabic translators allowed for increased surveillance and thus measurement of Arab-speaking populations, measurement is not just using a ruler to assess the length of an object. It is an intensely political act, an act where to 'reliably measure' this social in unqualified form is, at face value, unjust. If an algorithmic Social Credit System led to dystopia, I would not blame the algorithm but the 'topia' it comes from.

Qualifications for the Future

Ironic for an argument of abstraction, I end with an appeal for qualification. Much like an unqualified 'citizen' or 'civic virtue' – which strategically positions the rights and obligations of citizen away from its lived, differential reality, and thus affirms an ideal that likely will not apply in practice – the qualified examples of the U.S. and China show that not all types of people are citizens in the same way. The ideal of the nation, and the society that grounds that ideal, is itself exclusionary. While Arab people in the U.S. are formally citizens, their enacted practice of citizenship is ontologically conditioned through the

lens of white supremacy. The same, but in distinct ways, with China and its Uyghur population. When datafication reproduces the social, it is precisely this relationship of marginalization that is intentionally and unintentionally datafied.

We need to both qualify subjects of surveillance and the underlying assumptions of that surveillance. Neither capitalism nor socialism was discussed in Reijers or van 't Klooster – only an unqualified market – in debates about the efficacy of Social Credit Systems for enforcing authoritarianism or producing justice. I return to the original question: what is the 'social' of social credit? At the end of the day, these systems are not currently for the demos. The social is deployed as a source of information, turned against the demos by forces hegemonically empowered to maintain ideological control. If there is to be any justice, I would most intently ask: how do we make a Social Credit System democratically operated – not a social credit by a socialist Social Credit System?

The Docile Minds of Perfect Societies

Jelena Džankić[*]

Human behaviour has always been assessed and judged in some ways, be it through the most basic social interaction and benchmarking; through observation of norm abidance in transparent and less transparent ways; or, recently, through digital technologies. Wessel Reijers and Jens van t' Klooster look at Social Credit Systems as a rule by 'carrots' and 'sticks', based on structured observation and evaluation of individual behaviour. Illustrating their arguments with the example of China's Social Credit System (SCS), they come to opposite conclusions about this model of governance. I acknowledge the merits of van 't Klooster's idea of the desert, implying that a Social Credit System may have positive effects on distributive justice and offer appropriate alternatives for price mechanisms in market economies. However, I ardently oppose the use of surveillance mechanisms in regulating the relationship between individuals and governance structures. As a result of three interrelated dynamics, rather than creating 'perfect' citizens, Social Credit Systems are more likely to create calculated and passive subjects.

Discipline by Design

First, Social Credit Systems are tools for disciplining society. By design, they allow a centralised authority to observe and evaluate individual behaviour. Their power is in channelling human action towards the expected *award* and in avoidance of the possible *sanction*. Awareness of surveillance is central to Social Credit Systems, as it can curtail free will and individual liberty. Just as in Bentham's panopticon, individuals have no certainty of whether they are observed at any given moment or not. This motivates citizens to regulate their conduct under the assumption that they are being watched and scrutinised. Hence, Social Credit Systems have the potential to create 'docile minds' – an extension of the Foucauldian notion of 'docile bodies'- through normalisation of subtle surveillance and acceptance of *transparent* disciplinary mechanisms. A society composed of such 'docile minds' might lead to smooth governance and reduce conflict horizontally (among individuals) and vertically (between individuals and governance structures). If, for instance, a 'perfect society' is characterised merely by the absence of conflict, then 'docile minds' are conducive to it. These 'perfect citizens' will, however, have internalised the disciplining tools and forgotten their freedoms to such a degree that their 'perfect society' will resemble a golden cage. A prison, paradoxically. And in no panopticon, be it golden or digital, there is room for democracy because freedom is taken away by default.

A Society of Bystanders

Second, Social Credit Systems leave little room for valorising non-rewardable virtuous deeds or actions, but recompense utilitarian 'good' acts. Reijers rightly points out that the core idea behind an SCS is to 'make good citizens, *qua* citizens, through promoting certain civic virtues such as 'trustworthiness', but that the problem with such an approach is the externalisation of the score to 'virtuous action'. A 'virtuous action', in itself, does not require a reputational or financial reward. Hence, any rewardable 'good deed' would represent an action of instrumental conformity. Allocating scores that do not take into account the context of individual behaviour or action vis-à-vis rules will *not* create 'perfect citizens', but rather may give birth to a society of bystanders and individuals acting exclusively on the basis of personal utility. To illustrate this, imagine two scenarios where there exists a credit system with a fixed list of deeds for which an individual receives positive and negative scores.

---

[*]   RCAS Fellow, GLOBALCIT, Robert Schuman Centre, European University Institute.

Scenario 1: Jack goes for a stroll around the neighbourhood. All of a sudden, he sees an elderly woman having a heart attack on the other side of the street. To offer her help (an action bearing no reward in terms of social score), Jack needs to engage in civil disobedience - cross the street illegally (an act for which he will receive a negative social score). Under the Social Credit System, Jack would be a 'perfect' citizen if he kept on walking and succumbed to bystander apathy of a 'docile mind'. Yet, would it make him less 'perfect' if he committed a non-rewardable deed of highest moral value, which however presupposed a previous punishable action?

Scenario 2: Jill works as a taxi driver. She does not particularly like her job and is often rude to customers (an act bearing a negative social score). However, she knows that her blood donations (an action for which she would receive credits) will even out her behaviour towards her clients. Under the rules of the Social Credit System, Jill is a 'perfect' citizen, even though her 'good deeds' are a matter of instrumental conformism rather than expressions of virtue.

Slavish Behaviour

Third, sanctions inherent in Social Credit Systems have diffuse effects and as such are likely to induce individuals to future 'slavish behaviour'. Imagine Jack from the example above. His score has been lowered because he was caught on camera jaywalking to save the old lady. For the sake of the argument let's also imagine that because of his action Jack cannot travel to visit his relatives, has a 30 per cent higher electricity bill, and he has been denied a loan. This means that not only is Jack subject to a sanction, but also it is diffused over many segments of his life. Once experienced the repercussions of his action at different levels, Jack will be more likely to 'play by the rules'. Furthermore, in a context where surveillance of behaviour has been internalised, any public assessments of the quality of governance are likely to be framed through false consciousness. This shift towards a passive and obedient individual in the relationship with a central mandating authority implies a move from a free decision-making citizen to a mere subject.

Perhaps my scepticism of the democratic potential of Social Credit Systems is a direct outcome of my post-communist background. The lived past of surveillance by secret service that I experienced in communist times was rather dystopian. The present digital surveillance of online activities to assess and influence market behaviour (or, as the case of Cambridge Analytica has shown, political choices) is, to say the least, problematic. The 'brave new worlds' of literary works often depict gloomy digital futures, which are a direct product of normalisation of control and coercion by a central authority. To me, this points to the fact that one will find no 'perfect citizens' inside a panopticon, be it digital or not. Rather, it will be a golden cage hosting 'docile minds' - utilitarian bystanders and passive subjects.

Seeing Like an Authoritarian State

Cristie Ford[*]

The Social Credit System (SCS), as proposed in China, provokes an interesting philosophical question. Can human behaviour be moulded to create a better, fairer society by awarding 'credit' to those who behave in ways seen as positive by some broad-based authority, in this case, the state? Is the SCS a promising idea for developing a well-functioning society under conditions of complexity and diversity, or is it a massive exercise in government control over virtually all aspects of citizens' lives?

The SCS in China is not an abstract concept. If fully implemented, it would affect over one billion human beings in a context marked by an increasingly authoritarian and surveillance-obsessed government regime, extensive human rights abuses, political persecutions, and tragic 20th-century history of state-sponsored citizen-on-citizen violence. To be clear, the SCS is being proposed by the same government that is currently detaining hundreds of thousands, perhaps a million, of its ethnic minority citizens in 're-education camps' beyond the reach of its legal system[1].

Against this background, it is not 'wrong' to assess SCS 'merely in relation to the potentially nefarious tendencies of Xi Jinping's administration', as Wessel Reijers suggests. On the contrary, it is analytically problematic and perhaps amoral to proceed as if the SCS concept is a purely technocratic initiative that exists at some metaphysical separation from the regime that spawned it.

To be fair, we should be sceptical that any state is more interested in enhancing genuine collective virtue than in the surveillance and control that make its bureaucratic jobs easier. In *Seeing Like a State*, James C. Scott showed how a state values 'synoptic legibility', or the ability to deeply see its citizens, for its own purposes. Synoptic legibility makes it easier to tax people, conscript them, track them, and punish them. Scott describes how the state – along with other agents of 'homogenisation, uniformity, grids, and heroic simplification', including capitalism – tend to flatten outlived reality into forms that are legible for its purposes[2]. It is not a surprise that any state might see the promise in new data tools.

At the same time, it is not a coincidence that an authoritarian regime is advancing these particular, panoptic initiatives. The SCS is, at its root and in its purpose, inextricably entwined with a totalising and pervasive state-driven project of surveillance, homogenisation, and fear.

Should a State Try to Allocate Moral Deserts?

Respecting human dignity requires at least three things from the state: it requires that the state not offend – at a minimum – basic individual human rights; it requires the rule of law, when questions of individual rights and responsibilities intersect with coercive state power; and it requires that the state not be utterly panoptic, so that individuals have some scope for autonomy and privacy in their lives. President Xi's regime cannot be trusted to provide these things, but the problem with the SCS goes deeper than that.

For the sake of argument, let us imagine an SCS regime that respected at least the most basic human and equality rights, and accorded citizens something resembling the rule of law. What would this look like?

---

[*] Professor, Allard School of Law, University of British Columbia.

[1] Buckley, C. (2018). 'China Is Detaining Muslims in Vast Numbers. The Goal: Transformation.' *The New York Times*, 8 September 2018.

[2] Scott, J.C. (1998). *Seeing Like a State: How Certain Schemes to Improve the Human Condition Have Failed*. New York: Vail Ballou Press, p.8.

First, the regime observes fundamental values around anti-discrimination; human autonomy, dignity, and equality; and the ability to participate, through free speech and freedom of association, in public life. Within the broad SCS mandate, which covers not just formal rights but also the 'honest mentality' (that is, presumably, the virtue and morality, and perhaps thoughts and intentions) of individuals, equality and anti-discrimination provisions would have to be understood in correspondingly broad terms. So, even if divorce is considered morally wrong and punishable under SCS, a woman divorcing her abusive-to-the-point-of-homicidal husband would be exempt from punishment. She has done nothing 'wrong'. Homeless people are not further punished for sleeping under bridges, the old are not punished for failing to contribute their labour, dissidents are not punished for speaking freely, and mental illness does not algorithmically generate a worsening spiral of stigma and punishment. Being a follower of a disfavoured religious tradition, or being gay or transgender, does not, at a minimum, generate punishments that violate human rights. Republican thought, which Reijers alludes to, would go further to impose an ethos of anti-domination that does not simply prevent the state from punishing people for such personal and social characteristics; it would affirmatively acknowledge them as essential to an empowering and emancipatory conception of citizenship[3]. But let us go with minimum protections.

