Maintaining trust in a technologized public sector

Bodó, B.; Janssen, H.

DOI
10.1093/polsoc/puac019

Publication date
2022

Document Version
Final published version

Published in
Policy & Society

Policy and Society, 2022, **41(3)**, 414–429

DOI: https://doi.org/10.1093/polsoc/puac019
Advance access publication date 13 May 2022
Original Research Article

Maintaining trust in a technologized public sector

Balázs Bodó[1] and Heleen Janssen[1,2]

[1]Department of Law, University of Amsterdam, Amsterdam, The Netherlands
[2]Department of Computer Science and Technology, University of Cambridge, Cambridge, UK
Corresponding author: H. Janssen, Department of Law, University of Amsterdam, Roeterseiland campus, Building A, fifth floor, Nieuwe Achtergracht 166, Amsterdam 1018WV, The Netherlands. Email: h.l.janssen@uva.nl

Abstract

Emerging technologies permeate and potentially disrupt a wide spectrum of our social, economic, and political relations. Various state institutions, including education, law enforcement, and healthcare, increasingly rely on technical components, such as automated decision-making systems, e-government systems, and other digital tools to provide cheap, efficient public services, and supposedly fair, transparent, disinterested, and accountable public administration. The increased interest in various blockchain-based solutions from central bank digital currencies, via tokenized educational credentials, and distributed ledger-based land registries to self-sovereign identities is the latest, still mostly unwritten chapter in a long history of standardized, objectified, automated, technocratic, and technologized public administration. The rapid, (often) unplanned, and uncontrolled technologization of public services (as happened in the hasty adoption of distance-learning and teleconferencing systems during Corona Virus Disease (COVID) lockdowns) raises complex questions about the use of novel technological components, which may or may not be ultimately adequate for the task for which they are used. The question whether we can trust the technical infrastructures the public sector uses when providing public services is a central concern in an age where trust in government is declining: If the government's artificial intelligence system that detects welfare fraud fails, the public's confidence in the government is ultimately hit. In this paper, we provide a critical assessment of how the use of potentially untrustworthy (private) technological systems including blockchain-based systems in the public sector may affect trust in government. We then propose several policy options to protect the trust in government even if some of their technological components prove fundamentally untrustworthy.

Keywords: trust; public policy; emerging technologies; blockchain; risk-based policy

Trust at the intersection of technology and the public sector

The ongoing COVID-19 crisis has highlighted the complexity and fragility of trust relations at the intersection of society, technology, and government. Consider the following (by no means exhaustive) set of trust-related challenges that citizens, institutions, and civil servants were faced with due to the rapid incursion of various Corona Virus Disease (COVID)-related technologies into our bodies and socioeconomic relations:

Can I trust a facemask to protect me from infection? Should I trust the person (not) wearing a mask? Should I trust pandemic-related advice from governments and scientists? Can we, as a society, trust the brand-new Messenger Ribonucleic Acid (mRNA) vaccine technology? Can a traditional vaccine technology be trusted if it comes from an authoritarian regime such as Russia and China? Can I trust my government or European specialist agencies (such as the European Medicines Agency) to correctly assess the health risks associated with different vaccine technologies and not succumb to economic or political pressures (Henley, 2021)?[1] Can the government trust its citizens to voluntarily follow nonmandatory quarantine advice? Can a government be trusted with managing contact-tracing and vaccination-passport schemes? Do employers trust their workforce to work from home or do they prescribe the use of monitoring technologies? Can we trust our teleconferencing infrastructures, so they work reliably and protect our privacy? Can we have confidence in third-party teaching materials and e-learning environments, which schools are now forced to use to teach and assess our children? And where these systems malfunction, how does that affect trust in the government?

The pandemic did not itself create these questions around trust and technology, but it certainly highlighted and focused a set of more general questions about how digitization transforms societal trust relations. In that sense, COVID-related technological infrastructures are just a subset of other digital technological systems, such as artificial intelligence (AI), blockchain, automated decision-making (ADM), and recommender systems, which also deeply affect trust relations within society, and affect how citizens trust each other, how citizens and government bodies trust these technologies and affect the trust between citizens and their governments (OECD, 2017). In this article, we focus on one particular domain: the potential change in the (perceived) trustworthiness of various societal stakeholders, due to the use (or nonuse) of new technical infrastructures (Bodó, 2020).

The brief contours of a trust framework

Zucker (1985) suggests that trust rests on three distinct pillars: familiarity, control, and insurance. Trust is more likely to develop under the conditions of *familiarity*: situational normalcy, a shared and stable set of background knowledge and expectations, which trustor and trustee share. *Control* facilitates the emergence of trusting relationships as it gives some agency in the hands of the trustor to monitor and influence the behavior of a trustee, and possibly intervene in the event of unwanted behavior. Lastly, *insurance* mechanisms offer tools for the trustor to manage and minimize the risks and possible damage that are inherent in all trust relationships.

Luhmann and Giddens (Giddens, 1990; Luhmann, 1988) argued that while most trust relations are interpersonal, abstract societal systems and institutions also play an important role. On the one hand, these institutions, such as commercial firms or public bodies, are the object of trust relations (as a trustee). On the other hand, they are instrumental in facilitating the emergence of interpersonal, as well as societal, political, or economic trust relationships by providing frameworks of familiarity, control, and insurance.

In line with the institutional accounts of trust production (Bodó, 2020, 2021; Misztal, 1996; Shapiro, 1987; Sztompka, 1998, 1999; Zucker, 1985), we argue that the conditions of familiarity, control, and insurance can be produced by diverse institutional, social arrangements, which form different "infrastructures" of trust (Bodó, 2021). *Communal trust infrastructures* structured by shared religious, ethnic, or cultural roots, the norms and rules of professional associations, and shared epistemic frameworks or values offer a more traditional trust infrastructure based on interpersonal relations, often informal rules, norms, habits, rituals, and practices. In modern societies, *public trust infrastructures* offer similar frameworks through various activities and institutions of the state: public education, public service media, organization of democratic elections, democratically elected lawmakers, independent government agencies, oversight and enforcement agencies, and courts. A transparent, efficient, and disinterested public administration that executes the rules creates trust in societal relations beyond the reach of communal trust networks. Finally, private actors, such as banks, insurance companies, lawyers, accountants, or commercial brands offer *private trust infrastructures,* where trust is produced, traded, and accessed as a commodity.

In recent decades, novel techno-social systems have emerged, which also deliver the functions of a trust infrastructure, by facilitating trust-necessitating social-economic interactions.

[1] In Slovakia, the prime minister stepped down over a row about the use of Russian-developed Sputnik vaccine (Holroyd, 2021).

Vaccines and contact-tracing apps; teleconferencing and telecommunications systems; reputation-aggregating e-commerce platforms, such as eBay, Uber, or Airbnb; AI systems used in education, health diagnostics, or welfare management; social media platforms; and trust-minimizing blockchain-based systems are designed to facilitate the emergence of trust-requiring societal relations (Werbach, 2018).

In the tripartite system of communal, public, and private trust infrastructures, these technical infrastructures often follow the logic of private trust production: They are private parties, who sell trust as a commodity. They are developed by private parties and facilitate (or hinder) the emergence of trust relations under the logic of the markets, where one has to pay (with cash and/or with data) to have access to the trust services they offer (Bodó, 2021). For several reasons, such as high levels of automation, easy scaling, and ease-of-use, they are widely used by citizens, consumers, firms, and public bodies.

