PART III

Data Driven Operations

CHAPTER 11

A conceptual investigation of the trade-off between privacy and algorithmic performance

Job Timmermans & Roy Lindelauf

Abstract

Data-driven methodologies come with a huge benefit of optimising solutions for decision problems, but also potentially introduce new ethical risks, both in a general societal context and for defence and military organisations in particular. This chapter therefore conceptually investigates the trade-off between privacy on the one hand and algorithmic performance on the other, concerning the use of Ministry of Defence (MoD) relevant (bulk) datasets from a technical, moral and socio-political view. Because the MoD operates in a global context, the possible consequences of choices by the MoD are also considered in comparison with the choices made by other countries with respect to this trade-off.

Keywords: Privacy, Algorithmic performance, Machine learning, Security, Decision-making

11.1. Introduction

The last decade has seen an exponential development and integration of big data and machine learning algorithms within almost every domain thinkable, varying from informing bail and parole decisions[1], which patient is seen[2], college admissions[3] to filtering applicants for financial loans[4]. Within many decision-making challenges relevant to the Ministry of Defence (MoD), and defence and security in general, machine learning is gaining prominence too. Solutions exist for instance in automating tasking of technical collection platforms, in predicting the likelihood of terror attacks[5], in modelling terrorist and insurgent group behaviour[6], in deciding which surveillance image footage to focus on[7], fake news detection[8] and in operating (semi-) autonomous weapon systems such as tanks or drones[9], to name just a few.

This current exponential increase in data gathering and analysis capabilities confronts governments and organisations worldwide with ever-increasing *bulk* datasets to mine for information. Bulk datasets are sets of information about a large number of individuals, the majority of whom are not of interest to intelligence agencies. These bulk datasets, such as, for example, large-scale communications data or

names of passengers on airline flights are gathered by defence and intelligence agencies with the goal of providing both tactical and strategic foresight and warning[10]. These data-driven methodologies come with the huge benefit of optimising solutions for decision problems, but also potentially introduce new ethical risks, both in the societal context in general, and for defence and military organisations in particular. There is a fierce debate on several aspects related to mass surveillance using bulk datasets, such as the effectiveness of such surveillance programs in preventing terrorist attacks and serious crime[11], privacy concerns relating to surveillance and technology[12], fairness and transparency of the algorithms used to analyse such datasets[13], and concerning their judicial regulation[14]. In light of these developments, in 2017, the Dutch Parliament approved a revision of the 2002 Dutch Intelligence and Security Services Act, where the new version explicitly incorporates the use of modern technology and data gathering and analysis capabilities. An evaluation of this law concluded that the Intelligence and Security Services Act 2017 (WIV)[15] is not optimally synced with the technological complexity and operational practice of Dutch intelligence agencies and specifically stated that a new bulk data regime is required[16].

Incorporating rules and regulations on data and algorithms for their analysis, implementing process frameworks for non-discrimination by design and using explicit ethical algorithm design are all important ways to deal with the challenges of bulk data and should therefore be pursued. However, one trade-off that should not be overlooked upon implementing such measures, and which is especially important within the defence and security domain, is the trade-off between military advantages in the use of algorithms for decision-making and democratic values. The Dutch Defence Strategy on data science explicitly states that the use of complex algorithms should enhance military strength but also should be aligned with core democratic values. In this chapter, therefore we conceptually investigate the trade-off between one core value, namely privacy, on the one hand and algorithmic performance on the other, with respect to the use of MoD relevant (bulk) datasets from a technical, moral and socio-political view.[17] In addition, because the MoD operates in a global context, we consider the possible consequences of choices by the MoD in comparison with the choices made by other countries concerning this trade-off.

We start by introducing a coordinate system that aids in the analysis of this trade-off. In section 11.2 we use this system to present a conceptual mapping of the trade-off between algorithmic performance within a defence and security context and the value of privacy. Next, in section 11.3 we sketch the technical dimensions of and some solutions to this trade-off followed by the socio-political dimensions in section 11.4. We conclude the chapter by briefly looking ahead to ways to further tackle this multi-dimensional challenge.

11.2. Privacy vs. security trade-off

An insightful way to depict challenges that are faced when designing and deploying ethical algorithms is offered by van den Hoven et al.[18] based on earlier work by Kuran[19]. In their article, the authors propose a two-dimensional coordinate system with on each axis a different (moral) value that simultaneously has to be met by the design (see Figure 11.1).

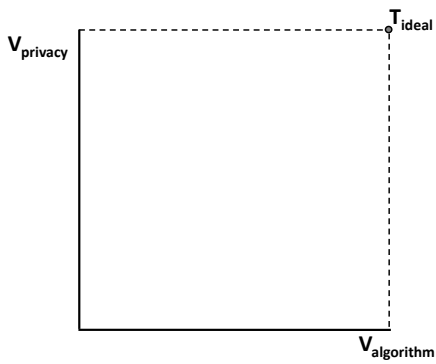

Fig. 11.1. Mapping of two moral values.[20]

In our case, the horizontal axis represents one or more moral values associated with the performance of the algorithm ($V_{algorithm}$), i.e. the benefits of deploying the algorithm, for example, its ability to discover potential targets in a large dataset. And, the vertical axis represents a moral value that may be negatively affected by deploying the algorithm, in our current example, the value of privacy ($V_{privacy}$).

The origin of the coordinate system then denotes the theoretical position in which none of the values is met. Moving upwards along the horizontal axis signifies an increase in the performance of the algorithm and hence a higher level of meeting the associated value commitment. Likewise, moving upwards along the vertical axis signifies a gradual increase in the level the value of privacy has met.

In an ideal world, both values $V_{algorithm}$ and $V_{privacy}$ can be fully met simultaneously (T_{ideal}). In that particular case, the performance of the algorithm and its associated values would not be impaired by meeting the requirement set by the opposing moral value such as privacy. In reality, however, in designing and implementing algorithms we will face situations in which different value commitments cannot (fully) be lived up to simultaneously.

It should be noted that the depiction of points and curves in our figure is a conceptual approximation of reality. In fact, a particular value may not be readily

quantifiable or be mapped onto an ordinal or a rational linear scale. It, however, does allow us to display the relative positions and trends of the possible solutions in dealing with the current challenges.

In this way the coordinate system allows us to map all possible ways a trade-off, conceptually at least, can be reached between these two values. In turn, this offers a canvas to display the technical and associated socio-political challenges and their interrelationships that the MoD faces when implementing algorithmic analyses on its data. Before turning to these challenges, first we will explicate two (partly) conflicting values and what is at stake for the MoD in more detail.

11.2.1. Algorithmic performance, security and safety

As already pointed out in the introduction, the use of algorithms for data analyses can be understood to contribute to the core values of the MoD.

In its mission statement, the Dutch MoD states that its main objective is to 'protect all that we as a nation cherish.'[21] To that end, it 'fight[s] for a world of freedom and security'[22]. To be able to meet these two core values in the near future the MoD finds that it is 'vital' that it becomes a 'technologically advanced organisation'[23]. Because the use of innovative technology lies at the heart of modern warfare, meeting the core values is thus tightly connected with the use of innovative technology. The MoD, therefore, acknowledges that it relies heavily on having access to advanced technologies such as 'Artificial Intelligence (AI) and Big Data' to meet its security interests[24].

