A Review of Value-Conflicts in Cybersecurity

An assessment based on quantitative and qualitative literature analysis

Markus Christen
University of Zurich

Bert Gordijn
Dublin City University

Karsten Weber
Brandenburg ***University*** of Technology

Ibo van de Poel
Delft University of Technology

Emad Yaghmaei
Delft University of Technology

Corresponding Author: Markus Christen christen@ethik.uzh.ch

Abstract

Cybersecurity is of capital importance in a world where economic and social processes increasingly rely on digital technology. Although the primary ethical motivation of cybersecurity is prevention of informational or physical harm, its enforcement can also entail conflicts with other moral values. This contribution provides an outline of value conflicts in cybersecurity based on a quantitative literature analysis and qualitative case studies. The aim is to demonstrate that the security-privacy-dichotomy—that still seems to dominate the ethics discourse based on our bibliometric analysis—is insufficient when discussing the ethical challenges of cybersecurity. Furthermore, we want to sketch how the notion of contextual integrity could help to better understand and mitigate such value conflicts.

Keywords: Cybersecurity, Moral Values, Value Conflicts, Privacy, Contextual Integrity

Introduction

The increasing use of information and communication technology (ICT) in all spheres of modern life makes the world a richer, more efficient and interactive place. However, it also increases its fragility as it reinforces our dependence on ICT systems that can never

be completely safe or secure. For example, it is difficult to keep industries and support systems functioning when there is a significant disruption of computer controls and monitors. Therefore, cybersecurity—the "collection of tools, policies, security concepts, security safeguards, guidelines, risk management approaches, actions, training, best practices, assurance and technologies that can be used to protect the cyber environment, organization and user's assets" (International Telecommunications Union, 2008, p. 3)—has become a matter of global interest and importance. Already more than 60 nations have officially published some form of strategy document outlining their official stance on cyberspace, cyber-crime, and/or cybersecurity (Klimburg, 2012; see also https://ccdcoe.org/strategies-policies.html for a more recent estimate). Given that the annual cost to the global economy from cybercrime alone is more than $400 billion (estimation from 2014; Center for Strategic and International Studies, 2014), the importance of cybersecurity is undisputed. Accordingly, one can observe in today's cybersecurity discourse an almost constant emphasis on an ever-increasing and diverse set of threat forms, ranging from basic computer viruses to cybercrime and cyberespionage activities, as well as cyber-terror and cyberwar (Dunn Caveltry, 2014).

This growing complexity of the digital ecosystem in combination with increasing global risks has created the following dilemma. Overemphasizing cybersecurity may violate fundamental values like equality, fairness, freedom, or privacy. On the other hand, neglecting cybersecurity could undermine citizens' trust and confidence in the digital infrastructure as well as in policy makers and state authorities. In order to increase our understanding of this dilemma, this contribution aims to provide an overview on value conflicts in cybersecurity based on a systematic review on the ethical, legal and technological literature on cybersecurity. In particular, we want to outline that a focus on the opposition of cybersecurity vs. privacy is insufficient for understanding the ethical complexity of cybersecurity. As cybersecurity affects ICT applications in all social domains, an exhaustive review of all value conflicts is not the goal of this paper. Instead, we will discuss some specific examples that go beyond the security-privacy dichotomy.

The contribution is structured as follows. First, we provide a bibliometric characterization of the cybersecurity domain since 1978 and we outline how major themes have developed with a focus on ethical topics. The bibliometric analysis is still ongoing, so we present a set of preliminary results. Second, we discuss in more detail specific examples of value conflicts. Third, based on the examples, we present a preliminary map that outlines the interrelation of (some of) the values that are involved. We also make a suggestion on how the framework of contextual integrity can help to mitigate possible value conflicts. Finally, in the conclusions, we discuss the relevance and limitations of the quantitative approach for creating a "value map" for cybersecurity and we outline the next step in our research endeavor.

Our paper aims to contribute to a better understanding of ethical issues in cyber-security. It emerges from a project funded by the European Commission (CANVAS—Constructing an AlliaNce for VAlue-sensitive cyberSecurity; see www.canvas-project.eu). CANVAS integrates the different perspectives of technology developers, legal and philosophical scholars as well as social scientists from the presupposition that technology development

in cybersecurity should be value-driven; i.e., should incorporate European values and fundamental rights.

A bibliometric characterization of cybersecurity

For our quantitative study, we identified the literature body on cybersecurity in two databases: *Web of Science Core Collection* (WoS) and *Scopus*.[1] We used the following methodology for identifying search keywords: In a first step, relying on literature, on the "keyword" function of *Scopus* (the database allows extracting keywords used in a publication body by the authors of publications) and on the cybersecurity experts in the CANVAS consortium,[2] we identified terms that characterize either general aspects of cybersecurity or specific aspects.[3] We also looked for terms that characterize ethical topics (see below) and terms that characterize the whole body of literature dealing with information technology, computation etc. Latter set is needed to relate the growth of the literature in the cybersecurity domain to an estimate of the overall growth of scientific literature in the domain in which cybersecurity is embedded. Otherwise, an identified growth could just be an indication of an increase in general publication activity. This first step generated 21 keywords groups of up to 92 keywords per group.