Trust in the government and its public institutions is a valuable resource, with often far-reaching impact on culture, social cohesion, and economic performance (Fukuyama, 1995). Yet, it is exactly this public trust that has been under significant stress in recent times (Delhey & Newton, 2003; Earle, 2009; European Commission, 2017; European Commission. Directorate General for Research and Innovation, 2017; European Foundation for the Improvement of Living and Working Conditions, 2018; Fukuyama, 1995; Organisation for Economic Co-operation and Development (Ed.), 2017; Rothstein, 2011; Zuckerman, 2021). The trust of citizens in their governments is threatened by many factors. Sometimes in foreign and domestic administrations, government officials themselves undermine trust in the state.[2] But governments, even if they enjoy high degrees of trust, face challenges that they may not be able to solve alone. Planetary scale challenges from environmental degradation, pandemics, or global economic crises pose challenges that nation-states are ill-equipped to address alone, and their failure can easily translate into distrust (Gilman & Blake, 2021).

To address the domestic and supranational challenges they face, public bodies increasingly rely on technical trust infrastructures in their governance processes and service provision. Technical systems, including complex AI and machine learning (ML) systems, are now often deployed to detect fraud, diagnose certain illnesses, allocate policing and other resources, and select targets in national security investigations or in warfare. Private technological trust infrastructures increasingly facilitate remote working and education. Private social media platforms increasingly mediate the discussions, societal debates, and election campaigns between elected officials and citizens. Smart cities often rely on private data collection and analytical infrastructures. Private companies offer electronic voting machines. The list of domains where private trust infrastructures penetrate public governance is growing every day.

While these private technology-based trust infrastructures have become deeply integrated into well-established governance procedures, most of them lack basic trustworthiness guarantees; therefore, their trustworthiness cannot be established or verified (Bodó, 2020). First, *ex ante familiarity*, however important it may be, may no longer be a resource as it is not always available while it takes time to develop (Karnow, 2020; Raymond & Connelly, 2020). The digital components of public services may often entail opaque digital black boxes, unexplainable algorithms, offering novel digital products and services, which get deployed on large populations without proper testing, and without (independent) assessment of short and long-term risks (Bodó et al., 2017; Rieder & Hofmann, 2020).

Secondly, *control* can take many forms in institutional settings: clearly defined rules, administrative procedures, enforcement rules, regulatory guidance, legal certainty, unequivocally spelled out expectations, assignments of roles and responsibilities across actors, liability regimes, as well as institutional frameworks, which monitor and intervene with the performance of technological infrastructures, are instrumental in building and maintaining trust in systems Nevertheless, instruments of control may also face limitations in the technology space. Monitoring infrastructures can be difficult to set up (Bodó et al., 2017; Rieder & Hofmann, 2020; Sandvig et al., 2014). Legislation usually takes time to be developed, implemented, and enforced. Institutional knowledge and competence to challenge, contest, and govern how their technological components work may simply be lacking or unavailable.

With inherent limits placed on familiarity and control, the role of the *insurance pillar* becomes even more important. Insurance covers all the tools that allow the trustor to assess the risks and potential damages associated with a possible breach of trust, offer ways to manage those risks, and minimize the possible fallout. Insurance works like a safety net, which allows *ex post* instruments to maintain trust in

[2] See Donald Trump's or Ronald Reagan's efforts: "Government is not the solution to our problem, government is the problem" or the Russian and US government's supposed meddling with foreign elections.

the face of knowns and unknowns. Within the newly emerging hybrid of privately provisioned technological and public, institutional trust infrastructures, a good risk and insurance-based policy approach could include clearly defined values to be safeguarded, regulatory sandboxes, careful implementation, or outright prohibition of high-risk technical systems both in public administration and beyond.

Establishing the trustworthiness of the private, technical components of public trust infrastructures is one of the key challenges for those public institutions that rely on them because ultimately it is their own trustworthiness that is at stake. According to recent studies (European Foundation for the Improvement of Living and Working Conditions, 2018; Organisation for Economic Co-operation and Development (Ed.), 2017), the two most important determinants of citizens' trust in public institutions is the quality of public services and the level of social tensions as perceived by the citizens. Most, if not all, of the conflicts around the public-sector use of untrustworthy technical infrastructure have an impact on these dimensions. Take, for example, the political crises triggered by the inappropriate use of technologies, such as the stepping down of the Dutch government following the scandal around supposed fraud in the childcare benefit system, or the embarrassment of the UK government in the AI exam grading scandal (both cases are explained in more detail below). Both cases revolve around the fact that the government was providing subpar service to its citizens due to the unequal, biased treatment of certain groups in society. There is a clear and present danger of citizens losing confidence in the government whenever the technical component fails.

In an era of growing uncertainty, distrust in governments may have devastating consequences. While some theorists argue that distrust in the government can actually be a sign of increased civic engagement and control (Sztompka, 1998), the currently dominant forms of distrust point to a different pattern: citizen disengagement, an increasingly hostile attitude toward forms of knowledge, information, and services provided by public institutions, reverting to communal trust infrastructure networks organized around shared, often fringe epistemic, communal ideological, or political frameworks, and a corresponding fragmentation of society into mutually incommensurable sections (National Intelligence Council, 2021).

The objective of this article is to identify public-sector mechanisms that may help avoid such an outcome. To that end, we pinpoint important drivers that push public-sector institutions to adopt untested and potentially untrustworthy technical components in their procedures and services and identify those considerations that are crucial in preserving or regaining trust in public institutions in a highly technological environment. We do this by exploring the justifications and potential loci of technical trust in the public sector (*Justifications and potential loci of technical trust in public sector* section), by identifying policy mechanisms that may help prevent trust gaps to occur (*Policy tools to achieve trust in public-sector technology* section), and by applying the findings to the use of blockchain technology in the public sector (*Blockchain as a policy instrument and as a policy challenge* section). This work is grounded in legal and sociological scholarly literature, recent case law, and relevant national and EU policy documents.

Justifications and potential loci of technical trust in public sector

Whenever technical infrastructure is used by government, local authorities, police forces, health services, and other public bodies, they seek to take decisions that affect the lives of citizens. Who should get what universal benefits? Whose insurance premiums should be heavily weighted? Who should be denied entry to a country? Whose knee or cancer operation should be fast-tracked? Who should get a loan? Who should be stopped and searched? Whose children should get a place in what primary or secondary school or university? Who should get bail or parole, and who should be denied these?

To solve issues and to assist decision-making, public administrations have broadly embraced the use of technical infrastructures, including that of ADM, ML techniques, and other technical infrastructures. Self-sovereign identity systems, distributed ledgers, and various blockchain-based services are in the more experimental stage, yet face high expectations, and correspondingly high public support in terms of funding, political, and institutional attention. The rationale for the use of machines is often repeated: they promise more efficient, cheaper, and time-saving service; they deliver judgments by "impartial" algorithms rather than prejudiced, hungry, or fallible judges; they provide value for money in the public sector; and so forth. The purported benefits of technical infrastructure in terms

of cost-saving, efficiency, speediness, or accuracy, to name but a few, have been much pronounced and led to swift internal public sector acceptance, building, and use.

Factors accelerating public-sector acceptance of technical trust

In the early 1980s, an international trend started in which the public sector increasingly adopted approaches to government and public service delivery termed "New Public Management." They sought to govern the public sector more like private businesses (Bureau Woordvoering Kabinetsformatie, 2021, p. 8; Hood, 1991; Prins et al., 2011, p. 30). By introducing market logic—emphasizing competition, private management, automation, performance metrics, efficiency, and cost-saving—it was felt that public administration could better fulfill its tasks (Cobbe et al., 2020, p. 48). At the time, technical infrastructure was rationalized by efficiency, cost-savings, and measuring (human) production. While the sharp edges of these ideas have partly been removed, they continue as key motives of the public sector (Cobbe et al., 2020, p. 48; Prins et al., 2011, p. 31). For many governments, the transition to the electronic government was presented as not only desirable but also as unavoidable (Prins et al., 2011).

