SARA DOBOSZ

Adam Mickiewicz University, Poznań, Poland
https://orcid.org/0009-0008-3365-3397
sardob3@st.amu.edu.pl

MIKOŁAJ WIDLIŃSKI

Adam Mickiewicz University, Poznań, Poland
https://orcid.org/0009-0004-6818-7788
mikwid@st.amu.edu.pl

Social Scoring: Opportunities and Threats for Contemporary Society Using the Example of the People's Republic of China

Abstract: The purpose of this article is to analyse the social assessment system in the context of its potential opportunities and threats, with particular reference to the example of the People's Republic of China. The study is based on an analysis of the literature and legislation governing the use of artificial intelligence systems in the European Union and internationally. The paper outlines the mechanism of the social rating system, its impact on social stratification, and its ethical and legal implications. The analysis also covers the Chinese Social Credit System, its structure, functioning and consequences for citizens, indicating its impact on social and economic life. Conclusions show these systems can deepen inequalities and violate rights, stressing the need for strong regulation.

Keywords: social scoring, AI Act, Social Credit System, China

Introduction

Over recent years, the role of artificial intelligence has grown significantly in many areas of life. While the rapid development of new AI technologies brings many advantages, it also entails numerous risks related to breaches of privacy, and the protection of personal data, as well as potential impacts on citizens' rights and freedoms. In particular, the uncontrolled use of artificial intelligence systems to profile citizens and then to issue automated administrative, business or financial decisions may raise concerns.

The publication focuses on the theoretical underpinnings and functioning of social scoring systems, using the Social Credit System introduced by the government of the People's Republic of China as a case study, and argues that while such systems may improve administrative efficiency and behavioural compliance, they pose serious threats to individual rights, reinforce social inequalities, and raise fundamental ethical and legal concerns.

1. Explanation of the Concept and Mechanism of the Social Scoring System

Social scoring can be defined as an algorithmic regulation system that uses scoring mechanisms to shape collective behaviour through the use of grades as incentives.[1] The premise of the system is to elicit reactions in individuals that are positive from the point of view of the entity that implemented the system. Behaviour categorised as positive boosts a citizen's overall rating, while negative behaviour lowers it. A higher rating facilitates access to goods and services and special privileges, such as a better job, the opportunity to gain an education at the best universities, or obtain credit on preferential terms. A low score, on the other hand, can isolate an individual from the rest of society by limiting access to services reserved for those with higher scores.

The basic mechanics of a social rating system involves each citizen being given a certain initial score, which is then altered through certain actions and behaviours engaged in by the individual. Even at the initial stage, it is possible for a child to be discriminated against and indirectly held responsible for the actions of the parents, as their initial rating may depend on their past activity. In the event that parents have a high rating, their children could already be in a privileged position at the start and receive a higher score.

In addition to monitoring the citizens' behaviour as individuals, a social rating system could also analyse interactions with other people, the social environments in which the individual rotates, his/her habits, customs, preferences. Such a mechanism would serve to create a profile of a person that takes into account his or her lifestyle and level of civic engagement, thus creating a comprehensive picture of the individual in a social context.

2. Social Stratification

Introducing a social rating system can lead to existing social divisions being perpetuated and deepened. According to sociologist Max Weber, the key elements of social stratification are wealth, power and prestige, each of which may be significantly affected by the functioning of such systems.

Wealth plays a crucial role in determining an individual's access to opportunities. Individuals originating from families with low social scores may encounter barriers in accessing high-quality education and elite universities. Limited educational opportunities, in turn, reduce their chances of obtaining well-paid employment, which perpetuates lower income levels and restricts upward mobility. Conversely,

---

[1] Nello Cristianini and Teresa Scantamburlo, "On Social Machines for Algorithmic Regulation," *AI & Society: Knowledge, Culture and Communication* 35, 2020: 651.

those with higher social ratings benefit from preferential access to prestigious institutions and career advancement, which further consolidates their privileged position—sometimes reinforced through activities such as philanthropy, which positively affect their scores.