In a second step, we assessed the sensitivity of the keywords, i.e. their ability to discriminate as good as possible publications that deal with cybersecurity from other publications that deal with completely different topics. For doing this, for each group, we started with an unambiguous set of terms that likely have a clear reference to cybersecurity (e.g. "cybersecurity or "data security") and we checked how much each term contributes to the number of hits. Terms that had almost no impact on the number of hits (i.e. they contribute less than 0.1% of the detected publications) were excluded. Then, to the remaining set, we added terms that we considered being more ambiguous and we checked which publications additionally joined the set. If more than 20% of the additional publications clearly had no relation to cybersecurity (or the sub-group under investigation) based on a check of the title for the first 50 hits, the keyword was considered too unspecific. This was done sequentially for each keyword and for each

1 *Web of Science Core Collection* (WoS): https://apps.webofknowledge.com, the search was performed for the category "topics", which searches in title, abstract, author keywords and so-called 'keywords plus' (added by WoS); *Scopus*: http://www.scopus.com, the search was performed for the category "title-abstract-keywords". Date of the search: 22-29/01/2017.

2 The consortium consists of 11 academic and non-academic institutions as outlined on https://canvas-project.eu/canvas/index.php/what-is-canvas/consortium-partners/

3 We looked for keywords for the following specific aspects (only a selection of them is discussed in this contribution due to space restrictions): Authentication, Bots, Crime, Critical Infrastructure, Cryptography, Cyber Conflict, Distributed Denial of Service, Defense (means), Forensics, Hacking, Identity Theft, Malware (general and specific), Phising, Spam, Surveillance. This is not a comprehensive list of all issues relevant to cybersecurity. It includes only topics where one can expect to some degree keywords with sufficient specificity (e.g., the notion of "Integrity"—a common goal within cybersecurity—is too unspecific for a search).

aspect under consideration—latter has been done within the set of papers that contain cybersecurity keywords. In this way, the keyword sets CYBER[4] and ICT[5] were generated, as well as keyword-sets for the specific aspects (not outlined here due to space restrictions; some examples are described below).

In a first step, we determined the annual growth of the general cybersecurity literature body estimated by CYBER relative to the literature body estimated by ICT. The search was performed end of January 2017 (to ensure that the year 2016 is covered as completely as possible). The result is shown in Figure 1. The basic trend is reproduced in both databases, namely a first peak in 1983, a substantial increase starting in the 1990s and local maxima around 2005, 2009 and 2015. The databases may be further complemented with papers published in 2016 in the next weeks (due to differences in the database curation processes), so the results for the last year should be taken with a grain of salt. Overall, the database *Scopus* includes substantially more cybersecurity papers than *WoS* (266,343 vs. 78,446 papers). This is because *Scopus* includes more conference abstracts and technical journals compared to *WoS*. Such differences are an important reason that bibliometric arguments usually should rely on searches in more than one database.

Figure 1: Growth of the cybersecurity literature relative to the publication activity in general ICT, estimated for the databases *Scopus* and *Web of Science Core Collection* (WoS).

In order to better characterize the overall cybersecurity literature body, we looked also to the subject categories to which the papers are attributed.[6] Figure 2 shows the result of this analysis. We find that—compared to the overall ICT literature body—cybersecurity is

4 The Boolean search expression is: botnet* OR "computer crim*" OR "computer security" OR cryptography OR cyberattack OR "cyber attack" OR cyberconflict OR "cyber conflict" OR "cyber crim*" OR cyberdefense OR "cyber defense" OR cybersecurity OR "cyber-security" OR "cyber security" OR cyberterrorism OR "cyber terrorism" OR cyberthread* OR "cyber threat*" OR cyberwar* OR "cyber war*" OR "data leak*" OR "data security" OR "denial of service" OR DDoS OR firewall OR "hardware security" OR "information security" OR "internet security" OR "IT security" OR malware OR "mobile security" OR "network security" OR "non-repudiation" OR "security breaches" OR "security of data" OR "security requirement*" OR "security software" OR "security system*" OR "security threat*" OR "security vulnerabilit*" OR sigint OR "system security" OR "voting system" OR "web security"

5 The search term included all terms from the set CYBER and the Boolean expression: algorithm* OR "artificial intelligence" OR "big data" OR "communication system*" OR comput* OR cyber* OR "data mining" OR "data processing" OR "database system*" OR digital* OR "e-mail" OR hardware OR "information and communication technolog*" OR "information management" OR "information processing" OR "information retrieval" OR "information science" OR "information system*" OR "information technology" OR "intelligent system*" OR internet OR "machine learning" OR "mobile phone*" OR "network protocol*" OR robot* OR "search engine*" OR "signal processing" OR smartphone OR "social media" OR "social networking" OR software OR telecommunication OR "virtual reality" OR website* OR "world wide web"

6 The subject categories characterize the disciplinary background to which a journal is attributed in which a paper is published; a journal can be attributed to more than one subject category.

more strongly discussed in computer science and engineering; in particular, the application domains "science" (which includes physics, chemistry, geosciences, etc.) and "life sciences" are less prominently represented. Interestingly, the relative weight of social sciences and humanities papers is even a bit larger in the CYBER literature body compared to the ICT body. Furthermore, in both databases, the majority of papers are conference proceedings papers (*Scopus*: 58%, *WoS*: 56%), followed by journal articles (*Scopus*: 34%, *WoS*: 39%); the other categories are negligible.

Figure 2: Relative fraction of subject categories in the ICT and CYBER literature bodies estimated for the databases *Scopus* and *Web of Science Core Collection* (WoS).