Additional factors have increasingly advanced public administrations' use of technical infrastructure. "Trendsetting countries" such as the Nordic states, the US, the UK, or the Netherlands sought to lower the threshold for citizens to government by improving the quality and speed of its service delivery and by increasing the efficiency of its internal processes (Ministerie van Binnenlandse Zaken en Koninkrijksrelaties, 1999; United Nations, 2010). "Connected" governance hoped to connect information from across several government agencies, enabling the multiplication of government responses to pressing, complex challenges, including a rapidly expanding welfare state (Noordegraaf & Ringeling, 1995). Some countries, such as Estonia, with its e-Estonia program, or Malta with its blockchain-friendly environment, saw the technologization of government as a way to propel their antiquated and often dysfunctional public administrations into the 21st century (Demary & Demary, 2021; Pérez-Morote et al., 2020).

Finally, supranational bodies, such as the UN, see potential in novel technology infrastructures to improve or bypass nonexistent, corrupt, or dysfunctional elements of public administration in developing countries, with blockchain-based aid, and ID programs (Dumitriu & United Nations Joint Inspection Unit, 2020), or land registries. The EU also tries to foster market and competition-driven approaches by enticing national governments to share their data with businesses and other parties, to help grow economic business models in their regional markets, and to be able to compete with other regional powers such as China or the US (European Commission, 2020).

As of the 2010s, many national digitalization strategies started to work toward digital-by-default strategies, further driving automation trends (UK Cabinet Office, 2012). Governments are increasingly "administering by algorithm"—usually based on two main types of systems: automation systems and augmentation systems (Veale & Brass, 2019). With all its "promises," opportunities to embed technical infrastructure in public administration were rapidly accepted and implemented. With this, internal and external institutional pressure to change public services from "street level bureaucracy" to system-level bureaucracy grew, whereby the role of technology rapidly altered from mere supportive (automation systems, such as data registration) into leading roles (such as case assessment; Bovens & Zouridis, 2002). Technology rapidly developed into an augmented, decisive role (e.g., automated execution, control, external control, pattern finding, and predictive systems). This pushed administrative considerations to the foreground, rather than citizen-focused governance approaches that require a significant amount of discretion and flexibility to find the most appropriate solution to the problems for that person in that context.

Good intentions

Public administrations have for the most part embraced the use of technical infrastructures with good intentions. Technical infrastructures serving public-sector uses generally attempt to *improve the quantity* or *efficiency* of routine public-sector operations. They may be used to enable the *automation of tasks,* which may have some complex elements, but that produces a relatively uncomplicated and objective outcome (Veale & Brass, 2019, n. 8).

Recent technological innovations subsequently led to public administrations' adoption of *augmentation* systems, where technology was intended to assist public administration with the production of *better decisions* (Hildebrandt, 2016). These expectations of better decision-making appear largely built on

centuries-old logics of governance and governmentality, whereas the power of the modern state rests on abstract, objective, statistical, and numerical forms of knowledge (Foucault, 1991; Scott, 1998), and on more recent beliefs that technologies based on such foundations *reduce human subjectivity* and *avoid arbitrary decisions*, as statistics-based information is often perceived as *objective* (Broeders et al., 2017, p. 24). Detection of deviating patterns is supposed to help *identifications of corruptibility* and *fraud detection*. Statistics-based analytics leading to automated profiles and decisions and augmented pattern findings and predictions are, moreover, often *believed to increase transparency* in decision-making.

Bad outcomes from the public sector's technical trust infrastructures

While intentions might be good, the outcomes of technical and private trust infrastructures are not necessarily so. Over the past decade, negative consequences from public administrations' use of technical infrastructures have begun to materialize.

Wrong choice of performance indicators

Firstly, bad outcomes may be produced by a public organization's *wrong choice of performance indicators*. Such occurred, for instance, in a US-based regulation that sought to predict a person's algorithm-based predisposition to recidivism (Angwin et al., 2016). The rationale behind this regulation was that US prisons are generally overcrowded, with a disproportionate number of Black people. If computers could accurately predict which suspects were likely to commit new crimes, the criminal justice system could become fairer and make better selections of who is in prison, and for how long. The predictions would assist judges to make more accurate decisions about the risk of recidivism of a person who was in custody.

The predictions, however, ultimately undermined the public administration's efforts to ensure individualized and equal justice, because the historical data with which the system had been trained would continue to make unjustified and unjust distinctions. It proved virtually impossible to develop a technical methodology that can help take bias out of a predictive algorithmic system because it is often deeply rooted in data, weight allocation, algorithms, analysis, and in society itself (Angwin et al., 2016).

While in the recidivism case where performance indicators were wrong, performance indicators can also be *absent*. This turned out to be the case in the Dutch government's "System risk indexation" (SyRI), which entailed a risk calculation model that was developed over the past decade by the Ministry of Social Affairs and Employment (ECLI:NL:RBDHA:2020:1878). It sought to predict the likelihood of an individual committing benefit or tax fraud or violating labor laws. In a recent landmark case, an immediate halt to SyRI was ordered, as it violated privacy rights. The law did not provide *any* information about the objective or any other information that could lead to the conclusion that an increased risk of fraud exists in certain geographic areas or groups of people (or that the risk model was functioning). In other words, SyRI proposed a nontransparent and nonreviewable mode of decision-making, where effective control was excluded. The case was seen as an important case to the controversial, but growing use by governments of profiling, ADM (and AI), and risk modeling in administering welfare benefits and other core services.

Insensitivity to parameters outside technical infrastructure: ecosystem awareness

A technical infrastructure may also aggravate a public administrator's miscomprehension of and insensitivity to parameters *external* to the technical infrastructure. Sociologists have long understood that technical infrastructures can only be appropriately considered within their broader socioeconomic, cultural, political, legal, and organizational contexts (Cobbe et al., 2020, p. 48). A technical design contains a designer's assumptions, including the specific purposes for which it is developed, deployed, and used, and the results it is intended to achieve. It should therefore be considered as a *sociotechnical* process, where technical and organizational priorities related to their design, deployment, and use are involved (Cobbe et al., 2020, p. 48).

This problem of miscomprehending how technical infrastructure sits in a broader sociotechnical ecosystem recently occurred with the development of COVID-19 vaccination passports that were intended to help societies ease lockdown measures while still being able to limit the spread of the virus. The idea of vaccination passports is not new and so far had not been controversial. These positive past experiences gave public administrators some confidence that such a solution would meet little

social resistance and would be an effective way to handle the pandemic. Yet, what they failed to foresee was that the mandatory use of digital vaccination passports in everyday life would not just expose and amplify existing, although relatively marginal resistance to vaccination in general, but also would prove to be the perfect vehicle to focus the sentiments of all those citizens who distrust their government or distrust science. What was designed to be a seamless and uncontroversial technical solution to verify an individual's immunity to the virus instead prompted widespread claims of discrimination and oppression, and violent resistance. Although it was hard to foresee, the digital vaccination passport is the perfect illustration of how a generic, dispersed distrust in government can find a way to express itself through the distrust in and rejection of a particular technological solution.

Public administrations' services and activities, in their relations and interactions with citizens, consumers, and organizations, often require a diversified, individualized, tailor-made approach. Taxation decisions, or decisions over social benefits, do not permit equal treatment of unequal cases. Where such equalizing solutions, answers, or activities have been embedded in or replaced by technical infrastructures, risks of under- or over-delivery, or worse, of discrimination in relation to a specific person's or group's needs can occur. In some public sectors, certain types of decisions lend themselves to being automated, thereby equalizing decisions toward all citizens alike.

For instance, in the Netherlands, the Ministry of Justice's automated collection system for speeding tickets automatically sends fines to offenders who were detected by the Ministry's detection systems, ordering them to pay before any objection can be made. Where no real complexity exists in the facts, the metrics, and the applicable rules, such ADM may not immediately give rise to major issues. Yet, where rules are of a more open nature, where relevant facts may vary considerably between two citizens, and where metrics can be discussed, automated decisions may be less appropriate. The obligation to treat equal cases equally and to treat unequal cases unequally may urge the public decision-maker to take a case-by-case decision, based on the specific context to which the rule applies, thereby accounting for how the metrics are involved and applied.