A good example to illustrate this point is the focus on information-based military decision- making ('Informatiegestuurd optreden') (IGO), which is regarded as the foundation of the future defence organisation[25]. IGO can be defined as the use of information to reduce uncertainty in (military) decision-making. Increasingly, this information comes from the (semi) automatic analysis of data using algorithmic techniques. By analysing such large quantities of data, IGO enhances the capacity to obtain insight, understanding and foresight of complex situations and operations[26]. This enables the military to detect and identify threats early on and take preventive measures when necessary.

The effective use of technology such as data analytics not only affords to meet the overarching value of security of the MoD, but also carries over to underlying values such as safety. Working with the 'most modern equipment' such as AI and big data ensures that all MoD's personnel 'can do their work as safely as possible.'[27] Similarly, by enabling more precise targeting, access to large datasets may help to reduce the risk of avoidable (civilian) causalities[28]. Additionally, there are a plethora of applications of AI and data science useful for the MoD: think of helicopter mission optimisation techniques (as described in chapter 7), predicting cache

locations of IED's based on historical attack data and cultural variables[29], predicting terrorist group behaviour and optimising policy options for their destabilisation[30], predicting border incursions[31] and understanding the cyber vulnerability equity process[32] to name just a few.

As an illustration, we briefly discuss another recent example of the use of data analytics relevant to defence and security, in this case concerning the proliferation of nuclear material. There exist clandestine supply markets consisting of individuals, businesses, research institutions and governmental organisations that participate in the illicit proliferation of nuclear material. Starting from an existing list of sanctioned entities, and based upon open-source data (LinkedIn and Bloomberg) alone, Andrews et al.[33] build a machine learning model (SPINN: Suspiciousness Prediction in Nuclear Networks) to analyse a proliferation network consisting of more than 74k nodes and over 1M edges and to predict the suspiciousness of those nodes. Based upon novel network-centric features they were able to predict who is suspicious with high accuracy[34]. They were able to identify suspicious entities that were not known before, like a company in the UK that specialises in mining zinc and other base metals. Models like SPINN, provide defence intelligence organisations with decision support systems to optimise their fight against nuclear proliferation by discovering unknown targets, which enables them to rank them according to their suspiciousness and as such contribute to increasing safety and security. In the next section, we discuss the competing value of privacy in the trade-off dilemma.

11.2.2. Privacy

Besides offering the benefit of enhanced security and safety, the use of algorithms to analyse data also has several potential drawbacks, not least the risk of privacy infringement, which gives rise to challenges ranging from legal constraints to ethical information use[35]. Although there is consensus that the value of privacy is important, the concept remains hard to explicate or pin down[36]. In general, it aims to constrain access to certain types of personal data and prevent persons to acquire and use information about other persons[37]. This is necessary because contemporary information technologies threaten privacy by reducing the amount of control over personal data thereby opening up the possibility of a range of negative consequences that result from access to personal data.'[38]

The use of intelligent systems for data analytic purposes makes the concern for privacy even more acute as they offer powerful mechanisms that could increase real and anticipated assaults on individual privacy[39]. There are many examples of privacy violations that could occur when large datasets containing information about individuals are analysed. Note that a privacy violation in our sense does not entail unauthorised access to and breaches of a raw database that contains sensitive

data (this is a security violation). With privacy in general the challenge is that only a surprisingly small number of seemingly unimportant facts about an individual in a large dataset is enough to identify him or her among the millions of individuals appearing in the dataset. In addition, we are concerned with the fact that specific details of an individual can be determined from the release of the neural network that is trained on the data. Consider for instance the well-known 2006 case of Netflix. The company publicly released the Netflix Prize dataset containing more than 100 million movie recommendations of almost half a million of its users as part of a competition. The goal of the prize competition was for researchers to come up with better algorithms that Netflix could use to improve their movie recommendation system. Cognisant of the Video Privacy Protection Act, Netflix replaced all user identifiers with unique but randomly selected IDs. Within a short period however, researchers from the University of Texas presented a de-anonymisation methodology and showed how to de-anonymise movie-viewing records in Netflix's dataset[40].

Even if an individual cannot be *uniquely* identified from a small set of seemingly trivial facts in a large dataset, this does not mean that no privacy violation occurs. For instance, consider the case where knowing some trivial facts about an individual leads a researcher to conclude that the specific individual is a member of a (possibly large) subset of the original dataset that all share some unique property (a certain disease for instance). Even though the individual's unique record cannot be pulled from the data, knowing that he/she has a certain disease is clearly a privacy violation. It has been shown that this can even be done if the data under consideration only contain aggregate data, for instance, when it does not contain *any* individual data at all[41].

In the case of bulk datasets used by defence organisations, similar privacy considerations can arise. It is claimed that large-scale private personal data collection efforts for counterterrorism (such as those revealed by Snowden in 2013) led to privacy violations[42]. Others suggest that bulk dataset collection is legitimate *on balance* because of its operational utility[43]. It is argued that the extent to which privacy matters for a given bulk communications dataset depends on whether the individuals in the dataset are foreigners or not and whether the data consists of mere meta-data only or also includes content (for instance mail messages). European law and international human rights treaties state that these are privacy intrusions, which might be overridden only in the interest of national security and serious crime[44]. European law states that if someone loses *control* of his/her data without their consent this is a privacy violation[45]. Both in the US and the UK, the Snowden revelations led to a reconsideration of bulk collection by the intelligence services and the adoption of a more stringent approach.

Overall, when collecting and mining bulk data, the MoD is faced with a trade-off between the desired or required operational utility based on algorithmic

performance, which contributes to meeting its core values such as security and safety, and living up to moral and legal obligations to protect personal privacy. To be able to make informed decisions on this matter, a clear perspective is required on the performance of algorithms applied to those datasets (and the value they provide) as a function of the level of privacy that they guarantee. To gain some insight into this trade-off we conceptually present several privacy algorithmic designs in the next section and their corresponding plots in the value mapping (Figure 11.1).

11.3. Technical dimension of the trade-off

One way to tackle the trade-off between privacy and algorithmic performance is by technical innovation, i.e., designing algorithms that allow for privacy protection. However, although certain technical solutions may be successful in addressing the privacy issue, this comes at a price, namely a reduced performance of the algorithm (and hence its corresponding values). To illustrate this point, this subsection discusses non-exhaustively several solutions that are currently available. By plotting the technical solutions in the value mapping, we compare the trade-offs between the values that are offered by the different solutions.

In the extreme case, technical solutions to the trade-off would entail that we give up our concern for one of the rivalling values completely. This implies that we would either give up all considerations for privacy to reap maximum algorithmic performance (T_{max} in Figure 11.2) or refrain from using the technology altogether and attain maximum privacy (T_{none}). In practice, however, these solutions are not desirable or acceptable and a more nuanced position in which both values can be attained simultaneously is sought after.

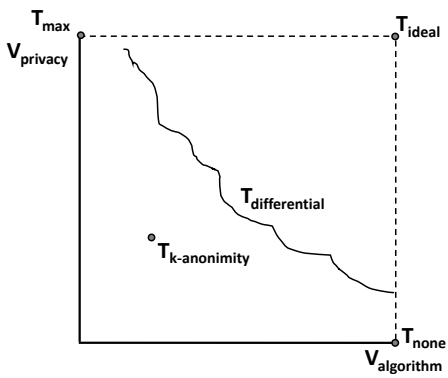

Fig. 11.2. A conceptual plot of different technical solutions to the value dilemma.

A generic solution to deal with privacy issues in big data is offered by the Privacy by Design (PBD)[46]. PBD consists of seven principles that ensure privacy (mitigation) measures are embedded throughout the design process[47], for example, requiring the design to be proactive and not reactive, privacy as the default setting and end-to-end security. An important principle of PBD that we will focus on here is the demand that privacy is to be embedded into the design itself and thus becomes an integral part of the technology.

