Mediating Technomoral Care: Taking Cyberwarfare Beyond State Level

by Anna Melnyk

MSc Thesis / November 2018

Mediating Technomoral Care: Taking Cyberwarfare Beyond State Level

Master Thesis

(23.900 words)

This thesis is dedicated to my grandmother, great grandmother, and all the innocent people who were mistakenly persecuted and sent to Siberia in twentieth century.

Anna Melnyk (1851497)

MSc Philosophy of Science, Technology and Society (PSTS)
Faculty of Behavioral, Management, and Social Sciences
University of Twente
Enschede
The Netherlands
2018-2019

Supervisor: Dr. Nolen Gertz
Second Reader: Prof. Dr. Philip Brey

Table of Content

List of Abbreviations

IL	International law
IHL	International humanitarian law
TM 2.0	Tallinn Manual 2.0
JWR	Just War Theory
DDoS	Distributed Denial of Service
JCW	Just Cyber War
CIA	Critical infrastructure attack
CDA	Confidential data acquisition
IPA	Interpretative Phenomenological Analysis
R1, R2, R3,	Respondent 1, Respondent 2, Respondent 3,
R4	Respondent 4
GDPR	General Data Protection Regulation
CCDCOE	NATO Cooperative Cyber Defense Center of Excellence

Summary

The Tallinn Manual 2.0 (TM 2.0, 2017) derives from a comprehensive academic research on how international law applies to cyber conflict. Subsequently, it aims to provide the set of instructions for state behavior in cyberspace. The cornerstone of the TM 2.0, Just War Theory, corresponds so far to the core state-centered values that underlie international law and thus starts with the idea of state security, which mainly focuses on the territorial integrity, protection of state borders, territory, and citizens from external threats. Although some researchers have pointed out, implicitly or explicitly, the limitation of such a state-centered account of values, no single study that focuses particularly on this as a matter of concern exists yet. By prioritizing state-centered security over human security values, the TM 2.0 lacks clarity on the other-than-state-centered values, in particular, *care*. Here, I define care following the philosopher of technology Shannon Vallor (2016) as: "a skillful, attentive, responsible, and emotionally responsive disposition to personally meet the needs of those with whom we share our technosocial environment." After a thorough examination of the TM 2.0 and the ethical debate around cyberwarfare, I suggest using a technomoral virtue ethics framework and focus on the value of care as a practice, that, while it appears in a couple of rules, remains undefined and untackled by the TM 2.0. The aim of this project is to reveal 1) the importance of the value of care in TM 2.0, 2) potential ways of reading of care in TM 2.0, and 3) the challenges that remain. Hence, the main research question that this project aims to address is the following: *In what way could an ethical theory of technomoral virtues fill in the value gap that currently exists in the Tallinn Manual 2.0 and, more broadly, in the debate surrounding the development of regulations for cyberwarfare?*

Key words: cyberwarfare, Tallinn Manual 2.0, state-centered values, human-centered values, technomoral care.

Foreword

This master thesis turned out to be more challenging than it initially seemed. There are many reasons for that, among which the biggest one is the multidisciplinary nature of cyberwarfare. That is why it would be fair to characterize this thesis as a product of an interdisciplinary collaboration based on supervision, conference presentations, participation in workshops, colloquia, and long-long hours of discussions. First and foremost, I want to thank my supervisors Dr. Nolen Gertz and Prof. Philip Brey for their feedback, critical comments, and guidance. I am also grateful for the input from my subject expert Dr. Wolter Pieters and external supervisor Prof. Shannon Vallor, and for every minute they spent thinking about and commenting on my work. Additional thanks to all the attendants of the colloquia in Delft Technical University for their comments and concerns that contributed to the last minute updates.

Another significant contribution to the debate analysis in my thesis was given by the participants of MANCEPT Workshop (2018) "Modern Warfare and Just War Theory." Especially, I want to thank Dr. Lonneke Peperkamp, Dr. Patrick Taylor Smith, Dr. Ronald Tinnevelt, Prof. James Pattison, Emil Archambault, Anh Le, Sara van Goozen and other participants who attended the event. It was an excellent opportunity for me and I was incredibly honored to present and discuss my work on the same stage with such researchers in Manchester. Also, I want to thank all the participants of the International Conference "Human-Technology Relations: Postphenomenology and Philosophy of Technology" that took place here at the University of Twente, who came to our panel and with whom we had interesting discussions afterwards.

The PSTS Master program, that became like a home for me, is a very meaningful experience in my life. I am very thankful to have met my classmates and professors. We had a great class, every lecture was engaging and interesting because of the discussions we had. I appreciate every minute we spent together with Prof. Lissa Roberts, Dr. Marianne Boenink, Prof. Saskia Nagel, Prof. Philip Brey, Dr. Kevin Macnish, Dr. Owen King, Prof. Ciano Aydin, Dr. Michael Nagenborg, Prof. Peter-Paul Verbeek, Dr. Kornelia Konrad, Dr. Michael Kuhler, Dr. Miles MacLeod, Dr. Koray Karaca, Dr. Nolen Gertz, Dr. Lantz Fleming Miller, Dr. Annalisa Pelizza and Ada Krooshoop for their help, support and advice throughout these years at Twente. Separate thanks to Prof. Petra de Weerd-Nederhof, Dr. Anne Dijkstra and all the

members of Research Honors Program 2018 for providing me with the possibility to acquire important knowledge about a research career, finding new friends and get exciting experiences together.

Special thanks to Dr. Sven Nyholm, Dr. Lantz Fleming Miller and Prof. Saskia Nagel for playing essential roles in the process of my development as a researcher. They were continually strengthening my belief in the ability to grow in this profession. Their feedback developed the strongest motivation in me when there were many challenges and obstacles.

Also, thanks to my friends, who believe in what I am doing and support me on the different levels. In the Philosophy Department, I met amazing people like Olya Kudina, Melis Bas and Mayli Mertens who were always there for me and were cheering me on along the road. With gratitude, I perceive all the help with editing and proofreading that Patricia Reyes Benavides did for me. Huge thanks to famous Ukrainian painter Andrey Babchinskiy for allowing me to use a beautiful piece of art on the cover page.

And most importantly, my family, who always believes in me and supports my duty in sharing what in my opinion are important messages and attempting to make this world a better place. To my brother who did not get quite what exactly I am studying but is always excited about my life. The biggest thanks goes to my parents, my grandma, and Alex. They did the most for me, and every single evening they were next to me physically and digitally, blindly believing in the significance of my dream. Without their strong belief, I would not be here now. Thank you!

Anna Melnyk
November 2018

1. Introduction

> *"A battle being waged by many, to defend the interests of many. (...) Just look at the Netherlands. (...) We're now one of the world's most digitally advanced countries. (...) And digital growth still hasn't reached its peak. So it's no surprise that the Netherlands is at the forefront of efforts to advance the debate on keeping cyberspace stable and safe."*
>
> *17 August 2018*
> *Mr. Stef Blok, Dutch Minister of Foreign Affairs*
> *1 Year Anniversary of the Tallinn Manual 2.0* [1]

Throughout four days in May 2017, hospitals, telecommunication companies, power plant grids, critical infrastructures, and many individual and corporate devices that used a Windows Operating System were infected by a ransomware cryptoworm. According to security researchers from the Kaspersky lab, "WanaCrypt0r 2.0" or "WannaCry" caused approximately 45,000 cyber attacks in 99 countries within those four days (Wong & Solon, 2017). This is just one of the cases that caused massive panic around the world about the coming era of cyberwarfare and raised concerns about the cyber defense of vulnerabilities of crucial IT systems.

At the same time, with a continuous increase of cybersecurity threats (which are projected to become one of the top future global concerns, *see* Appendix C), the knowledge on how to address them has stayed relatively underdeveloped. The Tallinn Manual 2.0 (TM 2.0) was published in 2017 as a response to the challenges raised by cybersecurity threats in the economic, military, and social domains. This document is constituted by a set of instructions for state behavior in cyberspace. It was written to support and inform decision-making processes of governmental and legal authorities with formulated guidelines on how international law applies to cyber conflict (Schmitt et al., 2017).

The TM 2.0 is an important starting point on the development of solid legal frameworks that will address different concerns and responsibility issues in the cyber domain. Currently, it is possible to point

[1] *See report:* http://puc.overheid.nl/doc/PUC_248137_11

out that the state, instead of people, is the central object of security in TM 2.0. Although the attempt to shift from cyberwarfare to the regulation of peacetime cyber operations was made in TM 2.0, similarly as in International Law (IL), the value of state security is privileged over the value of respect for human rights (Meredith & Christou, 2009, p. 6). This implies that the core values behind TM 2.0, as in the preceding version TM 1.0, are sovereignty-oriented and state-centered.

Before the TM 2.0 was published, many objections about the comprehensiveness of its main component, Just War Theory (JWT), and concerns about the potential applicability of the principles *jus ad bellum* (right to go to war) and *jus in bello* (right to conduct in war) to cyber issues have been raised by philosophers and ethicists of war (Rid, 2012; Taddeo, 2012, 2014, 2016; Smith, 2017, 2018; Stone, 2013, Barrett, 2013, 2015; Sleat, 2017; Eberle, 2013; Lin et al., 2014). So far, JWT, while being the cornerstone of the TM 2.0, corresponds to the core state-centered values that underlie IL and are promoted by the 'moderates' who directed and guided the development of TM 2.0 (Schmitt et al., 2017). At a very basic level, they argue that the principles *jus in bello* and *jus ad bellum* provide a sufficient ethical ground to approach cyberwarfare (Barrett, 2013, 2015; Eberle, 2013). This account, similarly to IL, is tied with the underlying idea of state security that mainly focuses on the territorial integrity, protection of state borders, territory, and citizens from external threats (Meredith & Christou, 2009).

Although some researchers have pointed out the limitation of such a state-centered account of values, implicitly as Lin et al., (2014) or explicitly as Taddeo (2012, 2014, 2016), currently, no single study exists which focuses particularly on this limitation as a matter-of-concern. 'Skeptics' implicitly indicate the disadvantages of the state-centered values since new problems raised by cyberwarfare are left "on the shoulders of individuals" (Lin et al., 2014). Whereas 'radicals' explicitly suggest that the technological component and related values should be emphasized and integrated into JWT. According to them, by merging JWT with the onto-centrism of Information Ethics (IE), it is possible to encompass techno-social assemblages of animate and inanimate actors like human, data, and infrastructures as assemblages of informational beings with an equally relevant moral status (Taddeo, 2012, 2014, 2016).

However, following some authors, I object to the 'radicals' assumption that the integrity of information systems is an intrinsic value which is equally acknowledged in different cultural contexts (Brey, 2007; Ess, 2008). Due to "the digital divide,"[2]- the idea that the world is divided between the online and offline - the distribution of risks and benefits of new and emerging technologies are spread

[2] Furthermore, as DiMaggio & Hargittai (2001) claim, "digital divide" leads to "digital inequality, by which we refer not just to differences in access, but also to inequality among persons with formal access to the Internet."

with a different amplitude around a culturally heterogeneous world (DiMaggio & Hargittai, 2001). Therefore, without a particular 'human perspective' behind the integrity of an information system, this value is too abstract for the practical implications in regulatory practice. Thus, a key argument in this thesis is that the focus on state-centric values in TM 2.0 raises serious concerns, whereas a shift to onto-centric values does not address these concerns, and the value gap that results from these approaches requires thorough examination.

1.1 Thesis Statement, Objective and Research Question

From a philosophy of technology standpoint, due to the rapid production of new and emerging technologies, the need for a critical appraisal of the values and practices informed by such technologies is actualized. Although the TM 2.0 is a comprehensive guide that encompasses numerous cyber operations, by prioritizing state-centered security (IL) over human security values, it lacks clarity on the other-than-state-centered values, in particular, care. Thus, after thorough examination of the TM 2.0 and the ethical debate around cyberwarfare, I suggest to use an aretaic ethical theory - technomoral virtue ethics framework - where technomoral virtues are considered as "a new alignment of our existing moral capacities, adapted to a rapidly changing environment that increasingly calls for collective moral wisdom on a global scale" (Vallor, 2016). When relying on technomoral virtue ethics and shifting the focus to the human-centric values such as the value of care as a practice, there is a possibility for a human-centered approach to contribute to the discussion. While care appears in a couple of rules in the TM 2.0 (92, 59, 114), it remains largely undefined and untackled by such document.

Hence, the aim of this thesis is: 1) to reveal the importance of the value of care in TM 2.0, 2) to suggest the potential ways of reading care, and 3) to reflect upon the challenges that remain after these interpretations. I define care following philosopher of technology Shannon Vallor (2016) as: "a skillful, attentive, responsible, and emotionally responsive disposition to personally meet the needs of those with whom we share our technosocial environment" (p. 138). Some further clarifications about what I mean by 'the ways of reading care' are needed. First, adopting the idea of technomoral change - alterations on the level of values and morality caused by technologies through time (Swierstra, 2016; Boenink et al., 2009) - is crucial for a comprehensive understanding of the new type of daily practices that people face and the subsequent corresponding values. In this thesis, I claim that reading the TM 2.0

with a technomoral virtue ethics lens makes it possible to unpack what are the practices informed by these values.