It is in this space between easy and complex contexts where accidents may easily happen. In Australia and in the Netherlands, citizens were wrongfully assessed under so-called "Online Compliance Intervention" (Australia) or by unfair algorithms in the childcare benefit system fraud scandal (*Toeslagenaffaire*) in the Netherlands.[3] In both situations, citizens were suspected of fraudulent activities, the suspicion being based on ADM. In the Dutch case, one of the most important factors determining whether a person would be a fraudster was a person's double nationality, which in itself would lead to discriminatory suspicion. Another important factor was that all citizens were screened in exactly the same manner while personal situations were different. Most salient, however, was that the Dutch system deliberately did not leave any room for discretion: The lawmaker's objective was to set a clear example in that it would not tolerate fraud with taxpayer's money.

Given knowledge, experience, and expertise to build technical systems largely reside with private companies, many public administrations invite private organizations in tendering procedures in order to buy their technical solutions "off the shelf," or ask these companies to adjust the private technical solutions, to make them fit for use in the public sector. ADM and AI systems are tasked by legislators and policy makers, whereby these tasks are designed, implemented, and deployed by systems designers, statisticians, computer experts, etc.

In that process, the technology is explicitly designed to "decontextualize" people, and to analyze, based on their data, whether they must be put in predefined categories and contexts. Claims that citizens under scrutiny by public sector organizations (that are assisted by these techniques) should be treated "in a humancentric way"; or that their identity has always been a crucial component in how they have been (mis)treated by the state, and, therefore, various (racial, gender) components of their identity should be acknowledged and respected, may from this perspective remain hollow, as decontextualization is the central objective of pattern finding, automating decisions, or predicting future behavior. Humans-in-the-loop may not offer trustworthy solutions, as humans are known for not contesting the metrics spat out by such systems—they often lack time to contest the outcome or perceive the results as objective (Broeders et al., 2017). Interestingly, (public, unpermissioned) blockchain-based systems

[3] For useful references to Online Compliance Intervention and the problems it created see, e.g., https://en.wikipedia.org/wiki/Robodebt_scheme; for references to the *Toeslagenaffaire* in the Netherlands, see https://nl.wikipedia.org/wiki/Toeslagenaffaire (in Dutch).

take this decontextualization to the extreme, as they do not assume any knowledge on their users, and were designed to facilitate transactions between users without any known identity attributes.

Tech solutionism beliefs

In 2020, the aforementioned COVID-19 contract-tracing apps were often presented as the ultimate solution to get us out of lockdown. Governments and their political leaders in charge of health care were looking for a fast, technical answer to the complexity and newness of the pandemic, including social unrest and uncertainty over the partly very restrictive lockdown measures that were implemented. Presenting technology as solutions to societal issues has been minted as tech "solutionism"—the belief that for every problem there is a technological answer (Morozov, 2014; Naughton, 2020). At its best, a technology that is well-balanced with the ecosystem in which it functions could potentially become part of a solution. But the COVID-19 tracking apps proved ineffective and potentially operationalized toward a fundamental rights dystopia.

Another example of where tech solutionism appeared was the cancellation of A-level and General Certificate for Secondary Education examinations in the UK. Rather than giving pupils grades that had been predicted by their teachers, they decided to use a deterministic algorithm. As a result, more than 35% of English pupils were downgraded (Hill et al., 2020). Additionally, the ratio of private-school students receiving A and A* was more than twice as high as the proportion of students at state comprehensive schools (McGregor, 2020; Naughton, 2020). After growing criticism, the government reversed the results and condemned the "mutant" algorithm (Stewart, 2020). However, the deterministic algorithm did exactly what it promised to do, and it was comprehensible for any competent expert (Naughton, 2020).

Other forms of collateral damage

When public organizations rely on private technological components to fulfill their tasks, they not only outsource certain parts of their mandate but often also the knowledge and expertise needed to operate the institutional–technological hybrid responsible for that particular task. While the deputization of complex technical systems would require the development of that technological mastery, which would allow public bodies to control their technical components, in practice often the exact opposite happens: Due to costs reasons, the difficulties of finding, paying, and retaining in-house technical expertise, public servants, policymakers outsource not just the operation of technical components but also whole swathes of institutional competency.[4] This carries the risk of losing institutional knowledge and expertise while also failing to build the necessary technical capacities to oversee the technical components.

Another potential risk is the upsetting of the government–citizen trust relationship (Moyson et al., 2016). Technical systems promise the automatic, disinterested, objective, often *ex ante* enforcement of rules embedded in them. In that sense, they are disciplinary technologies (Foucault, 1979; Opsahl et al., 2021), which are based on the fundamental *distrust* in the subject. The less a government trusts its citizens, the more likely it is to rely on technologies, which keep citizens under surveillance, control, and automatic enforcement (Eubanks, 2017). This is a vicious circle: As the more governance of citizens relies on such disciplinary technologies, the less it needs (or has room) to trust its subjects in the first place.

Negligence of trade-offs in commercially driven technology design

Recent quantitative research has demonstrated that citizens generally dislike approaches that give commercial organizations control of personal data in return for the digital services they provide (Hartman et al., 2020). Against that background, it is interesting that public governance often seems to neglect the *trade-offs* that (particularly) commercially developed technologies may entail. Where such technologies are deployed on citizens, these can significantly impact citizen trust in government.

[4] The Dutch Council of State warned in their Annual Report of 2020: "The weight of official expertise in policy preparation and implementation has become more limited in recent decades, with the result that expertise and continuity 'leaves' the ministries. When there is a decline in expertise within government bodies, the demand for external experts arises. This outsourcing of expertise may temporarily strengthen its supporting character, but in time it wears out the institutional memory that is a precondition for governments to be able to continue to deliver." (Raad van State, 2021, p. 27).

Public administrations' unfamiliarity with these trade-offs might cohere with their—disputable—prioritization of cost-saving and efficiency over fostering and maintenance of the public values they are primarily responsible for.

Public administrations are primarily responsible for the development of appropriate educational systems and for providing equal access to education, which are often part and parcel of a government's vision on education. As primary and secondary schools often did (and do) not have expertise to build their own technical systems, they turned to private technological infrastructure providers, such as Google Classroom, Google Docs and Gmail, and Microsoft Education, and Microsoft Teams technologies. Both companies see "education [as a] market." More generally, a leading vision among corporations is that schools are messy, slow, chaotic, that teachers are not competent, and that computers can do things better: more efficiently, faster, far more cheaply, with more fun, whereby everything can be measured (Microsoft Education & McKinsey & Company's Education Practice, 2021; Remie & Sedee, 2020; Williamson & Hogan, 2020).

Yet, these visions lack any scientific basis—in fact, no real proof exists of what "best education methods" are, making the field prone to pseudo-scientific (commercial) offerings. Corporate and public values are conflicting, where public values exist in slowness, chaos, and messiness, as these may represent a social value. Pupil learning curves are often not linear, but unpredictable and bumpy. While societal- and public-value-driven "education visions" are generally hard to reconcile with the commercially driven Silicon Valley ideologies, the latter still entice schools to purchase their tools—for reasons of efficiency, cost-savings, and thinking that less work may entail less pressure on teachers. Pupil achievements can be measured and compared, allowing companies to build their own pseudo-scientific, commercially driven visions and feeding today's pupils, who will become tomorrow's citizens, with their views on how citizens should be educated.

Policy tools to achieve trust in public-sector technology

Novel digital technologies create new, often unforeseen, risks. When the public sector decides to rely on novel technological components, these risks relate to both the technology's users and its subjects: the public institution as much as the citizen. Social scientists, such as Beck (1992), Douglas and Wildavsky (2010), Luhmann (2017), or Giddens (2003), pointed out that modern societies, their citizens, and public and private actors are subjected to increasing levels of risk and uncertainties. Beck noted that under such conditions of uncertainty and risk, the role of public policy and the state has increasingly become that of distributing those inherent risks of modernity across society.