Second, the SCS's doling out of 'encouragements' and 'constraints' would have to be responsive to some plausible version of the rule of law, meaning that the state is precluded from exercising power arbitrarily. Procedural safeguards and meaningful appeal mechanisms are in place for citizens, and some mechanism for state accountability is operating. For example, laws are administered by courts in which the accused are entitled to procedural fairness and impartial judges give reasons, which can be appealed, based on established methods of justification and by reference to basic human rights and constitutional commitments. This too is necessary if basic human rights are to be protected not only in theory but also in practice.

President Xi's China will not meet the requirements above, and this is not a coincidence. It is difficult to imagine President Xi's SCS *not* discriminating against minority religious groups or political dissidents, or financially irresponsible, and we cannot imagine people being accorded meaningful procedural fairness rights. SCS is part and parcel of this regime's priorities: a state that does not respect individual human rights and does not provide procedural fairness is also the kind of state that would conceive of a panoptic state-based regime for allocating moral deserts across a broad swath of human action, thereby denying individuals scope for autonomy and privacy in their lives. An open state that acknowledges diversity and perhaps even contestation, and respects individuals' ability to decide for themselves what constitutes a good life, is not the kind of state that dreams up SCS.

The State is Not the Market

Mathias Siems is correct in observing that whether SCS will be a 'digital republic' or a 'digital dictatorship' will depend on context, and he is correct in observing that rating systems are proliferating in the West as well. These systems also generate real risks, which we should be wary of[4]. It is also true that, as Jens van't Klooster notes, capitalist societies do not do well in 'giving individuals what they morally deserve'. Defined in that way, the market is clearly not 'fair'. Yet these comparisons miss a crucial point.

We should not equate state-directed surveillance and rating systems with the decentralised, partial, and often commerce-based rating systems being developed in the West. It is not valid to compare the market – a private, fundamentally decentralised system, which for all its considerable flaws no one yet has managed to corner – with a centralised state apparatus, as mechanisms for allocating 'moral deserts'.

[3] See, e.g., Michelman, F. (1988). 'Law's Republic', *Yale LJ,* 97, 1493-1537; Pettit, P. (1997). *Republicanism: A Theory of Freedom and Government.* New York: Oxford University Press.

[4] See, e.g., Baker, T. (2002). 'Risk, Insurance, and the Social Construction of Responsibility'. In: Baker, T. and Simon, J., eds., *Embracing Risk: The Changing Culture of Insurance and Responsibility*. Chicago: University of Chicago Press.

The market is not a mechanism for allocating moral deserts; it is a mechanism for *exchange*. And the market exists in China too, where it is no more 'fair' than it is elsewhere.

Wealth accretes across generations in the West, but it also accretes across generations all over the world and across history, except when war, or revolution, or large-scale tragedy disrupts the pattern[5]. In this sense, SCS is not a 'moral deserts'-based alternative that has much chance of supplanting other structures for allocating power and rewards.

Learning from History

The West has some experience with state regimes, often associated with the provision or withholding of social service benefits, which try to some degree to allocate moral deserts. These have often disproportionately infringed on the dignity and privacy of the poor and marginalised[6], which suggests two things: that across societies, state-based moral deserts mechanisms may be likely to treat non-conforming or marginalised individuals and communities as objects of contempt; and that it is a good thing that the people, who hand out social service benefits while asking invasive questions about a person's personal life, do not have more power and information than they do. But these are small potatoes. Other parts of the world, and China above all, have had more extensive and far more painful experience with virtue testing. The ideological purity tests and corresponding confessions, purges, and mass violence that characterised the Cultural Revolution are hardly two generations behind us.

In our post-Foucaultian, decentralised, financialised world, it may seem old-fashioned to focus on the state as a primary agent of oppression. It is provocative and intellectually interesting to equate Chinese state action with market action under capitalism. But presumably, we can learn something from those mid-20th-century western communists who favourably compared Soviet communism to liberalism well through Stalin's regime, ignoring its brutal reality in favour of intellectual engagement with its possibility under Marxist theory[7]. Engaging with theory while ignoring context is itself a moral choice. We should not be blind to authoritarianism because it is clothed in the rhetoric of civic virtue and progress.

---

[5] Piketty, T. & Goldhammer, A. (2014). *Capital in the Twenty-First Century*. Cambridge: Harvard University Press, pp. 146-55.

[6] See, e.g., Bridges, K.M. (2011). 'Privacy Rights and Public Families'. *Harvard J.L. & Gender*, 34, p.p. 113-74.

[7] Judt, T. (2003). 'The Last Romantic'. *The New York Review of Books*, 20 November 2003.

An Illusion of Western Democracies

Alberto Romele[*]

Wessel Reijers criticises the Chinese Social Credit System for two reasons. Firstly, because it cannot promote virtuous actions. According to the Aristotelian tradition to which Reijers refers, virtuous actions cannot be determined by the pursuit of reputation, money, pleasure, etc. Acting virtuously is not acting according to social conventions, but 'standing out distinguishing oneself'. Secondly, the Social Credit System denies the principle of virtuous governance by creating a master-slave relation between government and citizens that undermines the Greek-western principle of political freedom.

The Illusion of Political Freedom

The thesis I propose is that the reason why the Social Credit System so scandalises westerners is not because it is contrary to 'our' Aristotelian and Arendtian liberal political tradition. Rather, it is precisely because it shows the illusion upon which this tradition is founded. This consists in believing that there is a void at our disposal between people as 'free' citizens and the political as a set of laws.

Imagine that there was a large empty square, the mythical Athenian agora – not the real one, in which there were at least twenty buildings and structures, each with its complex access and behaviour rituals. Imagine further the mythical Athenian citizens – no children, women, or slaves. In this mythical empty square, supported by a few basic rules, virtuous action entirely depends on the single citizen's free will.

The Social Credit System scandalises westerners because it shows that such an ideal type does not exist. On the one hand, no political or social actor acts virtuously, at least if by virtue one means the self-contained search for the eudaimonia. On the other hand, all governments, at least from the 16th century, have tried to fill the gap between the laws of government and the governmental practices by making rules more granular. These are the two points I want to explain further.

The Absence of Disinterestedness

Firstly, I am convinced that there is no disinterested action, neither in political acting nor in other forms of acting and speaking. Human acts are always 'interested', which means that one plays a social game only if she believes that the game is worth playing[1]. Being interested in a specific social game is not the result of cold calculation. On the contrary, it rather implies a belief in its value, and it also means ultimately forgetting the game qua game. According to this perspective, virtuous is a person who best acts according to the implicit rules of a social game he or she is playing.

According to the Aristotelian model of civic virtue, a virtuous person aims to act without being influenced by the pursuit of ambition, money, or reputation. But this same virtuous action is in fact determined by a series of rules that do not arise from the individual herself, but from the definition that a certain tradition has given, which in the western context is primarily the liberal tradition. The act of believing in *this* conception of virtue, and of reiterating the same rules, leads to the creation of a certain habitus, a form of acting, thinking, and desiring common to all the people educated at the school of the Aristotelian and liberal tradition. This allows the participant to this specific game to recognise, appreciate, and judge each other according to the best and the worst. The Social Credit System does nothing different, but it does so by starting from another premise; it privileges, probably for historical and cultural reasons, social conventions rather than the standing out from the mass.

---

[*] Associate professor, ETHICS Lab, Lille Catholic University.

[1] Bourdieu, P. (1998). 'Is a Disinterested Act Possible?' In: Bourdieu, P., *Practical Reason: On the Theory of Action*. Stanford: Stanford University Press, p.p. 75-91.

Incidentally, I think, as Jens van 't Klooster underlines, that western societies are already full of indicators that publicly signal our value as members of a liberal society. I do not mean by that that, by centralising the 'rules of the game', the Social Credit System is fairer. Instead, it is based on dynamics of domination, discrimination, and social exclusion (for instance, in terms of access to basic services such as education, health, and food). Furthermore, the Social Credit System does not show the intention to compensate for injustices – for example, by giving more credit for 'virtuous' actions carried out by those coming from the segments of the populations that are least favoured in economic and cultural terms.

Weak Normativity

As for the government, the Belgian political theorist Thomas Berns affirmed that at least from the 16th-century, two forms of normativity have been pursued[2]. First, the normativity of the laws, which however is often too abstract and detached from concrete citizens' actions. Second, the 'discrete' and 'weak' normativity of numbers and quantification of the population.

The purpose of this second form of normativity is not to sanction crimes, but to correct small infractions and deviant behaviours. '*De minimis non curat lex*', 'the law does not concern itself with trifles', and for this reason, another normativity is needed to deal with such trifles. Berns refers especially to the works of the French political jurist and philosopher Jean Bodin (1530-1596) and to the manner in which he, with other thinkers of that century, rehabilitated a non-Greek, but a Roman form of politics, in particular the figure of the Roman censor. The normative action of the censor derived exclusively from his administrative practice. His decisions did not imply punishments, but forms of downgrading. The effectiveness of his activity mostly lied in provoking 'blushing' – *ruborem* is the Latin word used by Cicero. This form of normativity became more and more important as states developed statistics and other methodologies for the government of their populations.

The Social Credit System is not so different from what Bodin had in mind. It is the same matter of developing and practising a normativity, whose strength is not that of the law, but rather of a mode of quantification that better adheres to the reality and whose force depends precisely on this adherence. It is a more granular kind of normativity; whose main goal is to occupy the empty room for manoeuvring that the law is incapable of occupying. The Social Credit System does not aim (at least, not primarily) to punish individuals but to bring them to the self-correction and the self-control over their habits, behaviours, and actions. Its ultimate scope is to intervene where the law cannot, that is, in the diffused corruption of the mores that afflict contemporary Chinese society.

The techniques and technologies at disposal to the Chinese government are different from those that states had at disposal until a few years ago. It must also be acknowledged that in the Roman tradition to which Bodin referred, the censor was well distinguished from the legislator, while in the case of the Social Credit System, the two regimes of normativity are merged. However, I wonder if this does not correspond precisely to what has already happened for a long time in western democracies. We are scandalised by the Chinese Social Credit System not because it is contrary to 'our' political tradition, but because it brings forward the grey zone about which we usually prefer to delude ourselves.

---

[2] Berns, T. (2009). *Gouverner sans gouverner: une archéologie politique de la statistique*. Paris: PUF.