A common way to meet this principle is by using anonymisation. This mitigates the risk of disclosing personally identifiable information (PII), for example, by removing all or a selection of such information, aggregation of data or masking PII in data[48].

One such approach that was first introduced by Samarati and Sweeney[49] is called *k*-anonymity. The idea behind it is that the information of each person in the data cannot be distinguished from *k-1* others in the data. However, later critiques of *k*-anonymity showed that neither *k*-anonymity nor enhancements of it are entirely successful in preventing privacy leakage while keeping a reasonable data utility level[50]. This can easily be seen in the case where the data consists of medical records. Even if it satisfies *k*-anonymity, it might still be possible that all *k* individuals that a certain individual of interest belongs to, have the same disease, i.e., one can still infer highly private information of this individual even though their data is privacy protected by *k*-anonymity ($T_{\text{k-anonimity}}$).

Another innovative method that is gaining attention is a probabilistic anonymising technique called differential privacy[51]. This technology ($T_{\text{differential}}$) is especially fit for the current discussion because it allows for a gradual increase or decrease of the level of privacy[52], which helps us to illustrate how technical choices affect the level of realisation of moral values.

In the 2000s, the mathematical concept of differential privacy was developed by a group of computer scientists to solve the privacy challenges sketched in the second part of this chapter. The basic idea of differential privacy is that the deletion of a specific individual's data record from a large dataset does not influence the outcome of algorithmic analysis of the data very much. It is interesting to note that Google and Apple started using differential privacy on some of their applications after the first decade of the 2000s.[53] Differential privacy has an exact mathematical definition, and the amount of privacy can be 'tuned' in the sense that the stricter the privacy settings the less likely that *any* deviating analytical outcome will occur when a person's records are not included in the data compared to when they are included[54]. Thus, differential privacy guarantees for instance that the probability of getting annoying telemarketing calls does not increase when you agree to include your data in a study. It promises that the probability of getting picked out of line at an airport does not increase by very much if you allowed your past travel records

to be stored. Additionally, differential privacy ensures that *for every* individual in a dataset the beliefs on anything that an analyst, who runs an experiment on the data, holds are very close to when the same analyst runs the same experiment on the data with the records of *any* individual removed. Put differently: differential privacy ensures that if someone queries the output of a differentially private algorithm to search if individual X is in the dataset, he/she will get the same result notwithstanding if individual X is included in the dataset or not. Because differential privacy provides a 'tuning knob' for the amount of privacy that is required in a given setting (based on exact mathematical arguments), $T_{differential}$ is plotted in Figure 11.2 as a curve instead of as a point. Note that the exact form of this curve is not known and depends on the particulars of the situation at hand (type of algorithm, dataset structure, etc.). However, it is clear that $T_{differential}$ is expected to be monotonically decreasing in every setting, as an increase in privacy will never yield better algorithmic performance. Hence, ethical restrictions on algorithms do not come for free.

As a consequence, the level of privacy protection implemented in an algorithm is not solely a technical challenge but above all a socio-political decision. In the next section, therefore, we will take a conceptual look at the influence of non-technical privacy demands and the corresponding implications as illustrated by the value map.

11.4. Socio-political dimension of the trade-off

As discussed in the previous section, opting for a particular data analytical solution entails a particular trade-off between the values associated with algorithmic performance such as security and safety, and the value of privacy. However, even if a particular combination of these values is practically attainable it may not be socially desirable or acceptable from a legal or moral perspective.

On the one hand, national and transnational regulation and legislation set a 'hard' limit that may not be crossed by the value of privacy. For instance, the Dutch Data Protection Authority (DPA) supervises the processing of personal data to ensure compliance with laws that regulate the use of personal data. The tasks and powers of the DPA are described in the General Data Protection Regulation (GDPR), supplemented by the Dutch Implementation Act of the GDPR[55].

On the other hand, national and organisational culture, sometimes voiced by public opinion or civil society organisations (CSOs) reflects (partly overlapping) moral standards that set further demands on the level of privacy that is considered (socially) acceptable. For instance, a survey of public opinion in the US by RAND indicates that citizens are concerned about the ethical risks of military use of AI[56]. Likewise, the Dutch digital rights organisation Bits of Freedom, for example, aims

to safeguard online freedom by voicing privacy considerations in the media[57]. Additionally, on an organisational level, nowadays, the code of conduct of many large companies typically includes a statement on the handling of personal data.

The choice for a technical solution that 'solves' the trade-off dilemma thus not only depends on the technological feasibility of meeting the values but is also limited by socio-political factors such as public opinion, (organisational) culture and legislative regimes. The socio-political factors can therefore be interpreted to set a threshold value ($\underline{V}$) that represents the minimum level of the value to meet the legal requirement and be morally and socially acceptable.

Technical solutions that are located in the coordinate system below the threshold may be technically feasible but are considered unlawful and/or not considered acceptable socially and morally. In the fictitious case plotted in Figure 11.3, for example, $T_{k\text{-anonimity}}$ is plotted below the level of privacy that generally would be considered acceptable in the Netherlands ($\underline{V}_{Netherlands}$). The use of this solution therefore would be ruled out in the Netherlands because it does not guarantee the level of privacy as required by $\underline{V}_{Netherlands}$, despite its algorithmic performance.

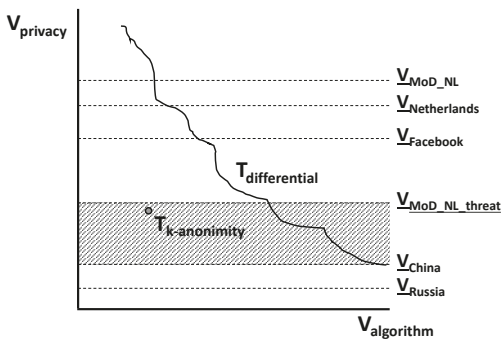

Fig. 11.3. A conceptual plot of different value thresholds.[58]

The conceptual privacy threshold value of the Netherlands differs from that of other countries due to cultural differences that also carry over to legislation and policy. Although all EU member states adhere to the GDPR, supervision and accompanying interpretations of the GDPR are the responsibility of national authorities.

In a similar vein, private and public organisations have a degree of freedom in interpreting and complying with legal and regulatory demands[59]. Organisational culture, therefore, affects the way companies address normative issues such as privacy. Alphabet Inc. and Facebook ($\underline{V}_{Facebook}$), for instance, intently push the boundaries of what is permissible in terms of privacy and sometimes even cross them based on the hefty fines recently imposed by France[60].

In contrast, due to the nature of the environment it operates in, the Dutch MoD may opt for a more cautious approach toward the value of privacy. Apart from being a public organisation, the MoD is tasked with national security and has a monopoly on the use of force[61]. As a consequence, it is under close public and political scrutiny[62]. What is more, due to this special position, the culture and tradition of the MoD aspire to abide by a higher moral standard than is customary in society in general[63]. Based on this, the argument can be made that the MoD's threshold ($\underline{V}_{MoD\_NL}$) lies above that of $\underline{V}_{Netherlands}$.

At the same time, however, because of its tasks and operational requirements the MoD functions across a unique spectrum of operations – e.g. using violence on behalf of the state- in comparison with any other public or private organisation[64]. Because, as argued above, MoD's use of advanced technology is increasingly considered necessary to meet the values of security and safety[65], this also may carry over to the deployment of complex algorithms. This does not imply, however, that deployment is not bound by legal and moral restrictions. The legitimacy of taking extreme measures still depends on the general public at home and the international community, and endorsement by law and human rights thinking[66].