In this regard, the main research question that will be addressed in this thesis is the following: *In what way could an ethical theory of technomoral virtues fill in the value gap that currently exists in the Tallinn Manual 2.0 and, more broadly, in the debate surrounding the development of regulations for cyberwarfare?*

In order to provide a comprehensive answer to this research question, I suggest to explore four sub-questions in Chapters 2 and 3. The research sub-questions are formulated in two groups as follows:

- *What are the key components that constitute the Tallinn Manual 2.0? What are the challenges that arise when the Tallinn Manual 2.0 is applied? What are the underlying values? How does the ethical debate, that preceded the issue of the Tallinn Manual 2.0, constitutes and addresses concerns regarding morality and other-than-state values?*

- *What does an aretaic reading of Tallinn Manual 2.0 imply? What is the relevance of technomoral care in the Tallinn Manual 2.0? To what extent does the analysis of the dynamics in the technomoral care reveal and conceal ethically relevant aspects in the Tallinn Manual 2.0?*

1.2 Definition of Cyberwarfare

After suggesting a thesis statement and research question, I propose to bring some conceptual clarity on the notion of 'cyberwarfare.' When referring to cyberwarfare as a multidisciplinary issue (*see Appendix A*), I define it, following the ethicist Mariarosaria Taddeo (2012), as "uses of ICT's within an offensive or defensive military strategy that is endorsed by a state and aiming at the immediate disruption or control of the informational environment, with agents and targets ranging both on the physical and non-physical domains and whose level of violence may vary upon circumstances" (p. 114). In this definition, the explicit emphasis is put on two aspects: 1) informational nature of cyberwarfare, and 2) transversality as a qualifying couple like 'physical — non-physical,' 'violent — non-violent,' 'human agent

— artificial agent' (Taddeo, 2012, p. 210). Thus, since there is no universally recognized definition (Hughes & Colarik, 2017; *see Appendix A*), I consider Taddeo's definition to be the most comprehensive and suggest to refer to it in this thesis.

1.3 Methodology

After clarifying the objectives of the research, in this section, the methodological techniques and strategies for collection and analysis of relevant data will be suggested. For the theoretical part of this thesis, cross-disciplinary literature review was made (digital libraries *Web of Science* and *Scopus)*. Military Studies, International Relations, Security Studies, Legal Studies, Ethics and Philosophy are just among several domains that discuss cyberwarfare. Considering the multidisciplinary nature of the phenomena, on the theoretical level, this interdisciplinary research proposes a contribution from philosophy of technology to this cross-disciplinary debate.

For the empirical analysis, the data collection was based on semi-structured interviews as the type of interviews that enable a researcher's flexibility to ask interviewees different questions depending on particular situations while working towards the same research goal (Kvale, 1983; Noor, 2008; Spradley, 1979). The way in which these interviews were conducted was approved by the Ethics Committee Behavioral Science (#18281) on April 6, 2018 and all the data is anonymized and used according to the GDPR rules. The questions are formulated in the developments and tendencies that took place in particular contexts (*see questions in Appendixes E, F, G, H*). The empirical observations are divided into three parts, each according to its methodological tool:

1) Interviews with experts from the leading discussion domains. In order to reflect upon ethically controversial issues that appeared during literature review, *an interview was conducted with R1, the expert in Computer Ethics*. To reflect upon the complexities of the cross-disciplinary debate an *interview was conducted with R2, the expert in International Relations.*

2) Interviews with Estonian citizens for the analysis of the Estonia case, detailed below. It was particularly difficult to come to the decision about who should be considered as a relevant stakeholder in this exemplary case. The decision was made in favor of two citizens who

represent both ethnic groups populating Estonia, that is, from Russian ancestry and from Estonian ancestry *(anonymized as R3 and R4 respectively)*. The criteria for the search of subjects was straightforward: they were present in the cyber attacks of 2007, and, since it was eleven years ago, they had to be adults during those events.

These stakeholders, as representors of their ethnic group, were chosen to provide empirical insight from both sides of the conflict with the help of the methodological toolbox of Interpretative Phenomenological Analysis (IPA). It is a qualitative approach that is mainly used in psychology to focus on the content of how people try to make sense of various phenomena (Smith, 2011). Since IPA admits that humans are sense-making organisms, during the interpretative endeavor “researcher is trying to make sense of the participant trying to make sense of what is happening to them” (Smith & Osborn, 2015). The analysis is based on Husserl’s account of phenomenology “which aims to produce an account of lived experience in its own terms rather than one prescribed by pre-existing theoretical preconceptions” (Smith & Osborn, 2015).

Thus, according to IPA there is no direct way to engage with the experience of people and each interview should be qualified as a set of “particular instances of lived experience” (Smith et al., 2009, p. 37). In-depth semi-structured interviews support the commitment to detailed personal accounts of participants. Furthermore, the particular advantage of IPA for this thesis is that it favors small size to large-scale groups of research participants that would inevitably require significant simplification of data. This advantage is linked with IPA’s idiographic commitment that relies on the idea that prior to more general claims, it is necessary to examine the detailed experience of each case.

In this view, IPA is not just relevant for this study as it favors small size research participant groups, but it also fits with the theoretical basis, because it shares virtue ethics’ emphasis on the relevance of individual experiences. To systematize the process of analysis in IPA, Smith, Flowers and Larkin (2009) suggest approaching it in *six steps*: reading, initial noting, developing emergent themes, searching for connections among emergent themes, moving to the next case, and looking for patterns among cases. Based on this sixth step, the two core processes that constitute an IPA analysis can be distinguished: a) the shift from the descriptive accounts of the participants to the interpretative pieces; and b) the identification of particular patterns in the shared themes across multiple participants, while retaining the commitment to the participants’ choice of words (Reid, Flowers and Larkin, 2005).

3) Extractions from legal manuals, newspapers, journalists' articles, and governmental reports about one of the classical case studies: the cyber-attacks on Estonian government in 2007.

The context of the events were depicted from news reports, academic publications, and published interviews with public and governmental authorities. The purpose of this analysis is to demonstrate that cyberattacks always appear in a particular techno-social order where distributions of technological risks and benefits are not equal. By observing the dynamics in the Estonia case and employing qualitative analysis, I aim to contribute to the cross-disciplinary discourse with approaches from philosophy of technology (*see* Chapter 3).

1.5 Case Study

It was reported that this case combined informational warfare with cyber attacks, thus brought together physical with digital domains of actions on the political level (Radin, 2017; Tikk & Vihul, 2010). These cyberattacks in Estonia were commonly recognized as an escalating case that boosted the discussions around cyberwarfare both in public (McGuinness, 2017; Tamkin, 2017) and academic debates (Rid, 2012; Shackelford, 2010; Herzog, 2011).

According to the former Estonian President's speech at the NATO Cooperative Cyber Defense Center of Excellence (CCDCOE) conference in 2018, the Estonian case changed the way cyberattack was perceived. Before, cyberattacks were perceived as "espionage, and not to damage adversaries and make a political point,"[3] as he further notices "It was averted and public," as "continuation of policy by other means." For more details, I introduce in box 1 a depicted narrative of the case study.

[3] *See* Keynote speech retrieved from: https://www.youtube.com/watch?v=j_Fx5TPwQ4E

Picture 1 and 2. Soviet soldier monument before (left) and after (right) relocation.

Box 1. Exemplar case: Estonia 2007

Context: The case starts with a decree from the local Estonian government about the relocation of the monument dedicated to the fallen in WWII in Estonia (pic. 1, 2). This monument has a semiotic meaning for approximately 400.000 people from the post-soviet community who use this memorial as a public holiday location. Governmental authorities were forced to find a solution to prevent tensions between hooligans who, it was said, were using the same monument as location for violent anti-Estonian protests. This relocation resulted in the explosion of adverse reactions of Russian minority's protesters including riots with Molotov bombs, burned cars, and unfortunately, even death (Ruus, 2008).

Cyberattacks: The night before the monument was relocated, a substantial attack on the information systems - DDoS [4] - overloaded and penetrated numerous governmental networks with the unfeasible concentration of connections, thus it made the systems unresponsive to users. Security experts have divided the attack into two phases: 1) amateurish attacks defaced websites or posted fake letters of apology from Estonian officials online, 2) a more sophisticated DDoS attack assessed as the work of expert hackers (Chloe, 2009). This phase requires the deployment of a botnet of over a million computers to flood the country with more than one thousand times its regular internet traffic (Ruus, 2008). At the same time, websites hosted on servers located within Russia posted instructions on how and what to attack (Leder et al., 2009).

[4] Distributed Denial of Service (DDoS) is a large number of coordinated computers that aim to overwhelm the targeted systems in two ways: 1) a massive number of individuals attacking simultaneously or 2) a network of computers, known as a botnet, covertly controlled through a backdoor to execute commands unknown to their owners (Leder et al., 2009).

Response: Eventually, the Estonian CERT developed a three-pronged approach to fixing the problems: 1) to increase the server capacity for their systems to handle more traffic, 2) to develop a filtering system to separate good message traffic from bogus message traffic associated with the attack, 3) to work with the authorities responsible for the root Domain Name System servers to take the identified botnet computers offline (Ruus, 2008). This required to sacrifice the country's connectivity to the rest of the world, since Estonia had to close its computer systems off from access outside of the country. After two years, it was revealed that the pro-kremlin youth organization "Nashi" took the responsibility for the cyberattacks (Shachtman, 2009).

1.5 Outline

It is significant for any in-depth research that requires the combination of empirical and theoretical analyses to identify and elaborate the following: state-of-the-art for the thesis, the definition of the theoretical framework, and methodological guidance. After I stated the problem, introduced research questions and provided detailed methodological guidelines in Chapter 1, the purpose of Chapter 2 is to inform the reader about the state-of-the-art for the thesis by supporting the analysis with several exemplifications of the tensions that occur during its application to well-known cases of cyberwarfare. Based on that, in Chapter 2 I give a critical appraisal of academic research on how IL applies to cyber conflicts in order to illustrate the advantages and disadvantages of a state-centered account of values originated from the state-centered security paradigm that is hold by TM 2.0. Further I seek to identify the other-than-state values in the debate around cyberwarfare, that in turn, preceded the issue of TM 2.0. In the Chapter 3, the definition of the theoretical framework and contribution from philosophy of technology will be suggested to facilitate a principally different reading of TM 2.0 and respectively JWT. Here, the notion of 'technomoral care' in the intersection of technomoral virtue ethics and the theory of techno-moral change will be introduced. Basing on this theoretical and conceptual analysis I propose a broader depiction of the other-than-state values that are at stake in cyberwarfare nourished with the empirical insights. The combination of the theoretical and empirical analysis proposed in this chapter is aimed at making a preliminary step in the shift towards a human-centered perspective on security and move to an understanding of the regulations of cyberwarfare beyond state-centered security paradigm level. Finally, I will integrate all the components of the analysis proposed by this thesis and address some objections to end with a few general concluding remarks where all the points and findings will be systematically recalled.

2. Tallinn Manual 2.0: Key Components and State-Centered Security

This chapter provides a preliminary overview of TM 2.0, its basic components and existing conceptual and regulatory challenges that currently surround legal, military, and ethical discussions about cyberwarfare. The group of research sub-questions that are addressed here is: *What are the key components that constitute the Tallinn Manual 2.0? What are the challenges that arise when the Tallinn Manual 2.0 is applied? What are the underlying values? How does the ethical debate, that preceded the issue of the Tallinn Manual 2.0, constitutes and addresses concerns regarding morality and other-than-state values?*

In the first section of this chapter, I introduce the TM 2.0 and give critical appraisal of academic research on how International Law (IL) is applied to cyber conflicts. Here, I will demonstrate how TM 2.0, IL, and Just War Theory (JWT) are interwoven, and what are the challenges to the applicability of the rules in the manual. Then, I will elaborate on concrete examples in regard to the conceptual and responsibility gaps that occur when TM 2.0 and principles of JWT are applied. The combination of these challenges will be referred as the 'regulatory gap'. Based on this, I will elaborate on another relevant, yet less covered point: that the state-centered values underneath TM 2.0 are linked to a state-security paradigm, and thus, representation of other-than-state values is opaque. In regard to the components of TM 2.0 that are displayed in sections 2.1 and 2.2, the sections 2.4 and 2.5 are focused on the underlying values that were put forward in the broader ethical debate around cyberwarfare. In each section, I will turn to counter-arguments 1) against the 'moderates' position and the analogy approach, and 2) against 'radicals' adopting and merging a 'moderates' position with philosophy of technology. Respectively, in each section, I will elaborate on the matter of explicitly or implicitly incorporated values.

2.1 Unraveling the Legal Resource

In 2017, the document called Tallinn Manual 2.0 (TM 2.0), an influential resource for states and legal advisers authored by twenty international law (IL) experts was published. This collaboration was facilitated and led by the NATO Cooperative Cyber Defense Center of Excellence with the aim to qualify and tackle cases of different conflicts and operations in the cyber domain. TM 2.0 "reinterpret[s] IL through the drawing of analogies between kinetic (physical) and cybernetic domains" and consists of

154 rules, that are based on legal regimes as the law of state responsibility, the law of the sea, air and space, international telecommunications law, diplomatic law, and human rights law (Efrony & Shany, 2018, p. 583).