We now present the result for a selection of the various aspects of cybersecurity investigated so far. We estimated the fraction of literature related to cryptography (CRYPTO[7]), cyberwar (WAR[8]), hacking (HACK[9]) and malware (MAL[10]) relative to CYBER, starting in 1991 (due to the low number of papers before the 1990s). The results are displayed in Figure 3. As a general remark, we see that the relative weight of these aspects within the set CYBER is generally larger in *WoS* compared to *Scopus*; the exact reason for this is currently under investigation by us. Qualitatively, however, the trends are similar in both databases. The CRYPTO literature body covers a considerable fraction of CYBER, however its relative importance decreases in time. In contrast, WAR, HACK and MAL show an increase, but at different moments in time. The discourse on hacking starts growing substantially in the 1990s, the discourse on malware in the early 2000s and the discourse on cyberwar in the late 2000s. More subtle differences between the two databases are currently under investigation.

7 Boolean expression: CYBER AND (certificate OR cryptanalysis OR cryptography OR decryption OR "digital signature" OR encryption OR "perfect forward secrecy" OR "public key infrastructure" OR "zero knowledge")

8 Boolean expression: CYBER AND ("advanced persistent threats" OR "cyber attack*" OR "cyber conflict*" OR "cyber defense" OR "cyber terrorism" OR "cyber threat*" OR "cyber war*" OR cyberattack* OR cyberconflict* OR cyberdefense OR cyberterrorism OR cyberthread* OR cyberwar* OR "digital sabotage" OR espionage OR "information warfare")

9 Boolean expression: CYBER AND ("amplification attack*" OR "black hat" OR "bot detection" OR botmaster OR botnet* OR "bot herder" OR "command and control server" OR "denial of service" OR DDoS OR fastflux OR hacker* OR hacking OR hacktivism OR "reflection attack*" OR spoofing OR "script kiddies" OR "white hat" OR "identity theft*" OR "identity hack" OR phisher* OR phishing OR spam* OR vishing)

10 Boolean expression: CYBER AND (adware OR backdoor OR "browser hijacker" OR "crypto-locker" OR "drive-by" OR dropper OR "exploit kit" OR keylogger OR malicious OR malspam OR malvertising OR malware* OR ransomware OR rootkit* OR scareware OR spyware OR trojan* OR virus* OR worm*)

Figure 3: Growth of selected cybersecurity topics estimated for the databases *Scopus* and *Web of Science Core Collection* (WoS).

Finally, we present the results for a selection of topics related to the ethics of cybersecurity. Here, we only compare the temporal development of papers that include the terminology of privacy (PRIVAT[11]) compared to papers that include a selection of other value terms (VALUE[12]). Here, we see for the first time conflicting results when comparing the two databases: whereas *Scopus* identifies a downward trend in the PRIVAT literature body, *WoS* identifies an upward trend. A first analysis indicates that this is not due to errors in the search procedure, but refers to differences between the two databases. We are currently investigating possible reasons for this conflicting result. The same holds for the peak in 2012 in *WoS* in the VALUE set, that is not present in *Scopus*. Overall, however, the set PRIVAT is still more than 4 (*WoS*) respectively more than 6 (*Scopus*) times larger than VALUE, indicating that the ethics of cybersecurity is still dominated by the privacy debate.

Figure 3: Growth of selected ethical aspects of cybersecurity estimated for the databases *Scopus* and *Web of Science Core Collection* (WoS).

In summary, the results provide a first step in a quantitative description of the cybersecurity literature body. The relevance of these findings for the overall scope of this paper will be discussed in the conclusion.

Paradigmatic value conflicts in cybersecurity

It is undisputed that cybersecurity has an ethical legitimation, because cybersecurity problems have ethical consequences such as economic damage due to loss of data (e.g. stolen data, a form of informational harm) but also physical harm (e.g. when critical physical systems are breached like the electricity grid). Furthermore, confidentiality breaches can violate intellectual property rights or privacy rights, i.e. lead to informational harm (Brey, 2007). However, while most cybersecurity measures aim at preventing harm, they at the same time can cause harm and violate human rights; for instance, by limiting personal freedom in order to counter cyber threats. Furthermore, some specific aspects of cybersecurity raise difficult ethical questions. Examples include the ethics of cyberwar; prominent topics here are to determine what counts as an act of war in cyberspace, how to deal with collateral damage and to determine the possible consequences of cyber-attacks (Rowe, 2008). Other scholars focus on the issue of "ethical hacking", thereby targeting the people who both instantiate the main threat for cybersecurity (by attacking vulnerable systems) and the main opportunity (by identifying weak spots in a system's security like so-called "zero days"). For example, Pike argues

11 Boolean expression: CYBER AND (privacy OR "data protection")

12 Boolean expression: CYBER AND (autonomy OR dignity OR equality OR fairness OR justice OR responsibility)

that the teaching of ethical hacking (being trained in how to support the protection of systems by knowing their weaknesses) is important to deal with current challenges but that one should be careful with the destructive nature of some of the skills that are taught (Pike, 2013). Related to this problem are ethical consequences of the political economy of cybersecurity. For instance, it is argued that the large sums of money that are paid for finding zero days in software can be incentives for unethical behavior (Egelman, Herley, & van Oorschot, 2013).

In order to provide a more structured introduction into value conflicts in cybersecurity, we will investigate examples of four domains in which ethical issues of cybersecurity can arise: business, health, politics as domains of application of cybersecurity technology, and general issues related to the design of such technologies.

Cybersecurity value conflicts in business: The question of how businesses ought to deal with cybersecurity can be analysed from different normative perspectives. Let us have a brief look at four examples. In 1962 Milton Friedman famously claimed that in a free economy "(…) there is one and only one social responsibility of business—to use its resources and engage in activities designed to increase its profits so long as it stays within the rules of the game, which is to say, engages in open and free competition, without deception or fraud" (Friedman, 1982, p. 133). In 1970 Friedman restated his idea in an influential New York Times article: The dominant obligation of managers, according to Friedman, was to look after the interests of their companies' shareholders whilst complying with legal requirements (Friedman, 1970).