The Citizen, the Tyrant, and the Tyranny of Patterns

Costica Dumbrava[*]

I agree with Wessel Reijers that social scoring systems limit political freedom and instrumentalise citizenship to impose social control. While technologies have always been used for political ends, the latest technologies relying on big data and complex algorithms offer uniquely powerful and highly effective tools to survey people, quash dissent, and reinforce an authoritarian rule. What is new is a wide appeal of technologies as 'fixes' for pressing social and political issues. Building on their 'success' in commercial sectors (banking and marketing), predictive algorithms and scoring systems are enthusiastically adopted by governmental agencies throughout the world to help making decisions in areas such as criminal justice, welfare, and border control. The Chinese Social Credit scheme is nevertheless unique because of its ambition to aggregate data from a wide variety of sources to provide a set of prescriptive algorithms for 'good citizenship' that is backed by state coercion.

Good citizenship cannot be captured or fixed by an algorithm, because: (1) people genuinely disagree about what good citizenship is; (2) there are limits to how any conception of good citizenship can be enforced in states that uphold the rule-of-law; and (3) even the best scheme of algorithmic citizenship would fail to achieve its objectives due to the inherent weaknesses of applying algorithms to social affairs.

Algorithms vs Citizenship

Firstly, there are presumably as many conceptions of good citizenship as there are citizens. Some focus on rights, others on duties; some emphasise a shared history or a sense of belonging, yet others point to civility, solidarity, sacrifice, etc. Any attempt to design a citizenship algorithm will have to deal with this pluralism, either by settling on a minimalistic and generic version of citizenship (making a rather toothless algorithm) or by prioritising certain views over others (through democratic processes or not).

Secondly, it is one thing to figure out what good citizenship is and another thing to establish what citizens should be required to do, believe, or express following such conception. Helping your neighbour instead of binging on alcohol might be the more celebrated behaviour in many conceptions of citizenship, but, unless explicit laws are proscribing such behaviours, the state should not penalise people for not helping neighbours and/or for alcohol binging. Using state coercion to enforce moral norms and social expectation, as the Chinese scheme seems to do, unwarrantedly transgresses the boundary between legality and morality. The Chinese citizens' scores are built by aggregating and inferring from a wide variety of data (such as online shopping data, use of services, administrative, financial and educational records, social media activity) without a public (open to contestation) explanation about why any of this data is relevant and/or how the inferences work.

Thirdly, assuming that we agree on a normatively acceptable scheme of algorithmic citizenship, the success will depend on its capability to represent and make sense of complex social issues. This criticism concerns algorithms more broadly but might be particularly relevant for the Chinese social scoring scheme given its lack of transparency and accountability. There is growing evidence about widespread bias and discrimination in algorithms applied in social contexts (e.g., policing, sentencing, hiring) [1]. The bias is partly due to the dirtiness of data, a problem that is particularly acute in the case of 'social' data, which is often (self)reported or recorded by people and thus, inevitably, contains errors, inaccuracies, inconsistencies, or simply untruthful information. Yet, a bigger problem with algorithms is that they are

---

[*] External associate, Maastricht Centre for Citizenship, Migration, and Development, Maastricht University.

[1] O'Neil. C. (2017). *Weapons of Math Destruction: How Big Data Increases Inequality and Threatens Democracy*. London: Penguin Random House.

bad at dealing with social and normative issues. While big-data algorithms could be useful in physical science (as in, say, inferring physical attributes of stars using data-driven models[2]), they are problematic when it comes to 'understanding' the social world, deliver social justice, and foster good citizenship. Even if we fix the many problems raised by applying algorithms in social affairs (dirty data, end of privacy, black-box algorithms, unreasonable inferences, etc.), we are still left with what I call the problem of the tyranny of patterns. This problem comes in two parts.

Tyrannies of the Past and the Similar

The first part is *the tyranny of the past*, which means that algorithms are used to assign risks based on past behaviours and social facts. Briefly, this locks individuals into patterned categories that deny them the capacity to redeem themselves and act freely. You may have a crime-stricken from your criminal record, but you may never be able to hide from previously recorded data about your social background, place of birth, and education. The second part of the problem relates to *the tyranny of the similar*. Algorithmic assessments of a person are not only based on data about their past, but also on data about other individuals and their past, who are deemed similar in some statistical way to the person in question. One may, therefore, be treated as a risky or trustworthy individual, and thus be barred or awarded access to certain resources and opportunities, because of who her fellow 'similars' are and how they behave. Algorithmic bias and unfairness are not just technical issues that could be fixed by more and better data or by sharper algorithms; it is a 'mathematical certainty'[3], given that they mirror the real world, which is itself biased and unfair. It is thus highly misguided to entrust algorithms (and people behind them) with the task of fixing citizenship and building more just societies.

---

[2] Folk, D. (2019). How Artificial Intelligence Is Changing Science. *Quantamagazine*, March 2019. Accessed at https://www.quantamagazine.org/how-artificial-intelligence-is-changing-science-20190311

[3] Fry, H. (2019). *Hello World: How to be Human in the Age of the Machine*. London: Penguin Random House.

Putting 'Good Citizens' in 'The Good Place'?

Jiahong Chen[*]

Debates over the Chinese Social Credit System frequently end up with either utopian or dystopian scenarios. While the western media often present the image of an Orwellian totalitarian future[1], the system is sometimes considered an innovative way of social governance[2]. In this contribution, I aim to answer the question as to whether a Social Credit System will be more likely to lead a society to a 'digital republic' or a 'digital dictatorship'. After analysing how the Chinese Social Credit System exhibits an enormous gap between policymaking and policy-execution, I argue that instead of a utopia or dystopia, such a system is more likely to lead us to a future of 'digital bureaucracy'.

List-Based Measures vs Score-Based Measures

Other contributors are aware that, despite the important uniqueness of *the* Chinese Social Credit System, the discussion should be kept more generally on the desirability of *a* Social Credit System. As Siems and Mac Dithigh have pointed out, the Chinese Social Credit System takes three different forms in practice: nationwide blacklists, compliance scores in pilot cities, and social credit scores by financial institutions. Both Reijers and Van't Klooster have chosen a more 'score-based' definition of a Social Credit System. Such a score-based image of a Social Credit System is probably also fuelled by the popularity of the *Black Mirror* episode 'Nosedive', where everyone is assigned a social rating.

Certain practices under the umbrella notion of 'Social Credit System' must be distinguished, as they might have a significantly different impact on individuals and be subject to very different practical or theoretical justifications. For example, China's most oft-cited social credit policy documents – the 2014 Outline[3] and the 2016 Opinion[4] – do not include measures involving a 'score' or 'rating'. Rather, both documents are overwhelmingly dominated by list-based measures, i.e., blacklists for trust breakers, and 'red'-lists for trust keepers. This is also the case for Shanghai's Social Credit Ordinance 2017.

The prevalence of list-based approaches results largely from the fact it was the Supreme Court who introduced some of the first initiatives to deal with disobedience of court decisions. In the absence of individual insolvency law (and contempt of court law) in China, it has always been a challenge for courts to enforce judgments in civil cases effectively. In response, the Supreme Court decided to publish a list of those who wilfully refuse to obey court decisions (as a reputational punishment) and to restrict their access to certain services (akin to bankruptcy restrictions), which turns out to be rather effective. This partly explains why most of these initiatives did not meet much resistance – or were even welcome by the public[5] – in the first place.

This is not to say, however, that these measures are intrinsically desirable, or that they are essentially different from the score-based measures. Yet, the relatively higher acceptability of these measures mirrors certain important features among different approaches to a Social Credit System. A blacklist released by courts may arguably be less worrisome in that it is subject to judicial control, applies only

---

[*] Research Fellow, Faculty of Science, University of Nottingham.

[1] https://nationalpost.com/news/china-bars-millions-from-travel-for-social-credit-offences

[2] https://www.washingtonpost.com/news/theworldpost/wp/2018/11/29/social-credit

[3] Legal Reform Agenda of the Fourth Plenum (2014), Translated version accessed at https://chinacopyrightandmedia.wordpress.com/2014/10/28/ccp-central-committee-decision-concerning-some-major-questions-in-comprehensively-moving-governing-the-country-according-to-the-law-forward/ on 16-11-2018.

[4] Guiding Opinion of the State Council on Establishing and Improving the System of Joint Incentive for Keeping Trust and Joint Punishment for Breaking Trust (2016).

[5] https://www.bloomberg.com/opinion/articles/2019-01-24/why-china-s-social-credit-systems-are-surprisingly-popular

to serious and wilful disobedience, and does not involve opaque or arbitrary algorithms. These claims are open to challenges, but the point is that certain practices falling within the scope of a Social Credit System are not necessarily unjustifiable, or at least not far from mechanisms that are commonly accepted across various jurisdictions[6]. The less discussed downside of a Social Credit System is its natural propensity to be exploited for bureaucratic purposes.

Policy on Paper vs Policy in Reality

One of the reasons why a Social Credit System is susceptible to bureaucratic abuses lies in its potential to be repurposed, which is what has happened in China. While the two key policy documents have not mandated score-based measures, such approaches are embraced by local authorities. In fact, it is well-documented that a large number of Chinese cities are already experimenting with a variety of schemes whereby local residents are treated differently based on their social credit scores[7]. These schemes concern matters ranging from the right to relocate to a megacity to prioritised access to public services. Points can be gained or lost through a wide spectrum of behaviour, including volunteering or blood donation, something certainly not covered by the 2014 Outline or the 2016 Opinion, but implemented in the name of 'contributing to the construction of the Social Credit System'.

Therefore, a crucial point to raise is that, regardless of the arguably good intention and reasonable design of the overarching system, the mere fact that local or sectoral enforcers may easily distort such schemes will on its own warrant serious precaution. The 'function creep' of the Chinese Social Credit System is a case in point: the list-based approach targeting punishable behaviour as advanced by the central policies has ended up as a score-based, local implementation that encourages 'desirable' behaviour.

One might argue such abuse has less to do with the system itself, but more with who controls the system and what constraints are put in place – the 'regulatory environment' in Siems and Mac Sitigh's words. Indeed, under a totalitarian regime, a Social Credit System will be readily available for suppressing dissidents. However, even in a democratic society, there remains a high risk of such a system being repurposed for a fully legitimate policy goal. The criteria of what should be rewarded and what should be punished under a Social Credit System may be exploited to serve a new policy objective that is not necessarily related to the 'social creditworthiness' of an individual. For example, a local government may decide to encourage the consumption of local products by incentivising such purchase with credit points. While it is debatable whether such practice is justifiable, it shows how a Social Credit System, irrespective of the government it serves, is bound to create a strong motivation for bureaucrats to expand the scope of the system to areas without democratic debates or scrutiny.