Nevertheless, in extraordinary circumstances, such as situations that represent a threat to national security, the threshold of the MoD ($\underline{V}_{MoD\_NL\_threat}$) may be set lower than in 'normal' – i.e. peacetime and routine conditions[67]- circumstances ($\underline{V}_{MoD\_NL}$), to allow an increase in the algorithmic performance. This is conceptually illustrated by the fictitious situation displayed in Figure 11.3, where a lower threshold allows for a higher algorithmic performance of differentially private algorithms as presented by the monotonically decreasing curve $T_{Differential}$.

Finding the right solution to the trade-off dilemma given certain circumstances can be precarious. A recent experiment by the Dutch land forces with bulk data exemplifies this. At the start of the COVID-19 pandemic, the Land Information Manoeuvre Centre (LIMC) experimented with analysing action groups by collecting and analysing data in times of crisis[68]. After attention by the media about its alleged illegitimacy, followed by political debate, the Minister of Defence decided to decommission this experiment[69].

This resonates with the broader debate on the mandate of civil services operating in the intelligence domain that is also centred around the privacy versus security argument[70]. A public consultation by the Ministry of Internal Affairs about the proposed intelligence law, for example, raised sensitivities that are felt about the necessity and proportionality of technological deployment, and the technical impossibility of compliance[71].

In the military domain, finding a solution to the trade-off dilemma is further exacerbated as increased algorithmic performance is expected to positively correlate with military power[72]. In meeting the values of safety and security, the MoD

competes with adversaries that have a different culture and political system, and hence abide by different standards concerning moral values such as privacy. This lowers their privacy threshold with regards to the MoD's and hence allows their algorithms to potentially perform better. To further substantiate this point we will look into the position of China and Russia based on investigations by the RAND Corporation and (professional) online media organisations.

Both China and Russia invest heavily in the use of AI in the military domain. The Chinese People's Liberation Army (PLA) is aggressively developing AI applications such as bulk data analytics to improve its position to win military conflicts in the future[73]. The main areas of interest of bulk data use by the PLA include command & control, surveillance & reconnaissance and cybersecurity[74], but also enabling rapid operational decision-making, akin to IGO in the Netherlands[75]. Equally, Russia regards AI as essential for the future of the military and is devoting much energy to catching up with the US and other competitors[76]. Like China, Russia has taken initiatives on both the tactical and strategical levels, for example, aimed at improving command & control, decision-making, training and procurement[77].

In addition, Russia and China share a lack of sensitivity to ethical issues such as privacy compared to countries in the West[78]. Acting as a precedent, both countries have a bad track record with regards to privacy and their own citizens, using technology to censor behaviour and speech, and curtail counter-regime activities[79].

However, similar to Western countries such as the Netherlands, privacy concerns are being voiced by the Chinese public[80]. In fact, China has recently issued stringent privacy protection standards that are akin to those of Western countries[81]. However, these restrictions on data use are aimed at private enterprises and do not affect the PLA[82]. Moreover, at the same time, this new legislation has increased the power of the government to access and control private data both in China and overseas via Chinese firms[83].

In contrast, in Russia evidence for these mitigating factors is less present. Conversely, to compensate for lagging behind in the AI arms race, Russia may 'feel that ethics is a luxury it cannot afford' and be more inclined to use its technology more aggressively[84]. In addition, the weakening position of the Putin faction may increase the risk tolerance for unlawful and immoral applications[85].

Summarising, it appears that adversaries of the Dutch MoD are more willing to sacrifice privacy to attain higher algorithmic performance than Western democracies. It can be expected that the corresponding thresholds set by China ($\underline{V}_{China}$) and Russia ($\underline{V}_{Russia}$) are significantly lower than those of the Netherlands and other Western countries (see Figure 11.3). Moreover, although the societal threshold of China evidently is higher than that of Russia, this does not seem to apply to the PLA.

The area between $\underline{V}_{MoD}$ and the thresholds of the adversaries represents the opportunity cost of the Dutch MoD in terms of algorithmic performance due to

its legal and moral standards. Depending on what level the algorithms are being used, these costs have to be paid out strategically, tactically and/or operationally. This raises the question of to what extent and under which circumstances Western countries are willing and can afford this potential and conceptual loss of algorithmic performance and hence of meeting the core values of security and safety?

Obviously, there are no simple answers to the trade-off dilemma between privacy and the values presented by algorithmic performance. Solutions will require due consideration and a deeper insight into all of the factors involved: technological opportunities, (moral) values, laws and regulation, public opinion, etc. Consequently, several disciplines need to be consulted such as law, engineering, ethics and the social sciences.

11.5. Conclusion

In this chapter, we explored the trade-off between privacy and security/ safety that the MoD faces when designing and implementing complex algorithms that aid in the analysis of large and complex datasets. Using a conceptual two-dimensional coordinate system with the value of privacy on the vertical axis and the value associated with algorithmic performance on the horizontal allowed us to clarify the technical, moral and socio-political dimensions of this challenge and the interwovenness of these dimensions.

First, we plotted and discussed several trade-off solutions and discussed some of their consequences. Trivially, for instance, maximum privacy yields zero algorithmic performance and hence reduces the value of safety and security to a large degree. Additionally, we showed that algorithmic design solutions to the trade-off dilemma exist – such as differential privacy – and can provide a 'tuning knob' for the amount of privacy that is desired whilst at the same time providing a certain amount of algorithmic performance. However, when one takes the socio-political context into account, only certain subsets of the possible solutions are allowed. In turn, these possible subsets of solutions are country-dependent as moral values vary around the globe. To illustrate this, we conceptually presented the cases of western countries versus China and Russia. Which solution to the trade-off is picked depends on a complex of social, political and societal factors and can have serious consequences for military power and effectiveness with respect to potential adversaries. To get agreement on the political and social mandate, it is not only recommended to involve experts such as lawyers, engineers, data scientists and foreign policy experts but also to reach out and engage with relevant stakeholders such as the general public, CSOs, military staff & troops and policy-makers. As a consequence, further investigation and eventually tackling this challenge in practice requires an interdisciplinary approach.

Notes

1 Alexandra Chouldechova and Aaron Roth, "A snapshot of the frontiers of fairness in machine learning", *Communications of the ACM* 63, no. 5 (2020): 82–89.

2 Jonathon Stewart, Peter Sprivulis, and Girish Dwivedi, "Artificial intelligence and machine learning in emergency medicine", *Emergency Medicine Australasia* 30, no. 6 (2018): 870–74.

3 Kanadpriya Basu et al., "Predictive models of student college commitment decisions using machine learning", *Data* 4, no. 2 (2019): 1–18.

4 Siddharth Bhatore, Lalit Mohan, and Y. Raghu Reddy, "Machine learning techniques for credit risk evaluation: a systematic literature review", *Journal of Banking and Financial Technology* 4, no. 1 (2020): 111–38.

5 Botambu Collins et al., 'A Survey on Forecasting Models for Preventing Terrorism', in *Advanced Computational Methods for Knowledge Engineering: Proceedings of the 6th International Conference on Computer Science, Applied Mathematics and Applications, ICCSAMA 2019*, vol. 1121 (Springer Nature, 2019), 323.

6 V. S. Subrahmanian et al., *Computational Analysis of Terrorist Groups: Lashkar-e-Taiba: Lashkar-e-Taiba* (New York: Springer, 2013); V. S. Subrahmanian et al., *A Machine Learning Based Model of Boko Haram* (Cham: Springer, 2021).