In response to Friedman, other normative ideas were developed, stipulating moral responsibilities for business going beyond his fairly minimal requirements. One approach focused on sustainability. In the late 1980s the *World Commission on Environment and Development* chaired by Gro Harlem Brundtland advocated the idea of sustainable development defined as "(…) development that meets the needs of the present without compromising the ability of future generations to meet their own needs" (World Commission, 1987). In the 1990s John Elkington developed an adjusted conception of sustainability for the business sphere taking into account the economic and social dimension in addition to the environmental one. All three together constitute what Elkington calls the Triple Bottom Line (Elkington, 1997).

A second alternative approach towards the moral obligations of business centres on the idea of corporate social responsibility (CSR). Archie Carrol advanced an influential conception whereby CSR is understood as a pyramid having four layers: the bottom layer consists of economic responsibilities, the second layer involves legal requirements, the third ethical obligations, and the top layer includes a philanthropic component (Carroll, 1991).

Edward Freeman advocated a third alternative: stakeholder theory. He defined a stakeholder as a "(…) group or individual who can affect or is affected by the achievement of the firm's objectives" (Freeman, 1984, p. 25). For stakeholder theorists the idea that business has a wide range of stakeholders with certain interests and

entitlements that have to be taken seriously does not only make business sense; it is an ethical requirement as well.

Without going into any thorough normative analysis, we will point out some important examples of ethical considerations based on the above-mentioned four approaches. From Friedman's minimalist point of view, business has an obligation to avoid breaches of cybersecurity as they can easily cause a drop in share value thus harming the shareholders. On the other hand, it is important not to invest a disproportional amount of time and money in cybersecurity, because excessive costs for mitigating cybersecurity risks would negatively affect profitability and thus, again, the shareholders' interests. Since corporations' main moral obligations concern their shareholders it is important to come up with the most efficient methods to reduce the most probable cyber risks.

The idea of the Triple Bottom Line introduces additional normative considerations. Let us exclusively concentrate on economic sustainability broadly understood, i.e. the long-term survival and maintenance of the context in which business flourishes. If we accept that corporations should contribute to economic sustainability thus understood, this would imply more stringent obligations to stress cybersecurity even for those companies that have no particularly enhanced short-term risk profile (because they do not deal with sensitive data or critical infrastructure, for example). We could make the argument that they should nevertheless contribute to herd immunity thus lowering long-term cyber risks for the economic infrastructure as a whole. Thus, a long-term systemic and collectivist risk management approach emerges instead of the individualized and short-term approach associated with Friedman's perspective.

Carroll's CSR model introduces ethical responsibilities "to do what is right" (p. 42) and "to avoid harm" (p. 42) over and above exclusively economic and legal responsibilities (that had arguably already been stipulated by Friedman, albeit in a somewhat diluted fashion). This opens up pathways for yet other arguments. Consistent with Carroll's ethical responsibilities, for example, it could be argued that even if cyber threats would not pose any economic risk and absent any legal requirements (e.g. a strong monopoly in a low regulation developing country), it would still be imperative to rigorously protect sensitive customer data to avoid information related harm.

From a stakeholder theory perspective, finally, a complex web of ethical considerations materializes. On the one hand, customer protection demands strong cybersecurity. Yet this should not lead to disproportional security measures reducing the freedom and autonomy of employees in an effort to identify 'malicious insiders' and exclude all internal threats to cybersecurity. Both customers and employees are important stakeholder whose rights and interests ought to be taken into account. Similar value conflicts arise when looking at reporting security breaches and data leaks. Whilst public notification might harm shareholders in the short term, it might benefit society in the longer term. Again both the shareholders and the community at large are legitimate stakeholders that both ought to be considered. As cybersecurity challenges for business are likely to increase for the foreseeable future, it is imperative to thoroughly analyse the

relevant normative considerations in order to strike a good balance between competing values.

Cybersecurity value conflicts in health: Conflicts with regard to cybersecurity in health are often related to privacy and data protection, i.e. securing health data against unauthorized access. However, there are other types of conflicts. For instance, reaching a high level of cybersecurity might be very costly and therefore, only a small amount of people might be able to afford strong cybersecurity. Cybersecurity also might contradict usability and accessibility.
These problems can be demonstrated with the paradigmatic example of the German eHealth Card (eHC): "As part of the German health-care reform, the current health insurance card is being upgraded to an electronic health card. On it, data on patient investigations, drug regulations, vaccinations and emergency data are stored. The aim is among other things to improve medical care and the prevention of drug incompatibilities and duplication of investigations" (Jürjens & Rumm, 2008). Initially, it was planned to disseminate the eHC to every person insured through the German health insurance system. However, due to strong opposition from various stakeholders, only at least 10 per cent of insured person should receive an eHC (Fox, 2010). Furthermore, due to security considerations concerning data protection some of the functions (electronic prescription and electronic health record; the latter can be used on a voluntary basis) of the eHC were not realized. Particularly German physicians are quite skeptical with regard to the eHC, since it is feared that its deployment will cause huge costs and will increase the workload of physicians and health care personnel. At the same time, the benefits, e.g. in terms of security, are less clear: "The efficiency of the system is considered as critical by the physicians, particularly in terms of data security and potential misuse of data. The primary concern of the physicians is the unauthorized access of a third party to stored data." In addition, "[r]egarding the introduction of the eHC to date, most physicians have criticized the very opaque communication and poor instruction on the subject" (Wirtz, Mory, & Ullrich, 2012). From the point of view of at least some stakeholders, it seems not to be satisfactory to only claim, for instance by state authorities, that cybersecurity and efficiency can be increased—more information is requested. Given the existing literature regarding security issues of the German eHC, many of the concerns that were mentioned by physicians seem to be correct (Sunyaev, Leimeister, & Krcmar, 2010; Winandy, 2010).