Fairness vs Efficiency

Even if the Social Credit Systems are tightly restricted to certain areas, they would nevertheless be subject to a more fundamental criticism in relation to the theoretical and practical limitations of rational control over individual behaviour. Along with the rise of modern bureaucracy, this challenge arises from the long-standing debate in political philosophy over the extent to which collective efficiency should be prioritised over individuality. The ratings of individuals may be employed as a 'convenient' or 'efficient' way of allocating scarce resources by the government. In cities where social housing is in shortage, for example, local councils may decide to give priority to those with higher scores as it would be easier for them than assessing the degree of individual needs on a case-by-case basis. While this is still better than

---

[6] Such as the U.S.'s 'no-fly lists'. For more examples, see the contribution by Siems and Mac Sitigh.

[7] https://foreignpolicy.com/2018/04/03/life-inside-chinas-social-credit-laboratory/

price mechanisms (Van't Klooster), it is not necessarily fairer than the 'old-fashioned' model that involves qualitative human judgment.

The functioning of a Social Credit System carries characteristics of a bureaucratic regime: (1) rationalisation and optimisation of efficiency (quantitative evaluation and influence of behaviour of citizen); (2) specialisation and division of labour (credit recorders, calculators, enforcers and auditors from different sectors); (3) clear formal rules and limited discretion (criteria of 'positive' and 'negative' behaviour); and (4) impersonality (insensitivity to individual circumstances). The term 'bureaucracy' is used in a neutral sense and does not necessarily carry any desirable or undesirable connotations. That said, bureaucracy enabled by a Social Credit System would likely have some unintended consequences, such as the scenario of 'computer says no', where citizens are denied services or resources when personal situations are not taken into account. The mundanity of the activities covered by such a system and the potential diffusion of responsibilities would also subject a Social Credit System to further criticisms.

The next question would be, are these characteristics essential to the concept of a Social Credit System, or are they simply the results of the (ab)uses of such a system? While I agree with other contributors that the actual effectiveness and harm of a Social Credit System would depend on the regulatory environment (Siems and Mac Sithigh), a Social Credit System, as a social engineering project, brings forth changes in technological configurations and power dynamics by altering the regulability[8] of individual behaviour. Such potentials are independent of the conditions of the regulatory environment and should not be dismissed. A Social Credit System may serve a totalitarian regime or a democratic one, but either way, it will bring existing bureaucracy to another level.

Conclusion: The Good Place and the Future of Bureaucracy

A Social Credit System faces two primary challenges regardless of the regulatory environment: first, it is prone to unchecked expansion of scope by bureaucrats; second, it strengthens the bureaucratic propensity pre-existing in a society. In answering the question of the debate – whether Social Credit Systems will lead to a 'digital republic' or a 'digital dictatorship' – one would have first to answer the question whether an ultimate bureaucratic society could be considered a utopia or a dystopia. To me, it could go either way, depending on the role of bureaucracy in that society. For a hypothetical society sitting perfectly between a republic and a dictatorship, a Social Credit System is most likely to reinforce bureaucracy – amplifying its merits, and also its flaws. The characteristics of a Social Credit Systems raised above – the tendency to be transferred for an incompatible policy objective or simply employed for administrative efficiency – are unlikely to drive the society to a utopian future but, at the same time, they will not necessarily lead the society closer to a totalitarian future, either.

While the idea of the Orwellian 'Big Brother' certainly captures dangers of a Social Credit System, a better metaphor seems to be the world of *The Good Place*, where one's afterlife will end up in 'The Good Place' or 'The Bad Place' based on the accumulated 'absolute moral worth' of their every single action, and the value of an 'undefined action' to be decided by a relevant department. Rather than a digital republic or a digital dictatorship, a Social Credit System is more likely to lead to a 'digital bureaucracy', a future where social credits are produced and consumed as a way of social indifference (rather than social suppression) imposed on individualism. On the macro-level, Social Credit Systems might be intended to make 'good citizens' (Reijers) – whether for the right or wrong reasons – but, on the micro-level, they might be executed to make it easier for bureaucrats to make mundane decisions of who should go to 'The Good Place', and who should go to 'The Bad Place'.

---

[8] For further discussion of the notion of 'regulability', see Lessig, L. (1998). 'The laws of cyberspace.' *Readings in Cyberethics*, 134, p. 136.

The Social Credit System as a New Regulatory Approach:
From 'Code-Based' to 'Market-Based' Regulation

Primavera de Filippi[*]

In 2014, China's State Council issued a notice for the establishment of a nationwide Social Credit System (SCS) by 2020. Thus far, 43 municipalities have been experimenting with a local implementation of the SCS – each with their own distinctive characteristics. Despite the different approaches adopted by various municipalities, the common denominator is that all pilots are designed to promote trustworthy behaviour by creating incentives to comply with social norms and economic obligations. Traditional institutions and social structures have been unable to catch up with the rapid pace of China's economic development over past years. This led to the establishment of an advanced market economy, characterised by an underdeveloped institutional and social framework, with local markets filled with counterfeited goods, scams and fraudulent activities, as individuals are becoming more individualistic and self-serving. As a result, the level of interpersonal trust in China has radically dropped, creating widespread discontent among citizens. The SCS is an attempt by the government to restore trust in people and society, by leveraging digital tools in order to govern society better and more effectively.

By collecting data about individual activities, online and offline, China relies on big data analytics to profile citizens into specific categories, depending on their attributes and past actions. Every municipal pilot has identified a series of desirable or undesirable actions, each associated with a particular payoff structure (either positive or negative). Desirable actions are rewarded with a high social credit score, resulting in a particular privilege or compensation, whereas undesired actions are punished, via blacklists or lower social score, resulting in a loss of privileges or economic sanctions. As such, the SCS can be regarded as an alternative form of regulation, which is not driven by law, but rather by a set of incentivization mechanisms powered by market dynamics and social norms.

In 1999, Harvard professor Lawrence Lessig identified four different forces that can be used to regulate individuals: law, social norms, markets, and architecture (i.e., technical infrastructure or code). Lessig underlined the fact that we cannot focus solely on legal rules to understand the forces that govern or influence individual behaviours; we must take instead a more systematic approach that properly accounts for the relative influence of each of these different levers. Against this background, the SCS presents a shift away from the law towards a new approach to regulation, which essentially consists of code (responsible for collecting data and translating it into a social score) combined with market dynamics (i.e., a series of economic incentives or disincentives to act in a particular manner) and social norms (as a result of status, privileges, and social pressure to increase one's social score).

On the Democratic Legitimacy of the SCS

To what extent does the SCS comply with the fundamental principles of democratic legal systems and human rights values?

First, because it operates independently of (or as a complement to) the legal system, the SCS can cover activities that are generally left outside of the legal scope. As argued by van 't Klooster, the SCS can promote behaviours that are not mandatory, but nonetheless desirable, while condemning activities that are not illegal *per se* and yet considered potentially harmful. For instance, in China, some of the SCS pilots are designed to punish citizens by reducing their credit score not only as a result of legal offences and misdemeanours, but also in response to lawful activities such as purchasing too much junk food or video games, which are not forbidden by law. Similarly, with regard to the incentivization

---

[*] Permanent Researcher, National Center for Scientific Research.

mechanisms, while it would be difficult to legally require citizens to perform certain types of activities (such as buying Chinese products or visiting elders more frequently[1]), the SCS incentivizes them in this direction by rewarding citizens with a higher score.

As opposed to traditional legal rules, the SCS provides greater granularity on the types of behaviours that can be incentivized or disincentivized, 'occupying the empty room for manoeuvring that the law is incapable to occupy', as Romele has shown. By stipulating different types of rewards or punishments on activities that are not strictly speaking regulated under the law, the SCS operates where the law is unable to intervene – introducing a new type of normativity, based on metrics, quantification and statistical analysis[2] (Berns, 2009) that has progressively infiltrated the legal and political system, rewarding or condemning individual behaviours independently of whether they are legal or illegal, or even right or wrong. The SCS, as underlined by Dumbrava, thus 'unwarrantedly transgresses the boundary between *legality* and *morality*'.

Second, because it does not rely on existing laws and regulation, the SCS is also not subject to the same type of constraints as the legal system. Traditionally, in most western or democratic societies, any legal system is subject to a series of substantive and procedural constraints, including the need for generality, publicity and predictability of rules, as well as the requirements of equitable judgement, due process, separation of powers, or the respect of private property and individual liberty[3]. Hence, China's SCS stands in contrast with the democratic concept of the *rule of law* – geared towards the separation of law and politics - and rather implements a *rule by law* – whereby law is used as a tool to command and discipline people, as it is generally done in authoritarian regimes. True, the SCS may theoretically comply with the rule of law. Yet, as highlighted by Ford, its current implementation in the various regions of China is more likely to run contrary to the basic premises of the rule of law and may even violate fundamental human rights values.

One important issue with these systems is that citizens can be punished multiple times for the same wrong because a lower social score could impact many different facets of their lives. For instance, individuals with lower social credit scores might simultaneously lose the privilege to obtain a visa, purchase property and secure a mortgage or a loan; they might be prevented from buying tickets for high-speed trains or air travel, booking hotel rooms, or even from enrolling their children in a good private school – all at the same time[4]. This aspect is often presented as a feature of the SCS, with the oft-touted slogan that '*whoever violates the rules somewhere shall be restricted everywhere*'. Yet, this aspect goes counter to one of the basic principles of the rule of law, according to which *no individual should be punished twice for the same offence*.

Moreover, because the consequences of an individual's social credit score are automatically defined and directly enforceable within the information system of the relevant authorities (e.g., a public administration, but also private actors like banks, transportation companies, or schools) with whom that individual interacts, the rules enshrined in the SCS are directly applicable without the need for any third-party enforcement authority. And because there is no specific procedure for appeals, individuals who believe in having been unfairly punished or wrongly blacklisted by the system are left with little to no recourse. This runs counter to several rule-of-law principles, including 'separation of power', 'due process' and 'presumption of innocence' – all designed to limit potential abuses of power by one governing authority over its citizens.

---

[1] https://www.dailymail.co.uk/sciencetech/article-6183893/Real-life-Black-Mirror-Chinas-moves-introduce-scorecards-ranks-behaviour-citizens.html

[2] Berns, T. (2009). *Gouverner sans gouverner: une archéologie politique de la statistique*. Paris: Presses universitaires de France.

[3] Waldron, J. (2011). Are Sovereigns Entitled to the Benefit of the International Rule of Law? *European Journal of International Law*, *22*(2), p.p. 315-343.

[4] https://www.thenation.com/article/china-social-credit-system/

On the Efficiency of the SCS

With the stated purpose of instilling societal trust, by incentivizing citizens to act virtuously while penalizing them for socially undesirable behaviour, the SCS is perhaps one of the most ambitious attempts by a government at producing ideal citizens. To what extent can the SCS fulfil these underlying objectives, and what is the overall efficiency of the system?