7 C. V. Amrutha, C. Jyotsna, and J. Amudha, "Deep Learning Approach for Suspicious Activity Detection from Surveillance Video", in *2020 2nd International Conference on Innovative Mechanisms for Industry Applications (ICIMIA)* (IEEE, 2020), 335–39.

8 Marina Danchovsky Ibrishimova and Kin Fun Li, "A machine learning approach to fake news detection using knowledge verification and natural language processing", *Advances in Intelligent Networking and Collaborative Systems.*, ed. L. Barolli, H. Nishino, and H. Miwa, vol. 1035 (International Conference on Intelligent Networking and Collaborative Systems, Cham: Springer, 2020), 223–34.

9 Stanislav Abaimov and Maurizio Martellini, 'Artificial Intelligence in Autonomous Weapon Systems', in *21st Century Prometheus*, ed. M. Martinelli and R. Trapp (Cham: Springer, 2020), 141–77.

10 Michelle Cayford and Wolter Pieters, "The effectiveness of surveillance technology: what intelligence officials are saying", *The Information Society* 34, no. 2 (2018): 88–103.

11 Hugo M. Verhelst, A. W. Stannat, and Giulio Mecacci, "Machine learning against terrorism: how big data collection and analysis influences the privacy-security dilemma", *Science and Engineering Ethics* 26, no. 6 (2020): 2975–84.

12 Glenn Greenwald, *No Place to Hide: Edward Snowden, the NSA, and the US Surveillance State* (New York, N.Y.: Metropolitan Books, 2014); Samuel A. Morgan, 'Security vs. Liberty: How to Measure Privacy Costs in Domestic Surveillance Programs' (Monterey, CA, Naval Postgraduate School, 2014).

13 Kiana Alikhademi et al., "A review of predictive policing from the perspective of fairness", *Artificial Intelligence and Law*, 2022, 1–17.

14 Plixavra Vogiatzoglou, "Mass surveillance, predictive policing and the implementation of the CJEU and ECtHR requirement of objectivity", *European Journal of Law and Technology* 10, no. 1 (2019): 1–18.

15 Ministry of Internal Affairs, 'Wet op de inlichtingen- en veiligheidsdiensten 2017' (2017), https://wetten.overheid.nl/BWBR0039896/2022-05-01.

16 R. Jones-Bos et al., 'Wet op inlichtingen- en veiligheidsdiensten 2017' (Den Haag: Rijksoverheid, 2021).

17 There are many ways for evaluating the performance of an algorithm. With respect to machine learning classifiers for instance common metrics are: accuracy, recall, precision, F1 score and the Matthews Correlation Coefficient, to name just a few.

[18] M.J. van den Hoven, G.J. Lokhorst, and I. van de Poel, "Engineering and the problem of moral overload", *Science and Engineering Ethics* 18, no. 1 (1 May 2011): 143–55, https://doi.org/10.1007/s11948-011-9277-z.

[19] T. Kuran, 'Moral Overload and Its Alleviation.', in *Economics, Values, Organization*, ed. A. Ben-Ner and L. Putterman (Cambridge: Cambridge University Press, 1998), 231–66.

[20] Based on van den Hoven, Lokhorst, and van de Poel, 'Engineering and the Problem of Moral Overload'.

[21] Dutch Ministry of Defence, 'The Dutch Ministry of Defence: Protecting What We Value – Tasks and Future – Defensie.Nl', onderwerp (Ministerie van Defensie, 2 September 2016), https://english.defensie.nl/topics/tasks-and-future/corporate-story.

[22] Dutch Ministry of Defence.

[23] Dutch Ministry of Defence, 'Defensievisie 2035. Vechten Voor Een Veilige Toekomst' (The Hague: Ministry of Defence, 2020), 3, https://open.overheid.nl/repository/ronl-cf4bd18b-15e0-4eff-9803-ca88a86e1122/1/pdf/defensievisie-2035-vechten-voor-een-veilige-toekomst.pdf.

[24] Dutch Ministry of Defence, 29.

[25] Dutch Ministry of Defence, 23.

[26] Dutch Ministry of Defence, 40.

[27] Dutch Ministry of Defence, 23.

[28] Dutch Ministry of Defence, 23.

[29] Paulo Shakarian, V. Subrahmanian, and Maria Luisa Spaino, 'SCARE: A Case Study with Baghdad', in *Proceedings of the Third International Conference on Computational Cultural Dynamics. AAAI*, 2009, 1–9.

[30] Subrahmanian et al., *Computational Analysis of Terrorist Groups: Lashkar-e-Taiba: Lashkar-e-Taiba*.

[31] Kevin T. Greene et al., "Understanding the timing of Chinese border incursions into India", *Humanities and Social Sciences Communications* 8, no. 1 (2021): 1–8.

[32] Haipeng Chen et al., "Disclose or exploit? A game-theoretic approach to strategic decision making in cyber-warfare", *IEEE Systems Journal* 14, no. 3 (2020): 3779–90.

[33] I.A. Andrews, S. Kumar, and F. Spezzano, 'SPINN: Suspicion Prediction in Nuclear Networks.' *IEEE International Conference on Intelligence and Security Informatics* (ISI). (2015), 19–24.

[34] Standard classifiers as Random Forest, Gaussian Naïve Bayes and support vector machines were used and analysed using the Matthews Correlation Coefficient.

[35] Christopher B. Davison et al., "Data privacy in the age of big data analytics", *Issues in Information Systems* 22, no. 2 (2021): 177–86.

[36] J.F.C. Timmermans, R. Heersmink, and M.J. van den Hoven, 'Normative Issues Report', Deliverable (FP-7 ETICA project, 2010), https://www.etica-project.eu/deliverable-files.

[37] M.J. Van Den Hoven, 'Information Technology, Privacy and the Protection of Personal Data', in *Information Technology and Moral Philosophy*, ed. M.J. van Den Hoven and J. Weckert (Cambridge ; New York, NY: Cambridge University Press, 2008), 301–21.

[38] Jeroen van den Hoven et al., 'Privacy and Information Technology', in *The Stanford Encyclopedia of Philosophy*, ed. Edward N. Zalta, Summer 2020 (Metaphysics Research Lab, Stanford University, 2020), https://plato.stanford.edu/archives/sum2020/entries/it-privacy/.

[39] R.S. Rosenberg, "The social impact of intelligent artefacts", *AI and Society* 22, no. 3 (2008): 367–83, https://doi.org/10.1007/s00146-007-0148-8.

[40] Arvind Narayanan and Vitaly Shmatikov, 'Robust de-anonymization of large sparse datasets', in *2008 IEEE Symposium on Security and Privacy (Sp 2008)* (2008 IEEE Symposium on Security and Privacy (sp 2008), IEEE, 2008), 111–25, https://doi.org/doi:10.1109/SP.2008.33.

41 Nils Homer et al., "Resolving individuals contributing trace amounts of DNA to highly complex mixtures using high-density SNP genotyping microarrays", *PLoS Genetics* 4, no. 8 (2008): e1000167.

42 David Lyon, "Surveillance, Snowden, and big data: capacities, consequences, critique", *Big Data & Society* 1, no. 2 (2014): 1–13.

43 David WK Anderson, *Report of the Bulk Powers Review* (London: HMSO, 2016).

44 Anderson.

45 Paul De Hert, "Identity management of E-ID, privacy and security in Europe. a human rights view.", *Nformation Security Technical Report* 13, no. 2 (2008): 71–75.