The deployment of the German eHC and similar ICT infrastructures in other countries might also be accompanied with potential discrimination. Due to security considerations, e.g. to protect medical data against misuse and unauthorized access, such infrastructures employ encryption and password protection of sensitive data. Laur mentions "[w]hile some people have already difficulty remembering a PIN (especially elderly and disabled people), having many more passwords that are intended to protect them could put them at risk of disclosure, loss or stealing" (Laur, 2015). Although Laur refers to electronic health records in general, the problem also applies to the German eHC in particular. The security measures employed in the case of the eHC are not designed in a way the idea of universal design and general accessibility is demanding. This raises questions regarding social justice and equality. It is quite likely that the affected stakeholders will create their own

work-arounds, for example by writing passwords or PINs on the eHC or by disclosing them to health care personnel, which certainly will reduce their level of data protection, privacy and security with regard to their medical record. In other words, cybersecurity measures that shall protect medical information of citizens but are ill constructed from a usability point of view, force at least some parts of the population to act in a way that reduces their security.

To sum up, at least the example of the German electronic health card and probably other instances of such cards or electronic health records show that cybersecurity can be in conflict with other values than privacy. In the above-mentioned cases, we see conflicts with regard to usability and accessibility, social justice and equality. Moreover, increasing cybersecurity almost always causes economic burdens which might not be fairly distributed.

Cybersecurity value conflicts in national security: Value conflicts with respect to cyber security in the political domain are regularly phrased in terms of security versus privacy, but at closer inspection they are often more complicated. Take for example the discussion about end-to-end encryption in *WhatsApp*. Governments and security agencies have argued that they need to be able to access such encrypted communication for security reasons, e.g. to be able to early detect possible terrorist attacks. Opponents of such access by governments and security agencies do not only point at privacy considerations, but also at the fact that encrypted communication that cannot be accessed by governments and their agencies might be important for the democratic process, and that it enables opposition movements in countries with totalitarian or suppressive regimes.

A similar issue has arisen in relation to the *Tor* network. "Tor is free software and an open network that helps … defend against traffic analysis, a form of network surveillance that threatens personal freedom and privacy."[13] The networks operates as "a group of volunteer-operated servers that allows people to improve their privacy and security on the Internet."[14] In the aftermath of the hacking of the Democratic Party during the US elections, it turned out that a Dutch private *Tor* server had probably been used in the hacking.[15] The Tor server was owned by Rejo Zenger, A Dutch *Bits of Freedom* employee. *Bits of Freedom* describes itself as "the leading Dutch digital rights organization, focusing on privacy and communications freedom in the digital age".[16] While Zenger recognized that *Tor* servers can be misused by hackers, and are in that sense a threat to cybersecurity, he believes that this is a price worth paying, not only for reasons of privacy but also because these servers may be crucial for whistle blowers to

13 See: https://www.torproject.org/ - Accessed 28/01/2017.

14 See: https://www.torproject.org/about/overview.html.en - Accessed 28/01/2017.

15 See: http://nos.nl/artikel/2151234-tor-helpt-hackers-maar-ook-klokkenluider-stoppen-heeft-geen-zin.html - Accessed 28/01/2017.

16 https://www.bof.nl/home/english-bits-of-freedom/ - Accessed 29/01/2017.

reveal abuses. Again, the value that is at stake here is not just privacy but also a range of civil liberties that are seen as crucial for democracy and the democratic process.

Another example is profiling. In this case, values like non-discrimination and absence of bias are at stake and are potentially conflicting with security. In profiling, people are approached, judged or treated in a certain way because these have characteristics that fit a certain profile and that are associated with certain other traits (i.e. traits other than by which they are identified as belonging to the profile). Profiling is used for a wide range of purposes. It may be used by the police or security agencies to find criminals or terrorists; by airports to decide who to check more carefully, by (internet) companies to target certain consumers, by banks in deciding who to give a loan (and against what percentage). As these examples already suggest sometimes profiling serves security objectives. At the same time, profiling may inflict all kinds of undeserved harm on people, from nuisance to false accusations to even, in extreme cases, imprisonment of innocent people. Although profiling may involve privacy violations, because personal information is gathered to fit somebody into a profile, the main issue at stake is not privacy. Rather the issue is that a generalization is made based on limited information about a person. This generalization is based on statistical information about a group to which a person belongs while, due to its probabilistic nature, this information may say nothing about that particular person. Profiling may lead to stereotyping and discrimination. For example, the use of facial recognition technologies by the police and security officers has led to such concerns. Some studies suggest that facial recognition cognition algorithms are less accurate for certain social groups or races (Klare, Burge, Klontz, Vorder Bruegge, & Jain, 2012), which may lead to racial bias in their use (Garvie, Bedoya, & Frankle, 2016; Introna & Wood, 2004).