One reason why the SCS may be deemed to produce better citizens than a capitalist market economy is that, as shown by van 't Klooster, markets often fail to properly reward good citizens who work all their lives diligently and follow the rules; but rather reward those who accumulate scarce resources and use these resources to accumulate more capital. The market system is geared towards productive and allocative efficiency and does not optimize towards morality or values. The SCS is thus presented as a better alternative to the market economy because it rewards citizens based on their social and political behaviours.

Yet, to the extent that it is possible to accumulate social credit, it is difficult to force people to always act virtuously (as Reijers puts it) provided that they have a sufficiently high credit score that they can afford acting in a harmful manner. Hence, while the SCS might require some people to act virtuously in some cases, it does not necessarily 'make' them virtuous. On the contrary, one could fairly argue that the SCS might introduce a new type of transactionality in the realm of social and political activities, whereby an individual may feel legitimised to act wrongly, provided that he or she has performed a sufficient number of good deeds in the past, so as to discount for the lower score.

Moreover, in addition to encourage or discourage specific behaviours by economic incentives, the SCS also rely on social pressure to regulate individual behaviour. In China, there have been rumours that some of the SCS pilots have implemented a system whereby people may see their social credit score drop by the mere fact of interacting with the wrong person, online or offline[5]. This is likely to lead to an outright marginalization of all those who do not act according to the 'social canon'. The reverse is also possible. While some implementations of the SCS have been using individuals' social network activities as a way of potentially reducing their score, other implementations – such as Alibaba's Zhima (Sesame) Credit – have been designed to increase the social credit score of people interacting with other high-scored individuals[6]. These systems bring the risk of replicating the capitalistic dynamics of the market economy, whereby people who accumulate social credit could leverage it to obtain even more social credit. This could pave the way for the establishment of a new society of 'elites' in which a small number of high-ranked rule-abiding individuals would only interact with other high-ranked rule-abiding individuals.

Another interesting point relates to how the SCS might ultimately legitimize a national surveillance system. In order to properly function, any SCS will require the collection of a vast amount of information about citizens, which might raise important privacy concerns. Yet, most Chinese citizens are drawn to the SCS[7] because it allows them to show to the government (and others) that they are 'good citizens', in addition to the economic or non-economic privileges they can gain. Because of these benefits, people are keen to participate to the SCS – often without questioning the source or legitimacy of the rules it enshrines – and thus willingly disclose information about their daily activities to private or governmental actors even if that comes at the expense of privacy. Even people who might be reluctant to give up their privacy are left with little to no choice. Indeed, individuals who try to mask their activities in order to protect their privacy will ultimately be at a disadvantage, since they will not be rewarded for being good citizens. Perhaps most importantly, the mere fact of not disclosing information, rather than being regarded as a legitimate choice motivated by a desire to preserve one's privacy, could instead qualify as

---

[5]   https://observer.com/2018/02/america-isnt-far-off-from-chinas-social-credit-score/

[6]   https://news.cgtn.com/news/3d3d774e7751444f32457a6333566d54/index.html

[7]   Adam Minter, writing for Bloomberg, stated that in a 2018 survey 80 per cent of respondents either somewhat or strongly approve the credit system, although it is admittedly difficult to gather the true sentiments of Chinese individuals.

an indicator that a person is, in fact, not a good citizen. The metrics results thus inverted: 'I'm a good citizen, because I have a good credit score' has replaced the previous 'I have a good credit score, because I am a good citizen'.

Finally, as described by Reijers, the SCS does not contribute to creating more virtuous citizens, because it does not allow them to *choose* to act virtuously, but rather prescribes them how to do so. This leads to a rationalisation and actual objectivization of what it means to be a good citizen. People no longer need to exercise their own judgement as to how to act; they just need to follow the rules stipulated for them – thereby preventing them from distinguishing themselves in the public sphere[8]. As a result, people no longer have the opportunity to think for themselves about what is right or wrong. Rather than taking the risk of acting in a way that the SCS will consider to be an undesirable action, they can simply delegate part of their morality to the system.

On the Terror of Metrics

An additional problem relates to the difference between *laws* and *metrics*. Laws are general rules that stipulate what *shall* or *shall not* be done. Given the inherent ambiguity and flexibility of natural language, laws need to be appreciated by a judge in light of specific facts before they can be applied to the particular case at hand. Instead, the SCS relies on a set of strict and specific rules (i.e., codified metrics) that can be automatically processed by a machine. Because they do not account for specific facts, metrics are mostly concerned with the ends, independently of the means used to achieve these ends.

With no opportunity for human appreciation, the SCS merely codify rules into a particular set of metrics, with no account for the underlying values that motivated the establishment of these rules. Citizens' most rational behaviour would, therefore, be to focus only on improving the metrics, rather than acting in line with what these metrics were originally intended to be an indicator of. While it might be true that – as famously said by Peter Drucker – 'what you cannot measure, you cannot improve', one should also not forget that the mere establishment of a particular metrics-based system might introduce a significant bias in behaviour, so that 'only what you can measure will improve'. Citizens might thus be tempted to only engage in activities that will boost their social credit, regardless of the virtue or morality of these actions. If the metrics are poorly defined, individuals might even be rewarded for doing something immoral (e.g., lying to their peers – an action currently bearing no punishment in terms of social scoring) to improve their chances of being selected for community services or other rewarded activity. Ultimately, because the system can only account for a limited amount of information, and data fed into the system is necessarily incomplete, the SCS might generate unfair outcomes by lack of contextual information on the *why* and the *how* of a citizen's individual behaviour. As any metrics-based system, the ends will always justify the means.

.

---

[8] Arendt, H. (1958). *The Human Condition*. Chicago: University of Chicago Press

Is a Social Credit System Good for Women?

Mirjam Müller[*]

In 2015 a group of feminists created the hashtag *#GiveYourMoneyToWomen*. The idea was to draw attention to the time and energy that women spend on providing uncompensated emotional care for men. As such, *#GiveYourMoneyToWomen* critiqued how goods, which are typically associated with women, such as compassion, empathy or moral attention, are often taken for granted and not adequately valued. Gender-stereotypical assumptions about women's natural disposition and desire to care hide the time and energy that caregiving typically involves. In a capitalist economy, the value of goods tends to be tied to their exchange value. By turning private caregiving into a marketable good, this feminist intervention aimed to make it visible, as something that needs to be valued.

However, there are reasons to be sceptical about turning caregiving into a marketable good. Imagine that for every five minutes of listening to a friend's worry, we would start cashing in. Rather than integrating caregiving into the price-mechanism, wouldn't it be more attractive to find different ways in which it could be valued? It is at this point that a Social Credit System becomes interesting. In this blogpost, I explore potential implications of a Social Credit System for gender equality. The claim is this: a Social Credit System is, in principle, able to integrate a wider set of behaviours and characteristics that merit reward than the price mechanism. It could hence turn out to be better at valuing feminine-coded tasks, such as care-work. Yet, I argue, feminists should be sceptical with regards to the emancipatory potential of a Social Credit System, as such a system might turn out to reproduce dominant forms of valuing rather than promoting real change.

Social Credit Systems: How Do They Work?

Following other contributors in the blog (e.g., Reijers or van 't Klooster), I understand a *Social Credit System (SCS)* as a centralised system of governance that uses surveillance techniques of the state and/or private entities to (1) trace multiple observable behaviours of citizens, (2) assign scores based on these behaviours, and (3) reward individuals with high ratings and punishes those with low ratings. As other contributors to this debate have pointed out, a SCS does not merely punish or reward citizens based on the lawfulness of their actions. Instead, it also evaluates the morality or social desirability of their actions. For instance, a person's score can go down as a result of parking in a non-parking area, getting a divorce, or failing to repay a loan. Low credit scores can lead to sanctions ranging from being blocked from a dating website to being banned from long-distance flights. As such, the SCS tracks a broad range of actions in order to incentivise socially desirable behaviours and character traits in citizens. The exact way in which a SCS is designed and implemented is highly context dependent. Yet, I want to abstract from any specific SCS and instead focus on the potential of such a system to contribute to gender equality.

Emancipatory Potential of a Social Credit System?

In his contribution to the debate, Jens van 't Klooster has argued that a SCS is potentially better at realising distributive justice than a price mechanism. As a SCS is sensitive to a wider range of actions and behaviours that merit reward, using social credit ratings to distribute benefits and burdens might be fairer than relying on the price mechanism. If the SCS does fare better in rewarding a wider range of behaviours and personal traits than a price mechanism does, this could be good news for feminists: tasks that are distributed along gendered lines could be valued differently if they are rewarded in the social credit scheme. For instance, in one of the local social credit schemes in China, purchasing diapers for a

---

[*] Postdoctoral Fellow, Justitia Amplificata, Free University Berlin.

child increases the social credit score[1]. Consequently, there is a potential that feminine-coded tasks that often go unnoticed, such as child-care, will become more visible and differently valued as a result. This, in turn, could have significant effects on how stereotypically female tasks are valued and understood more broadly.

There are, however, reasons to be sceptical.

Why a Social Credit System is Not (Necessarily) Good for Women

While a SCS can assign value to a wider range of behaviours and personal traits than the price-mechanism, it is unclear whether this, in and of itself, will promote feminist values. Following Elizabeth Anderson, what and how people value is tied to the social relations within which they stand[2]. Whether or not we ascribe value to certain goods is not independent of how these goods are produced, distributed, and used. For instance, the gendered division of labour according to which women bear primary child-care responsibilities entails specific modes of valuation, such as thinking about child-care as a 'labour of love'.

In a patriarchal society, what it means to be a 'good citizen' and, consequently, which behaviours and character traits are rewarded will be guided by dominant norms about gender and the institutions upholding these norms. If we think about values in this way, it becomes questionable whether, and to what extent, a SCS will significantly change our ways of valuing and increase the value of behaviours that so far fall outside the price mechanism. Rewarding tasks that often go unnoticed, such as buying diapers, might have effects on how people value more broadly. Yet, the extent to which a SCS will be able to change valuing practices to increase gender equality will depend on (at least) two conditions.

First, whether or not a SCS may affect valuing practices more broadly will depend on how open existing social practices are with regards to integrating new values and revising existing ones. Think again about the practice of caregiving. Assigning a social credit score to feminine-coded goods, such as emotional support or empathy, does not necessarily lead to a higher valuation of these goods or a change in the practice of caregiving more generally. For instance, taking care of a relative might increase someone's credit score. Yet, if nothing changes with regards to the division of labour, and hence if women continue to be the primary caretakers, the way in which these tasks are valued might merely take a different form. Visibility does not automatically lead to appreciation if there is no change in the underlying practices. Consequently, if human values are dependent on the social practices within which they are embedded, a change in values requires a change in social practice.