46 Giuseppe D'Acquisto et al., "Privacy by design in big data: an overview of privacy enhancing technologies in the era of big data analytics", *ArXiv Preprint ArXiv:1512.06000*, 2015.

47 Ann Cavoukian, 'Privacy by Design The 7 Foundational Principles' (Canada, 2009).

48 Christopher Graham, "Anonymisation: Managing Data Protection Risk Code of Practice", *Information Commissioner's Office*, 2012.

49 Pierangela Samarati and Latanya Sweeney, 'Protecting Privacy When Disclosing Information: K-Anonymity and Its Enforcement through Generalization and Suppression', 1998.

50 Josep Domingo-Ferrer and Vicenç Torra, "A Critique of K-Anonymity and Some of Its Enhancements", in *2008 Third International Conference on Availability, Reliability and Security* (IEEE, 2008), 990–93.

51 Cynthia Dwork and Aaron Roth, "The algorithmic foundations of differential privacy.", *Found. Trends Theor. Comput. Sci.* 9, no. 3–4 (2014): 211–407.

52 Cynthia Dwork, "Differential Privacy: A Survey of Results", *International Conference on Theory and Applications of Models of Computation* (Springer, 2008), 1–19.

53 Jun Tang et al., "Privacy loss in Apple's implementation of differential privacy on Macos 10.12", *ArXiv Preprint ArXiv:1709.02753*, 2016, 1–12.

54 Dwork, 'Differential Privacy: A Survey of Results'.

55 DPA, 'Tasks and Powers of the Dutch DPA', Autoriteit Persoonsgegevens, 2022, https://www.autoriteitpersoonsgegevens.nl/en/about-dutch-dpa/tasks-and-powers-dutch-dpa.

56 Forrest E. Morgan et al., 'Military Applications of Artificial Intelligence: Ethical Concerns in an Uncertain World' (RAND Corporation, 28 April 2020), xiv, https://www.rand.org/pubs/research_reports/RR3139-1.html.

57 Bits of Freedom, 'Bits of Freedom: Voor jouw internetvrijheid', Bits of Freedom, 2022, https://www.bitsoffreedom.nl/.

58 Note that the given thresholds represent an estimation of their relative rather than absolute positions for the sake of the argument.

59 Job Timmermans, 'Exploring the Multifaceted Relationship of Compliance and Integrity—The Case of the Defence Industry', in *NL ARMS Netherlands Annual Review of Military Studies 2021*, ed. Robert Beeres et al., NL ARMS (The Hague: T.M.C. Asser Press, 2022), 95–113, https://doi.org/10.1007/978-94-6265-471-6_6.

60 Siladitya Ray, 'Google And Facebook Hit With $238 Million Fines In France Over Privacy Violations', Forbes, accessed 19 May 2022, https://www.forbes.com/sites/siladityaray/2022/01/06/google-and-facebook-hit-with-238-million-fines-in-france-over-privacy-violations/.

61 Joseph Soeters, Paul C. van Fenema, and Robert Beeres, 'Introducing Military Organizations', in *Managing Military Organizations: Theory and Practice*, ed. Joseph Soeters, Paul C. van Fenema, and R.J.M. Beeres (New York, N.Y.: Routledge, 2010), 1–16.

62 Militaire Spectator, "De grenzen van veiligheid", *Militaire Spectator* 183, no. 3 (8 March 2014): 102–3; Peter Olsthoorn, Marten Meijer, and Desiree Verweij, 'Managing Moral Professionalism in Military Operations', in *Managing Militant Operations. Theory and Practice*, ed. Joseph Soeters and Paul C. van Fenema (New York, N.Y.: Routledge, 2010), 138–49.

63 J.F.C. Timmermans et al., "Dertien Jaar Integriteitsonderzoek: Een Zoektocht Naar Defensie-Identiteit. Van Incident Naar Onderzoek Naar Incident", *Militaire Spectator* 2 (2021): 72–83.

64 Soeters, van Fenema, and Beeres, 'Introducing Military Organizations'.

65 Dutch Ministry of Defence, 'Defensievisie 2035. Vechten Voor Een Veilige Toekomst'.

66 Soeters, van Fenema, and Beeres, 'Introducing Military Organizations'.

67 Soeters, van Fenema, and Beeres.

68 Dutch Ministry of Defence, 'Land Information Manoeuvre Centre helpt Defensie anticiperen – Nieuwsbericht – Defensie.nl', nieuwsbericht (Ministerie van Defensie, 16 November 2020), https://www.defensie.nl/actueel/nieuws/2020/11/16/land-information-manoeuvre-centre-helpt-defensie-anticiperen; Esther Rosenberg and Karel Berkhout, 'Alle alarmbellen gingen af, maar waarom greep dan niemand in bij Defensie?', *NRC*, 1 November 2021, https://www.nrc.nl/nieuws/2021/11/01/waarom-greep-niemand-in-bij-defensie-a4063866.

69 Dutch Ministry of Defence, 'Kamerbrief over uitvoeren moties Land Information Manoeuvre Centre – Kamerstuk', kamerstuk (The Hague: Ministry of Defence, 26 October 2021), https://www.rijksoverheid.nl/documenten/kamerstukken/2021/10/26/kamerbrief-over-uitvoeren-moties-land-information-manoeuvre-centre; Rosenberg and Berkhout, 'Alle alarmbellen gingen af, maar waarom greep dan niemand in bij Defensie?'

70 A. Claver, "The big data paradox: juggling data flows, transparency and secrets", *Militaire Spectator* 187, no. 6 (2018): 309–23.

71 Claver.

72 Morgan et al., 'Military Applications of Artificial Intelligence', xiv.

73 Morgan et al., 81; Derek Grossman et al., 'Chinese Views of Big Data Analytics' (RAND Corporation, 1 September 2020), 27, https://www.rand.org/pubs/research_reports/RRA176-1.html.

74 Grossman et al., 'Chinese Views of Big Data Analytics', viii–ix.

75 Morgan et al., 'Military Applications of Artificial Intelligence', 82.

76 Morgan et al., 83.

77 Andrew Eversden, 'A Warning to DoD: Russia Advances Quicker than Expected on AI, Battlefield Tech', C4ISRNet, 24 May 2021, https://www.c4isrnet.com/artificial-intelligence/2021/05/24/a-warning-to-dod-russia-advances-quicker-than-expected-on-ai-battlefield-tech/.

78 Morgan et al., 'Military Applications of Artificial Intelligence', xiv.

79 Alina Polyakova, 'Weapons of the Weak: Russia and AI-Driven Asymmetric Warfare', *Brookings* (blog), 15 November 2018, https://www.brookings.edu/research/weapons-of-the-weak-russia-and-ai-driven-asymmetric-warfare/; Grossman et al., 'Chinese Views of Big Data Analytics', 49.

80 Morgan et al., 'Military Applications of Artificial Intelligence', 79.

81 Grossman et al., 'Chinese Views of Big Data Analytics', 11, 81.

82 Morgan et al., 'Military Applications of Artificial Intelligence', 81.

83 Matt Pottinger and David Feith, 'Opinion | The Most Powerful Data Broker in the World Is Winning the War Against the U.S.', *The New York Times*, 30 November 2021, sec. Opinion, https://www.nytimes.com/2021/11/30/opinion/xi-jinping-china-us-data-war.html.

84 Morgan et al., 'Military Applications of Artificial Intelligence', 98–99.

85 Morgan et al., 98.

References

Abaimov, Stanislav, and Maurizio Martellini. 'Artificial Intelligence in Autonomous Weapon Systems'. In *21st Century Prometheus*, edited by M. Martinelli and R. Trapp, 141–77. Cham: Springer, 2020.