Public institutions ideally address uncertainty over a technology's purported qualities, its design, its actual functioning, its known and unknown characteristics, technical properties, and nontechnical potentialities and risks and harm they may cause, prior to their deployment. This policy approach requires some level of familiarity with the technology, its design, operation, and potential impact. When such a priori knowledge is unavailable or incomplete, other policy approaches such as risk-management focused policies must step in (Perrow, 2011; Sparrow, 2000). Hutter (2005) documented the rise in the popularity of risk-based policy instruments, such as cost-benefit analyses, quantitative risk assessments, or regulation formulated on a precautionary principle, which helps in assessing, managing, and reducing individual and societal risks and harms that characterize different policy domains.

Adaptive approaches, developed in the context of policy domains with a long temporal horizon and high levels of uncertainty, have also become dominant in certain domains, such as policies that need to take climate change into consideration (Bloemen et al., 2019). In their extensive study, Aven and Renn (2010) provide a detailed account and critique of the policy framework based on and focusing on risk governance, especially around the quantification of risks, the choice of metrics, and the ability to accurately conceptualize, capture and measure indirect, long-term, and societal risks and harms.

Adequacy and restrictions of current policy instruments and of public technology evaluators

The purported qualities of the technical components of public institutions include, for instance, that AI systems can provide accurate and objective predictions, that blockchain systems can transform the delivery of public services, and that they have the potential to redefine the relationship between government and the citizen in terms of data sharing, transparency and trust (European Parliament resolution of 26 May 2016 on virtual currencies, 2016; Hancock & Vaizey, 2016). The technical design of

such systems reflects these ambitions and aspirations, but they face certain limitations and trade-offs. An AI system may have to make compromises in the resources needed for training or have problems encoded in the training data; blockchain systems may face ecological costs; ADM can lead to under- or overinclusive decisions. The conception, design, and deployment of technical systems happen within current social, economic, and institutional frameworks, which further shape and modulate their functioning—vis-à-vis their original design.

Public institutions seeking to deploy these technologies in order to achieve their policy objectives are still compelled to consider the particular technology's (touted) qualities in relation to the envisioned policy objective and wider public values such as the rule of law, fundamental rights, and democratic principles. Also how and whether its design is able to fulfill the public policy's objectives and values, what its trade-offs vis-à-vis their policy objectives are, whether the technology will actually function as envisioned by the public institution, what the technology's known and unknown features are, and how potential risks and harms of these technologies can be mitigated.

Public institutions generally have a range of intent-agnostic policy instruments at their disposal that help them to bridge the gap between technology's aspirations in public policy contexts on the one hand and uncertainties concerning the particular technology's qualities, its design, its functioning, and/or its unknown, long-term consequences on the other. Potentially, useful instruments that may be of use can be often found in *ex ante* evaluation policy strategies. Such instruments usually permit public institutions to identify, assess, and mitigate problems with emerging technologies before actual harm is done. They usually include risk and impact analyses and assessments, analyses of where resistance against a certain policy may exist, or where a policy may be adopted with less resistance, social cost-and-benefit analyses, early performance analyses, or the identification of who should be involved in what phase of the process, and in what way (e.g., mere participation or co-decision; Dutch Ministry of Justice, 2021). Among legislative instruments, sunset clauses might for instance be considered, meaning that a regulation ceases to exist, unless the democratic legislator has determined that the regulation should be continued.

While *ex ante* policy instruments are among the more appropriate policy tools to assess public sector's use of a technology, we must accept that even in a best-case scenario, these instruments have limitations: They cannot prevent public sector officials using these tools in an incorrect, overoptimistic, or miscalculating manner, arriving at wrong perceptions of risks and harms. The persons involved may perhaps not consider certain forms of risk (SyrRI), underestimate potential harms and risk, or may not have sufficient expertise to interrogate a technology's inner workings, thereby making the policy instruments inaccurate, as they cannot identify how the technology's design might negatively affect the policy objective, or public values related to that objective (the Dutch childcare benefit system fraud scandal, the recidivism case). Reasons might also be less profound, as the public-sector officials using the instruments might be overenthusiastic over the cost-saving and efficiency benefits of a certain technology (UK AI exam grading) or believe that the technology solves several different complex and interlinked issues altogether. Sometimes, especially where technology adoption takes place in an unplanned and impulsive manner, as was the case with remote working and education technologies, those using the policy instruments in the institutional framework were unprepared to properly fulfill their task.

Ways forward?

The effective and trustworthy use of policy instruments to assess a technology's design, its deployment and use, often requires specialized expertise—not just technical expertise but also proficiency from broader organizational, legal, and economy angles.[5] This expertise requires time to build up and substantial resources to retain it in public organizations. The institutional response to blockchain-based crypto-assets is a point in case: Understanding how crypto-assets or smart contracts work requires complex technical expertise while understanding how they upset investment and saving patterns; money, debt, and asset markets and what kind of risks they inject into the global financial system requires novel institutional competencies. The same complexities apply to ML techniques, often used in AI, and in ADM and profiling.

[5] In the context of COVID vaccinations, the aforementioned annual report of the Dutch Council of State noted that "The strong dependence on external parties in areas such as testing and vaccination policy raises the question of whether governments have sufficient grip on the supply of essential facilities when the occasion arises" (Raad van State, 2021, p. 29).

Under such conditions, the public sector's current use of *ex ante* policy instruments for assessing the appropriateness and fit of a specific, envisioned technology for a particular public sector task may be in need of an additional trust securing mechanism. Our (Janssen's) decades-long work in the public sector concerning the use and regulation of novel digital technologies yielded a tentative set of policy mechanisms, which—in line with the relevant literature (Janssen, 2020; Leighninger, 2014; Sparrow, 2000)—can address these trust-related challenges. Such mechanisms include (1) being upfront about the public values to be preserved and safeguarded even at the cost of stifling innovation, (2) taking a mature approach to impact assessments, where public values, democracy, and security considerations are properly weighted and fully respected, (3) performing iterative, prescribed, technology agnostic impact assessments, (4) considering the wider social, cultural economic ecosystems on which the technology may have impact, (5) giving extra attention to valuing, breeding, and maintaining institutional expertise and competence to minimize competence outsourcing and loss, (6) building robust and effective monitoring mechanisms that also feed short- and long-term vision development, (7) pro-actively developing or re-evaluating technology agnostic policy visions in various societal domains (e.g., education and pandemic response) with extended public participation, (8) engaging in conversations with public officials and other societal stakeholders whether the use of a particular technology is necessary, and if so, how it impacts public values, (9) building independent assessment capacities to determine if the technology is part of the solution or part of a trust problem, and (10) create/task institutions with repository/observatory functions, where experiences and knowledge can be accumulated, shared, and disseminated.

That said, even if public institutions are optimally equipped and empowered to assess risks prior to the use of a technology, they cannot always do that correctly, and even if they try, the response may be inadequate. In those cases where the *ex ante* risk assessment may misjudge the actual risks and harms, public institutions must also be prepared to implement sometimes dramatic *ex post* damage-control policies.

Blockchain as a policy instrument and as a policy challenge

In this article, we have used multiple technical systems in public policy and administration as case studies. Now, we would like to focus on the use of blockchain-based systems because of their specific technical qualities as well as the policy challenge they pose.

