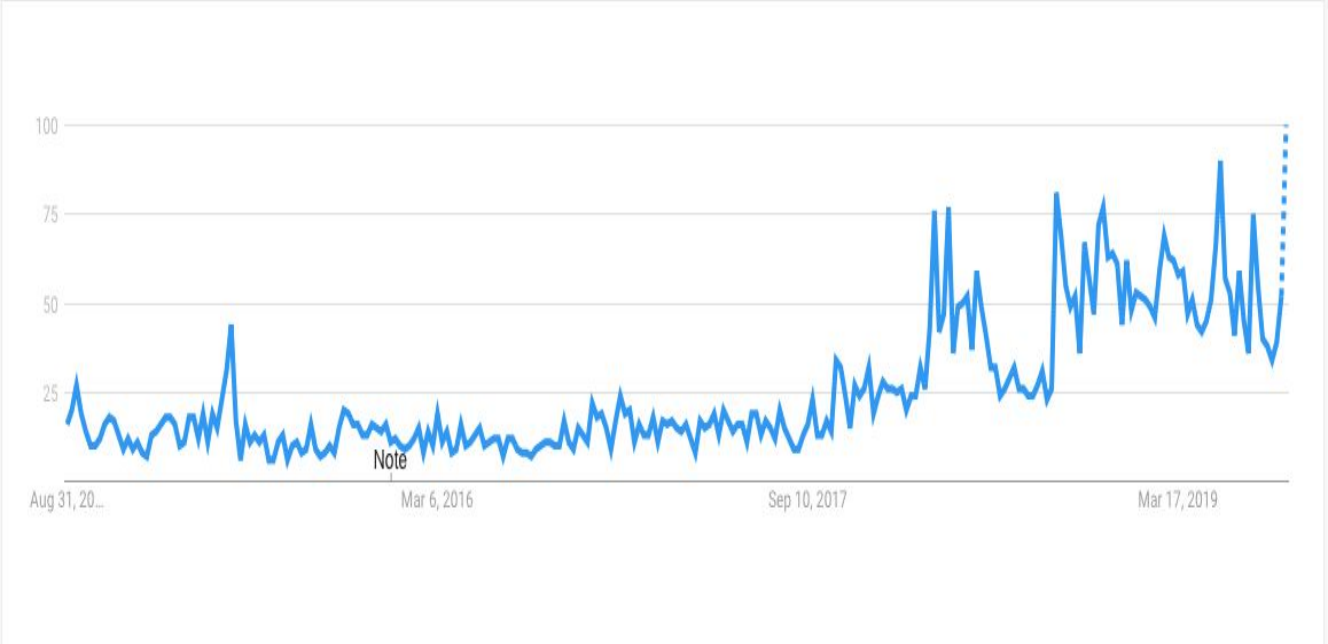
Surveillance Based Social Credit Score: Exposing and Elucidating the Ethical Issues

Anshumaan Dash
Anurag Agrawal
Bharatdeep Maan
Snehadeep Guha

Plaksha Tech Leaders Fellowship

Technology Analysis



This chart represents the relative search interest in the term 'Social Credit' 2014 onwards[5]. 100 represents peak popularity. As evident, since 2017 interest in the subject has grown quite a bit. So what is 'Social Credit' and why is it all the rage all of a sudden?

Before we get to social credit we have to understand the two main technology pillars that underpin social credit systems:

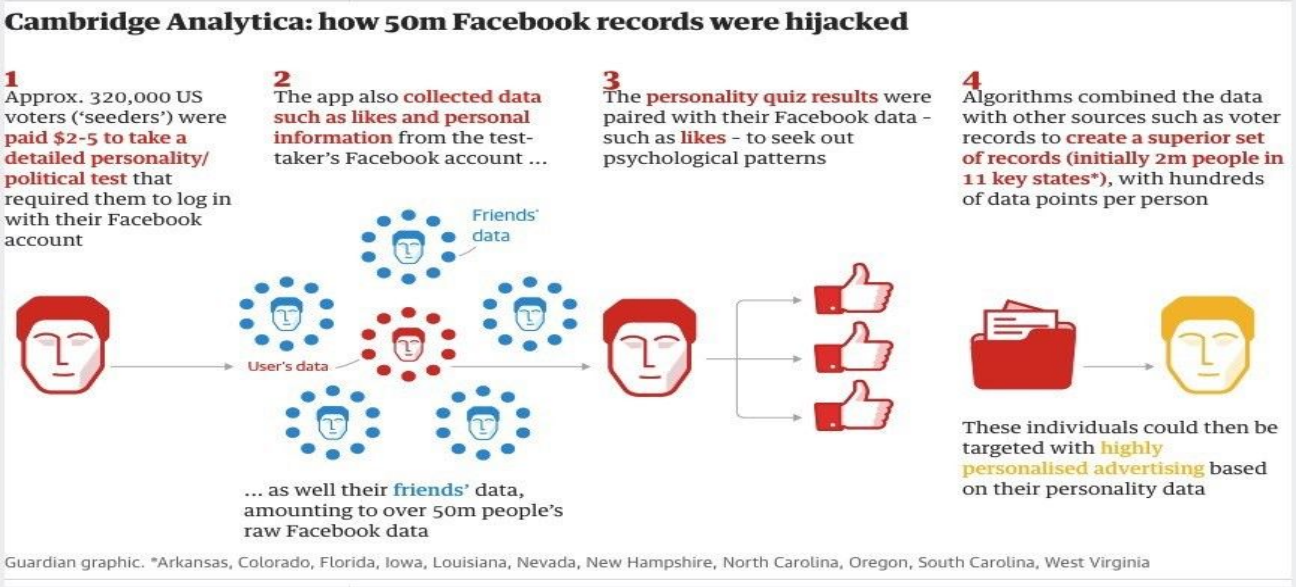
1) Computer Vision
2) Big Data Analysis

Computer vision is an interdisciplinary field that deals with how computers can be made to gain a high-level understanding from digital images or videos. Application of computer vision ranges from tasks such as industrial machine vision systems to visual surveillance systems[8]. Advancements in processing power and machine learning algorithms mean that computer vision systems are gaining a better and more thorough understanding of the world around us.

This first image that follows on the next page is taken from the Super Bowl XLVIII between Baltimore Ravens & San Francisco 49ers. In the next three images, we see how computer vision correctly identifies the vertical yard lines, the Baltimore Ravens players (ensconced in red bounding boxes) and finally the San Francisco 49ers players (ensconced in green bounding boxes)[9].

'Big Data' is a field focused on finding ways to systematically analyze and extract information from large, complex and oftentimes non-traditional datasets[10]. An excellent example of this would be the psychological profiling of individuals based on their Facebook interactions done by Cambridge Analytica.

Advances in computer vision for public surveillance coupled with improvements in big data