Referring to the International Court of Justice affirmation in the *Nuclear Weapons* advisory opinion, the TM 2.0 states that although cyber operations are not explicitly mentioned in the existing law of armed conflict treaties, "the general rules that determine the legality of weapons will also determine the lawfulness of cyber methods and means of warfare" (Schmitt et al., 2017, p. 451). In the Rule 103 (Definitions of means and methods of warfare), there are two conceptual specifications provided: 1) "'means of cyber warfare' are cyber weapons and their associated cyber systems," and 2) "'methods of cyber warfare' are the cyber tactics, techniques, and procedures by which hostilities are conducted" (Schmitt et al., 2017, p. 452).

Yet, based on the extracted rule, because of the development of new and emerging technologies, warfare can be extended to the new man-made domain 'cyberspace' by making use of cyber weapons. Cyber weapon,[5] while often being a computer program, cause damage when exploiting one or more vulnerabilities in the system. These can be vulnerabilities in software/hardware configurations and the so-called human factor.[6] Both cyber weapons and information about a system's vulnerabilities are not just developments of military agencies, but often sold on the black market (Maathuis et al., 2016). Cyber tools that are actively used can be one of, or a combination of, the following: distributed denial-of-service attack (DDoS), critical infrastructure attack (CIA), thief or destruction of hardware, cyber-espionage, confidential data acquisition (CDA), web vandalism and propaganda. Each of these tools require different levels of sophistication, and thus, each one is significantly different regarding technical and financial means.

Due to such diversity of cyber weapons, in TM 2.0, compared to the first version of the Tallinn Manual, a conceptual alteration from the 'cyberwar' to 'cyber operations' was accompanied with the extended focus from the international humanitarian law (IHL) governing cyberwarfare to the peacetime legal regimes. Such extension has been justified by observing that in the contemporary world, the vast number of cyber attacks would "fall beneath the threshold at which IHL would typically declare them to be a formal act of war" (Leetaru, 2017). In other words, cases involving the cyber weapons of theft or

[5] Maathuis et al., (2016) provide a comprehensive working definition for cyber weapons: "A computer program created and/or used to alter or damage (an ICT component of) a system in order to achieve (military) objectives against adversaries inside and/or outside cyberspace."

[6] Human factor - the type of vulnerabilities where users provide necessary details for access and control of confidential data on their computers or computers they use. For more detail *see*: https://seclab.stanford.edu/courses/cs203/lectures/humanfactor.pdf

destruction of hardware, cyber-espionage, confidential data acquisition (CDA), and sometimes also DDoS would all fall beneath the threshold into the area unregulated by IHL and called the "grey zone" (Schmitt, 2017). TM 2.0 makes an attempt to provide sufficient instruction for the regulation of the cases that fall into this 'grey zone.' However, the current version is still underdeveloped to guide legal experts in such type of the cases.

To provide some background on TM 2.0, IHL is a branch of IL that specifically deals with the armed conflicts. Renaissance jurist Hugo Grotius and philosopher Francisco di Victoria, the 'fathers' of IL, started a new chapter proposing their different account of ethics and law of war which "in part aimed at developing a codified law of war that can be applied and enforced across nations" (Beard, 2014). Both thinkers made an attempt to systematize the already existing knowledge about morality and principles of war and then develop an IL of war, foundationally based on deontology "as the element of morality that is most like in form to the law" (Beard, 2014).

The long tradition of debates around the justification of wars shows gradually consolidating perspectives throughout history, dating back to philosophers like Augustine de Hippo, Cicero and Thomas Aquinas (Beard, 2014).[7] But only after the 1977 Vietnam war, the "traditionalist" account of JWT with the principles [8] of *jus ad bellum* (right to go to war) and *jus in bello* (right to conduct in war) was suggested by Michael Walzer in the book "Just and Unjust Wars" and became a solid and recognized theory (Lazar, 2017). JWT seeks a middle path between pacifism (where neither state nor individual should engage in war-effort whatsoever), and realism (where the validity and effectiveness of moral rules that aim on regulation of the behaviors of states is denied) (Lazar, 2017).

The debate in the philosophy of war is largely polarized between "traditionalist" and "revisionist" approaches. The core disagreements are around the principle of noncombatant immunity and the moral equality of combatants (Pattison, 2018). These principles are largely what distinguish the "traditionalist" JWT of Walzer (2017), for whom discrimination meant you could not target civilians, and the contemporary JWT where discrimination means "noncombatant deaths are permissible only if proportionate to the military objective sought" (Lazar, 2017). According to McMahan (2010) the "revisionist" account holds the idea that it is permissible to target noncombatants in situations when they are legally responsible and, "conversely, the moral equality of combatants is mistaken because morally innocent combatants, such as those who make just contributions to a war, are not permissible targets" (Pattison, 2018). The "traditionalists" largely criticized "revisionists" in two ways: i) they

[7] For more details *see* Brekke (2005).

[8] *jus ad bellum* (just cause, good intentions, proportionality, likelihood of success, last resort and legitimate authority) and *jus in bello* (proportionality, necessity and discrimination). *See Appendix B*.

attempt to defend “claims that could undermine weakening compliance with international humanitarian law” or, ii) “conversely, they see “revisionists” as only tackling abstract philosophical puzzles that have little, if any, real-world applicability for the foreseeable future” (Pattison, 2018).

The majority of the rules from TM 2.0 are relying on the basis of Walzer’s “traditionalist” account of JWT and the principles of *jus ad bellum* and *jus in bello* (Efrony & Shany, 2018, p. 584). They do accept the new ethical, legal, and philosophical challenges that are raised by cyberwarfare. However, they are optimistic about the resources of JWT to comprehend and tackle these issues (Barrett, 2013, 2015; Eberle, 2013; Schmitt et al., 2017; Sleat, 2017). Even though cyberwarfare sometimes is referred to and compared with the idea of The New Cold War,[9] because of the absence (in most cases) of physical violence and ‘bloodlessness,’ it is often argued that the nature of soft and non-violent cyberwarfare can serve as a means to avoid traditional war and its bloodshed (Taddeo, 2012, 2014; Barrett, 2013; 2015; Rid, 2012). Whereas in traditional warfare where the evils are often referred to as casualties and physical damages, in the case of cyberwarfare, the balance is unobvious and problematic (*discussed by:* Smith, 2017; Stone, 2013; Rid, 2012; Barrett, 2015).

While being a tool of IHL, according to the proponents of the ‘analogy approach’, the principles of JWT are fulfilling their function and tackle ethical issues in cyberwarfare of different type (Schmitt et al., 2017). A good example of this is the Standard View (Schmitt, 1999). TM 2.0 holds this consequence-based account called the ‘Standard View’ (*proposed by* Schmitt, 1999; *criticized by* Smith, 2018) and prioritizes the larger cyber attacks on critical infrastructure (e.g. Stuxnet)[10] and military facilities where human death (or injuries) and object destruction (or damage) are potentially involved. Since such cases are qualified as ‘attacks,’ the substantial rules of the Additional Protocol I to the Geneva Conventions, which restrict the conduct of hostilities, are applicable. As a result, in TM 2.0, non-state actors and peacetime cyber espionage constitute cyber operations that are not regulated by the IHL and thus, require further discussions and negotiations (Schmitt, 2017).

This relates to perhaps the most discussed principle in the ethical debate - just cause or *casus belli.* This principle is the most important for TM 2.0 from the *jus ad bellum* principles, because the current cornerstone - the Standard View (Schmitt, 1999) - explicitly clarifies what cases are considered as acts of war and what cases are not. Thus, when operating, the principle is indicative for the cases that are regulated by IHL and are above the ‘war’ threshold (Schmitt et al., 2017). This is a particularly

[9] *See* https://www.theguardian.com/technology/2010/jan/28/cyber-attacks-hacking

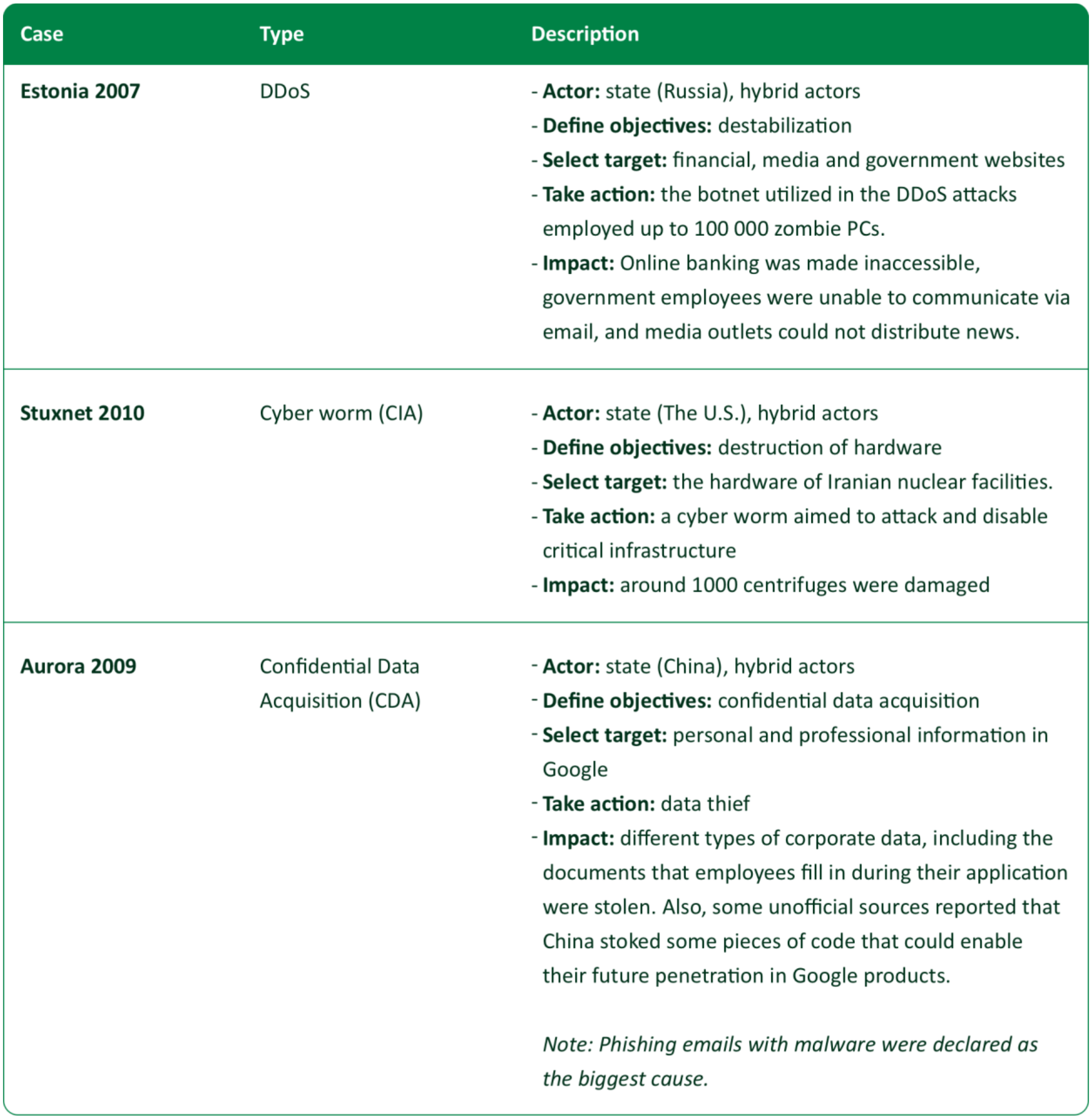
[10] *See* https://www.bbc.com/news/technology-11388018

important point for TM 2.0, because according to this principle, a state may justify their military response to cyber operations (*discussed by:* Smith, 2018; Sleat, 2017; Barrett, 2013).

However, while it became possible to clarify things with categorization of an act of war, cyberspace still raises many other concerns in relation with TM 2.0. Particularly important is the core disagreement between “traditionalist” and “ revisionist” accounts: the principal of discrimination of non-combatants immunity. Regarding this principle, there are two challenging issues that can be indicated in the “traditionalist” account of TM 2.0. First, there are no indicative possibilities to distinguish combatants from non-combatants, as cyber warriors can live regular lives at their homes, or be a state-sponsored activists (Schmitt, 2017, p. 9). Second, combatants can easily hide among the civil population, which might result in an increase in surveillance from the state and thus threaten the rights of the entire population (Taddeo, 2014). This dilemma concurs into the responsibility and liability gaps that are appearing when TM 2.0 is applied (Barrett, 2013; 2015). Thus, although by shifting the scope from cyberwarfare operations to peacetime operations TM 2.0 attempts to address these challenges, or at least enact further research, this principle challenges the analogy approach, since non-combatant immunity is a significant dilemma in TM 2.0.

Now, in order to be less abstract, let me provide some examples. I suggest to use the conceptual model of the context of use of a cyber weapon designed by Maathuis et al., (2016) that consists of five components: 1) actor: those who are responsible for cyber operation (state, non-state actors, hybrid actors); 2) define objectives: actor’s goals; 3) select target: an object, an entity, or a person engaged in taking advantage on the adversary; 4) take action: employment of a cyber weapon; and 5) impact: physical and non-physical result or effect (Maathuis et al., 2016). Referring to this model, in Table 1 I list three cases: one case that constitutes the act of war, one case of peacetime cyber espionage, and a case of DDoS.