Various public administrations within the Netherlands, the EU, and around the globe have been experimenting with the implementation of blockchain-based systems in diverse policy domains: public registries, Central Bank Digital currencies, digital identity solutions, to name but a few. Blockchain systems tick many of the boxes, which the public sector is sensitive to. The stated goal of blockchain developers is to create a system that is rule-based, transparent, efficient, automated, independent of human subjectivity, and bias. Most importantly, its trustworthiness guarantees differ substantially from those more traditional social, institutional, and economic guarantees that the public sector usually relies on. (Permissionless) blockchain systems' trustworthiness is supposed to be underwritten by its decentralized architecture, strong cryptography, crypto-economic incentive structures, the openness of its networks, transparent codebase, censorship resistant nature, low barriers to entry and exit, and its capacity to *ex ante* enforce rules encoded in its various layers from the protocol to the (smart contract based) application layer (Bodó et al., 2021). Its main design premise, rooted in anarcho-libertarian ideology and politics (Golumbia, 2016), is that one does not need to have trust in any of the systems' participants yet is still able to engage in the social, economic, political relations the system allows for, whether that is value transfer, voting, or property registration.

In social relations, the trustworthiness of various societal stakeholders, especially those with coercive powers, has been regarded as a key to and an indispensable prerequisite of good governance. The value proposal of blockchain systems is that this trustworthiness requisite can be eased or dropped altogether. If it is possible for public bodies to deploy a policy or engage with citizens without having to trust them, then this can raise efficiency and save a lot of cost and effort on the part of those public institutions, which would otherwise have to spend large amounts of public resources[6] to safeguard their own trustworthiness. In times where trust in public institutions is challenged, blockchain might seem a particular appealing proposition. With the use of blockchain-based services (and AI, and

[6] (Davidson et al., 2018) estimated that more than a third of the US employment in 2010 was occupied in positions, which directly or indirectly were involved in producing trust.

similar technical infrastructures), public institutions may think that they can *outsource trust*, or more precisely, outsource the task of creating trustworthy subsystems to private technological actors, without the trustworthiness of the public institution itself being negatively affected. The blockchain case may, in their perception, also be capable of *outsourcing distrust*: Citizens may not perhaps trust public sector institutions, but with blockchain-based services in place, citizens do not even have to. In this vision, blockchain-based public services offer citizens a place where they can work with that distrust productively while still being engaged with those institutions.

Yet, even in the best-case scenario, this logic can only work if the blockchain systems *themselves* are trustworthy. Scholarship on this issue points out that technical safeguards and design principles do not automatically create a trustworthy technology and reminds us that one still needs to put confidence in various system stakeholders that put values and politics in the blockchain-based systems: developers, infrastructure service providers (miners), smart contract providers, etc. (Becker & Bodó, 2021; De Filippi & Loveluck, 2016; Werbach, 2018). These stakeholders currently constitute a highly fluid, and often unstable, unregulated, self-governing techno-social system, which is often driven by Silicon Valley ideologies, libertarians, cypherpunks, moral entrepreneurs, fraudsters, hit-and-miss innovators, and speculative market-forces. This also means that none of the traditional public values that are usually embedded in public policy instruments to safeguard the trustworthiness of complex techno-social systems are in place.

It may seem paradoxical, but one may require some form of policy and state intervention to ensure the trustworthiness of a trust-minimizing techno-social system, which public sector institutions hope to use to address a growing distrust in them. This leads us to the question: how can those public bodies who face serious challenges in terms of them being trusted contribute to the trustworthiness of a system, which is built to operate under extreme conditions of distrust?

Earlier, we pointed out that emerging technologies pose specific policy challenges, which are often difficult to address with the usual set of policy tools. There are simply too many social, economic, technical, political, and cultural unknowns in these systems that can be easily addressed by careful planning, organizational, legal and procedural safeguards, and institutional checks and balances (Sztompka, 1998). These limitations highlight the need for a risk-based approach, where those institutions who are looking to address certain policy challenges, including their own trustworthiness with the use of blockchain-based technologies, have to invest in both a better understanding of all the possible (and sometimes impossible) ways something can go wrong, and the potential *ex ante* ways to address, mitigate, and correct these, if necessary.

This situation creates a unique policy conundrum. Public institutions that are looking to implement novel technologies in their services can only rely on risk-based policy instruments to ensure the trustworthiness of their technological components. If something goes wrong, however, the damage is already done, not just to the subject (the policy intervention in question) but also to the reputation and perceived trustworthiness of the institution itself. The Dutch childcare benefit fraud scandal or the UK AI exam grading debacle had two distinct classes of victims. On the one hand, citizens suffered immediate and direct harm by the technical component of the system. On the other hand, the public institutions concerned also suffered devastating consequences in terms of loss of trust in their ability or willingness (or both) to provide high-quality public services, to maintain the social cohesion of society, or act in a transparent, fair, and accountable manner. The corresponding loss of trust in the public institution can cause widespread, cascading, and unforeseeable consequences at every level of society.

This means that policy that incorporates technological components must be prepared to plan, manage, and mitigate two distinct types of risks and potential harms. One is direct and immediate and affects policy subjects. The other is indirect, long-term, and widespread and affects the public institution: the trust in it. These two types of risks require different approaches. The first one seems easier to manage: Even though it is difficult to enumerate all possible risks, it is easier to define those indicators, both qualitative and quantitative, by which harm is defined. The breach of the rule of law, fundamental rights, including discrimination, procedural, and substantive justice, quantifiable health risks, or safety metrics can define the thresholds at which corrective action becomes necessary. Reparative and restorative action can minimize damage; reforms can make adjustments to avoid the risk of repetition.

Meanwhile, the reputational and trust damage suffered by the institution is harder to remedy. Trust, once lost, is hard to regain. A growing mistrust in public institutions being able and willing to act competently, in the best interest of the citizen and society, can easily fragment society and convince

individuals and social groups to find alternative trust infrastructures. The alternatives: a replacement of trust in public institution by communal, often tribal, trust networks or by private trust suppliers seems equally problematic (Bodó, 2021).

Maintaining trust in highly technologized public institutions and regaining trust lost due to technological failures is therefore a policy task in and one which, in our view, cannot be solved through predominantly technological means. Maintaining trust and fighting distrust in government is not a job; public bodies should outsource to private parties offering various trustless blockchain applications, and supposedly very efficient AI systems. One cannot necessarily foresee when and how technical components of public administration will fail, but one can certainly better prepare for the possibility. The best insurance mechanism in the context of emerging technologies is to have safeguards—such as those embodied in the 10 mechanisms mentioned in the *Ways forward?* section —to maintain trust in the public institutions that use them. Investing in and using these mechanisms may help the public sector to identify and mitigate the negative implications that any new technology—blockchain-based or other—may entail and to prevent losses of public trust.

Summary

Writing in 2017 in her article "*Blockchains: Regulating the unknown,*" Finck (2018) suggested a number of policy options regarding the societal use and public uptake of blockchain-based services: a wait-and-see approach, informal guidance on how existing regulation may apply to emerging blockchain-based applications and practices, sandboxing, co-regulation, and issuing new legislation, as well as a regulator's experimenting with blockchain-based services for their own purposes. She highlighted this latter approach as something to be encouraged, in order to facilitate market innovation, and institutional capacity building, as well as to avoid the pitfalls of premature legislation and institutionalization.

In this paper, we argue for a more conservative approach. Through a number of case studies, we have pointed out the potentially devastating costs that the failure of private digital trust infrastructures can impose on society in general and on the trust in public institutions in particular. In the last few years, we as a society have had to learn a lot about the previously unknown, hidden, or known but dismissed, risks, and harms, which may come with overreliance on, and overconfidence in, technological infrastructures in the public sector. These failures should serve as a cautionary tale for the next era of technology policy, where techno-solutionism and techno-optimism must give way to a renewed focus on public values, trustworthy public institutions, and trustworthy private technological infrastructures.

Funding

The Blockchain and Society Policy Research Lab has received funding from the European Research Council under the European Union's Horizon 2020 research and innovation program (grant agreement no. 759681).

Conflict of interest

None declared.

References

Angwin, J., Larson, J., Kirchner, L., & Mattu, S. (2016). *Machine Bias*. ProPublica. https://www.propublica.org/article/machine-bias-risk-assessments-in-criminal-sentencing?token=_ZxdeE_wocC45nxQoi4TRYZ4jEdTGjPb.