Another value issue that might arise due to the collection of data by certain organizations for security reasons and that is not completely covered by privacy is the creation of power imbalances. Economic monopolies or oligarchies are often considered undesirable, and in democracies, balancing the (political) power between citizens and their government is an important concern. Maintaining certain power balances is therefore considered important by many for a healthy economy and for democratic politics. What seems to be less recognized is that in the information age, the possession of information about others and their behavior is increasingly a source of power. This also means that organizations that collect or possess large amounts of (personal) data may have increasingly power over other actors, which may lead to the disruption of existing power balances and the creation of new power imbalances. This applies to companies like *Google* or *Facebook* that collect large amounts of data about users and consumers, but also to governments and security agencies that may collect large amounts of data about citizens—and to providers of cybersecurity technologies as well, as they activities may involve the access to highly sensitive data. It should be noted that the accumulation of large amounts of data in the hands of a few may lead to new power imbalances and may be problematic even if such data is anonymized, or if people have given their informed consent for the collection, storage and use of their data. This means that even if privacy concerns are properly addressed, the accumulation of large amounts of data in the hands of a few may be considered problematic for economic as well as political reasons.

Value conflicts in cybersecurity design processes: Cybersecurity field offers opportunities for security service providers to be responsive to secure digital ecosystem but it must also challenge them to ensure that such opportunities are taken to reflect the need of embedding fundamental values in innovative security services and products. This mainly happens under Research and Development (R&D) initiatives of these companies to address real-world cyber-security threats and scenarios in addition ensure trust and confidence among their clients. Hence, security service providers must continue to build and protect their clients, their user machines, engage different stakeholders including businesses and customers to reflect their technical, political, social and ethical values and concerns, and respect functional values upon which our society was built on e.g. autonomy, equality, fairness, freedom and responsibility (Van den Hoven, 2008). They should then not only focus on the security vs. privacy dichotomy, but also highlight any other linkages between the core value of security in cybersecurity to other social and ethical values in order to understand the ethical problems of cybersecurity. To fulfill so, they need to address pro-actively future threats that come with the emergence of the next generations of technologies and services (e.g. IoT, 5G, etc.). But how can we make sure the R&D of security service providers today will meet the needs and requirements of tomorrow? How can we make sure relevant social and ethical values will be incorporated into the security service providers' R&D? The discussed issues are solved when security service providers understand their security initiatives consequences.

Some security service providers may detect an intrusion attempt through their independent security labs. The question is how aggressively they should block and stop an intrusion attempt. In most cases, a trade-off is seen as they could block a bit or aggressively in contrast. From a customer point of view, the rate of detection over the collection of files could be against the positive rate. Security vendors have then two choices: either apply for high intrusion detection rate or apply high positive rate. Choosing either of those choices has economic implications for customers, and has an impact on their fairness levels. In addition, cybersecurity operators (such as ICT security software providers) have considerable access power to the computer systems of their customers, which involves privacy risks, including the possibility of wrongly accusing an employer of the customer to be responsible for a certain cyber threat. The right split between the rate of detection and the positive rate is absolutely a subjective issue and depends on the specific context. For instance, a military environment comparing a health sector environment might ask a higher positive rate instead of only detecting data, and as this request fulfills by security service providers, one can argue different fairness level within different cyber space environments.

Another example is, when security vendors collect data from user machines, where they need to have user consent for different activities. In fact, security service providers ask their clients to fill out the consent forms and clients literally accept relevant consent forms. The question here is to what extent security service providers need to give their clients access to systems. Here is a value conflict between security measures and their impacts on customer' access. More control from customers cause less control by security service providers, in turn revealing more sensitive information. Hence, the level of

customer' access to user machines and digital infrastructure, which can involve discrimination (Custers, Calders, Schermer, & Zarsky, 2013), is twined to the security level. Hence, this value conflict between security and discrimination also must be taken into account within security service providers' R&D.

Accordingly, security service providers' R&D must also address social and ethical values rather only security while they take appropriate security measures. This is essential for the maintenance of critical societal or economic activities in different sectors (e.g., energy, transport, banking, financial services, health, and digital infrastructure).

Mapping and evaluating value conflicts in cybersecurity

However, how can cybersecurity service providers get a reasonable understating on the values involved in cybersecurity problems? Our suggestion is to provide a map on how key aspects of cybersecurity activities positively or negatively affects those values. Figure 4 shows a first draft of such a map based on values discussed in the previous sections. The map shows that cybersecurity is directly related to harm prevention values —both information harm (e.g., caused through disclosure of personal information) or physical harm (e.g., preventing damage on the critical infrastructure). Harm-prevention is supporting for a set of other important values such as privacy or personal freedom. However, cybersecurity measures usually involve some degree of monitoring, cause economic costs and require personal efforts. Those elements usually have negative impact on a whole set of values: Economic costs raise problems of resource allocation that can be in conflict with notions of social justice of equality. Personal efforts needed is confronted with the problem that individuals differ with respect on possessing the necessary (e.g., cognitive) resources. Surveillance not only increases the risk of discrimination and privacy violation; false negative results also can directly impact a core value cybersecurity usually upholds, namely preventing information harm (e.g., because the accused employee loses reputation). The problem is further complicated by the possibility that some values may be in a conflicting relation as well. Personal freedom can counteract social justice, requiring some kind of balancing—exemplified by John Stuart Mills famous quite in *On Liberty*: "The only freedom which deserves the name, is that of pursuing our own good in our own way, so long as we do not attempt to deprive others of theirs, or impede their efforts to obtain it."