Power is also shaped by social rating mechanisms. Individuals who receive negative evaluations may be excluded from leadership roles and decision-making positions. As a result, political, administrative and organisational influence remains concentrated among those who established or uphold the rules of the scoring system. This dynamic reinforces existing hierarchies and limits democratic access to positions of authority.

Prestige, understood as social recognition and respect, may become increasingly tied to one's social score. Evaluation systems contribute to the stigmatisation of low-rated individuals, whose social interactions can negatively impact the ratings of others, leading to their gradual exclusion from public life. In contrast, high scorers enjoy increased recognition and trust, often based not only on personal merit but also on inherited social capital.

As access to resources, power and respect becomes increasingly dependent on algorithmic outcomes, social mobility is gradually being replaced by systemic reinforcement of inherited advantage. Over time, these dynamics are likely to result in the entrenchment of existing inequalities, with the gap between high and low-rated individuals continuing to grow, thus reinforcing systemic social stratification.[2]

3. Potential Risks and Concerns

The introduction of a system of social assessments raises a number of ethical concerns and risks related to both the violation of citizens' fundamental rights and potential socio-economic impacts.

A key issue seems to be issues related to the processing of personal data and the violation of individual privacy. Social assessment systems rely on collecting and analysing large amounts of information about citizens. This data can range from basic information usually collected by public authorities to that related to monitoring an individual's activity in society. This makes the issue of the potential leakage of this data all the more serious, as social assessment systems will collect a comprehensive picture of an individual's activities, combining different sources of information into a single database.

Equally dangerous could be the psychological burden on citizens. The awareness of being constantly monitored and assessed can lead to increased anxiety

---

[2] "Social Scoring Systems: Current State and Potential Future Implications," Kaspersky Blog, https://www.kaspersky.com/blog/social-scoring-systems/, accessed 10 March 2025.

and people withdrawing from social life. The need to adapt one's behaviour to the expectations of the social evaluation system can result in self-censorship, reduced freedom of expression, and a refusal to take actions that could be negatively evaluated. In the long term, this can lead to social conformity, in which individuals adhere to certain norms not because they are right, but for fear of negative consequences. What emerges is not simply a surveillance apparatus, but a form of quiet governance that rewards submission and penalises deviation, regardless of actual harm or benefit to society.

Moreover, the mechanism of the social evaluation system and the way it is scored can be difficult to understand. Citizens may not have access to full information about the assessment criteria or the specific actions that influenced their score, which would make it difficult to challenge possible errors in the system. With this state of affairs, it is not difficult to be manipulated, both by users who wish to improve their ratings and by private and public institutions that may abuse the system to achieve their own ends. Such a lack of transparency would lead to a deepening sense of injustice and helplessness in the face of the system's arbitrary decisions.

4. Legal Regulations

The dynamic development of artificial intelligence and, at the same time, the resulting risks have been recognised at EU level. This is reflected in Regulation (EU) 2024/1689 of the European Parliament and of the Council of 13 June 2024, which lays down harmonised rules on artificial intelligence and amending Regulations (Artificial Intelligence Act). At the outset, it should be noted that the EU legislator does not provide a clear assessment of what role artificial intelligence plays nowadays, but recognises both its negative and positive aspects in an attempt to ensure the necessary protection of individuals without limiting technological development. The regulation indicates that it aims to improve the functioning of the internal market and promote the spread of human-centred and trustworthy artificial intelligence.[3]

Three key aspects relevant to social assessments can be extracted from the content of the regulation: transparency, non-discrimination and accountability. These safeguards suggest an awareness that non-transparent or biased evaluating systems, such as those based on social scoring, can distort the very idea of fairness that modern democracies seek to uphold:

---

[3] Regulation (EU) 2024/1689 of the European Parliament and of the Council of 13 June 2024 laying down harmonised rules on artificial intelligence and amending Regulations (EC) No 300/2008, (EU) No 167/2013, (EU) No 168/2013, (EU) 2018/858, (EU) 2018/1139 and (EU) 2019/2144 and Directives 2014/90/EU, (EU) 2016/797 and (EU) 2020/1828 (Artificial Intelligence Act) (OJ L, 2024/1689, 12.7.2024), Recital 1.