Second, it matters crucially how the catalogue of behaviours and character traits that merit reward/punishment is created and, especially, who gets a say in what it entails. In that sense, it will matter, who can give input with regards to which behaviours constitute a 'good citizen'. Suppose only certain groups of people, e.g., those that are already dominant in a given society, such as white men, are asked what they consider to be desirable behaviours or character traits. In that case, the catalogue might reproduce dominant forms of valuing. Contrary to that, if a wider group of people in different social positions gets a say in what should or should not go on the list, chances are that marginalised experiences will become more visible and, as a result, might be valued differently. Consequently, the extent to which a SCS can transform the way of valuing in a given society will also depend on how different voices are integrated into the process of determining which behaviours or character traits are socially desirable.

---

[1] https://www.forbes.com/sites/bernardmarr/2019/01/21/chinese-social-credit-score-utopian-big-data-bliss-or-black-mirror-on-steroids/#53e1fe0d48b8.

[2] Anderson, E. (1993). *Value in Ethics and Economics*. Cambridge, Mass; London, England: Harvard University Press, p. 12.

In sum, whether a SCS can promote feminist valuing practices will depend on how rigid the social practices within which it is embedded are and, importantly, who gets a say in determining which values matter. Therefore, feminists should not put too much hope on a SCS to change how feminine-coded tasks are valued. Instead, they should continue to fight for changing the social practices within which values are embedded. If a Social Credit System can be designed and implemented in a way that it provides resources for that purpose, even better.

Data Diets and Democracy: The Chinese Social Credit System From a Machine Learning Perspective

Joshua Fairfield[*]

Introduction

The current debate asks whether the set of technologies that make up the Chinese Social Credit System (SCS) trends towards a digital dictatorship or a digital republic. I make two points. First, Chinese AI may well become stronger than western AI because of differences in the data available to train machine learning algorithms. Second, the mechanisms of machine learning may tend towards a reversal of the long-standing advantages of information decentralisation in capitalist economies, with negative knock-on effects for liberal democracies worldwide.

My conclusions: the SCS trends against democracy. It is being built by a competent and motivated anti-democratic system with social control as one stated goal. The more important question though is whether the Chinese machine learning data diet will make Chinese AI stronger than western AI, and whether the realities of machine learning will undermine western-style capitalism and liberal democracy. As this essay argues, I think there is a real chance that both will occur.

These conclusions place the current piece in both agreement with and at odds with the framing arguments for this debate. Both Reijers and van 't Klooster move from talking about *the* SCS to a discussion of *a* SCS. That move is correct in that the surface question of whether the Chinese version of these systems will be turned to anti-democratic ends is not particularly interesting. But machine learning algorithms cannot be divorced from the data they study - data produced by citizens in a given context. Thus, I place the machine learning algorithms back into context by talking about the role that citizen-produced data plays in training machine learning algorithms. For three reasons, discussed below, Chinese machine learning algorithms may benefit from more and more varied data than do western versions and this, in turn, may impact the underpinnings of capitalism and democracy worldwide.

The Chinese Data Diet

The key to machine learning is data—its volume, completeness, and variety. Recent advances in algorithm design have been merely (or mostly) advances in the amount, variety, and quality of the data that is able to feed machine learning algorithms. To the extent that the SCS represents something different from the emerging western system of artificial-intelligence-driven pricing (of credit as of any other commodity or service), the differences will need to lie in the data channelled into these systems, and the differences that the AI will learn on a different data diet.

I then leave the reader with some musing about what it would mean if the SCS and similar systems generate superior machine learning results because of improved access to datasets, amount of data, and the new on-ramps created by the physical systems of the SCS (neighbour reporting and the like) [1]. In particular, I suggest a dangerous potential overturning of a longstanding principle of capitalist development. Capitalism in its late-20th-century idiom has succeeded largely because decentralised information processing is far more efficient than is centralised data processing. The 'Invisible Hand' of the market is a metaphor for the coordinating actions of decentralised economic decisionmakers, who organise production around price points, for example. The question is whether the strong synergies of

---

[*] Professor, School of Law, Washington and Lee University.

[1] http://time.com/collection/davos-2019/5502592/china-social-credit-score/ (2019).

centralisation inherent in machine learning will undermine this economic consensus, and thus undermine liberal democracies in general.

Machine learning and related technologies consist of a range of techniques, the full scope of which is outside the reach of this article. One shift, however, is worth mentioning. Many of the algorithms we rely upon have existed with minor updates for several decades. What changed in the early 21st century is that the raw volume of data available to train machine learning algorithms exploded.

Data matters because, for at least many of the most common and most effective machine learning programs, the machine does not understand anything. Rather, it finds connections in data by looking at a lot of it. Machine learning algorithms rely on correlations in the data that are relevant to the prediction the algorithm is trying to make. The more data a machine learning algorithm can ingest, the more correlations it can find. If one's concern is pure accuracy and not explainability, the weights used by a machine learning program can be quite extensive: perhaps my political reliability or creditworthiness correlates with my social media contacts, recent purchases, books read, social media commentary, outcomes of judicial processes, and so on. Each correlation may yield information gain, which will help in making a better selection.

The difference in the machine learning algorithms created for advertisement, election influence, and creditworthiness in the West and those that result from the SCS - *if* there are important and salient differences - can be usefully explained in terms of the differences in data diet that the SCS system affords. This point has gone underexplored in the debate. The following parts discuss how the SCS might differ from the western machine learning context in terms of access to data, amount of data, and physical on-ramping.

Access to Data

Data in the United States and Europe is often kept separate from other data pools[2]. This is usually for reasons more of competition than privacy. Facebook's knowledge of its users is its goldmine; likewise, Google's tracking of users as they interact with websites across the internet. A machine learning algorithm that had access to the combined Facebook and Google datasets would be able to extract out features and would undoubtedly be more accurate than one based on either browser-tracking or social media interactions alone. Unintentionally, this offers some protection against machine learning algorithms. The question is whether the Chinese political structure offers avenues for combining datasets that western countries do not.

First, the Chinese government has a stronger hand in the information industry than do western governments[3]. This government whip-hand sometimes merely increases raw government power, for example, in the form of country-wide censorship mandated by the government and carried out by industry[4]. Second, the Chinese government's overt encouragement of the SCS overlaps with its public commitment and large investment in artificial intelligence[5]. In addition, powerful consortia vie for the political and commercial capital at play. My view is that there is ample cause to believe that these political considerations may cause Chinese AI to benefit from larger and more varied datasets than do those trained by individual U.S. companies, no matter how large.

---

[2]   https://www.sas.com/content/dam/SAS/en_us/doc/whitepaper1/5-essential-components-of-data-strategy-108109.pdf

[3]   https://www.worldpoliticsreview.com/articles/23945/china-forces-big-tech-to-make-a-choice-play-by-beijing-s-rules-or-be-left-out

[4]   https://www.nytimes.com/topic/destination/internet-censorship-in-china.

[5]   https://www.cnas.org/publications/reports/understanding-chinas-ai-strategy.

Amount of Data

The key question is whether the Chinese legal, cultural, and technological frameworks permit access to more, different, and more granular data. More data is easy to imagine but hard to confirm. The population numbers are higher: some estimates put the number of mobile internet users in China at more than double the entire United States population[6]. That is a lot of smartphones recording a lot of data, which can be used for targeted behavioural advertising, political verification, and so forth.

However, overlapping data is often less useful than different datasets. Data about you as you surf the web is even more helpful when combined with data from your smartphone geolocator, or your smart television's parsing of your conversations and viewing habits. This is called sensor fusion: combining multiple datasets often results in information beyond their sum. Sensor fusion has a greater impact in deeply gadgeted societies. I am aware of no study on comparative western-Chinese sensor fusion. Such a study would prove deeply informative on the relative strengths and weaknesses of Chinese machine learning.

Municipal On-Ramping

The last element is whether the Chinese government support, especially in the form of model city initiatives and the structure of bureaucratic advancement within the Chinese Communist Party, has a strong impact on building out physical sensors, like smart cities or municipal sensors, that serve to on-ramp citizens data and make it available for machine learning.

I am not certain that government power and different legal and social norms surrounding privacy make much of a difference for physical on-ramping, at least when one compares China to the United States and the U.K. In those countries, the temptation to circumvent constitutional restrictions on government mass surveillance by routing the data collection through third-party firms has proven irresistible. In continental Europe, norms surrounding personal privacy in public spaces take on a different aspect, and we might expect AI trained on sparser European data to be at some disadvantage.

Physical on-ramping, in the form of municipal cameras, pressure sensors, and other so-called 'smart city' technology, seems to be developing similar amounts of coverage in China and the United States. The mechanisms may be slightly more commercially driven in the U.S., and the government plays more of a recipient and beneficiary of commercial datasets, which it turns to policing and intelligence purposes, but I have no sense that the amount and flow of data across physical sensor networks is less.

My understanding is that often party advancement is tied to the furtherance of party buzz words and goals. To the extent Social Credit is a party buzz word, we may expect municipal and regional officials to make efforts to develop facilities for the SCS in their areas of expertise and authority. If, then, party members who are in positions of municipal authority believe that physical on-ramping of municipal sensor data is valuable to their political careers - as I think the evidence indicates is the case at least at present - then we might expect data on-ramping through municipal sensors to strongly exceed at least European rates.

Concluding Thoughts

I have raised these points because they are salient to the two fundamental questions posed by the SCS. Those questions are political and economic (if one sees those as separate spheres). The political element, which has attracted the most attention, is the extent to which the SCS will be profoundly anti-democratic. Since the SCS was designed by and for a non-democratic political system, this is a bit of a

---

[6] https://www.forbes.com/sites/niallmccarthy/2018/08/23/china-now-boasts-more-than-800-million-internet-users-and-98-of-them-are-mobile-infographic/.

non-question. Of course, the system is being used to enforce political hygiene, and there is already more than enough evidence of these technologies being used to monitor and oppress.

The follow-on problem is harder. AI in general, and machine learning as its most currently successful instantiation, may destabilise the current consensus in economics that decentralised information processes in markets, coordinated primarily by price points, are the best way to run an economic system. It is not inconceivable that machine learning may change the playing field of capitalism. I do not suggest that centralised AI will be able to guide a national economy better than will some form of combined machine learning and decentralised firm-based decision-making, but I do suggest that the balance of those contributing elements may shift decidedly.

Scoring Systems: Levels of Abstraction

Francesca Lagioia[*] and Giovanni Sartor[**]

One of the contested issues in this debate relates to similarities and differences between scoring systems in China and in the West – how unique is China? In this post, we will try to reconcile the different perspectives, arguing that both commonalities and differences exist, depending on the adopted level of abstraction. Thus, we shall zoom in the Chines Social Credit System (SCS), examining the features it shares with other systems, and point to related issues: it is a scoring system, it is formal, it is ICT based, it is surveillance-based, it is opaque and unaccountable. This enables us to distinguish commonalities and differences.