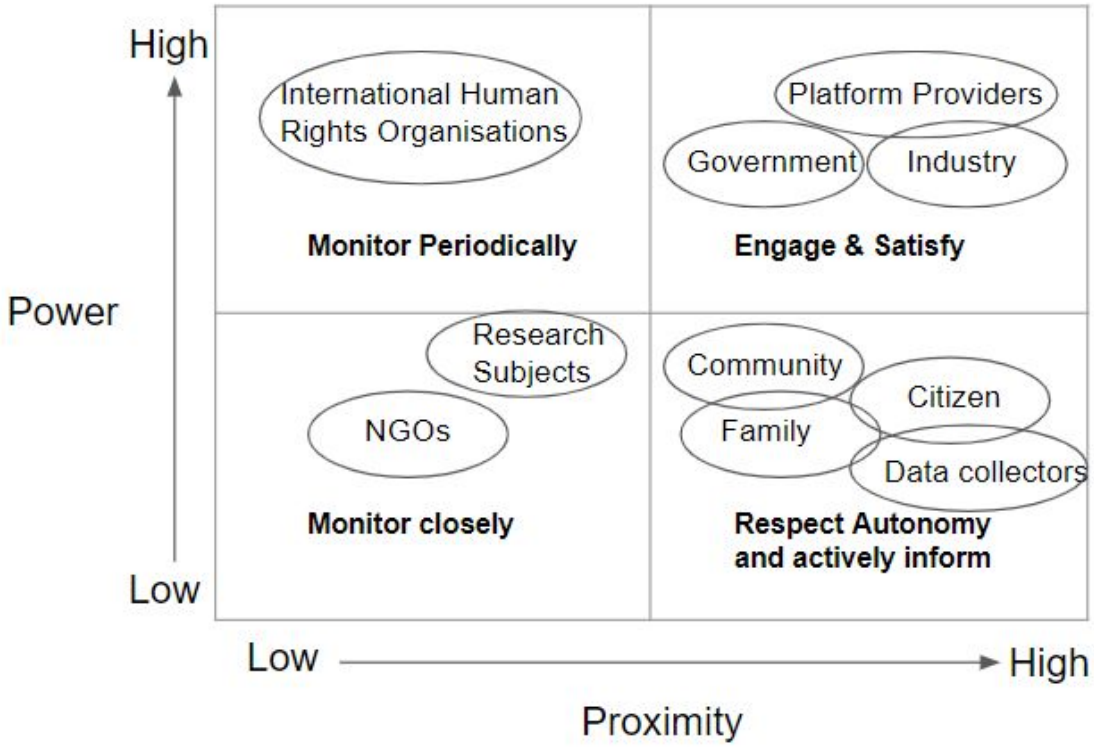


analysis on the personal, financial and medical data of an individual have led to proposals for building large-scale reputation systems intended to standardize the assessment of a person's economic and social standing. Such a system in its entirety is called social credit. Social credit systems are better understood as an overarching ideology: encompassing punishment and rewards, to improve governance and stamp out disorder and fraud. These systems, in effect, create a detailed socio-economic profile of each individual. The actions of each individual under the purview of the system are monitored and points are awarded for good behavior while engaging in practices frowned upon results in points being deducted from your score. Breaching a threshold on the lower end results in being blacklisted. Prospective punishments for blacklisted individuals include flight ban, exclusion from private schools, slower internet connections and public shaming. The most obvious adopters of social credit are governments, whether it be local, state or national.

While social credit systems are meant for the good of society, technology does raise a few ethical issues:

1) The ubiquitousness of mass surveillance required for social credit systems means that individuals don't have a choice when it comes to deciding on their participation. They have no privacy in the public or private sphere and no way of knowing when, where and how they are being monitored.
2) Social credit systems gamify our lives. It formalizes social conduct and incentivizes similitude. Furthermore, it takes away an individual's right to be forgotten.
3) The punishments handed out in social credit systems are extralegal. There is no presumption of innocence, no legal representation, no judge, no jury and often no system of appeal[4].

Stakeholder Analysis[23]



Citizen - The people who are being surveyed have the greatest impact of this technology. No point in talking about respect for autonomy as the deployment of the system itself provides a green flag to suppress the autonomy of individuals. The individual would always be under the fear that his actions could be misinterpreted. Since even the builder of these technologies may not be aware of the hidden layers of the algorithm you will always be scared.

Community - As is evident from the example of China, technology can be used to suppress the interests of some communities. The risk of non-maleficence is much greater in terms of a community perspective. For example - In India, which is home to several religions if the government lists a specific activity as punishable like praying to Allah in a large gathering at a public place. Now if points are being deducted for that from some people and the data is used to punish others in the same community then the entire community would be marginalized. Similarly is the case for blacks. Known for their open use of cuss words if the entire community is targeted then the repercussions are huge. Also, the concept of justice takes an entirely new form. Justice means that if something bad happens it should be distributed across all. Here also same happens but only restricted to that community.

Family - The issues of a family are similar to the community. The deplorable actions of one family member take a toll on others as well. For example in the world of credit, your credit score is not affected if you are not a co-borrower in a loan of a person. But in this system, if your father doesn't pay his bill, that hampers your chances of getting a loan.

Research Subjects - These are the people who would be providing the initial data to the algorithm to check the feasibility. The respect for autonomy takes a huge hit in these cases as the subject would never know what kind of their data is being taken. If a company tracks a person to see whether he broke the traffic signal when it was red, what happens to that data the subject would never know. The subject would receive no benefit as such, but the government may take action on him for some of his actions. If the same person is tracked continuously, then the government would have tons of data on him which if taken by an adversary could have implications.

International Human Rights Organisations - Organisations that operate under this domain have some power, especially if they are related to the UN. They can call for sanctions on a country if they find gross violations of human rights. However, a country like China got away with a concept like this so the jurisdiction of these organizations is also questionable.

NGOs - The local NGOs are neither directly impacted or have sizeable power. However, in a democratic setup, they may raise these issues through media and protests.

Platform Providers - The Alibaba group for example in China has been entrusted with the responsibility of making sense out of all this data. This organization thus had a huge impact and greatest proximity to this technology. The benefits of this company are obviously huge. The dataset can be employed to build different verticals for this company.

Government - The beneficence for the government is quite evident. A system like this provides them with unlimited power over its citizens. It will know your political beliefs/ views and may tamper with your credit accordingly. The long term impact is a dictatorial form of government even if it started off with democracy. On the flip side, it also provides the government with ways to identify and tackle with miscreants in society. It helps them allocate resources in a manner they deem fit.