– Transparency of systems, linked to Article 22 of the GDPR, refers to the right to obtain an explanation of the outcome of the algorithm. The scope of the algorithm and the mechanics of the system must be understood for high-risk decisions.

– Non-discrimination refers to the need to collect a variety of data in order to avoid bias and prevent uncertain decision-making and unfair use of that data against different groups in society.

– Accountability, on the other hand, relates to the need to identify those responsible for decisions made using artificial intelligence systems.[4]

Pursuant to Article 5(1)(c) of the AI Act, it is prohibited to place on the market, put into service or use AI systems for the evaluation or classification of natural persons or groups of persons over a certain period of time based on their social behaviour or known, inferred or predicted personal or personality characteristics, where such social scoring leads to either or both of the following:

(i) detrimental or unfavourable treatment of certain natural persons or groups of persons in social contexts that are unrelated to the contexts in which the data was originally generated or collected;

(ii) detrimental or unfavourable treatment of certain natural persons or groups of persons that is unjustified or disproportionate to their social behaviour or its gravity.[5]

It is important to note that the EU legislator has not chosen to prohibit social assessment systems altogether. The prohibition applies when such assessments lead to unfair or prejudicial treatment in contexts unrelated to the original purpose of the data collection or when such treatment is unjustified or disproportionate to the social behaviour of the person concerned. The doctrine emphasises that the prohibition should not affect the lawful practice of assessing individuals which is carried out for a specific purpose in accordance with Union and national law.[6]

Article 5(1)(h) of the Regulation also prohibits the use of real-time remote biometric identification systems in public spaces for law enforcement purposes, except when their use is strictly necessary for specific purposes, such as searching for victims of abduction, trafficking and sexual exploitation, preventing imminent danger to life or terrorist attacks, and identifying suspects of serious crimes punishable by a prison sentence of at least four years.[7]

In the case of real-time systems, biometric data capture, comparison and identification take place instantly or almost instantly (without major delays). Systems that use surveillance cameras to analyse images in real time and identify citizens

---

[4] Aviad Raz and Jusaku Minari, „AI-Driven Risk Scores: Should Social Scoring and Polygenic Scores Based on Ethnicity Be Equally Prohibited?" *Frontiers in Genetics* 14, 2023: 1169580, https://doi.org/10.3389/fgene.2023.1169580.

[5] Artificial Intelligence Act, Article 5.

[6] Els J. Kindt and Catherine Jasserand, "Commentary on Article 5. Prohibited AI Practices," in *The EU Artificial Intelligence (AI) Act: A Commentary*, ed. Ceyhun Necati Pehlivan et al. (Wolters Kluwer, 2024), 109.

[7] Artificial Intelligence Act, Article 5(1)(h).

on the basis of these images should therefore fall into this category. In the case of post-factum systems, this provision will not apply, as the biometric data have been captured beforehand and their comparison and identification occur at a later stage. This applies to material such as photographs or video recordings from surveillance cameras or private devices, which were generated before the biometric identification system was used.

In November 2021, the United Nations Educational, Scientific and Cultural Organisation (UNESCO) adopted Recommendations on the Ethics of Artificial Intelligence through a resolution of its 41st General Conference. Among the recommendations is that artificial intelligence systems should not be used for social scoring or the exercise of mass surveillance.[8]

5. Introduction to Social Credit System

For the foreseeable future, nearly one and a half billion citizens of the People's Republic of China (PRC) will be subjected to continuous evaluation under the Social Credit System (SCS). The scope of this endeavour is unprecedented and unmatched by any previous social engineering project in human history.[9]

Initial declarations by the Government of the People's Republic of China envisaged full centralisation of the system by 2020, but this proved unrealistic. In 2022, the Chinese authorities published the first version of regulations that systematised the operation of the Social Credit System, documenting the implementation process to date and setting directions for increasing its effectiveness in the future. Currently, there is no official information on the final date for the full implementation of the system at the national level.