Alikhademi, Kiana, Emma Drobina, Diandra Prioleau, Brianna Richardson, Duncan Purves, and Juan E. Gilbert. "A review of predictive policing from the perspective of fairness". *Artificial Intelligence and Law*, 2022, 1–17.

Amrutha, C. V., C. Jyotsna, and J. Amudha. 'Deep Learning Approach for Suspicious Activity Detection from Surveillance Video'. In 2020 2nd International Conference on Innovative Mechanisms for Industry Applications (ICIMIA), 335–39. IEEE, 2020.

Anderson, David WK. Report of the Bulk Powers Review. London: HMSO, 2016.

Andrews, I.A., S. Kumar, and F. Spezzano. "SPINN: Suspicion Prediction in Nuclear Networks." *IEEE International Conference on Intelligence and Security Informatics* (ISI)., 19–24. IEEE, 2015.

Basu, Kanadpriya, Treena Basu, Ron Buckmire, and Nishu Lal. "Predictive models of student college commitment decisions using machine learning". *Data* 4, no. 2 (2019): 1–18.

Bhatore, Siddharth, Lalit Mohan, and Y. Raghu Reddy. "Machine learning techniques for credit risk evaluation: a systematic literature review". *Journal of Banking and Financial Technology* 4, no. 1 (2020): 111–38.

Bits of Freedom. 'Bits of Freedom: Voor jouw internetvrijheid'. Bits of Freedom, 2022. https://www.bitsoffreedom.nl/.

Cavoukian, Ann. 'Privacy by Design The 7 Foundational Principles'. Canada, 2009.

Cayford, Michelle, and Wolter Pieters. "The effectiveness of surveillance technology: what intelligence officials are saying". *The Information Society* 34, no. 2 (2018): 88–103.

Chen, Haipeng, Qian Han, Sushil Jajodia, Roy Lindelauf, V. S. Subrahmanian, and Yanhai Xiong. "Disclose or exploit? A game-theoretic approach to strategic decision making in cyber-warfare". *IEEE Systems Journal* 14, no. 3 (2020): 3779–90.

Chouldechova, Alexandra, and Aaron Roth. "A Snapshot of the Frontiers of Fairness in Machine Learning". *Communications of the ACM* 63, no. 5 (2020): 82–89.

Claver, A. "The big data paradox: juggling data flows, transparency and secrets". *Militaire Spectator* 187, no. 6 (2018): 309–23.

Collins, Botambu, Dinh Tuyen Hoang, HyoJeon Yoon, and Ngoc Thanh Nguyen. 'A Survey on Forecasting Models for Preventing Terrorism'. In: *Advanced Computational Methods for Knowledge Engineering: Proceedings of the 6th International Conference on Computer Science, Applied Mathematics and Applications*, ICCSAMA 2019, 1121:323. Springer Nature, 2019.

D'Acquisto, Giuseppe, Josep Domingo-Ferrer, Panayiotis Kikiras, Vicenç Torra, Yves-Alexandre de Montjoye, and Athena Bourka. "Privacy by Design in Big Data: An Overview of Privacy Enhancing Technologies in the Era of Big Data Analytics". ArXiv Preprint ArXiv:1512.06000, 2015.

Davison, Christopher B., Edward J. Lazaros, Jensen J. Zhao, Allen D. Truell, and Brianna Bowles. "Data privacy in the age of big data analytics". *Issues in Information Systems* 22, no. 2 (2021): 177–86.

De Hert, Paul. "Identity Management of E-ID, Privacy and Security in Europe. A Human Rights View." Nformation Security Technical Report 13, no. 2 (2008): 71–75.

Domingo-Ferrer, Josep, and Vicenç Torra. "A Critique of K-Anonymity and Some of Its Enhancements". In *2008 Third International Conference on Availability, Reliability and Security*, 990–93. IEEE, 2008.

DPA. 'Tasks and Powers of the Dutch DPA'. Autoriteit Persoonsgegevens, 2022. https://www.autoriteitpersoonsgegevens.nl/en/about-dutch-dpa/tasks-and-powers-dutch-dpa.

Dutch Ministry of Defence. 'Defensievisie 2035. Vechten Voor Een Veilige Toekomst'. The Hague: Ministry of Defence, 2020. https://open.overheid.nl/repository/ronl-cf4bd18b-15e0-4eff-9803-ca88a86e1122/1/pdf/defensievisie-2035-vechten-voor-een-veilige-toekomst.pdf.

———. 'Kamerbrief over uitvoeren moties Land Information Manoeuvre Centre – Kamerstuk'. Kamerstuk. The Hague: Ministry of Defence, 26 October 2021. https://www.rijksoverheid.nl/documenten/kamerstukken/2021/10/26/kamerbrief-over-uitvoeren-moties-land-information-manoeuvre-centre.

———. 'Land Information Manoeuvre Centre helpt Defensie anticiperen – Nieuwsbericht – Defensie.nl'. Nieuwsbericht. Ministerie van Defensie, 16 November 2020. https://www.defensie.nl/actueel/nieuws/2020/11/16/land-information-manoeuvre-centre-helpt-defensie-anticiperen.

———. 'The Dutch Ministry of Defence: Protecting What We Value – Tasks and Future – Defensie.Nl'. Onderwerp. Ministerie van Defensie, 2 September 2016. https://english.defensie.nl/topics/tasks-and-future/corporate-story.

Dwork, Cynthia. 'Differential Privacy: A Survey of Results'. In: International Conference on Theory and Applications of Models of Computation, 1–19. Springer, 2008.

Dwork, Cynthia, and Aaron Roth. "The algorithmic foundations of differential privacy." *Found. Trends Theor. Comput. Sci.* 9, no. 3–4 (2014): 211–407.

Eversden, Andrew. 'A Warning to DoD: Russia Advances Quicker than Expected on AI, Battlefield Tech'. C4ISRNet, 24 May 2021. https://www.c4isrnet.com/artificial-intelligence/2021/05/24/a-warning-to-dod-russia-advances-quicker-than-expected-on-ai-battlefield-tech/.

Graham, Christopher. 'Anonymisation: Managing Data Protection Risk Code of Practice'. Information Commissioner's Office, 2012.

Greene, Kevin T., Caroline Tornquist, Robbert Fokkink, Roy Lindelauf, and V. S. Subrahmanian. "Understanding the timing of Chinese border incursions into India". *Humanities and Social Sciences Communications* 8, no. 1 (2021): 1–8.

Greenwald, Glenn. *No Place to Hide: Edward Snowden, the NSA, and the US Surveillance* State. New York, N.Y.: Metropolitan Books, 2014.

Grossman, Derek, Christian Curriden, Logan Ma, Lindsey Polley, J. D. Williams, and Cortez A. Cooper III. "Chinese Views of Big Data Analytics". RAND Corporation, 1 September 2020. https://www.rand.org/pubs/research_reports/RRA176-1.html.

Homer, Nils, Szabolcs Szelinger, Margot Redman, David Duggan, Waibhav Tembe, Jill Muehling, John V. Pearson, Dietrich A. Stephan, Stanley F. Nelson, and David W. Craig. "Resolving individuals contributing trace amounts of DNA to highly complex mixtures using high-density SNP genotyping microarrays". *PLoS Genetics* 4, no. 8 (2008): e1000167.

Hoven, Jeroen van den, Martijn Blaauw, Wolter Pieters, and Martijn Warnier. 'Privacy and Information Technology'. In *The Stanford Encyclopedia of Philosophy*, edited by Edward N. Zalta, Summer 2020. Metaphysics Research Lab, Stanford University, 2020. https://plato.stanford.edu/archives/sum2020/entries/it-privacy/.