Data collectors - They have close proximity but would not enjoy a lot of power.

Value Conflict Analysis

Computer vision technology is becoming very popular and its application ranges from surveillance, retail (Amazon Go[11][12]), the agriculture and food industry[13], social credit system, etc. The technological advancements[14] such as integrated hardware components, high computation power, deep learning and advent of data have made it possible to use the technology in real-time. With all these applications, there are conflicts of values between all the stakeholders, and it raises ethical questions.

The important conflicts of values and their reasons in different technology phases are:-

1. Analysis/Design and Research Phase: The technology must be designed taking engineering ethics into consideration rather than focusing mainly on monetary benefits. For companies, the key interests are collection and ownership of data (videos, images, etc) and using that data for commercial purposes but for users its privacy and justice. So, in the design phase, the flow and control of data must be defined properly.

2. Development and Testing Phase: In this phase, the development and testing shouldn't be done in public places or on real people without their consent. The user's interest that their data must be kept confidential and safe, otherwise their consent must be taken. Mostly, in the development phase, these interests are ignored as it slows down the whole development and testing life-cycle.

3. Marketing and Monitoring Phase: While marketing the product, the customers must be aware that how their data will be used? who will be monitoring that data? and what are the implications? Since this technology is still evolving, the monitoring phase is very important to make sure that the technology is not biased in any way, otherwise, it can raise serious questions and can have adverse effects on both the user and the tech company. The control and regulation on video surveillance are limited, so in the monitoring phase, companies have to make sure that video feeds are not abused in any way[15].

The primary reason for conflict between the stakeholders is the limited knowledge and awareness regarding the technology. The technology still hasn't proven effective[15] to be used in public spaces and that too in real-time.

The nature of the system results in a tension between the government and the populace. The major points of conflict between the people being surveilled and participants higher up the chain revolve around the following:

1) Social credit systems by design are not optional. The autonomy of the populace being surveilled is infringed upon.

2) The potential for abuse in a system such as this is greater. Fear is the currency underlying social credit solutions.

3) Every human should have the right to be forgotten. The reinvention of self is an essential characteristic of well-functioning societies.

4) Pre-existing biases manifest themselves in data. Training upon the same data results in

reinforcement of said biases rather than the ideal goal of impartiality.

5) Because of the fact that power in such a system is concentrated at the top, it can be used to carry out targeted discrimination against one portion of the society based on their ethnicity, religion, etc.

Below are the potential benefits and harms of the technology used in our case of 'Social Credit System' (especially China's implementation of the technology): [16] [17] [18][19][20]

Benefits	Harms
It will help the government to identify and monitor fraudulent activities and prevent theft [16]	Personal privacy will be violated as continuous monitoring by cameras[16]
The credit scores can be improved by increasing the consumption of certain items using certain platforms, so it will catalyze a rigorous business environment. [16]	The nature of technology itself doesn't leave room for people to behave naturally and make mistakes. [16]
Credit history: In a country like China where credit cards are not that old and there is very limited credit data. The Social Credit System will provide additional data points. [18]	The tech and its usage lacks legal regulations and also transparency issues as there's no known standard way for users to improve their scores [16][18]
Governance: The Social Credit System provides additional real-time data points such as economic data, personal activities, behavioral patterns, etc, which will help the government in decision making and planning for the overall development of the country. [18]	Social Credit System provides too much power to the governments and organizations to judge and determine people's morals, beliefs, and ethics.[16]
Administrative Efficiency: The Social Credit System enables digital capabilities and creates a platform for administration with a lot of information. It will definitely help to increase administrative efficiency. [18]	Political victimization: Since the government has access to all the information. Monarchies and governments such as China's government can target specific people and organizations. The government opposition and people who are critical to the authority can be targeted. [18]
Real-time data collection: Since Social Credit System is a digital platform in itself, this will provide real-time data and information which can be used for efficient problem-solving.[18]	Autonomy: Individuals will not have the choice to opt-out of the system if it is implemented by the government. People who don't have their own opinions, they can be tailored to the national agenda.[18]
Integrated ecosystem: It will help cluster the number of old systems that governments were using into a single platform with all the data including the social credit system. It will help governments to	Gamification: There is a chance of gamification of the whole system. The powerful and rich people can have the ability to take advantage of the system and improve their personal ratings

meaningfully collate data. [18]	since hurting others. [18]
Anti-Corruption: Once the system is matured it can make difficult for people, especially public officials to practice corruption easily. The system will make it difficult to break the system hence decreasing the number of corruption cases. [18]	Fear of Associating: Individuals can be penalized and threatened with punishments even if their friends, family, and people in their social networks make mistakes. [18]