The idea of social engineering is not a new phenomenon in China. The Social Credit System is a development of previous social control mechanisms, but one enriched with advanced data analysis technologies. The foundations of the current concept of social governance were established under the Hu Jintao government (2002–2012) as part of the "small state, big society" strategy. The system therefore does not represent a radical change in the Communist Party of China's (CPC) approach to the governance of society, but rather an evolution in response to global technological change.[10]

---

[8] United Nations Educational, Scientific and Cultural Organization, "Recommendation on the Ethics of Artificial Intelligence," 2022, 20.

[9] Mateusz Bartoszewicz, "Chiński System Zaufania Społecznego," *Przegląd Geopolityczny* 32, 2020: 58–67.

[10] Alicja Bachulska, *Rozwój systemu oceny wiarygodności społecznej w ChRL: między "orwellowskim koszmarem" a technokratyczną utopią?* (Ośrodek Badań Azji, Centrum Badań nad Bezpieczeństwem, Akademia Sztuki Wojennej, 2019).

6. Structure of the Social Credit System

The Social Credit System aims to create mechanisms for rewarding and punishing individuals based on an assessment of their trustworthiness. It consists of three main components: a "blacklist" system, financial credibility rankings and local social rankings.

The first of these mechanisms, the "blacklist" system, is a nationwide register of individuals who have breached existing regulations. These lists are publicly accessible, made available through government websites or posted on local notice boards. Individuals on these lists are subject to a wide range of sanctions, including restrictions on the use of public transport, obtaining credit or employment in the public sector. These restrictions are enforced by the so-called Unified Punishment System.[11]

The second key element of the Social Credit System under discussion are the credibility rankings produced by Chinese financial institutions. Of particular note is the Zhima Credit (Sesame Credit) system, a collaboration between public and private entities. This system is being developed by Ant Financial Group Holding, a company owned by the Alibaba Group. Its main objective is to build users' credit history by analysing their activity on the Alipay app. Thanks to the huge amounts of data being harvested in real time and the use of Big Data-based algorithms, the system generates ranking ratings for all active users.

The last component of the SCS, local social rankings, has been introduced experimentally in more than 40 selected pilot cities. This system, also known as "Social Scoring", involves evaluating the citizens of a city based on points awarded that reflect their behaviour. For actions considered negative, citizens lose points. However, for behaviour judged positive, there is the possibility of regaining them.

This type of social credibility assessment system is the most controversial, due to its strongly normative nature. This type of arrangement may raise concerns about the possibility of manipulation and arbitrariness in assessing aspects of social life that are traditionally not subject to numerical valuation, such as the quality of human relations or activities for the public good. Currently, the implementation and oversight of pilot programmes in this area remains the responsibility of local administrations, primarily at city level. Depending on local circumstances, different urban centres implement evaluation systems differently. For example, Suzhou has an Osmanthus Points system, whereby units are rated on a scale from 0 to 200 points. In contrast, the city of Suqian has implemented the Xichu Points system, in which each individual starts with an initial score of 1,000 points and is then assigned to one of the rating categories ranging from AAA (highest) to D (lowest).[12]

---

[11] Bartoszewicz, "Chiński System Zaufania Społecznego," 60.
[12] Bachulska, *Rozwój systemu oceny wiarygodności społecznej w ChRL*.

The varied approaches to the design and implementation of social scoring systems indicate the lack of a uniform standard and highlight the significant challenges associated with their potential impact on social structures and individual attitudes.