Table 1. Exemplary cases of the use of means and methods of cyberwarfare.



Case	Type	Description
Estonia 2007	DDoS	- **Actor:** state (Russia), hybrid actors - **Define objectives:** destabilization - **Select target:** financial, media and government websites - **Take action:** the botnet utilized in the DDoS attacks employed up to 100 000 zombie PCs. - **Impact:** Online banking was made inaccessible, government employees were unable to communicate via email, and media outlets could not distribute news.
Stuxnet 2010	Cyber worm (CIA)	- **Actor:** state (The U.S.), hybrid actors - **Define objectives:** destruction of hardware - **Select target:** the hardware of Iranian nuclear facilities. - **Take action:** a cyber worm aimed to attack and disable critical infrastructure - **Impact:** around 1000 centrifuges were damaged
Aurora 2009	Confidential Data Acquisition (CDA)	- **Actor:** state (China), hybrid actors - **Define objectives:** confidential data acquisition - **Select target:** personal and professional information in Google - **Take action:** data thief - **Impact:** different types of corporate data, including the documents that employees fill in during their application were stolen. Also, some unofficial sources reported that China stoked some pieces of code that could enable their future penetration in Google products. *Note: Phishing emails with malware were declared as the biggest cause.*

Interestingly, according to Taddeo's definition and up to a certain extent, all three cases in Table 1 could be qualified as cyberwarfare. But only the case of Stuxnet it is constituting the act of war in the legal sense when appealing to the TM 2.0 core premise of the Standard View. In this case, a cyber worm successfully attacked and damaged approximately 1000 centrifuges on Iranian nuclear reaction (Maathuis et al., 2016). Considering the fact that the facility was not connected to the internet, this case changed the common understanding of the potentiality of cyber weapons. Moreover, it is the first digital weapon that physically destroyed its target. Due to these specifics, this case is sometimes recalled as a part of an armed race between the United States and the Middle East. Stuxnet is classified by TM 2.0 as an 'attack' to which the Additional Protocol I to the Geneva Conventions is potentially applicable. Similarly to Stuxnet, the more recent case of a cyber attack on the Ukrainian power grid in 2015 falls under the same category - act of war - and thus constitutes *casus belli* according to JWT.

In a legal sense, the other two cases are more problematic since they fall beneath the threshold at which IHL would be applicable and are in the 'grey zone' of IL. In the case of Estonia, although a vast majority of the pertinent scholars assure that Russia is behind these attacks, non-state actors (pro-kremlin activists) officially declared their responsibility for the attacks. At the same time, these activists did not reveal that they were specifically paid or motivated by the Russian government to deploy these attacks. This was also highlighted in the interview with R2 who made an emphasis on the unclarity of the link between pro-kremlin activists and the Russian government. Yet a hybrid actor, or in other words, a non-state actor who is potentially supported by a state actor, may remain unidentified or vice versa (or identified as terrorists, activists et cetera). Indeed, as in the Estonia case, even though the causality chain seems to pinpoint towards a hybrid actors' status, a retribution challenge unveils the failure to assign any legal responsibility on the state level, since cyber operations that are involving non-state actors are not regulated by IHL. In turn, such issues result from a tension in JWT and TM 2.0 on the matter of the applicability of the discrimination principle.

A similar non-state actor challenge is present in other type of cyber attack - confidential data acquisition - named Aurora (2009).[11] However, in this case, there is an additional constraint to the applicability of IHL. This case qualifies as peacetime cyber espionage and falls into the 'grey zone' of IL. In the TM 2.0, it is explicitly highlighted that "non-violent operations, such as psychological cyber operations and cyber espionage, do not qualify as attacks" (Schmitt et al., 2017, p. 415). Subsequently, such liability gaps result in a tension on the legal level between a justified military response to a given state actor according to IHL and law enforcement towards a hybrid or non-state actor which is further complicated by the chosen methods.

Considering the above-mentioned points, it is unclear what would be the legitimate response of an attacked state to its offender (Schmitt, 2017, pp. 249-251, 254). Whereas it is notable that such cases of cyber espionage may potentially result in an economic crisis due to corporate data about future and current projects being stolen and, thus might even potentiate future attacks on technological products such as robotics (data including pieces of codes for different technological products). The existence of non-state actors is also a challenge for the more recent cases that are actively discussed in the media:

[11] *See* https://www.theguardian.com/technology/2010/jan/28/cyber-attacks-hacking

The 2016 U.S. presidential elections (Schmitt, 2017), BREXIT, WannaCry, and NotPetya. However, diplomatic measures such as expelling diplomats, economic sanctions, and criminal indictments were enacted in some of these cases despite the existing liability challenge.

2.2 Values and State-Centered Security

In addition to the challenges that were raised from the application of TM 2.0, like responsibility gap (caused by the tension with the discrimination principle) and conceptual gap (*casus belli*, inapplicability of IHL to a vast majority of cyberwarfare cases according to the Standard view), another less covered point shall be raised. Specifically, the fact that the "traditionalist" account of JWT underneath TM 2.0, focuses mainly on the state-centered and sovereignty-centered values like territorial integrity and protection of state borders, territory, objects and citizens from external threats (Schmitt et al., 2017, *see for example* p. 11-29). In the webcast from the launch of the Tallinn Manual 2.0 on the International Law Applicable to Cyber Operations at the Atlantic Council in Washington, D.C., the panelist and director of TM 2.0 - Michael Schmitt, made a comment that confirmed the underlying idea that state-centered values are still very dominant in the project, despite the shift to peacetime cyber operations.[12] In his speech, he makes the following statement:

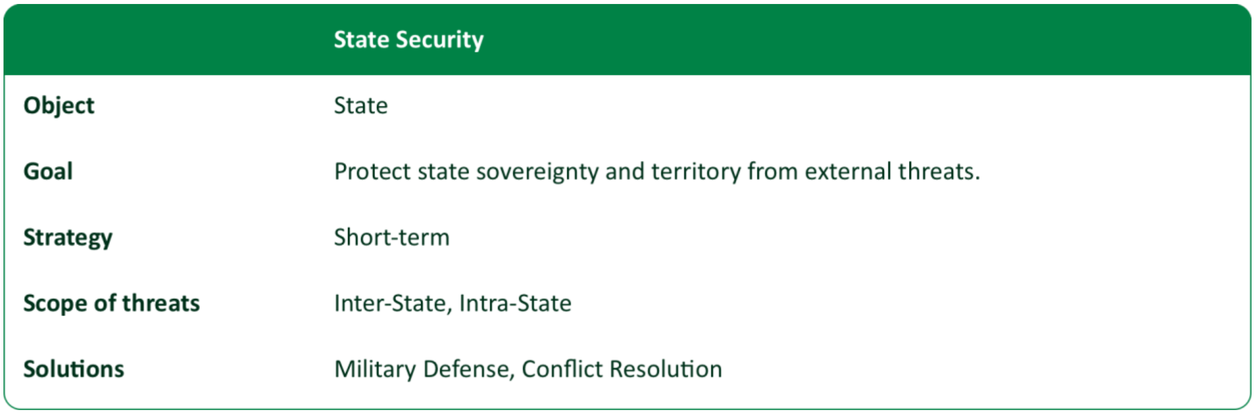
> We have some ground from which we can develop some norms because I believe that we all are operating for roughly the same perspective (he is talking about the other experts who represent their countries here). So, the big elephant in the room that we have in a project is: is this a helpful project? *From a state perspective?* (original accent here). We were not writing for academics, we are writing for countries, for legal advisors.[13]

In accordance with Schmitt's line of reasoning, the conceptual core of TM 2.0 - IL - starts with the idea of a state-centered security that mainly focuses on 1) the territorial integrity, and 2) the protection of state borders, territory, objects and citizens from external threats. To make this claim clearer, I explain this concept in Table 2 by referring to the initial domain of the current and ongoing operations since the end of the Cold War debate in International Relations Studies (Mahmud et al., 2008).

[12] *See* webcast. Retrieved from: https://www.justsecurity.org/37559/tallinn-manual-2-0-international-law-cyber-operations/

[13] This particular part of the Schmitt's speech can be found starting from 32 minute of the webcast.

Table 2. State-security paradigm.



	State Security
Object	State
Goal	Protect state sovereignty and territory from external threats.
Strategy	Short-term
Scope of threats	Inter-State, Intra-State
Solutions	Military Defense, Conflict Resolution

The position underneath state-security paradigm in TM 2.0 requires a critical appraisal. Even though TM 2.0 tends to indicate and locate the threshold for a broader range of cyber operations by specifying what ‘armed conflict’ in cyber domain entails, it implicitly prioritizes the cases that are classified as ‘armed conflict.’ In that regard, and also considering the objective emphasized by Schmitt in his speech, it is possible to point out in the current stage that in TM 2.0 the state, instead of the people, is the central object of security. Furthermore, by informing the short-term solutions, TM 2.0 provides guidance for the cases that justify military response and construct conflict resolution strategies like sanctions and other types of economic responses. Hence, the attempts to broaden its scope to the peacetime regulations and human rights resulted, in fact, merely in acknowledgement of the existence of the ‘grey zones’ of IL, and case-by-case discussions that follow state-centered objectives.

The potential reason behind such an emphasis on state-centered security was explained by von Heinegg’s (2014) in regards to TM 1.0: “If governments accepted the application of a unified body of international rules on cyber security that no longer distinguishes between State conduct on the one hand and the conduct of non-State actors on the other hand, the military option would simply be unavailable” (p. 16). Human-centered values are introduced. However, they are incorporated in a very narrow way, since the analogy to the traditional types of warfare ultimately bounds the scope to a state perspective and its objectives and concerns.

Moreover, just like in IL and correspondingly in JWT, the value of state security is regarded over the value of respect for human rights (Meredith & Christou, 2009, p. 6). A similar pattern can be identified in TM 2.0. As the Rule 22 states: “Countermeasures, whether cyber in nature or not, may not include actions that affect fundamental human rights, amount to prohibited belligerent reprisals, or violate a peremptory norm. A State taking countermeasures must fulfill its obligations with respect to diplomatic and consular inviolability" (p. 122). However, right after this rule is introduced, it is indicated in the point four that the convergence upon the requirements for this rule are rather problematic.

> The open question is the degree to which the prohibition extends to other human rights. For instance, cyber activities raise concerns regarding the right to privacy (Rule 35), thereby begging the question of whether a cyber operation that affects this right may qualify as a countermeasure or, instead, is precluded on the basis that the right is 'fundamental', as that term is understood with respect to Article 50. The International Group of Experts could not achieve a consensus on this point. (…) As discussed in Rule 34, whether or how human rights apply extraterritorially is unsettled and controversial. (Schmitt et al., p. 123).

Based on these generic observations of the correlation between state-security paradigm and objectives of TM 2.0, the conclusion follows that TM 2.0, similarly to TM 1.0, exclusively addresses state objectives of cybersecurity concerns and is thus mutually tied with the state-centered security paradigm. Because of such a focus in TM 2.0, the core values are sovereignty-oriented and state-centered, which makes the applicability of the manual for regulation of cyber operations rather problematic when other types of cyber operations are considered seriously, including the issues like threat to people and their well-being, education, health, democracy, income, and political freedom. On this note, let me turn to the sections 2.3 and 2.4 and the proposals that were made in the ethical debate around cyberwarfare that preceded the issue of TM 2.0.

2.3 'Skeptics': Against State-Centrism

The 'skeptics' are opponents of the analogy approach to the ethics of cyberwarfare. Skeptics have been explicitly acknowledging the significant alterations brought by the development of cyber weapons and their disparity with traditional types of warfare. While being seriously concerned about emerging and consolidating ethical challenges, they argue that JWT is unable to comprehensively accommodate these new inquiries. Subsequently, they are skeptical about the necessity to frame cyber attacks under the paradigm of warfare (Rid, 2012; Lin et al., 2014).

This position was suggested by Thomas Rid in the paper with the provocative title "Cyber War Will Not Take Place" (2012). He argues that cyber war has never happened in the past, does not take place in a present, and it is unlikely to occur in the future (p. 6). According to Rid (2012), this statement is true especially if one reads more thoroughly Clausewitz's conception of war and takes its focus on violence as a starting point (p. 6). In contrast to the reading that Arquilla and Ronfeldt (1993) make of Clausewitz (p. 32, *see Appendix A*), Rid puts an emphasis on the necessity of physical and lethal aspects of act in order to be considered cyberwarfare (Rid, 2012, pp. 7-8). Rid (2012) considers destruction and damage of the objects resulted from cyber operation as an insufficient criterion.[14] He claims that cyberwarfare is "a potentially lethal, extrinsic, and political act of force conducted through malicious code" (Rid, 2012, p. 6).