Ethics Matrix Analysis:

Stakeholders	Beneficence	Non-Maleficence	Autonomy	Justice
Platform Providers and Data Collectors	(+) (+) They will have access to data[24]	(+)	(-) (-) (-) Specifications concerning the working of the algorithms will be stipulated by the government	
Government	(+) (+) (+) Concentration of power in the hands of a single polity	(-)	(+) (+) They are the decision-makers	(+)
Citizens	(+) (+) Citizens are rewarded for positive behavior	(-) (-) The ease with which individuals can be targeted based on religion, ethnicity sex	(-) (-) The citizen has no voice whether he wants to be a party to this or how his data is being used	(+) The legal system will have concrete data/evidence, especially in democratic countries
Society	(+)(+) Less crime and more ordered and civilized society	(-) This leads to the modification of societal behavior. Uprisings and revolts can be quelled easily	(-) (-) Same as Citizen	(+) (+) Assuming that the algorithm deployed is unbiased every member of the society is treated the same.
Human Rights Organizations		(-)		

Proposed Innovative Options

One of the key issues that one faces with the social credit system is the way it is implemented, i.e. surveillance. We collect data about the citizens from the video feeds that are collected using the ubiquitous surveillance system. These videotapes are then analyzed using computer vision, wherein we extract a myriad of information about the person involved and the scene s/he is in[21]. These details may include public information like the number plate of a vehicle, the number of people at a place, recognizing their activities, etc. These can also detect real-time security threats like unidentified objects, theft, burglary, sexual assaults, battery, etc. Till this point, one would think this technology is the best choice for humans to have in order to form a safe and crime-free society, and they should given the advantages and the benefits this technology provides us with. However, let's not forget all the other information we can extract from these scenes that might make someone uncomfortable. This information can include the identities of the people involved, their past criminal records, any physiological deformities they might have, their shopping habits, their smoking and drinking habits, their mood on a certain day, and their probability of conducting a crime on that day. It can be clearly seen that there is a very thin line one can cross while implementing this technology, to violate not only someone's personal space but their fundamental rights as well. Even with this flaw, we can still claim that all of this is required to create a crime-free and utopian society. But, this claim rests on top of one fundamental assumption that the inference made by this technology is absolute and there is no flaw in the working. On the contrary, any system based on Artificial Intelligence actually works on probabilistic computing, meaning, the results from the system are given with a probability number attached to them.

Hence, even if this system gives the verisimilitude of perfection, it has many underlying flaws that can cause an ethical dilemma. To begin with, there lies a fundamental flaw in the testing and deployment of the surveillance system. We can never test the system enough to claim that it is one hundred percent correct in predicting the outcomes. Also, unless we deploy such a system in a ubiquitous manner and act on the results, we will never know how efficiently it works. This creates a deadlock situation which successively puts unnecessary constraints on the engineering teams.

Second, the data is often used without the knowledge of the citizens. To elaborate, people only get to know that they are under surveillance. They have no idea what all information about them is extrapolated from these surveillance tapes. This is a clear violation of their autonomy. To account for this, the engineering teams would have to come up with a list of parameters that the AI system will use in order to predict the credit score. This problem is not as simple as it seems. Technologically, the modern computer vision algorithms use intricately advanced deep neural networks to classify, detect, extract pose, segment, understand an activity, and predict an outcome[21]. The mere fact that deep neural networks are used to perform these actions, makes the mechanism unexplainable. This has been the issue with neural nets ever since they were introduced. Significant research is going on to make their results more explainable. Irrespective of the fact that they outperform any other existing algorithm to perform these tasks, exacting how they do it is still a mystery to us. With this in the way, the list of parameters that the engineering teams have to come up with will always, in some way, be incomplete. This makes the task extremely difficult and strengthens the ethical questions surrounding the technology.

Pathway & Justification

Instead of all the technological ambiguities and ethical dilemmas, this system does have a myriad of benefits that can lead to the making of a very safe and just society. The team has come up with some measures to ensure that the process is made highly ethical. First, we should respect the autonomy of the citizens by informing them how the system works. The more informed they are, the easier it will be to deploy the system and gain acceptance. To bolster this claim, we can take the example of the 'Aadhar' system, currently deployed by the Indian government. The 'Aadhar' is a unique identification number that maps your personal information and biometric information to all the public accounts that you possess[22]. This took a year and a half to enforce and once people started seeing the benefits of a transparent system, they accepted it eventually. Now, all the transactions are secured, tracked and transparent. Hence, once the citizens are aware as to what information can be used from the surveillance tapes, it will be easier to ask for their autonomy as they can see the reduced number of criminal activities and frauds in the society.