7. Case Study: Rongcheng City

One of the most advanced Social Credit System pilot programmes was implemented in Rongcheng, located in Shandong province with a population of 650,000. The city has recorded above-average results in implementing the SCS.

In 2014, the local city government issued a series of regulations setting out the rules for the operation of social ranking: *Rongcheng Municipal Regulations for Rewarding or Punishing Individual Citizens Based on Credibility Rating and Rongcheng Citizens' Credibility Rating Standards*. According to the rules set out in the documents, each resident of the city starts with a pool of 1,000 points, which can be added or subtracted depending on their actions. For example, false accusations result in a deduction of 30 points, while participation in charitable activities results in an award of 10 points.

The point grading system is based on a scale from "A" to "D". Model reliability ("AAA") is assigned to individuals who have accumulated more than 1050 points. In contrast, individuals whose score falls below 599 points are given a grade of "D" and are considered unreliable.

Those with an "A" rating are eligible for certain incentive measures. Among the most significant privileges is inclusion on the "red list" of creditworthiness, which is a confirmation of their financial reliability. In addition, these individuals may benefit from priority access to enrolment in educational institutions, basic social benefits, social assistance, as well as preferential employment conditions and career advancement. In addition, they may be subject to limited or completely abolished supervisory and control procedures. In turn, individuals whose score places them in the "D" category are placed on a "blacklist", resulting in public disclosure of information about their low reliability. As a consequence, they may become subject to detailed surveillance and monitoring and lose previously awarded honorary titles. In addition, they may be deprived of government financial support, face restrictions on access to credit, and experience the suspension or cancellation of their professional qualifications.

According to data provided by the local government of Rongcheng City, the introduction of the citizen assessment system in 2018 contributed to a significant decrease in the number of administrative cases. In addition, there was a 22.8% decrease in the number of social conflicts and disputes, as well as a reduction in the rate of non-performing loans at financial institutions. The enforcement rate of court decisions increased by 10% compared to the same period in the previous

year. At the same time, the credibility of public administration and the level of trust between government representatives and citizens improved. The public satisfaction index remained high, not falling below 96%.[13]

8. Control or a New Form of Freedom?

The government of the People's Republic of China's assumptions about the Social Credit System cover a much wider scope than just the Chinese territory. The main objective of the system is to classify and monitor as many citizens and institutions operating in the Chinese market as possible. The implementation of this mechanism may put pressure on individual operators who wish to maintain a high ranking and thus remain competitive in the huge Chinese market. It is worth noting that despite the formal "impersonality" of the rating algorithm, the decision to award a particular ranking remains with the PRC government, which has an overriding legal role.[14]

The functioning of the Social Credit System depends on its wide public acceptance. According to research to date, the level of approval among citizens remains high, especially among representatives of the middle class living in large urban areas.[15] The prevailing perception is that the system expands their freedom, especially in the area of consumption and economic opportunities. From the perspective of the average PRC citizen, the Social Credit System can be seen as a tool to facilitate decision-making by setting unambiguous patterns of behaviour. Through an elaborate structure of rewards and sanctions, users of the system are provided with precise guidelines for actions conducive to maintaining a favourable socio-economic position. In addition, the rankings enable individuals to better monitor the consequences of their decisions. But when everyday actions are filtered through a logic of quantifiable reward, freedom itself risks being reduced to a matter of scorekeeping.

Summary

The social scoring system is based on the collection and analysis of data on citizens in order to influence individuals' behaviour through reward and punishment mechanisms. Its introduction can lead to a widening of social inequalities, limiting access to education, the labour market and prestigious positions for those with low grades, as well as reinforcing divisions in society. In addition, the system risks breaching

---

[13] Mi Zhang, "Analiza Systemu Oceny Wiarygodności Społecznej z wykorzystaniem teorii Nudge" (Master's thesis, Szkoła Główna Handlowa w Warszawie, 2021).