Aven, T., & Renn, O. (2010). *Risk management and governance: Concepts, guidelines and applications*. Springer.

Beck, U. (1992). *Risk society: Towards a new modernity*. Sage Publications.

Becker, M., & Bodó, B. (2021). Trust in blockchain-based systems. *Internet Policy Review*, 10(2), 2.

Bloemen, P., Steen, M. V. D., & Wal, Z. V. D. (2019). Designing a century ahead: Climate change adaptation in the Dutch Delta. *Policy and Society*, 38(1), 58–76. https://doi.org/10.1080/14494035.2018.1513731.

Bodó, B. (2020). Mediated trust: A theoretical framework to address the trustworthiness of technological trust mediators. *New Media & Society*, 23(9), 146144482093992. https://doi.org/10.1177/1461444820939922.

Bodó, B. (2021). *The commodification of trust* (SSRN Scholarly Paper ID 3843707). Social Science Research Network.

Bodó, B., Brekke, J. K., & Hoepman, J.-H. (2021). Decentralisation: A multidisciplinary perspective. *Internet Policy Review*, *10*(2), 2. https://policyreview.info/concepts/decentralisation.

Bodó, B., Helberger, N., Irion, K., Borgesius Zuiderveen, F. J., Moller, J., van der Velde, B., Bol, N., van Es, B., & de Vreese, C. H. (2017). Tackling the algorithmic control crisis – the technical, legal, and ethical challenges of research into algorithmic agents. *Yale Journal of Law & Technology*, *133*(33).

Bovens, M., & Zouridis, S. (2002). From street-level to system-level bureaucracies: How information and communication technology is transforming administrative discretion and constitutional control. *Public Administration Review*, *62*(2), 174–184. https://doi.org/10.1111/0033-3352.00168.

Broeders, D., Schrijvers, E., & Ballin, E. H. (2017). *Big data and security policies: Serving security, protecting freedom* (wrr-Policy Brief no. 6) [Beleidsnota]. Wetenschappelijke Raad voor het Regeringsbeleid. https://www.wrr.nl/publicaties/policy-briefs/2017/01/31/big-data-and-security-policies-serving-security-protecting-freedom.

Bureau Woordvoering Kabinetsformatie. (2021, April 30). *Eindverslag informateur Tjeenk Willink* [Verslag]. Bureau Woordvoering Kabinetsformatie. https://www.kabinetsformatie2021.nl/documenten/verslagen/2021/04/30/eindverslag-informateur-tjeenk-willink.

Cobbe, J., Lee, M. S. A., Janssen, H., & Singh, J. (2020). Centering the law in the digital state. *Computer*, *53*(10), 47–58. https://doi.org/10.1109/MC.2020.3006623.

Davidson, S., Novak, M., & Potts, J. (2018). The cost of trust: A pilot study. *The Journal of the British Blockchain Association*, *1*(2), 1–7.

De Filippi, P., & Loveluck, B. (2016). The invisible politics of Bitcoin: Governance crisis of a decentralised infrastructure. *Internet Policy Review*, 5(3). https://doi.org/10.14763/2016.3.427.

Delhey, J., & Newton, K. (2003). Who trusts?: The origins of social trust in seven societies. *European Societies*, 5(02), 93–137. https://doi.org/10.1080/1461669032000072256.

Demary, M., & Demary, V. (2021). *The European blockchain centers* (Research Report No. 9/2021). IW-Kurzbericht. https://www.econstor.eu/handle/10419/231388.

Douglas, M., & Wildavsky, A. (2010). *Risk and culture: An essay on the selection of technological and environmental dangers* (1. paperback printing, 1983, [Nachdr.]). University of California Press.

Dumitriu, P., & United Nations Joint Inspection Unit. (2020). *Blockchain applications in the United Nations system: Towards a state of readiness* (JIU/REP/2020/7). United Nations.

Dutch Ministry of Justice. (2021). *Kenniscentrum Wetgeving en Juridische zaken—Beleidsinstrumenten op categorie*. https://www.kcwj.nl/kennisbank/integraal-afwegingskader-beleid-en-regelgeving/6-wat-het-beste-instrument/61/categorie%C3%ABn.

Earle, T. C. (2009). Trust, confidence, and the 2008 global financial crisis. *Risk Analysis*, *29*(6), 785–792. https://doi.org/10.1111/j.1539-6924.2009.01230.x.

ECLI:NL:RBDHA:2020:1878. ECLI:NL:RBDHA:2020:1878 (District Court of the Hague March 6, 2020). https://uitspraken.rechtspraak.nl/inziendocument?id=ECLI:NL:RBDHA:2020:1878.

Eubanks, V. (2017). *Automating inequality: How high-tech tools profile, police, and punish the poor*. St. Martin's Press.

European Commission. (2017). *"Perception of Key Institutions" in Designing Europe's Future: Trust in Institutions, Globalisation, Support for the Euro, Opinions about Free Trade and Solidarity*. Special Eurobarometer 461.

European Commission. (2020). *A European strategy for data* (COM(2020) 66 final). European Commission. https://eur-lex.europa.eu/legal-content/EN/TXT/?qid=1593073685620&uri=CELEX:52020DC0066.

European Commission. Directorate General for Research and Innovation. (2017). *Trust at risk: Implications for EU policies and institutions*. Publications Office. https://data.europa.eu/doi/10.2777/364327.

European Foundation for the Improvement of Living and Working Conditions. (2018). *Societal change and trust in institutions*. Publications Office. https://data.europa.eu/doi/10.2806/736845.

European Parliament resolution of 26 May 2016 on virtual currencies. (2016). Pub. L. No. 2016/2007(INI). https://www.europarl.europa.eu/doceo/document/TA-8-2016-0228_EN.html.

Finck, M. (2018). Blockchains: Regulating the unknown. *German Law Journal*, *19*(04), 665–692. https://doi.org/10.1017/S2071832200022847.

Foucault, M. (1979). *Discipline and punish: The birth of the prison*. Vintage Books.

Foucault, M. (1991). Governmentality. In G. Burchell, C. Gordon & P. Miller (Eds.), *The Foucault effect: Studies in governmentality* (pp. 87–104). University of Chicago Press.

Fukuyama, F. (1995). *Trust: The social virtues and the creation of prosperity*. Free Press.

Giddens, A. (1990). *The consequences of modernity*. Polity Press.