Hoven, M.J. van den, G.J. Lokhorst, and I. van de Poel. "Engineering and the problem of moral overload". *Science and Engineering Ethics* 18, no. 1 (1 May 2011): 143–55. https://doi.org/10.1007/s11948-011-9277-z.

Ibrishimova, Marina Danchovsky, and Kin Fun Li. 'A Machine Learning Approach to Fake News Detection Using Knowledge Verification and Natural Language Processing'. In *Advances in Intelligent Networking and Collaborative Systems*., edited by L. Barolli, H. Nishino, and H. Miwa, 1035:223–34. Cham: Springer, 2020.

Jones-Bos, R., L. van den Herik, B. Jacobs, W. Nagtegaal, and S.E. Zijlstra. 'Wet op inlichtingen- en veiligheidsdiensten 2017'. Den Haag: Rijksoverheid, 2021.

Kuran, T. 'Moral Overload and Its Alleviation.' In *Economics, Values, Organization*, edited by A. Ben-Ner and L. Putterman, 231–66. Cambridge: Cambridge University Press, 1998.

Lyon, David. "Surveillance, Snowden, and big data: capacities, consequences, critique". *Big Data & Society* 1, no. 2 (2014): 1–13.

Militaire Spectator. 'De grenzen van veiligheid'. Militaire Spectator 183, no. 3 (8 March 2014): 102–3.

Ministry of Internal Affairs. Wet op de inlichtingen- en veiligheidsdiensten 2017 (2017). https://wetten.overheid.nl/BWBR0039896/2022-05-01.

Morgan, Forrest E., Benjamin Boudreaux, Andrew J. Lohn, Mark Ashby, Christian Curriden, Kelly Klima, and Derek Grossman. "Military Applications of Artificial Intelligence: Ethical Concerns in an Uncertain World". RAND Corporation, 28 April 2020. https://www.rand.org/pubs/research_reports/RR3139-1.html.

Morgan, Samuel A. 'Security vs. Liberty: How to Measure Privacy Costs in Domestic Surveillance Programs'. Naval Postgraduate School, 2014.

Narayanan, Arvind, and Vitaly Shmatikov. 'Robust De-Anonymization of Large Sparse Datasets'. In: 2008 IEEE Symposium on Security and Privacy (Sp 2008), 111–25. IEEE, 2008. https://doi.org/doi:10.1109/SP.2008.33.

Olsthoorn, Peter, Marten Meijer, and Desiree Verweij. 'Managing Moral Professionalism in Military Operations'. In: *Managing Militant Operations. Theory and Practice*, edited by Joseph Soeters and Paul C. van Fenema, 138–49. New York, N.Y.: Routledge, 2010.

Polyakova, Alina. 'Weapons of the Weak: Russia and AI-Driven Asymmetric Warfare'. Brookings (blog), 15 November 2018. https://www.brookings.edu/research/weapons-of-the-weak-russia-and-ai-driven-asymmetric-warfare/.

Pottinger, Matt, and David Feith. 'Opinion | The Most Powerful Data Broker in the World Is Winning the War Against the U.S.' The New York Times, 30 November 2021, sec. Opinion. https://www.nytimes.com/2021/11/30/opinion/xi-jinping-china-us-data-war.html.

Ray, Siladitya. 'Google And Facebook Hit With $238 Million Fines In France Over Privacy Violations'. Forbes. Accessed 19 May 2022. https://www.forbes.com/sites/sidadityaray/2022/01/06/google-and-facebook-hit-with-238-million-fines-in-france-over-privacy-violations/.

Rosenberg, Esther, and Karel Berkhout. 'Alle alarmbellen gingen af, maar waarom greep dan niemand in bij Defensie?' NRC. 1 November 2021. https://www.nrc.nl/nieuws/2021/11/01/waarom-greep-niemand-in-bij-defensie-a4063866.

Rosenberg, R.S. "The social impact of intelligent artefacts". *AI and Society* 22, no. 3 (2008): 367–83. https://doi.org/10.1007/s00146-007-0148-8.

Samarati, Pierangela, and Latanya Sweeney. 'Protecting Privacy When Disclosing Information: K-Anonymity and Its Enforcement through Generalization and Suppression', 1998.

Shakarian, Paulo, V. Subrahmanian, and Maria Luisa Spaino. 'SCARE: A Case Study with Baghdad'. In *Proceedings of the Third International Conference on Computational Cultural Dynamics.* AAAI, 1–9, 2009.

Soeters, Joseph, Paul C. van Fenema, and Robert Beeres. 'Introducing Military Organizations'. In *Managing Military Organizations: Theory and Practice*, edited by Joseph Soeters, Paul C. van Fenema, and R.J.M. Beeres, 1–16. New York, N.Y.: Routledge, 2010.

Stewart, Jonathon, Peter Sprivulis, and Girish Dwivedi. 'Artificial intelligence and machine learning in emergency medicine'. *Emergency Medicine Australasia* 30, no. 6 (2018): 870–74.

Subrahmanian, V. S., Aaron Mannes, Amy Sliva, Jana Shakarian, and John P. Dickerson. *Computational Analysis of Terrorist Groups*: Lashkar-e-Taiba: Lashkar-e-Taiba. New York: Springer, 2013.

Subrahmanian, V. S., Chiara Pulice, James F. Brown, and Jacob Bonen-Clark. *A Machine Learning Based Model of Boko Haram.* Cham: Springer, 2021.

Tang, Jun, Aleksandra Korolova, Xiaolong Bai, Xueqiang Wang, and Xiaofeng Wang. "Privacy loss in Apple's implementation of differential privacy on Macos 10.12". *ArXiv Preprint ArXiv:*1709.02753, 2016, 1–12.

Timmermans, J.F.C., M.P. Bogers, R.M.M. Bertrand, and R.J.M. Beeres. 'Dertien Jaar Integriteitsonderzoek: Een Zoektocht Naar Defensie-Identiteit. Van Incident Naar Onderzoek Naar Incident'. Militaire Spectator 2 (2021): 72–83.

Timmermans, J.F.C., R. Heersmink, and M.J. van den Hoven. 'Normative Issues Report'. Deliverable. FP-7 ETICA project, 2010. https://www.etica-project.eu/deliverable-files.

Timmermans, Job. 'Exploring the Multifaceted Relationship of Compliance and Integrity—The Case of the Defence Industry'. In: *NL ARMS Netherlands Annual Review of Military Studies 2021,* edited by Robert Beeres, Robert Bertrand, Jeroen Klomp, Job Timmermans, and Joop Voetelink, 95–113. NL ARMS. The Hague: T.M.C. Asser Press, 2022. https://doi.org/10.1007/978-94-6265-471-6_6.

Van Den Hoven, M.J. 'Information Technology, Privacy and the Protection of Personal Data'. In: *Information Technology and Moral Philosophy*, edited by M.J. van Den Hoven and J. Weckert, 301–21. Cambridge ; New York, NY: Cambridge University Press, 2008.

Verhelst, Hugo M., A. W. Stannat, and Giulio Mecacci. "Machine learning against terrorism: how big data collection and analysis influences the privacy-security dilemma". *Science and Engineering Ethics* 26, no. 6 (2020): 2975–84.

Vogiatzoglou, Plixavra. "Mass surveillance, predictive policing and the implementation of the CJEU and ECtHR requirement of objectivity". *European Journal of Law and Technology* 10, no. 1 (2019): 1–18.