[14] Bartoszewicz, "Chiński System Zaufania Społecznego," 62–63.

[15] Bachulska, *Rozwój systemu oceny wiarygodności społecznej w ChRL*, 63.

privacy, increasing surveillance and self-censorship, and lacking transparency in the setting of grades and their impact on citizens' lives.

In response to these risks, the European Union introduced the AI Act, which prohibits the use of AI systems for social assessment of individuals if it leads to disproportionate or prejudicial treatment. The regulation further prohibits the use of real-time remote biometric identification systems in public spaces, apart from certain security exceptions. At the international level, UNESCO has issued recommendations emphasising that AI systems should not be used for social scoring or mass surveillance.

The introduction of a social scoring system carries serious risks for individual rights and the social fabric. While EU and international regulations emphasise the need to protect privacy and prevent abuse, the challenge remains to enforce these rules effectively and adapt them to the dynamic development of artificial intelligence.

Assessing the Social Credit System in the People's Republic of China requires adopting a specific perspective. From an internal perspective, encompassing the perception of the Chinese Communist Party, its interests and the cultural conditioning of Chinese society, the SCS appears as a natural continuation of existing trends and an effective management and control tool. In contrast, the external perspective, in particular as represented by Western countries considered democratic, highlights significant divergences in the understanding of individual freedom, its relationship with the state or civil rights.[16]

It should be noted that the PRC does not have a unified, nationwide SCS system, but only about a collection of systems that often differ in terms of scoring, assessment methods and the aspects of civic life they influence. The common elements that dominate the Social Credit System are three key areas: the public classification of individuals into "red" and "black" lists, local social rankings and the assessment of the trustworthiness of financial institutions. As time goes on, the number of issues to be analysed and algorithmised is bound to increase. An important milestone in the development of this system is its planned unification and centralisation, covering the entire PRC territory.

Given the growing complexity of the system and China's importance in the global economy, the possibility of adapting selected mechanisms of the SCS outside the country should be considered. However, in light of the risks to individual rights, social equality, and legal safeguards identified throughout this analysis, it remains uncertain whether democratic states would be willing to reconcile such systems with the fundamental principles of their political and legal cultures. The further evolution and long-term consequences of social scoring thus remain a subject of ongoing debate, raising essential questions about the balance between technological efficiency and the preservation of civil liberties.

---

[16] Bachulska, *Rozwój systemu oceny wiarygodności społecznej w ChRL*, 64.

BIBLIOGRAPHY

Bachulska, Alicja. *Rozwój systemu oceny wiarygodności społecznej w ChRL: między "orwellowskim koszmarem" a technokratyczną utopią?* Ośrodek Badań Azji, Centrum Badań nad Bezpieczeństwem, Akademia Sztuki Wojennej, 2019.

Bartoszewicz, Mateusz. "Chiński System Zaufania Społecznego." *Przegląd Geopolityczny* 32, 2020: 58–67.

Cristianini, Nello, and Teresa Scantamburlo. "On Social Machines for Algorithmic Regulation." *AI & Society: Knowledge, Culture and Communication* 35, 2020: 645–62.

Kindt, Els J., and Catherine Jasserand. "Commentary on Article 5. Prohibited AI Practices." In *The EU Artificial Intelligence (AI) Act: A Commentary*, edited by Ceyhun Necati Pehlivan, Nikolaus Forgó, Peggy Valcke. Wolters Kluwer, 2024.

Raz, Aviad, and Jusaku Minari. "AI-Driven Risk Scores: Should Social Scoring and Polygenic Scores Based on Ethnicity Be Equally Prohibited?" *Frontiers in Genetics* 14, 2023: 1169580. https://doi.org/10.3389/fgene.2023.1169580.

Zhang, Mi. "Analiza Systemu Oceny Wiarygodności Społecznej z wykorzystaniem teorii Nudge." Master's thesis, Szkoła Główna Handlowa w Warszawie, 2021.