Giddens, A. (2003). *Runaway world: How globalization is reshaping our lives*.
Gilman, N., & Blake, J. S. (2021). *Francis Fukuyama: Will we ever get beyond the nation-state?* NOEMA. https://www.noemamag.com/francis-fukuyama-will-we-ever-get-beyond-the-nation-state.
Golumbia, D. (2016). *The politics of Bitcoin: Software as right-wing extremism*. University of Minnesota Press.
Hancock, M., & Vaizey, E. (2016). *Distributed ledger technology: Beyond block chain*. Government Office for Science. https://www.gov.uk/government/news/distributed-ledger-technology-beyond-block-chain.
Hartman, T., Kennedy, H., Steedman, R., & Jones, R. (2020). Public perceptions of good data management: Findings from a UK-based survey. *Big Data & Society*, *7*(01), 2053951720935616. https://doi.org/10.1177/2053951720935616.
Henley, J. (2021, May 25). Influencers say Russia-linked PR agency asked them to disparage Pfizer vaccine. *The Guardian*. http://www.theguardian.com/media/2021/may/25/influencers-say-russia-linked-pr-agency-asked-them-to-disparage-pfizer-vaccine.
Hildebrandt, M. (2016). *Smart technologies and the end(s) of law: Novel entanglements of law and technology (Paperback edition)*. EE Edward Elgar Publishing.
Hill, A., Davies, C., Halliday, J., Obordo, R., & Obordo, R. (2020, August 13). A-level results day 2020 live: 39.1% of pupils' grades in England downgraded - as it happened. *The Guardian*. https://www.theguardian.com/education/live/2020/aug/13/a-level-results-day-2020-live-students-teachers-government-ucas-mock-exams-triple-lock-nick-gibb.
Holroyd, M. (2021, March 28). *Slovakia's prime minister steps down amid Sputnik V vaccine scandal*. Euronews. https://www.euronews.com/2021/03/28/slovakia-s-prime-minister-to-step-down-amid-sputnik-v-vaccine-scandal.
Hood, C. (1991). A public management for all seasons? *Public Administration*, 69(1), 3–19. https://doi.org/10.1111/j.1467-9299.1991.tb00779.x.
Hutter, B. M. (2005). *The attractions of risk-based regulation: Accounting for the emergence of risk ideas in regulation (Vol. 33)*. CARR.
Janssen, H. L. (2020). An approach for a fundamental rights impact assessment to automated decision-making. *International Data Privacy Law*, *10*(1), 76–106. https://doi.org/10.1093/idpl/ipz028.
Karnow, C. E. A. (2020). The opinion of machines. In W. Barfield (Ed.), *The Cambridge Handbook of the Law of Algorithms* (pp. 16–46). Cambridge University Press.
Leighninger, M. (2014). Want to increase trust in government? Update our public participation laws. *Public Administration Review*, *74*(3), 305–306. https://doi.org/10.1111/puar.12208.
Luhmann, N. (1988). Familiarity, confidence, trust: Problems and alternatives. In D. Gambetta (Ed.), *Trust: Making and breaking cooperative relations* (pp. 94–109). Basil Blackwell.
Luhmann, N. (2017). *Risk: A Sociological Theory*. Taylor and Francis Group. http://public.ebookcentral.proquest.com/choice/publicfullrecord.aspx?p=5017817.
McGregor, S. (2020). *Artificial Intelligence Incident Database*. Artificial Intelligence Incident Database. https://incidentdatabase.ai/about.
Microsoft Education & McKinsey & Company's Education Practice. (2021). *The class of 2030 and life-ready learning: The technology imperative*. Microsoft. https://info.microsoft.com/ww-landing-McKinsey-Class-Of-2030-Whitepaper.html?lcid=en-us.
Ministerie van Binnenlandse Zaken en Koninkrijksrelaties. (1999). *Actieprogramma Elektronische Overheid; Brief minister met het Actieprogramma Elektronische Overheid* [Officiële publicatie]. Tweede Kamer der Staten-Generaal. https://zoek.officielebekendmakingen.nl/kst-26387-1.html.
Misztal, B. A. (1996). *Trust in modern societies: The search for the bases of social order*. Polity Press.
Morozov, E. (2014). *To save everything, click here: The folly of technological solutionism*. PublicAffairs.
Moyson, S., Van de Walle, S., & Groeneveld, S. (2016). What do public officials think about citizens? The role of public officials' trust and their perceptions of citizens' trustworthiness in interactive governance. In J. Edelenbos and I. van Meerkerk (Eds.), *Critical reflections on interactive governance* (pp. 189–208). Edward Elgar Publishing.
National Intelligence Council. (2021). *Global trends 2040* (NIC 2021-02339). Office of the Director of National Intelligence. https://www.dni.gov/files/ODNI/documents/assessments/GlobalTrends_2040.pdf.
Naughton, J. (2020, April 25). Contact apps won't end lockdown. But they might kill off democracy. *The Guardian*. http://www.theguardian.com/commentisfree/2020/apr/25/contact-apps-wont-end-lockdown-but-they-might-kill-off-democracy.
Noordegraaf, M., & Ringeling, A. B. (1995). *De ambtenaar als publiek ondernemer*. Coutinho.

Opsahl, G. G., & Galperin, E., and Kurt. (2021, May 20). Fighting disciplinary technologies. *Electronic Frontier Foundation*. https://www.eff.org/deeplinks/2021/05/fighting-disciplinary-technologies.

Organisation for Economic Co-operation and Development (Ed.). (2017). *Trust and public policy: How better governance can help rebuild public trust*. OECD.

Pérez-Morote, R., Pontones-Rosa, C., & Núñez-Chicharro, M. (2020). The effects of e-government evaluation, trust and the digital divide in the levels of e-government use in European countries. *Technological Forecasting and Social Change*, 154, 119973. https://doi.org/10.1016/j.techfore.2020.119973.

Perrow, C. (2011). *Normal accidents*. Princeton University Press.

Prins, C., Broeders, D., Griffioen, H., Keizer, A.-G., & Keymolen, E. (2011). *IGovernment*. Amsterdam University Press.

Raad van State. (2021). *Jaarverslag 2020*.

Raymond, A. H., & Connelly, C. (2020). Governance of algorithms: Rethinking public sector use of algorithms for predictive purposes. In W. Barfield (Ed.), *The Cambridge Handbook of the Law of Algorithms* (1st ed., pp. 233–250). Cambridge University Press. https://doi.org/10.1017/9781108680844.013.

Remie, M., & Sedee, M. (2020). *Techreuzen willen de school hervormen*. NRC Handelsblad. https://www.nrc.nl/nieuws/2020/07/19/techreuzen-willen-de-school-hervormen-a4006368.

Rieder, B., & Hofmann, J. (2020). Towards platform observability. *Internet Policy Review*, 9(4), 1–28.

Rothstein, B. (2011). *The quality of government: Corruption, social trust, and inequality in international perspective*. University of Chicago Press.

Sandvig, C., Hamilton, K., Karahalios, K., & Langbort, C. (2014). Paper presented to "Data and Discrimination: Converting Critical Concerns into Productive Inquiry," a preconference at the 64th Annual Meeting of the International Communication Association (pp. 1–23). Seattle, WA, USA. https://doi.org/10.14763/2020.4.1535.

Scott, J. C. (1998). *Seeing Like a State: How certain schemes to improve the human condition have failed*. Yale University Press.

Shapiro, S. (1987). The social control of impersonal trust. *American Journal of Sociology*, 93(3), 623–658.

Sparrow, M. K. (2000). *The regulatory craft: Controlling risks, solving problems, and managing compliance*. Brookings Institution Press.

Stewart, H. (2020, August 26). Boris Johnson blames "mutant algorithm" for exams fiasco. *The Guardian*. http://www.theguardian.com/politics/2020/aug/26/boris-johnson-blames-mutant-algorithm-for-exams-fiasco.

Sztompka, P. (1998). Trust, distrust and two paradoxes of democracy. *European Journal of Social Theory*, 1(1), 19–32. https://doi.org/10.1177/136843198001001003.

Sztompka, P. (1999). *Trust: A sociological theory*. Cambridge University Press.

UK Cabinet Office. (2012). *Government Digital Strategy 2012* (p. 52). Cabinet Office. https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/296336/Government_Digital_Stratetegy_-_November_2012.pdf.

United Nations. (2010). *United Nations E-Government Survey 2010*. United Nations. https://publicadministration.un.org/egovkb/en-us/reports/un-e-government-survey-2010.

Veale, M., & Brass, I. (2019). Administration by algorithm?: Public management meets public sector machine learning. In M. Veale & I. Brass (Eds.), *Algorithmic regulation* (pp. 121–149). Oxford University Press.

Werbach, K. (2018). *The Blockchain and the New Architecture of Trust*. MIT Press.

Williamson, B., & Hogan, A. (2020). *Commercialisation and privatisation in/of education in the context of COVID-19*. Education International.

Zucker, L. G. (1985). Production of trust: Institutional sources of economic structure, 1840 to 1920. In L. L. Cummings & B. Staw (Eds.), *Research in organizational behavior* (pp. 1840–1920). JAI Press.

Zuckerman, E. (2021). *Mistrust: Why losing faith in institutions provides the tools to transform them* (1st ed.). W. W. Norton & Company.