Commonalities between the Chinese SCS and Other Systems

First, the Chinese SCS shares with other social *rating systems* the feature of assessing people based on their skills, character, motivation, etc. Rating is, in fact, a common feature of social contexts in which it contributes to people's reputation. For instance, in commercial relations, merchants are rated based on their sales-volume, quality of service and reliability, which determine their reputation – a key asset for a successful business. Workers are rated based on their performance, attitude, and motivation; their ratings determine employment opportunities. University teachers are rated based on student's assessment, publications, and papers' citations.

Second, China's SCS shares with other rating systems the fact of being *institutionalised and formalised*: methods are defined for recording data and determining rankings, which have to be expressed in predefined formats. Some formal scoring systems also share with the Chinese SCS the synthetic expression of rankings through numeric values or positions in predefined scales (for instance, this traditionally applies to students' achievements). This enhances the possibility of assessing and comparing individuals; each is labelled through his/her reputational values, which can be used across different contexts and through time and have persistent objectivity. Numeric values in predefined scales are common to several systems – from archives of criminal records to credit recordings, managed both by humans and by automated ones.

Third, China's SCS is *ICT-based*. ICT-based scoring systems present peculiar features. They enable scores to be grounded on massive sets of data and be obtained by any kind of complex computations. In modern societies, ICT-based scoring systems have been used in various domains: ranking job applicants, assigning credit scores, assessing risks of recidivism, etc. ICT-based ratings present common aspects: they persist in time, have an apparent degree of objectivity, can be accessed and communicate at a distance, and can trigger computations.

Fourth, China's SCS shares with other recently developed scoring systems the fact of being *surveillance based*. As activities and interactions become computer-mediated, observable sources of information can be stored, merged, and intelligently processed to assess individuals according to past behaviour and anticipate and direct future actions. The collection of useful information is increased by the extraction of data from the physical world through ubiquitous sensors. This leads to scoring systems whose functioning is based on the massive collection of personal data that is automatically collected and processed through AI technologies. Consider, for instance, how users are surveyed and scored in services such as eBay, Uber and Airbnb.

[*]   Postdoctoral Research Fellow, CIRSFID, University of Bologna.

[**]   Professor, Department of Law, European University Institute.

Peculiarities of the Chinese SCS

After presenting the commonalities between the Chinese SCS and other SCSs, we address its peculiarities.

A first distinctive feature of the Chinese Social Credit system, as compared to scoring systems in the West, is that *the government* drives it, rather than private actors. The leading surveillance paradigm in the West is represented indeed by private actors, acting for profit, and implementing the economic model of the so-called 'surveillance capitalism', as exemplified by leading internet companies, such as Google, Facebook and Amazon[1]. These companies assess and quantify the relevant attitudes and propensities of their users, support various kinds of public rankings, interact with companies providing specialised assessment and quantifications (e.g., in the credit or insurance domain). It is true that in western countries scoring may be used in the public domain (e.g., assessing risks of criminal behaviour or measuring the performance of civil servants). Examples are predictive systems to determine the risk of recidivism, which in turn contribute to determine sentences[2]. Although the 'surveillance state' also exists in the West – an example is the PRISM system set up by the U.S. government, which merged data collected from private companies and telecommunications – thus far, no systems run by a government in western countries reach the pervasiveness and intrusiveness of the Chinese SCS.

A second distinctive feature of the Chinese SCS is the fact that this system provides for a single score based on *data covering all aspects of personality*. The Chinese SCS aggregates data coming from different data platforms and collected by various private and public stakeholders into a centralised data infrastructure, expanding the traditional commercial rating into financial, political, judicial, and social domains. This is an important difference from the West, where surveillance systems and the corresponding scoring models look at citizens as 'individuals', namely, each system focusing on a particular face of a person (e.g., as consumer, debtor, gamer, worker, insured, patient, doctor, etc.) and ranking individuals with regard to aspects pertinent to that faces. Thus, it seems that a fragmented evaluation should have a lower impact on self-esteem and social ranking since each individual can distinguish his or her identity from each sectorial scoring, the impact of each being is to the corresponding context.

A third distinctive feature consists in the fact that the Chinese SCS aims to *assess the moral and civic merit of citizens* and provide corresponding punishments and rewards. On the contrary, western systems tend to focus on aspects related to economic performance, or risk of illegal behaviour. The single score provided by the system tends to become an indicator of each person's social value and can trigger effects pertaining to all aspects of individual's life: from work opportunities, to access to transport, and up to health care. In such a context, individuals may react to the fact of being surveyed and scored by adapting their behaviour to the system's rules, that is, acting in order to avoid sanctions and gain rewards. This may substitute and crowd out authentic moral/legal experience, reasoning and choice, which requires the individual to address the conflict between self-interest and moral/legal requirements. In a Kantian spirit, we may say that the exercise of morality requires the possibility of immortality, which is somehow excluded by pervasive scoring, even when scores are effectively linked to pro-social behaviour (which may not always be the case given economic and political imperatives). Reijers also argues that scoring may induce conformity rather than distinction and creativity. The extent to which it is true depends on the scoring methods: if scores are based on scalable factors (e.g., school results, response time, users' ratings), they may incentivise excellence, though according to established criteria. On the other hand, scoring leads to comparisons, competition, stress, and possibly to social action problems, where each individual invests too much to get a higher rank than others (e.g., in school or work effort). Since not everybody can excel, some will fall behind, which may lead to negative outcomes, such as dismissal or

---

[1] Zuboff, S. (2019). *The Age of Surveillance Capitalism*. New York: Hachette.

[2] Larson, J., S. M. and Lauren Kirchner, and J. Angwin (2018). 'How we analysed the Compas recidivism algorithm'. Accessed at https://www.propublica.org/article/how-we-analyzed-the-compas-recidivism-algorithm

missed opportunities, and the loss of self-esteem and marginalisation. The difficulty of improving a score (of rehabilitation or redemption) may lead to disinvestment and renunciation. The impact of scoring on individual motivation, while not been limited to the Chinese SCS, is brought to the extreme by its operation.

A fourth distinctive feature of the Chinese SCS is that it has been *designed and implemented by an authoritarian government*. Though the Chinese SCS has been presented as an implementation of a model of social virtue, inspired by the Chinese Confucian traditions, there is no warranty that it is not driven by goals of political control. There is no institutional mechanism that may ensure that the rankings correspond to an acceptable notion of civic virtue. The centralisation of data collection and scoring infrastructure rules out direct influence by civil society, the absence of which enhances the information asymmetry between state authorities and individuals; it diminishes the possibility of citizens to hold the political power accountable by a counter-surveillance mechanism, for instance, through licensing and audit requirements.

China's SCS, while presenting commonalities with other scoring systems, is characterised by a cluster of features – often extreme versions of general trends – that make it both unique and exemplary.

Untrustworthy: Social Credit Isn't What You Think It Is

Jeremy Daum[*]

Technology is redefining citizenship, governance, and individuality in China and elsewhere. We face challenges of securing personal data, limiting collection and use of online information, and controlling the spread of misinformation. Many of these issues may be global, but China is different. It lacks a healthy civil society or free media to foster debate and draw attention to potential abuses, it lacks independent courts that might review police power, and it has a government that views individual rights as inherently subordinate to obligations to maintain social order and stability.

This means there will be fewer brakes on government experimentation with new technologies in the pursuit of stability and order. There will be less resistance to invasive data collection, wrongful arrests from facial recognition misidentifications, or over-inclusive profiling of potential criminals by opaque algorithms. Stories from Xinjiang already confirm these fears[1].

Despite these differences, or because of them, there can be great comparative value for democracies in watching China's integration of technology, governance, and society, but meaningful comparison requires accurate understanding. Social credit should not be used as a stand-in for all Chinese government uses of new data technology, nor should China be used solely as a foil for our fears about technology's creeping reach into privacy and decision-making.

In his blog post, Reijers defines the essential characteristics of a Social Credit System as including tracking and surveillance, scoring, and gamification through rewards and punishments. He further distinguishes 'Social Credit Systems' from other recordkeeping or rating schemes based both on the breadth of *inputs* (including 'a great many kinds of everyday behaviours') and the 'integral character' of the *output* 'score', which creates a generalised, holistic profile assessment connecting seemingly unrelated activities.

These criteria highlight core concerns about how information from our rapidly expanding digital presences might be used for control by corporations and governments, but they are neither closely derived from Chinese law nor reflective of its practice.

How Social Credit Works

Generally speaking, China's Social Credit System includes three discrete pillars: (1) a financial credit reporting component similar to that found in many nations, which seeks to evaluate one's ability to repay loans; (2) an administrative enforcement component involving a series of industry blacklists and joint punishment agreements for those who violate laws and regulations; and (3) a moral education component that focuses on teaching values of honesty and integrity[2]. While there is some overlap between these areas (e.g., a conviction for fraud would be relevant to all three), they have different goals, methods, and even terminologies, and thus should be considered individually.

Consider the *inputs* of surveillance and behaviour tracking. China is widely regarded as a surveillance state, and its law enforcement has sweeping investigative powers with minimal external restraints. Social credit could be said to facilitate tracking by consolidating existing information records under citizen and

---

[*]   Senior Fellow, Yale Law School Paul Tsai China Center, Yale University.

[1]   Mozur, P. (2019). 'One Month, 500,000 Face Scans: How China Is Using A.I. to Profile a Minority'. *New York Times*, April 14, 2019.

[2]   http://www.gov.cn/zhengce/content/2014-06/27/content_8913.htm；translation: State Council Notice concerning Issuance of the Planning Outline for the Establishment of a Social Credit System (2014-2020), available at https://www.chinalawtranslate.com/en/socialcreditsystem/

organisation identification numbers and publicly disclosing them for a limited time, but it does not really expand the scope of government data collection. It is, instead, concerned with public, government-created records, such as for administrative punishments, permits, or licenses that reflect natural persons' and organisations' 'compliance with legal and contractual obligations in social and economic activity[3]'.

As for Reijers' expected *output* of holistic scoring or ranking, neither the 5-year planning document initiating the establishment of China's Social Credit System nor any other higher-level social credit legal authority mentions a social credit *score*. A few dozen regions have created local 'scoring' systems for 'personal integrity' (the language of the moral education component) but point deductions in these systems result from violations of legal obligations, not an integrated behavioural assessment. Moreover, the rewards and punishments are nominal in comparison to existing punishments for those same violations. The criminal penalty for drunk driving, for example, is likely of more concern than that the conviction will lower one's score. Even the most prominent point systems have so little impact on employment, education, and daily life that most citizens remain totally unaware of them[4].

Within certain fields and industries, regulatory agencies reward consistent compliance with facilitated and priority services ('green channels') while violations can result in heightened scrutiny or restrictions. This type of 'ranking' is industry-specific rather than generalised to other areas, meaning a company viewed as a compliant or non-compliant 'food producer' would only face the consequences in the area of food safety regulation.