Thus, Rid (2012) takes his own stands in reading Clausewitz while deriving the conclusion of what should be count as cyberwarfare and what not. Such reading of Clausewitz, however, fails to address the significance of information (Clausewitz, 1997, pp. 64, 149) and the remark about the diversity of wars (Clausewitz, 1997, p. 22). At the same time, as Stone points out, the value of Rid's argument "rests not so much on whether it is right or wrong, but on its capacity for provoking debate around cyberwar and, by extension, about the fundamentals of war itself" (Stone, 2013, p. 107). By critically examining Rid's position and exploring the three concepts of force, violence, and lethality, Stone (2013) demonstrates that cyber attacks can be constructed as acts of war that are aligned with Clausewitz's conceptual underpinning.

Another skeptical point of view that is more relevant in the context of other-than-state values is questioning the relevance of JWT in regulation of cyber issues. In their paper "Is Warfare the Right Frame for the Cyber Debate?" (2014), Lin et al. observed that cyberwarfare will remain an ethically difficult issue until solid and responsible cyber policy is developed. As they suggest: "We may plausibly reframe the problem not as warfare but as a private defense, i.e., self-defense by private parties, especially commercial companies, as distinct from a nation-state's right to self-defense" (Lin et al., 2014, p. 40).

In this regard, the authors pointed out the relevant aspects and asymmetry in the measure of justice to victims and deterrence for aggressors. They claim that new ethical and philosophical problems that are created by cyber attacks are currently "on the shoulders" of private individuals and will be so until policymakers encompass these issues in regulating such new forms of aggressions (Lin et al., 2014, p. 56). According to the authors, JWT, which was relevant in the context of traditional types of warfare,

[14] To consider: since 2012 there were quite a lot of physically harmful attacks, especially in the industrial settings. *See* Moreno (2018).

is to blame as a cause of this regulatory gap. A similar theme emerged in the interview with R1, whose point was also objecting the necessity of the military context and warfare as a frame.

Indeed, a large amount of security issues in the cyber realm are currently on the shoulders of individuals and private parties. However, even though the regulations and liability for smaller types of attacks, as it was shown in the previous chapter, are still insufficiently covered in TM 2.0, certain directions for future research were put forward. Since 2014, the year when Lin et al. published their article, an important step towards establishing international regulations for cyber operations was made, even though it does not yet provide a full-fleshed account of regulatory practices for the cases that fall below the threshold. In other words, military context is no longer applicable to the types of cyber operations that are below the threshold level and are recognized by IHL as an act of war and are deemed as the 'grey zone' of IL. Thus, all-things-considered, the explicit message of the critique provided by Lin et al., (2014) is slightly outdated. Nevertheless, the key message of their critique is aligned to what this thesis attempts to make an explicit point: since people cannot grant sufficient level of security for themselves in order to be resistant to cyber attacks, there are always third parties involved which are providing security as service.

To be more precise, neither Rid's (2012) nor Lin's et al., (2014) offer constructive comments on the missing values, which cannot be accommodated when framing the regulation of cyber issues through the lens of military context. Rid disregards this point at all and focuses only on the provocative skepticism about the seriousness of cyber attacks and the necessity to be cautious about them as if they were acts of war. Lin et al., (2014), however, appear to implicitly indicate some of their worries about the disregarded human-centered values. In particular, while addressing their worries about 'private individuals' and resting 'on their shoulders' problems caused by cyber attacks, the underlying message they pinpoint is the clash between state-centered values and human-centered values in advantage of the first.

Thus, a 'beyond line' reading of Lin et al.'s argument (2014) would imply the idea that JWT and military context disregards the human-centered perspective on values like justice, liberty, and others in favor of valuing the protection of state borders, territory, and citizens from external threats. In such way, the skeptical argument proposed by Lin et al. appears to make a stronger critical claim towards the applicability of JWT to cyber attacks. This claim is aligned with my observations, which I share in the next chapter. But, before that, let us make a critical appraisal of the approach proposed by 'radicals' in the next subsection.

2.4 ‘Radicals’: Onto-Centric Values

The ‘radicals’ argue that JWT is solely insufficient to comprehend this relatively ‘new’ cyber domain. Thus, an analysis of the theoretical foundations of JWT is required to adjust its “traditionalist” account to the novel challenges brought by the transversality of cyberwarfare (Taddeo, 2012, 2014; Sleat, 2017, pp. 2-3). The core proposal suggested by ‘radicals’ is to revise such theoretical foundations to advance it with adjustments to the numerous manifestations of technological developments and pinpoint towards new values (Taddeo, 2012, 2014, 2016). Taddeo (2012) claims that the establishment of a functional model that will encompass informational infrastructures and data as relevant aspects in an ethical analysis based on the JWT is the right way to go. Correspondingly, in several papers, Taddeo (2012, 2014, 2016) proposes a collaboration between JWT and Informational Ethics (IE) towards one new approach for the comprehensive ethical analysis of cyberwarfare: Just Cyber War (JCW).

Taddeo (2012) largely criticizes the ‘moderates’ position by arguing that JWT is an anthropocentric theory, which in the case of cyberwarfare leads to the framework’s insensitivity because of the involvement of information and digital infrastructures and conforming to these technological expansion values. Her solution is brought by the claim that to enhance the framework’s sensitivity, JWT should be enriched by the ontological basis of macroethical theory - Informational Ethics (IE), proposed by Luciano Floridi (1999). In other words, to provide ethical and legal frameworks that will encompass all the specifics of cyberwarfare, it is necessary to incorporate the significance of informational beings into the principles of JWT and make a shift from an anthropocentric to a non-anthropocentric theory. This way Just Cyber War (JCW) framework emerged as onto-centric theory.

Even though I consider Taddeo’s intention to incorporate the idea that values are changing with a dramatic technological expansion as a relevant point to consider, I find her account of values as misguiding. In order to increase clarity, two inquires need to be set aside: 1) the deontological question of whether cyberwarfare will ever be able to satisfy the principle of JWT and the requirements of IHL; and 2) the utilitarian question of whether cyberwarfare will ultimately prevent or cause more suffering in combatants and civilians. Instead, I suggest delving deeper in IE to address a question that lacks attention, that is, the significance of the values behind its onto-centric ontology in cyberwarfare.

Taddeo justifies her choice by relying upon Floridi’s (1999) assumption that IE is based on an ontology of information, and therefore, is non-anthropocentric but onto-centric. Because it is an ontology that sees reality as an “Infosphere” that consists of not just human beings and non-human beings, but of so-called “informational beings” (Floridi, 1999), it shifts from merely human agents to the inclusion of both humans and non-humans actors and actants. One might ask here, what is the difference between “Infosphere” and “cyberspace?” Following Taddeo’s clarification (2012), “Infosphere” is not the same phenomena as cyberspace because it is not a domain, but rather an ontological dimension of the contemporary world. Computers, robots, networks and even pieces of code are similar to human agents and objects from environment like grass and flowers, they are all equally considered as inhabitants of the “Infosphere”:

> As remarked in (Taddeo and Vaccaro 2011), the Infosphere is the totality of what exists. The Infosphere includes agents and objects, relations and processes, as well as the space within which they act. It is not to be confused with cyberspace, as it includes online as well as offline and analogue domains. (...) The Infosphere is the environment in which animate and inanimate, digital and analogue informational objects are morally evaluated. (Taddeo, 2012, p. 214)

In other words, IE considers humans as informational beings similarly and equally to any other animate or inanimate being such as wind and water, or AI and smartphones. Hence, as it is argued, every informational being has a basic moral right to flourish, and thus, the ontological equality is a core conception in IE (Floridi, 1999). Considering the idea of ontological equality, Taddeo attempts to cover the gaps that result from the application of traditional anthropocentric theories of warfare to cyberwarfare. Taddeo (2012, 2014) demonstrates that such combination of the frameworks increases sensitivity to different types of actors: human, non-human, and informational actants and actors. By doing so, it discloses the ethical issues that are at stake in cyberwarfare.

Taddeo (2012, 2014, 2016) emphasizes that the main benefit brought by this perspective is that the framework’s sensitivity provides the possibility to ethically encompass the damage and destruction of various targets like information, access to personal data, and digital governmental infrastructures. Following the author’s own words: “A human being, who suffers the consequences of a cyber attack and an informational infrastructure that is disrupted by a cyber attack are both to be consider the receiver of the moral action. The morality of that action will be assessed on the basis on its effect on their rights to exist and flourish” (Taddeo, 2012, p. 215).

Thus, the value that goes hand-in-hand with ontological equality and grants flourishing, in Taddeo’s sense, is ‘informational integrity.’ Taddeo (2014) claims it to be central in the JCW framework (p. 132). Moreover, following Taddeo’s addition, on the one hand, it is possible to claim that she proposes a "revisionist" account of JWT in such onto-centric version that significantly affects the principle of noncombatant immunity. This point is vital to Taddeo's (2012, 2014, 2016) argument, since JCW wants to target humans, even civilians, if they are a threat to the Infosphere.

While challenges of false positives and false negatives remain open for discussion in TM 2.0, Taddeo's perspective will not make any beneficial improvement to solve the regulatory gap. Rather, it will create even more tension with the principle of discrimination. Take for instance the involvement of hybrid actors, who are state-sponsored actors (like in the case of attacks on Estonia). Before their relation to the state is officially confirmed, they are accounted as non-state actors by IL. Thus, according to the principle of discrimination, they cannot be targeted. Therefore, Taddeo's point to disregard non-combatant immunity when threat is imposed to the Infosphere (2014) is a rather dangerous shift that might potentially result in non-state actors holding legal responsibility instead of state actors, who sponsored the attacks.

However, while this point is at least discussed by 'moderates,' 'skeptics,' and 'radicals,' the question of other-than-state values remains opaque. Does 'information integrity' sufficiently represent the value that proves to be relevant in ethical analysis of cyberwarfare and cyber conflict at large? Even though information integrity probably can be considered as an extrinsic[15] value that supplies TM 2.0 with novel values, it still serves nearly as a state-centered concern. Indeed, it is definitely not as an intrinsic value in a human-centered sense. In order to support this claim, I suggest exploring two objections to the Taddeo's onto-centric proposal of JCW. While the first objection will be raised to IE as a solution in general, the second objection will tackle the value of 'information integrity' in particular.

2.4.1 Challenging the Genuine Value of 'Information Integrity'

Although JCW seems to overcome the anthropocentrism and confusion of traditional theories, it opens up foundational concerns about a whole branch of new questions that make the applicability of the framework even more complex and unproductive. Sleat (2017) argues that such a modification of JWT was "unable to get a sufficient grasp on the questions that a theory should enable us to ask" (p. 5). He supports this point by stating that "any adequate ethical theory for regulating even conflict in the cyber realm needs to retain what I shall call the 'human perspective'" (Sleat, 2017, p. 5). In other words, 'a

[15] Intrinsic value: IIF is worthy of desire in and for itself. Extrinsic value: the value that something has in virtue of its extrinsic, relational properties. *See*: Lemos (2015) and Zimmerman (2015).

view from the universe' suggested by Floridi (1999) as a core aspect of IE is insufficient in capturing the relationally that is always bounded with someone's perspective. Thus, such 'a view from the universe,' leads nowhere since humans will still interpret the cases that are waiting to be regulated.

The critique also relates to the meaning and relevance of the information that is given and constituted only by human perceivers. Volkman (2011), for instance, goes a bit further than Sleat in his critique of IE, suggesting that the value of information can be obtained only through a 'human perspective' in the unfolding story of actual life. The central message that should be extracted from Sleat's (2017) and Volkman's (2011) objections to IE is the following: because IE fundamentally relies on 'the ontological equality' and 'a view from the universe,' it suggests a universalized generalization and presuppositions without a critical questioning of their coherence with actual cases and relevant values from the lives of actual human agents in each particular cultural context.

The latter point was extensively discussed in different regards by philosophers Brey (2007) and Ess (2008). The vagueness and totality of values like 'ontological equality' and 'informational integrity' are imposing the challenge of unintended prejudices and discrimination in the cultural diversity of the contemporary world. So far, IE proved to be limited in its potential to encompass the variety of meanings that loaded the interpretations of the values in each particular cultural context around the globe. This was pointed out by Brey (2007) who critically examined IE in a cross-cultural context. The author based his premise on moral relativism and the fact that the values are profoundly different across societies and throughout history. Correspondingly, he claims that, on the meta-ethical level of analysis of moral relativism, truth of moral judgment is relative to societies and histories.

Furthermore, despite globalization and liberal initiatives, some values tend to consolidate into universally recognized and accepted worldwide intrinsic values. It is unlikely that these are 'ontological equality' and 'informational integrity.' A similar point, on a meta-ethical level, was raised by Brey (2008) in response to Floridi's argument that "information objects have intrinsic value and are therefore deserving of moral respect." Brey (2008) objects that, even though information objects do deserve some respect, it does not make them intrinsically valuable, but rather "the value at issue is instrumental or emotional value for others."

This objection was addressed by Floridi himself, where he indicated that this critique appeared because "Brey confuses a causal with an inferential reasoning" (p. 192). Floridi (2008b) continues with clarifications to this claim: "The actual argument seeks to establish that entities deserve respect *because* they have intrinsic value, not that *if* entities deserve respect *then* they have intrinsic value" (p. 192). Nevertheless, the claim itself is further justified in a quite vague way, where Floridi (2008b) concludes on a rather rhetorical note suggesting that “we should not fear to respect the universe too much” (p. 193). This, however, seems to be weak justification that instead of directly addressing Brey’s objection, finds a way to reinterpret it as merely a wrong understanding of the underneath logic.