Further, this system, to be effective needs to learn and be monitored on a constant basis. To accommodate this, the team has come up with the solution that we would go for pilot deployment and beta testing for a couple of years. During this phase, even though predictions will be made by the system, no actions will be taken based on it. This will allow the engineers to test the system's efficacy and make the public be comfortable being under surveillance.

Third, this mechanism of credit score calculation, to be made legal, has to go through constitutional means. For this to happen in a country like India, it has to be proposed as a bill in the parliament and the bill needs to be passed in the houses of the parliament, post which, there can be a constitutional amendment. To add, the team has proposed that the system only detects the anomalies and extracts the data. It has to be a specialized council authorized by the government that calculates the Credit Score, rather than the technology.

Fourth, the team has come up with an ethical solution to the problem with what are the impacts of the change in social credit. Any person, who is accused of being a miscreant, s/he should have the option to appeal in the legal court and hence, we propose that the change in the social credit score should be allowed to be questioned and this needs to go through the judicial system.

The team has also proposed that to be ethically sound, the social credit score need not only do down. We can employ a rewarding system, i.e. a positive credit score for all the good deeds they are doing in society. This would further lead to citizens having a fair chance to have a right to be forgotten.

Needless to say, we could further resolve our ethical dilemmas if we had access to some extra information. Though we have found a circumvent around the problem, we still do not know how to make the decisions of a neural net explainable. If this is resolved in the near future, the question of the autonomy of the citizens can be marked resolved as well. Also, since all the data will reside with the government, there is a chance that this data, even if the law restricts it, can be misused within the government, or hacked by some external organization. Therefore, without assurance about these factors, this system still possesses an ethical threat.

References

1) https://www.fastcompany.com/90394048/uh-oh-silicon-valley-is-building-a-chinese-style-social-credit-system
2) https://www.privateinternetaccess.com/blog/2015/10/in-china-your-credit-score-is-now-affected-by-your-political-opinions-and-your-friends-political-opinions/
3) https://www.ft.com/content/726905b6-c8dc-11e9-a1f4-3669401ba76f
4) https://time.com/collection/davos-2019/5502592/china-social-credit-score/
5) https://trends.google.com/trends/explore?date=today%205-y&q=%2Fm%2F01dn4s
6) https://www.theguardian.com/technology/2017/sep/07/new-artificial-intelligence-can-tell-whether-youre-gay-or-straight-from-a-photograph
7) https://medium.com/analytics-vidhya/analyzing-medium-posts-to-understand-impact-of-cambridge-analytica-scandal-5841f46703d6
8) https://en.wikipedia.org/wiki/Computer_vision
9) http://cvgl.stanford.edu/teaching/cs231a_winter1415/prev/projects/LeeTimothy.pdf
10) https://en.wikipedia.org/wiki/Big_data
11) https://en.wikipedia.org/wiki/Amazon_Go
12) https://www.geekwire.com/2016/amazon-go-works-technology-behind-online-retailers-groundbreaking-new-grocery-store/
13) https://link.springer.com/article/10.1007/s00217-012-1844-2
14) https://zbigatron.com/the-reasons-behind-the-recent-growth-of-computer-vision/
15) https://www.aclu.org/other/whats-wrong-public-video-surveillance
16) https://digitaleconomyforum.org/chapter-2-chinas-social-credit-system
17) https://en.wikipedia.org/wiki/Social_Credit_System#Implications_for_citizens
18) http://almostism.com/social-credit-system-china-implementation-benefits-criticism/
19) https://www.bloomberg.com/opinion/articles/2019-01-24/why-china-s-social-credit-systems-are-surprisingly-popular
20) https://www.merics.org/en/blog/chinas-social-credit-systems-are-highly-popular-now
21) B. Zhou, A. Andonian, and A. Torralba. Temporal Relational Reasoning in Videos. European Conference on Computer Vision (ECCV), 2018
22) https://uidai.gov.in/
23) Emerging and Readily Available Technologies and National Security A Framework for Addressing Ethical, Legal, and Societal Issues - Chameau
24) https://www.economist.com/leaders/2017/05/06/the-worlds-most-valuable-resource-is-no-longer-oil-but-data