Entities with more serious violations may be blacklisted by industry authorities, which does result in 'joint-enforcement' penalties imposed by other regulatory bodies in unrelated fields, based on inter-agency agreements[5]. These punishments often include consideration of the blacklisting when applying for professional credentials and licenses, and limits on bidding in government projects or receiving public funding. Blacklisting by the FDA for selling tainted foodstuffs, for example, can impact applications to run a securities trading business.

More exotic penalties, such as restrictions on the purchase of airplane tickets and other expensive goods, are essentially unique to the court systems' blacklist for those refusing to perform on valid court judgments. The appropriate scope and legality of such restrictions, and how long they should be enforced, remain hotly debated topics in Chinese legal circles.

In short, it is questionable whether China's emerging Social Credit System described above would qualify as 'social credit' under Reijers' criteria of what is a Social Credit System.

A Well-Informed Citizen is a Vigilant Citizen

Posing questions about how technology can be used to shape citizens, and change what it means to be a citizen, is of critical and immediate importance, but using China as a blank slate on which we project hypotheticals causes more confusion than clarity. It can distract us from more pressing concerns regarding China, technology or both.

As to China, the distraction may be direct, in shifting our focus from other crises like that in Xinjiang, or indirect, creating confusion as to the true causes for concern about social credit. If social credit is largely a mechanism for enforcing laws, is the real issue the enforcement or the unjust laws such as those that limit speech? Is the real concern China's Social Credit System, or a society perceived as so

---

[3] See   http://fj.cma.gov.cn/xmsqxj/xxgk/flfg/201905/t20190509_447094.html;   English   translation   available   at https://www.chinalawtranslate.com/xiamen-social-credit/

[4] https://www.scmp.com/magazines/post-magazine/long-reads/article/3012574/village-testing-chinas-social-credit-system

[5] https://www.creditchina.gov.cn/zhengcefagui/zhengcefagui/zhongyangzhengcefagui1/201803/t20180312_110465.html. Translation available at https://www.chinalawtranslate.com/en/

undemocratic that it could not meaningfully resist the imposition of a system like the one that Reijers envisions?

For understanding our own situation, overstating China's use of technology can be just as harmful as underestimating its use by our own governments and corporations. Because China is often held up as embodying one extreme on a spectrum, misrepresentations of its law and practice can shift the boundaries of the global discussion, making anything less invasive seem acceptable by comparison. This can lead to complacency as technology shapes our societies back home. If we only envision a caricature of China, with citizens labouring under an all-seeing algorithm, we ignore how subtle, and even comfortable, a control-based model can be; we lower our guard to more realistic threats to freedom and autonomy that surveillance capitalism is bringing to us all.

Edited by Liav Orgad and Wessel Reijers

Social Credit Systems: China and the West

Liav Orgad*

Reijers and Van 't Klooster have provoked a fascinating debate on social credit systems. The disagreement among the authors is fundamental – factual and normative. There is no agreement on the definition of a social credit system, the normative benchmark to evaluate it, the implications on the institution of citizenship, the function of the Chinese version of a Social Credit System, and the lessons to be learned from China's system to western democracies. The reader is left with no clear answer on whether a Social Credit System, in one version or another, is normatively 'good' or 'bad' (in terms of legitimacy, fairness, or efficiency). Is the 'problem' of China's Social Credit System related to its design, technological function, and the regime in which it operates, and can there ever be a liberal-democratic version of a Social Credit System? I want to address below one aspect of the debate: the comparison between China and the West.

Credit Systems in the West

Scoring systems are well known in western societies. On the *private* sector, ranking is everything – Airbnb, Uber, eBay, and even university professors (e.g., Rate My Professors platform). We rate one another (as passengers, clients, students), products, locations, and situations; by using technology or without it. Examples are all around us: agencies that rate job applications and provide employment and credit checks (e.g., 'Experian'); insurance companies that offer better deals for people who are willing to give access to personal data like fitness trackers (e.g., Vitality's Fitness Tracker Offers) and tracking devises to record driving patterns (e.g., Allstate or State Farm); landlords who decide on potential tenants based on 'tenant blacklists' in the housing rental market; and firms that base their decisions on data bought from 'data brokers'. The consequences of such ratings can be severe – a refusal to a job application and health insurance policy or inability to return a product in case of being on a 'store-returns blacklist' (for customers who often return products). The most known examples are private credit scores in the United States (e.g., FICO) and other states (e.g., Schufa in Germany), which assess the 'creditworthiness' of a person based on financial criteria (e.g., payment history, debt burden, credit used, length of credit history, and geosocial data), and whose score are frequently used by banks, insurance and credit card companies. Such methods are expanding in the West.

Scoring systems are utilised not only in the private sector but also in the *public* sector. In Israel, for instance, every citizen who serves in the military gets a personal score (a quality index, *kaba*), which reflects his/her value as a soldier. The numerical figure (41-56) influences the placement in military positions, which, in Israel, is likely to affect the future prosperity of the individual. Reported criteria include personal data (education or motivation) as well as the result of a psychometric test and an interview. At Ben Gurion Airport, every passenger is assigned a score according to the security risk he or she presents, and there are 'blacklists' and 'redlists' according to the security risk.[1] Scoring systems are used by public authorities in almost every western country. Not all scoring methods are tech-based, and some of them are unnoticed, but the public sector regularly rates the valuableness, reliability, and credibility of both citizens and non-citizens (think, for instance, of a points-based system for immigration).

At the municipal level, ideas of social credit are on the rise. In the past years, there have been attempts to use technology to motivate 'good citizenship' through the creation of a catalogue of 'good deeds'.

---

* EUI Florence, Director, 'Global Citizenship Governance'; WZB Berlin Social Science Center, Head, 'International Citizenship Law'; IDC Herzliya, Associate Professor, Lauder School of Government.

[1] Hasisi, B., Margalioth, Y., & Orgad, L. (2012). 'Ethnic Profiling in Airport Screening: Lessons from Israel, 1968-2010', *American Law & Economics Review, 14*(2), p. 517.

Examples of deeds include voting, helping the elderly, taking first-aid courses, organising cultural events, and undergoing self-employment workshops. Citizens can choose and implement good deeds from a municipal catalogue, which will then be recorded and scored; based on the score, each citizen will receive a reward from a parallel catalogue of 'municipal benefits'. Rewards may include free public transportation and bicycle rental, tickets to cultural events, and reduced municipal housing. The idea is to use financial incentives to motivate civic engagement, social solidarity, volunteer work, and social interaction with the hope that, over time, these incentives will create a culture of communal activism. Different versions of this system exist in European cities (e.g., Innowave CityPoints in Cascais, Portugal).

China and the West

In spite of the rise of social credit systems in the West, the case of China differs significantly for at least four reasons. First, *consent and choice*: with some exceptions, western rating systems are voluntary, not mandatory. Participation on private platforms such as Airbnb and Uber, and even getting a credit score at a credit card company, are voluntary – one does not have to sign in and can always opt-out. In China, however, according to the State Council Agenda, the system will become mandatory in 2020 and people will be included in the scope of the system regardless of personal choices (existing regional pilots in China are mostly voluntary). This will create a coercive scoring system in which participation is the default.

Second, *scope*: rating systems in western societies do not usually score people qua citizens, but qua drivers, landlords, and clients. By targeting people as citizens, China's Social Credit System does not target them according to a particular, limited profession, but according to being members of the political community. By scoring people *qua* citizens, every individual and every single aspect of the public life becomes conducive for being scrutinised. If a person has a low credit score in the United States, she may not be able to get a loan, but this will not affect to which schools her child can be admitted or whether she can travel on high-speed trains. In other words, the Chinese system is not limited to one field but all-encompassing. Due to that aspect, the implications of doing wrong radiate beyond the context of the wrongness into other spheres of life. The system is comprehensive for another reason. In western democracies, the rule of law entails that everything which is not forbidden is allowed; this creates a division, which is not always clear, between law and morality. In China, Social Credit Systems target not only legally impermissible actions but also socially and professionally undesirable actions. By giving rewards and impose sanctions to activities that are not illegal per se but, at most, are socially undesirable, China extends 'citizen-making mechanisms' to the moral improvement of people.

Third, *authority*: China is developing a database whose sources are both public and private and whose outcomes will be implemented by the public and the private sectors. Data sharing between government institutions in China, and between them to private companies – such as Alibaba, Tencent, WeChat, and Baidu – will produce a centralised database where many aspects of a person's public and private life are recorded: commercial data (e.g., mobile phone purchases), social data (e.g., social media contacts), and digital data (e.g., internet search history). As an analogy, think of a data-sharing mechanism between the U.S. Government, Facebook, Google, and Amazon. Such a database, and the system of carrots and sticks attached to it, will allow China to extend the system's logic from 'good citizenship' to 'good personhood'.

Fourth, *human rights*: having a high or a low score, or being on a blacklist, can have far-reaching implications. The affected human rights are not peripheral but fundamental. If one has a low score as a passenger in Uber, he or she may not be able to get an Uber ride, but can still take a taxi from a different provider or take a bus or a train or drive by car; except for cases of discrimination, one has no 'right' to get Uber services or a Facebook account. In China, the affected human rights are fundamental: education, health, housing, as well as freedom of movement, human dignity, and privacy issues. Of course, the implications of being blacklisted in the West can also be severe – think of a person who is

classified as a high-risk at the airport or has a low credit score in the United States – and may infringe upon fundamental human rights. Still, western societies have judicial review in which a person can challenge the assigned score (and the system as a whole) and can vote in a free election on the future of the system.

This is not a fundamental case against the social credit systems but, rather, against their design, application, and consequences. In other words, it is not a matter of principle, but degree. The challenge, with which political theorists are likely to struggle in the next years, is how to design a democratic version of social credit. Eventually, a social credit system mirrors all the good and the bad things of the legal system and the socio-economic structure that implements it.

About the editors

Liav Orgad
EUI Florence, Director, "Global Citizenship Governance"
WZB Berlin Social Science Center, Head, "International Citizenship Law"
IDC Herzliya, Associate Professor, Lauder School of Government

Wessel Reijers
EUI Florence, Max Weber Fellow, 'Global Citizenship Governance'
WZB Berlin Social Science Center, Visiting Researcher

"Global Citizenship Governance" Research Team

Liav Orgad
Director
EUI / WZB / IDC

Jean-Thomas Arrighi
Research Associate
EUI Florence

Johanna Hase
PhD Fellow
WZB Berlin

Jules Lepoutre
Research Associate
EUI Florence

Ashley Mantha-Hollands
PhD Fellow
WZB Berlin

Wessel Reijers
Max Weber Fellow
EUI Florence

Émilien Fargues
Max Weber Fellow
EUI Florence

Federico Tomasello
Research Associate
EUI Florence