Let me demonstrate in my own words why I consider Brey’s objection to be cogent. Although in the contemporary world interconnectedness and the internet are widely spread, the affordance of such technological developments is not equally distributed around the globe (DiMaggio & Hargittai, 2001). Certain countries benefit from technological progress and the expansion of information technologies more than others. Therefore, in certain countries, the value of ‘information integrity’ is more likely to be accepted reluctantly as an extrinsic value (at least in connection to concerns about privacy) than in the countries where intrinsic values (e.g. care, justice, civility) are violated on a daily basis. Now, ten years after Floridi replied to Brey, ‘information integrity’ is still far from being recognized as intrinsic value even in technologically developed countries such as the Netherlands, Estonia, and the U.S. One of the reasons for this might be that this value falls apart from traditional values. While being recognized around the world, it still implies different meanings and practical interpretations.

Following Brey (2007), a fair critical point to IE can be provided when distinguishing contextual, institutional, and behavioral differences as a key aspect of the informational environment. Based on this critique, he concludes that, although it is hard to present cogent arguments for the universal truth of particular moral principles and beliefs, it is still possible to develop a rational argument for/against particular moral values and overarching conceptions of ‘the good’ across moral systems. His particular suggestion was to enhance IE to the level of Intercultural Information Ethics (IIE) where “IIE should aim to interpret, compare and critically evaluate moral systems in different cultures regarding their moral attitudes and behavior towards information and IT” (Brey, 2007, p. 9). However, this fruitful future trajectory for the further extension of IE into IIE has not yet manifested in a more advanced framework.

A similar critique, though in a less constructive manner, was proposed by Charles Ess (2008). The author claims that IE is limited in its ability to include diverse cultural and ethical traditions and unify them merely within its approach. To put it into the author's own words:

> The equally significant divergences make the point that any effort to establish such universal norms and values in the domain of information ethics may remain only partly successful and delimited by irreducible differences between diverse cultures that must be acknowledged in the name of respecting cultural integrity. In other words, *although the relativist’s story is not the whole story, neither is it a false story* (emphasize added). (Ess, 2008, p. 201)

Therefore, to sum up, neither IE radically solves the challenges of JWT in the context of cyberwarfare, nor does it fill in the existing value gap with the suggested onto-centricity. By proposing instrumental values as the intrinsic one, JCW makes the applicability of JWT even more abstract and vague. Moreover, since Taddeo's proposal rests beyond mere concerns about the basic philosophical core, but is put forward as an updated moral theory for regulation of cyberwarfare, the clarity it brings to the other-than-state values does not compile with the actual values that are at stake.

Thus, the challenges of IE that were demonstrated above resonate with the JCW proposed by Taddeo. In other words, although 'radicals' tend to advance the existing approach according to informational relevancies of the contemporary world, the value gap still cannot be sufficiently filled in resting solely on 'information integrity' as the cardinal value that is put forward as intrinsic. Particularly, JCW disrespects the importance of the questions about the distribution of benefits and risks in online-offline divide that currently exist.

Because the value of 'information integrity' is too abstract, it does not pinpoint towards intrinsic values that are relevant in each particular cultural context. Despite Floridi and Taddeo claiming it to be intrinsic, it does not follow that it is actually recognized as such, or will be in the future. This way, apart from the contribution to the discussions around ethics of cyberwarfare, the 'radicals' onto-centric account may suggest the enhancement of the state-centered security, where the information systems and other state property can be protected for the sake of the state's objectives and aims.

Moreover, another reason to consider is that JCW, if applied, is not concealing moral issues raised on the matter of the principle of noncombatant immunity. Rather, JCW will violate the traditional anthropocentric focus of that principle which protects civilians and leaves some hope that human-centered values will be one day incorporated. That is why the critique I suggest does not indicate Taddeo's depiction of values in the scope of onto-centrism as wrong, but it rather considers onto-centrism as a misleading 'lens' when proper regulation of cyberwarfare is aimed to be achieved.

2.5 Conclusion

In this chapter, a critical examination of the TM 2.0 was presented. Key issues on the conceptual and regulatory gaps were introduced, definitions and elaborations were provided, and from this basis, challenges in the legal, military, and ethical discussions around cyberwarfare were illustrated. Since the

TM 2.0 is a set of guidelines on how IL applies to cyber conflict that predominantly relies on JWT, there are certain conceptual and legal limitations in its applicability which were framed here as the regulatory gap.

In particular, these so-called ‘grey zones’ - when IL does not cover the types of cyber operations which do not constitute an act of war - are actually the area within which the vast majority of the currently known cyber operations fall, and subsequently, remain untackled. Moreover, a brief discussion of “traditionalist” and “revisionist” accounts of JWT was elaborated in regards to TM 2.0 and the ethical issues raised there on the matter of noncombatant immunity. These challenges were introduced under the overarching scope of the tension that appears when JWT - which is fulfilling its role with traditional types of warfare - is applied to cyberwarfare yet remains unproductive and limiting in that context.

Then, after I indicated the dominant state-centered security paradigm that underlies IL and JWT, while at the same constituting the basis of TM 2.0, a critical examination of two main ethical lines of reasoning considering JWT were offered. In the section 2.3, I started with the position of ‘skeptics,’ who are largely critical about the necessity of the military context and JWT as a frame for the regulatory debate around cyberwarfare. Some of the proponents of this position implicitly object to the state-centered values which was highlighted, since it is aligned with the core idea of this thesis.

At the same time, in the section 2.4, the third group was introduced - the ‘radicals’ - who integrate philosophy of technology with JWT in order to adjust the latter to the novel ethical issues raised by cyberwarfare. Proposed by them, the JCW framework was elaborated and further challenged in the final section of this chapter in order to show that the genuine value of ‘integrity of information systems,’ despite being onto-centric orientated, is still serving the state-centered security paradigm when applied. Finally, it was shown that due to the relevance of the distribution of risks and benefits of new and emerging technologies and the correspondence with cultural heterogeneity in the regulation of cyberwarfare, the onto-centric core of JCW appears to be context-insensitive and ultimately does not fill in the existing value gap.

3. Aretaic Reading of Technomoral Care in the Tallinn Manual 2.0

In this Chapter, I suggest a different reading of TM 2.0 which enables a broader depiction of other-than-state values that are at stake. In particular, in section 3.1.1, I offer an aretaic reading of JWT to zoom in to the virtues required for the 'selfless service' of military personnel. In doing so, I explain the contributing framework of technomoral virtue ethics and stress its relevance in the light of new and emerging technologies compared to deontic ethical theories. In section 3.1.2, I focus on 'care' in TM 2.0 and the opportunities that are brought by potential ways of reading it as different types of practices. Based on this analysis, in section 3.2 I elaborate on *care* as a value and *care* as a virtue, of which its importance is largely overlooked in the ethical debate around TM 2.0 and cyberwarfare more broadly. Finally, in section 3.3, I will discuss the findings about technomoral care in TM 2.0 in the scope of human-centered values shared by a human security paradigm. The research sub-question that I address here are following: *What does an aretaic reading of Tallinn Manual 2.0 imply? What is the relevance of technomoral care in the Tallinn Manual 2.0? To what extent does the analysis of the dynamics in the technomoral care reveal and conceal ethically relevant aspects in the Tallinn Manual 2.0?*

3.1 Ethical Underpinning and Human-Centered Values

As it was shown in Chapter 2, apart from its state-centered values, the ethical ground of TM 2.0 is essentially based on a deontological (IL and JWT) combined with a consequential (the Standard view) moral ground. According to Swierstra (2016), the rule-based ethical arguments such as deontological and consequential are typically favorable in contemporary societies. Swierstra (2016) critically observes: "In modern, pluralist and liberal, societies public debate is conducted mainly on the basis of this kind of rule ethical arguments. The reason for this is that rule ethics is considered as sufficiently objective and impartial to allow for a reasonable consensus in many cases" (p. 5). However, since in the case of TM 2.0 these two normative theories are not comprehensive enough to fully grasp its content, an additional contributing perspective is required. Let me elaborate on the reasons behind this claim by first putting aside the consequential perspective.

The Standard view is a consequence based account that is central for TM 2.0. It holds the consequential position that considers the expected outcomes or consequences of an act to determine whether or not that act is morally permissible (Bentham, 1890). Although a consequentialist perspective on the classification of the outcomes presumes their commensurability, it is impossible to distinguish, according to the consequentialist point of view, whether some of the outcomes should be marked as special in comparison with others, especially when technologies are at stake (Reijers & Coeckelbergh, 2017, p. 98). Following John (2007), “it is possible to argue from a consequentialist standpoint, which views all outcomes as ultimately commensurable, to a view that certain outcomes ought to be treated with particular care if we can show that a policy which does not treat those outcomes as special is likely to lead, over time, to worse outcomes than would have occurred had we adopted a policy which treated that class of outcomes as special." Thus, turning back to TM 2.0, I put aside a consequential perspective as non-exhaustive since it does not satisfy the demands of good quality policy making and legislation due to the absence of stable requirements or gradation of special outcomes (Held, 2011, p. 180).

3.1.1 The Tension Between Deontic and Aretaic Moral Ground

Even though it is a less popular point of view, there were initially two main components in the morality of war: deontic - the absolute moral law, and aretaic - the virtues of individuals who were forced to fight in war (Beard, 2014). A deontic ethical basis of JWT relies on the following normative argument: an action is right IFF it is consistent with a universal maxim (law, principle etc.) as referring to the Kantian formula - an action is morally right only if it can serve as a universal law (Kant, 1998, p. 14). Walzer (2017), who puts an emphasis on the deontological account of JWT by aiming to develop a universally applicable and accepted morality of war that will prescribe or forbid practices in accordance to the existing standards of communities around the world.

Indeed, as it was demonstrated on Chapter 2, Walzer’s perspective resulted in the “traditionalist” account of JWT - a set of deontological rules which ought not to be violated (Walzer, 1977). Walzer wants a recognition of a human, “common” morality that would deny any claims of a “duty” to do anything which is supposedly more important than humanity like “reason” or “nation” (Walzer, 1977). Nevertheless, deontology is still merely explaining basic principles which may not be violated intentionally. But it does not provide sufficient instructions when one particular law is applied to different cases with different outcomes. There are many historical examples of the dramatic losses resulted from blind obedience to orders that seem to be justified by ‘universally’ agreed principles and laws (*see example in* Anscombe, 1958).

Thus, the significant obstacle of deontology is the following: in order to arrive at the universal maxims that will capture and encounter the extensive technological impact on politics, society, and individuals in a single law, one should be familiar with the diverse, visible, and invisible implications of new and emerging technologies. As Reijers & Coeckelbergh (2017) critically remarked: “Kantian ethics suffers from the restraint to offer universal rules to concrete and complex practical contexts” (p. 98). Since “when two or more moral duties clash, we have to look at individual situations in order to determine which duty will override another,” without such detailization, it would be particularly difficult to anticipate the rightness of action. Potential prioritization becomes ethically and morally problematic (Tavani, 2013, p. 59).

Just war principles create particularly problematic issues because in the cases of moral dilemmas caused by novel technologies, like drones, weaponized autonomous robotics and cyber weapons, there is no proper mechanism for resolving such moral dilemmas (Gertz, 2014, Vallor, 2016). This was illustrated in the movie “Eye from the sky,” where a drone operator was given the order in correspondence with the law to target civilian during operation. This civilian was a small innocent girl. The duty clash that raises significant ethical concerns has been afforded by technology in this case. The abilities to see, understand, and realize provided by the drone to its operator led to a traumatic experience of murdering a small girl. As it was suggested by sociologist Bruno Latour (2005), technologies can “authorize, allow, afford, encourage, permit, suggest, influence, block, render possible, forbid, and so on” human actions (p. 72).

A similar idea is shared by proponents of postphenomenology. Following Don Ihde’s (2010) technological mediation, this case resulted in the human-technology relation that lead to a change in perception, and correspondingly, a drone operator’s mediated experience of the situation. Even though it might seem that drone operator or cyber warrior are “cubicle warriors,” philosopher of technology Nolen Gertz (2014) uses mediation theory to show why technologies might not “shield” from trauma but actually make it easier to experience it and contribute to PTSD that they have been diagnosed with (p. 109). Thus, it is important to notice that “technologies affect our actions not just by altering the course of action (like billiard balls do to each other) but by mediating our reasons or motives to act in a particular way” (Swierstra & Waelbers, 2012).

From this view, I propose to recall the tradition of the aretaic reading of JWT, which will be however suitable for application in the complex technological context. An aretaic ethical theory is technomoral virtue ethics. It is an agent-oriented, technology sensitive, and ‘subjective’ framework with emphasis on the experience of human-technology interaction.[16] In accordance to the idea that technologies are not neutral, in the book "Technologies and virtues," (2016) American philosopher Shannon Vallor adjusts virtue ethics to the irrevocable changes brought by technological advancements. She elaborates on the empirical cases that evidently point out on that dynamic. Vallor (2016) emphasises the theory of technomoral virtues as beneficial in two ways: it is pluralistic (open to more than one mode of expression of well-being) and malleable (adapted to the needs and affordances of the present human condition and environment) (p. 44).