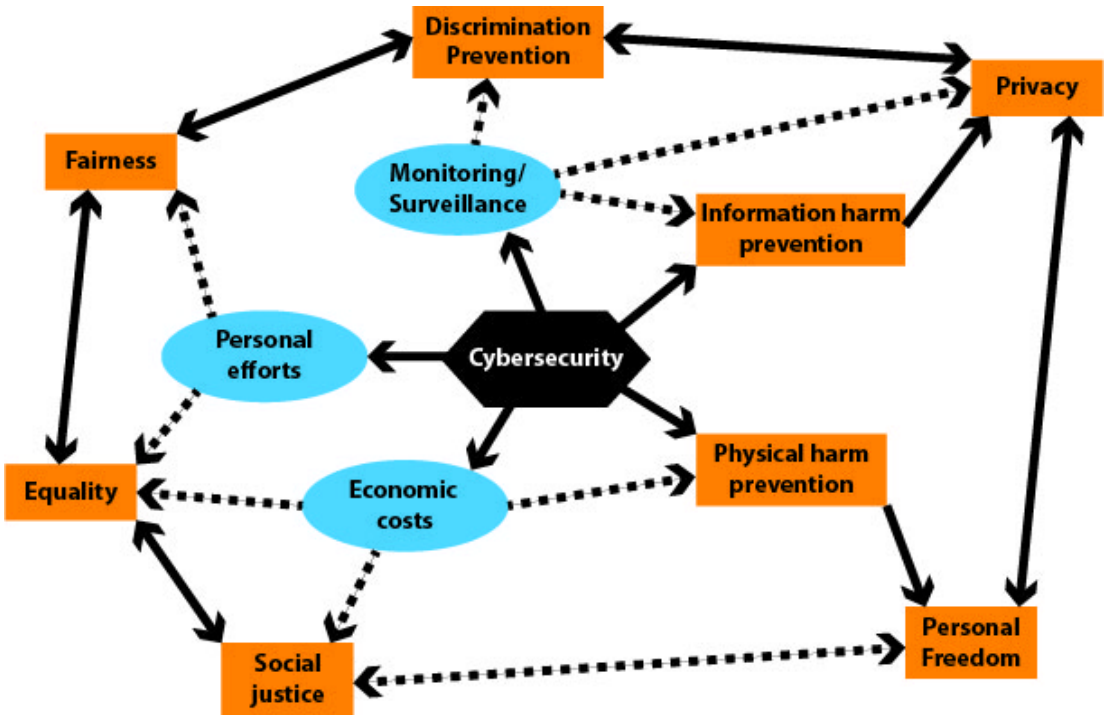

Figure 4: Outlining a first draft of a map on value conflicts in cybersecurity. Arrows with continuous lines show positive (i.e., supporting) relations, whereas arrows with dotted lines show conflicting relations (the map does not show all possible relations). Orange squares refer to values.

So, how can a map help to mitigate these conflicts? For doing this, it is important to recognize that the context strongly influences the framing of those value conflicts. By "context", we do not refer to the details of each single case, but to the notion of "contextual integrity", that emphasizes the importance of distinct social spheres and context sensitive norms in data flows (Nissenbaum, 2004). Contextual integrity forms a main moral reason for data protection (Van den Hoven, 2008); it is legally expressed in terms of purpose specification, use limitation, and data minimization. Contextual integrity refers to the fact that the human environment is structured in social spheres that provide important reference points for human beings. Humans expect to be treated differently in a family context compared to, for example, in a governmental organization. They accept inequality in treatment in the economic sphere that they would not accept in the health, legal or education sphere. Thus, the interpretation of moral values such as justice or autonomy, and the rules related to these values—for example in the case of justice different allocation rules such as "an equal share for everyone" compared to "sharing ac-cording to needs"—differ along these social spheres. The analysis of value conflicts in cybersecurity will have to consider such context effects.

The following example might be useful for explaining the role of contextual integrity. Let us assume that a cybersecurity service provider has considerable access to a computer network of a customer for allowing intrusion detection. In performing its monitoring task, the cybersecurity provider detects suspicious activities within the company computer network that might indicate that some employees of the company are actually involved in offensive cyberattacks against some other target, i.e., the company is not a victim of an

external intrusion, but acts as an attacker towards third parties. What is the status of this information with respect to contextual integrity? At first sight, we are in a business context with a contractual obligation between the customer and the provider. Disclosing this information towards the customer is thus an obligation. The fact that employees of the companies themselves act as attackers, however, have implications that go beyond the business domain. One possibility is that the employer acts on behalf of the company in order to disturb the activities of a foreign company that practices industrial espionage. This "hacking back" would follow the conception of "self-defense" that might also morally justify offensive actions, but its legality is questionable (Lin, Allhoff, & Abney, 2014). The other possibility is that the employee acts on his own, e.g. he tries to hack into a governmental website of an oppressive state as a form of hacktivism. This is surely illegal with respect to the internal rules of the company, but let us assume that the person emigrated from this oppressive country and that some of his family still lives there. If the information about this hacking activity leaves the business context, the ethical problems aggravate. In the first case, by approaching jurisdiction, the case likely becomes public—and a company that has a reputation of hacking back is likely to become an even more likely target, thus increasing the cybersecurity problem[17]. In the second case, a risk of actual physical harm to third parties (the family of the offender) is risked. Thus, in either case, the ethical solution might be that this sensible information does not leave the business context. In the first case, the company could be informed by the security provider about the substantial risks "hacking back" actually involves, whereas in the second case, the employee could be internally sanctioned for his behavior without disclosing the reason. This example illustrates that getting an understanding on how the changing context influences the ethical valence of information is important to understand value conflicts in cybersecurity.

Conclusion

In this contribution, we first demonstrated an (expected) growing importance of cybersecurity in general (measured by the number of publications) and an escalation in terms of describing the severity of the incidences (as increasingly war-like nowadays). We also found that the terminology of privacy still is the dominating ethical term in the debate, although there are indications that this is changing. By referring to exemplar cases, we found that—although cybersecurity has an ethical justification in terms of harm prevention—cybersecurity activities can induce conflicts with other values, some even counteracting the ethical legitimation of cybersecurity as such.