Vallor's definition of virtues is aligned with the Aristotelian tradition and interlinked with many contemporary aretaic ethicists. On a very basic level, in almost every modification of virtue ethics the primary claim is that an action is right IFF it expresses some virtue (Aristotle, 1934/1996; Chappell, 2013). Initial tradition has its roots in Ancient Greek philosophy, where the main emphasis was on the individual agent who performs the actions and on her development and moral nature. In Nicomachean Ethics, Aristotle discusses different virtues as ways of seeing circumstances and actions with particular emphasis on those that are central for the excellence of character (Aristotle, 1934/1996).

A virtue - or *arete* - is a disposition of voluntarily and deliberate choice. It is the power that shapes an agent's action, but without force and pressure (Chappell, 2012, p. 48). Contrarily, vice is deficiency or excess of certain characters. Following the nature of virtues, in The Golden Mean rule, the moral behaviors are the mean between two extremes - at the one end excess, at the other, deficiency (Aristotle, 1934/1996; Porter, 2017, p. 87). Yet, only balanced perception-of and response-to each particular class of actions or feelings (i.e. without excesses and deficiencies) may correspond with the mean.

In contrast to traditional types of aretaic theories, Vallor puts an accent on the role of technologies (Vallor, 2016; Reijers & Coeckelbergh, 2017). This makes technomoral virtue ethics specifically advantageous in comparison with other modern aretaic developments. The starting point of her argument describes the essential reasoning behind the connection between ethics and technology where "technologies invite or afford specific patterns of thought, behavior, and valuing; they open up new possibilities for human action and foreclose or obscure others" (Vallor, 2016, p. 2). In other words,

[16] According to Chappell (2012), for Aristotle, in comparison with other moral theories, there is no formula for right action. Instead, the idea of becoming good, where the goodness of character determines the rightness of an action is central.

she argues that certain human behaviors and ways of perceiving are enabled, co-shaped, and guided by technological affordances.[17]

Vallor (2016) indicates the urgency of shifting mere virtues to technomoral virtues, which are engaged with the environment of the rapid development of new and emerging sciences and technologies (p. 32). Basing on Reijers & Coeckelbergh (2017) reading of Vallor's theory: "Virtues are defined as technomoral since they are consistent with the "basic" moral capacities of human beings but also result from new "alignments" of these capacities in line with our adaptation to a rapidly changing technological environment" (p. 98). Hence, Vallor in a very generic sense signifies the mediating role of technologies without explicitly referring to mediating role of technologies (Vallor, 2016, p. 27).

The first step in suggesting a technomoral reading of TM 2.0 is to recall some basics of the aretaic perspective on JWT. An important observation is that certain virtues of soldiers were emphasized dating back to Plato, Aristotle, Augustine, and Aquinas. In Clausewitz's "*On War*" (1997), for instance, the chapter five is dedicated to the military virtue of an army, with the emphasis on bravery, aptitude, powers of endurance, and enthusiasm (p. 153-157). Bravery, in this case, is often referred to as courage and, according to many authors, is the most relevant to military service in relation to fear and confidence (Aristotle, 1934/1996, p. 63; Robinson, 2006, p. 166; Vallor, 2016, p. 217; Howard, 2014, p. 158). Consider a soldier who faces a psychologically and morally problematic situation where she is required to risk her or someone else's life and take immediate action. In such cases, courage is a virtue that assists in doing the right thing in spite of the difficulties that are involved.

As it was demonstrated in the movie "Eye from the sky",[18] a soldier needed to follow the order and take an immediate action - to sacrifice a little girl in order to fulfill his professional duty. In these cases, a lack of courage could potentially cost hundreds of dead civilians if terrorists that were preparing an attack would not be targeted timely. Contrarily, excessive courage could result in rashness, which subsequently could lead to even more dramatic consequences with other civilians involved. The dynamic in virtue-vice continuity depends particularly upon the possibilities afforded by technology, like precisely counting the distance, using video cameras to scan the radius, and pressing the button to execute an attack in a dark and silent room thousands of kilometers away.

[17] Even though Vallor (2016) does not make a clear reference about this term, it is a solid concept in STS. Following Hutchby (2001), affordances can be defined as "functional and relational aspects which frame, while not determining, the possibilities for agentic action in relation to an object" (Hutchby, 2001, p. 448).

[18] *See* plot: https://en.wikipedia.org/wiki/Eye_in_the_Sky_(2015_film)

Hence, a certain degree of courage is necessary in order to fulfill professional duty in the military service. The lack of it, similarly to the excess, would lead to an incapability to function according to the required standards of military service. But as it was noticed many times, such preparation does not stand even close to the realities of war full of violence, innocent deaths, injuries and other terrifying experiences (Gertz, 2014). Thus, a preliminary practice of courage and care does not predict the excellence of character before a real situation that reveals a position in a virtue-vice continuity (Vallor, 2016).

Military personnel, at least in Western countries, voluntarily consent to a host of professional duties, at least such as the obligation to obey commanding officers and the duties to uphold the principles of JWT. This means that even in risky situations they should obey their duties and sacrifice themselves (or others) to, for instance, uphold the principle of proportionality. This is also part of the education and preparation processes that most military personnel are overcoming. That is why, in the education of military personnel, professionalism is deemed as a required excellence in the military personnel traits. An emphasis is put on the practice of a particular situation and the decisions that should be made there. As Vallor highlights in her book (2016):

> The core profession of the modern military vocation is the ideal of "selfless service." "Selfless service" expresses the moral virtues of care and courage, for military culture is rooted in our most basic moral understanding of our duty to care and make sacrifices for the security and well-being of those who depend on us most. A soldier is distinguished from a criminal, mercenary, or 'barbarian' only by virtue of this selfless service. (Vallor, 2016, p. 217)

The correspondence of security and well-being is not that straightforward as it may appear at first glance. Since the core basics of technomoral virtue ethics are tied with the Ancient Greek tradition, well-being, or *eudaimonia* is a central notion here. Referring to Aristotle, Vallor (2016) claims that the excellence of character leads to well-being, which consists of being rational and entails the development of both types of virtues, those of intellect and those of character. Oppositely to virtues, *eudaimonia* is a goal in itself, not a mean (Porter, 2017, p. 86). To put it differently, *eudaimonia*, or flourishing, is the natural end given in itself, but it is preceded by the reference to virtue as a way of seeing and acting, where it is up to the agent to give meaning and act in accordance with *eudaimonia* (Aristotle, 1934/1996).

Hence-wise, in order to sufficiently grasp the ultimate goal of ‘selfless service,’ one more significant point should be introduced. Aristotle initially argues that *eudaimonia* is not merely of ethical importance, but also of political relevance. This claim rests on the idea that individuals construct together a community, or to be more precise, a *polis*. Yet to have a prosper *polis*, each individual should contribute with their own excellence of character and be able to develop and practice advanced moral skills. Therefore, military personnel should ideally be trained to habituate moral and intellectual virtues that will not merely contribute but also provide and protect the possibility for habituation for others. Indeed, for Aristotle, *eudaimonia* is the ultimate objective of politics, which is also indicated by the very last words of the Nicomachean Ethics - “let us begin” (Aristotle, 1934/1996, p. 284).

However, as Vallor (2016) pointed out following MacIntyre, *eudaimonia* does not simply mean the total sum of individual efforts to live well. Rather, she highlights that “a moral practice is irreducibly social enterprise” (p. 49). Quoting Vallor’s observations: “If we consider the internal goods of global community, intercultural understanding, global justice, human security, and collective wisdom, it becomes immediately evident that each of these readily makes sense as part of a personal narrative in which an individual's beliefs and intentions are directed toward the joint achievement of these goods” (p. 49). Thus, following Vallor’s reasoning, the emphasis on the individual moral skills and habituation of virtues, in contrast to vices, should be promoted as the starting point in the discussions about social and political well-being and the ways to achieve the ultimate goal: *eudaimonia*.

Interestingly, similar ideas found their place in legal studies. Currently, legal theories are predominantly based on deontological ethics and nowadays face a so-called ‘legal turn’ towards aretaic ethical theories as a foundation of law. As Solum (2018) claims, the “fundamental aim of law should be the promotion of human flourishing, where a flourishing human life is understood as a life of rational and social activities that express the human excellences” (p. 6). I consider these preliminary observations serve as an arrow indicating the direction towards further investigations of an aretaic ethical theory beneath the legal frameworks that regulate cyberspace. With this in mind, before turning to the elaborations on technomoral care, let me provide an aretaic reading of TM 2.0 and propose my contribution that is potentially aligned with the tendency in legal studies known as the ‘legal turn.’

3.1.2 Referring to the Value and Virtue of Care in the Tallinn Manual 2.0

In this thesis, I consider 'care' as a particularly relevant notion since it is referred to in different rules of the TM 2.0 (Schmitt et al., 2017, pp. 279, 420, 476, 529). However, each mention of care is indicating different connotations for such notion, as well as a different actor who is responsible for expressing it by different means. I'll go through three out of four of the rules[19] where care is central in order to elaborate on how an aretaic reading of care as 'technomoral care' can resolve practical issues resulted from the existing conceptual unclarity in TM 2.0.

The Rule 59 (Respect for Space Activities) refers to a standard of care that requires states "to act in a manner that does not impede the exercise by other States of the rights they enjoy in outer space" (Schmitt et al., 2017, p. 279). So, in case of this rule, care is articulated as a standard of the state behavior and appears to be aligned with Walzer's account of the principle of discrimination. A problem in conceptual unclarity that leads to similar practical issues of the application of JWT can be found in the next rule. In Rule 92 (Definition of Cyber Attack), it is suggested that "care is required when identifying the originator of an attack" (Schmitt et al., 2017, p. 420). Here, except for the emphasis on the idea that the originator of an attack should not be confused with the intermediary, there are no specifications of what meaning should the notion of care entail.

While both these rules align to the proportionality and discrimination principles, without an aretaic reading of care, they would be uninformative. These rules raise questions such as: How is it possible to care in this case? Is care here only about the end result? Should intermediaries be checked without assigning them liability in a wrong manner? Or is the goal being cautious to not to approach the issue in a vicious manner by surveilling and violating the rights of other non-state actors that are suspects but innocent? According to Walzer's "traditionalist" view on the principle of discrimination, the notion would be care about outcome - no intermediary falls into a false positive category. When referring to these two rules, in Taddeo's sense, an onto-centric principle of discrimination, as found in JCW, would make the virtue of care unnecessary and even possibly unethical because it has no value from a state perspective, since its understanding is bounded by the deontological reading of the duty of care.

But is that all what we can extract from the notion of care here? Despite this rule being based on the principles of JWT, could an aretaic reading of care as a skillful practice here, without excess and deficiency, make practical guidance of this rule more transparent? Potentially, it could. First of all, the

[19] In the Rule 140 (Duty of care during attacks on dams, dykes, and nuclear electrical generating stations) it is stated that the "particular care must be taken during cyber attacks against works and installations containing dangerous forces" (p. 529). This rule will not be covered in this thesis.

clarity could be brought to the particular roles of involved stakeholders, since it is unclear who is responsible to act according to the duty of care: a cyber warrior, governmental entity, military institution, regular soldier or security agency? These are some of the issues that are raised when an aretaic reading of TM 2.0 is applied to the rule 92 and 59.

At the same time, in addition to the question “who should be responsible for the duty of care?,” I argue that the reverse questions “who provides security of the systems?” and “are those actors who provide security of the systems also responsible to address cyber attacks?” are legitimate and relevant in the scope of TM 2.0. This leads us to another rule that I consider to be more relevant for the scope of this thesis, Rule 114 (Constant care). It is formulated as following: “During hostilities involving cyber operations, constant care shall be taken to spare the civilian population, individual civilians, and civilian objects” (Schmitt et al., 2017, p. 476). Although previous rules were mainly referring to care in accordance with the principle of discrimination, this rule is no longer bounded by JWT. Even the opposite, despite the notion of care being abstract in a way here, it implies a certain excellence of character in order to be fulfilled. Due to the limited scope of this thesis, I propose in my discussions to solely focus on the latter rule and provide a comprehensive analysis in its regard. This rule is supported by seven elaborative points, two of which are particularly relevant:

> The law of armed conflict does not define the term ‘constant care’. The International Group of Experts agreed that, in cyber operations, the duty of care requires commanders and all others involved in the operations to be continuously sensitive to the effects of their activities on the civilian population and civilian objects, and to seek to avoid any unnecessary effects thereon. (Schmitt et al., 2017, p. 476)

> Use of the word ‘constant’ denotes that the duty to take care to protect civilians and civilian objects is of a continuing nature throughout cyber operations; all those involved in the operation must discharge the duty. The law admits of no situation in which, or time when, individuals involved in the planning and execution process may ignore the effects of their operations on civilians or civilian objects. *In the cyber context, this requires situational awareness at all times, not merely during the preparatory stage of an operation* (emphasize added). (Schmitt et al., 2017, p. 476)

In a narrow depiction of this rule, a state-centered reading of the value of care seems to diminish the categories ‘civil population’, ‘individual civilians’ and ‘civilian objects’ as if they were merely the objects of state security concern. In other words, this rule seems to make an implicit reference to the virtue of care, urging for care about civilians and their objects, but leaves ‘care’ here as a black-box. In comparison to the rules 92 and 59, the final sentence of the second point of the rule 114 is particularly pointing towards a value of care that is going beyond state-centered security, since “*In the cyber context, this requires situational awareness at all times, not merely during the preparatory stage of an operation”* (Schmitt et al., 2017, p. 476).