This work, however, should be seen as a preliminary result. First, because the quantitative analysis has pointed to some conflicting results that need further investigations—which is not easy given that the literature body is huge. Second, because the examples provided deserve a deeper normative analysis as done so far. Our draft of a value map requires more refinement. Furthermore, it should include a more intuitive way of visualizing the contextual aspects of possible conflicts. For doing this, we consider the framework of

17 See: https://business.kaspersky.com/hacking-back-ii/4556/ - Accessed 02/02/2014

contextual integrity as fruitful. However, more work is needed in order to explore this approach.

Acknowledgement
This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No. 700540 and the Swiss State Secretariat for Education, Research and Innovation under contract number 16.0052-1.

References

Brey, P. (2007). Ethical Aspects of Information Security and Privacy. In M. J. Carey & S. Ceri (Eds.), *Security, Privacy and Trust in Modern Data Management* (pp. 21–36). Heidel-berg: Springer.

Carroll, A. B. (1991). The pyramid of corporate social responsibility: Toward the moral management of organizational stakeholders. *Business Horizons*, *34*(14), 39–48.

CSIS - Center for Strategic and International Studies. (2014). Net Losses: Estimating the Global Cost of Cybercrime Economic impact of cybercrime II. Retrieved January 31, 2017, from http://www.mcafee.com/mx/resources/reports/rp-economic-impact-cybercrime2.pdf

Custers, H. M., Calders, T., Schermer, B. W., & Zarsky, T. Z. (2013). Discrimination and Privacy in the Information Society Berlin/London: Springer. In *Studies in Applied Philosophy, Epistemology and Rational Ethics* (Vol. 3). Berlin/London: Springer.

Egelman, S., Herley, C., & van Oorschot, P. C. (2013). Markets for zero-day exploits: ethics and implications (pp. 41–46). ACM Press. https://doi.org/10.1145/2535813.2535818

Elkington, J. (1997). *Cannibals with Forks. The Triple Bottom Line of 21st Century Business*. Mankato: Capston Publishing Ltd.

Fox, D. (2010). Elektronische Gesundheitskarte. *Datenschutz Und Datensicherheit – DuD*, *34*(12), 844.

Freeman, R. E. (1984). *Stakeholder management: A stakeholder approach*. Boston: Pitman.

Friedman, M. (1970). The social responsibility of business is to increase its profits. *The New York Times Magazine*, *13*, 122–126.

Friedman, M. (1982). *Capitalism and Freedom.* Chicago: The University of Chicago Press.

Garvie, C., Bedoya, A. M., & Frankle, J. (2016). *The perpetual line-up. Unregulated police face recognition in America.* Georgetown Law Center on Privacy & Technology.

Introna, L., & Wood, D. (2004). Picturing algorithmic surveillance: the politics of facial recognition systems. *Surveillance and Society*, *2*(2/3), 177–198.

International Telecommunications Union. (2008). ITU-TX.1205: series X: data networks, open system communications and security: telecommunication security: overview of cybersecurity. Retrieved January 31, 2017, from https://www.itu.int/rec/T-REC-X.1205-200804-I

Jürjens, J., & Rumm, R. (2008). Model-based Security Analysis of the German Health Card Architecture. *Methods of Information in Medicine*, *47*(5), 409–421.

Klare, B. F., Burge, M. J., Klontz, J. C., Vorder Bruegge, R. W., & Jain, A. K. (2012). Face Recognition Performance: Role of Demographic Information. *IEEE Transactions on In-Formation Forensics and Security*, *7*(6), 1789–1801.

Klimburg, A. (Ed.). (2012). *National cyber security framework manual.* NATO CCD COE Publications.

Laur, A. (2015). Fear of e-Health Records implementation? *Medico-Legal Journal*, *83*(1), 34–39.

Lin, P., Allhoff, F., & Abney, K. (2014). Is Warfare the Right Frame for the Cyber Debate? In L. Floridi & M. Taddeo (Eds.), *The Ethics of Information Warfare* (pp. 39–59). Berlin: Springer.

Nissenbaum, H. (2004). Privacy as contextual integrity. *Washington Law Review*, *79*, 119–157.

Pike, R. E. (2013). The "Ethics" of Teaching Ethical Hacking. *Journal of International Technology and Information Management*, *22*(4), 1–7.

Rowe, N. C. (2008). Ethics of Cyber War Attacks. In L. Janczewski & A. Colarik (Eds.), *Cyber Warfare and Cyber Terrorism* (pp. 384–394). Hersey: Information Science Reference.

Sunyaev, A., Leimeister, J. M., & Krcmar, H. (2010). Open Security Issues in German Healthcare Telematics (pp. 87–194). Presented at the Proceedings of the Third International Conference on Health Informatics (HealthInf 2010), Valencia/Spain.

Van den Hoven, J. (2008). Information technology, privacy, and the protection of personal data. In J. van den Hoven & J. Weckert (Eds.), *Information technology and moral philosophy* (pp. 301–321). Cambridge, New York: Cambridge University Press.

Winandy, M. (2010). A Note on the Security in the Card Management System of the German E-Health Card. In *Electronic Healthcare—Third International Conference* (pp. 193–203). Casablanca/Morocco.

Wirtz, B. W., Mory, L., & Ullrich, S. (2012). eHealth in the public sector: An empirical anal-ysis of the acceptance of Germany's electronic health card. *Public Administration*, *90*(3), 642–663.

World Commission. (1987). Report of the World Commission on Environment and Development: Our Common Future, Chapter 2: Towards Sustainable Development. Retrieved from http://www.un-documents.net/our-common-future.pdf