From these three rules that refer to the value of care, and especially, from the second elaborative points on the rule 114, it is clear that the notion of care appears as a black-box that is bounded by the state-centered focus on what the act of care as duty implies. On the one hand (Rule 59, 92), care could be understood in the sense of the principle of discrimination as mere duty of care without an in-depth insight into the different forms of its expression. On the other hand (Rule 114), the opaqueness of care could be avoided if such notion is understood in the sense of technomoral virtue ethics.

Here, elaborations on the notion of care could reveal in what way care, as both, the character trait, and as a practice, is developing to grant security to civil societies that face a dramatic expansion of interconnectedness of technologies with the minimal level of security. At the same time, the value of care does not have a homogeneous single meaning around the world, since technological risks and benefits are distributed unequally and require adjusting types of practices of care to the level of technological advancement in a particular context. This point shall be considered seriously when the legal guidelines provided are conceptually opaque. More discussion on how to read care in the rule 114 are introduced in the next section.

3.2 Technomoral Care - What does it mean?

For the purpose of this thesis, I found it relevant to gain some empirical insights about the connotations of the value of care and the practices informed by them in the “selfless service” provided by military, governmental authorities, and third parties to their citizens and users. The section 3.2 is aimed at introducing the notion of ‘technomoral care’ in the intersection of the theories within the domain of philosophy of technology: technomoral virtue ethics and the theory of techno-moral change.

Apart from the appearance of conceptual opaqueness around care in TM 2.0, I justify the focus on the value and virtue of care by relying on insights from empirical data. Although there were other emerging themes such as justice, honesty, civility and other values, and virtues informed by them, in both interviews R3 and R4, and during the interviews with experts R1 and R2, care had the biggest emphasis in current data sets. Since the emerging theme 'care' had the biggest pattern in the interviews with R4 and R3 and was explicitly mentioned a couple of times during the interviews with R1 and R2, I considered it to be the core issue for this thesis.

Aretaic ethics, which is the central theory on which I base my contribution, appears to be counterintuitive for many readers when compared with care ethics. Yet, I have a particular reason to use it instead. Caring relationships are the cornerstone of care ethics, where there is an emphasis on the dependency where one who is less vulnerable is required to care for one who is more vulnerable (Held, 2006). At the same time, care ethics is a feminist approach deeply rooted into historical contingencies surrounding the role of unpaid work by women in the past and present (Held, 2006). Care ethics is thus holding a very important political message about economic arrangements in favor of more just social conditions, which value we should not undermine.

Care ethics and aretaic ethics share many similarities, but as a prominent scholar in care ethics Virginia Held (2006) claims: "in its focus on relationships rather than on the dispositions of individuals, the ethics of care is, I argue, distinct" (p. 4). Aretaic ethics, in contrast, focuses on the dispositions of individuals, which makes it sensitive to the traditions and diverse components like trust, honesty, civility, courage, justice and others virtues that are necessary for the moral habituation. Most importantly, these two traditions differ on the question of "whether human flourishing or the caring relationships is the ultimate goal" (Vallor, 2016, p. 224). And this, in turn, contradicts the objective of this thesis in seeking a paradigm shift in legal studies towards a framework which ultimate goal would be human flourishing.

Therefore, the overarching scope of this thesis does not intend it to be a standing alone work on care in cyberwarfare, where care ethics could achieve the freedom of exploring the extensions of care in the relationality between physical and cyber, military and civilian domains. Rather, this thesis offers a holistic and comprehensive perspective on the relevant habituation that are corresponding with the requirements of technologically induced moral change. In this regard, quite a number of relevant correlations between virtues, vices, and values can be made based on TM 2.0.

Therefore, this thesis should be considered as research in philosophy of technology of one of the correlations that is often rooted in traditions that are in place in different communities. I consider this particular aspect to be relevant when suggesting *eudaimonia* to be an ultimate goal that the further revisions of TM 2.0 should aim to address. In this regard, I consider aretaic ethics as providing more flexibility and accentuation to continue to focus on the different technosocial traditions and support with these observations the constructive comments about the future directions for TM 3.0 based on the idea of human flourishing.

Before explaining the theoretical underpinning and providing some examples, let me clarity first why I use both connotations of care: *care as value and care as virtue*. I consider the correlation between values and virtues to be quite straightforward: *values evolve throughout history as something in principle, whereas virtues are conformity with these principles in actual daily experiences. In other words, values are ideals and goals that tend to be more aspirational, whereas virtues, on the very basic level, are values in action, or lived values.* Let me put it differently. We, people, tend to value care - we tend to consider the value of care as an important one. However, it does not necessarily follow that we always act in conformity with the value of care. The value of care is often something theoretical or even ideal, whereas the virtue of care is practical and requires training and improvements so that an excellence of the character trait can be achieved.

Indeed, in this thesis I define care following the philosopher of technology Shannon Vallor (2016) as: “a skillful, attentive, responsible, and emotionally responsive disposition to personally meet the needs of those with whom we share our technosocial environment” (p. 138). While Vallor (2015, 2016), as it was shown in the previous section, claims that technologies contribute to the habituation of virtues and vices, other philosophers of technology link the idea of moral habituation with the dynamic nature of values over the time (Swierstra, 2016, Boenink et al., 2009). Techno-moral change is the notion that holds the idea that, throughout a historical epoch, the values that are heterogeneously shared among people had changed according to different stages of technological development (Swierstra, 2016, Boenink et al., 2009). As Swierstra notices in the “Introduction to the Ethics of New and Emerging Science and Technology” (2016):

> We date back the origins of the human species as far back as – well – the earliest artifacts we manage to find. However, technological artifacts have become much more ubiquitous since the so-called Scientific Revolution of the seventeenth and eighteenth centuries and more

> particularly since the Industrial Revolution of the nineteenth and twentieth centuries. How we relate to the natural world around us, to our fellow human beings, and even to ourselves, all these relations are co-shaped[20] by the dynamics of a rapidly evolving technology. (Swierstra, 2016)

This quote provides a significant insight that tracks the acceleration of technological developments and the techno-moral change brought by them. In this view, I situate the value of care in the context of techno-moral change as a value that evolves over the time in order to highlight the dynamic nature of values and make explicit the correlation between the value of care and virtue of care.

However, one more point shall be made about intersection of technomoral virtue ethics and the theory of techno-moral change. Importantly, even though this concept relies on a temporary dimension, it does not imply cultural relativism (Swierstra, 2016; Boenink et al., 2009). I suggest then that technomoral care, that evolves as a value and as a practice due to the rapid technological expansion, should not only hold a temporal dimension, but also a spatial one. To a certain degree, the idea of cultural relativism existing in the aretaic traditions enables the theory of technomoral virtues to reveal what are the different roles that technologies play in different cultural contexts.

Nevertheless, a small remark should be made about technomoral virtue ethics and cultural relativism. On one hand, it's true that technomoral virtue ethics appeal to internal standards of moral excellence in the agent's community as a practical guide. So it stands to reason that these standards will vary somewhat in expression from culture to culture. But one well-known form of 'cultural relativism' is incompatible with virtue ethics, since it makes ethics depend merely on the assent/agreement of the relevant community. On the other hand, for Aristotle, Aquinas, and many other virtue ethicists, there are certain objective conditions of human flourishing that must be met by any virtuous society, so there could be cultures whose notions of virtue are shared but in fact are false or corrupted because they objectively fail to enable and foster human moral, political, intellectual, and/or spiritual cultivation (Chappell, 2013).

Let me now turn to the case study and elaborate on the relevance of care based on the suggested theoretical underpinning. The central issue in the narrative about the events in Estonia in 2007 is the conflict of values. It is particularly relevant because, as the reader could notice in the description of this case, the monument that the Estonian government wanted to relocate had an

[20] Swierstra does not provide a clear definition of co-shape. But in this thesis, I define it as a transformative processes resulted from interrelatedness of human, society and technology.

important meaning to those whose ancestors were taken as victims of war. Intrinsic values like care, dignity, and justice are some of the relevant values which patterns emerged in the interview with R3 and R4. Furthermore, it is important to notice that these values might have a slightly different meaning for the native Estonian population than for those with Russian ancestry.

For instance, as R4 pointed out, the intrinsic value of dignity was less likely to be shared by the native Estonian population since the WWII for them was not about victory, but rather, the occupation of Estonia by the Soviet Union, which came with the discrimination of the native culture and language. So, the monument, for the majority of the native Estonian population, is not about dignity but about discrimination of cultural diversity and years of occupation. The cyber attacks, within this complex cultural and historical context, could reveal to the citizens' the incapability of the governmental authorities to deal with challenges. The danger was that Estonian government would express their inability to fulfill the promises of care by technological means in form of provision and maintenance of resilient system while encountering the needs of both ethnic groups. As R2 commented:

> These attacks produced some inconveniences in some structures to destabilize the situation. But the civilian population, in fact, did not suffer directly from them. The conflict was already existing in Estonia since the independence war and some problems between nations were used to promote gradually, step by step, the escalation of this conflict. And this is one of the success stories. Of course, the situation was already there, different communities watched different TV programs, used different languages, shared different cultural and historical values. And indeed, Russia did not succeed back in 2007 by innovatively using the combination of actions: this monument, disinformation, and cyber attacks, even though they had used a weapon that no one expected or was prepared for. [...] Thus, the cyber attack was the supportive strategical part that aimed to destabilize the positions of the Estonian government in the already complicated situation.

Hence, as it is pointed out in the quote, cyber attack is always entering a context that is ultimately never empty. In these contexts, *a specific technological and social order with its set of standards and norms require significant adjustments in what care as practice would imply to correspond to diverse meanings of the value of care.* As it was revealed in three of the interviews (R4, R3, R2), the Estonian government managed not only to fix the problems in their digital infrastructure fast enough so that their citizens did not feel the threat as something they need to worry about, but also care about their citizens by educating them, empowering them, and building capacities to mitigate cyber threats. Following R3 statement:

> Among all these events of the "Bronze night," as citizens, we did not really feel these cyber attacks. And just after a couple of days, everything we needed to use (banks, administrative websites) were functioning as if nothing had happened. Notably, after these events almost every year we learn how to use a new and up-to-date security practice for our digital daily needs.

These forms of expression of care are significantly different from how care could be expressed just fifty years ago. What is essential, is that *the Estonian government not only managed to fix the problems technically, but it was able to create a common ground to bring together and promote a dialogue between the indigenous Estonian population and the population with Russian ancestry who happened to stay in Estonia after the USSR collapsed.* This achievement is notable due to the fact that the meaning of the intrinsic values of care may vary or even be opposite among the two groups. For instance, for the Russian population, potentially, these cyber attacks (at least in the way activists justify them) were an attempt to respond to the violation of an intrinsic value like care (also justice and dignity). While for the native Estonian population, up to a certain degree, these attacks were rather about the violation of different values. Nevertheless, technologies partly had their role in this conflict resolution.

During the interview, R4 mentioned that the level of prosperity and welfare in Estonia increased with all these advancements that the government implemented and distributed equally. R4 also mentioned that both "*Estonian and Russian population could have equal access to the educational procedures.*" Thus, the conflict in 2007 was resolved in an exemplary way since *the government could express technomoral care (skillful and empathetic) and, by doing so, they promoted trustworthy relationships with both social groups and equipped them equally with similar attention and care.* As it was pointed out in the interview with R4 and R3, the care about the inclusiveness (and dialog) of different social groups is somewhat a traditional value in Estonia. Not surprisingly, *the Estonian e-governance platform is operating in three languages: Estonian, Russian, and English.* This can be considered as an evolution of the traditional values in Estonia resulting in new forms of expression of care. Quoting upon this point from the interview with R3:

> From the very beginning, when Estonia had separated from the USSR and the question of future trajectory was sharp because the country is small and seriously lacking natural and financial resources, there was a tied connection with the EU and the support they could offer. Thus, the Estonian government decided to develop information technology and tourism as two main directions. Later, they prepared everything and connected the whole country to the internet. They have a hamlet system of living and they are proud to have such a system. The Estonian government cares about us. And thus, the internet connection was provided all over the country to every house. The government put effort into teaching citizens, and people themselves were very open to learning because everything was digitized and this appeared to be very convenient.

As it was emphasized in this quote, *the Estonian government gradually advanced the forms of expressions of care with technological means and security measures suited for each individual's needs.* Apart from the idea that technologies afforded new ways of practicing care, if a balanced dynamic in technomoral virtue care is obtained, then the virtue of other democratic values (dignity, liberty, equality and justice) and the virtues informed by them can be successfully promoted along. As R4 pointed out in the interview:

> I think, maybe, of course, this is the only delusion, that the security in Estonia is so strong that we, as citizens, do not feel any danger, but rather opposite, we feel more protected, more secured, with all these technologies implemented by the government; or at least, on the similar level as other European countries. *The fact that Estonia is the leading country in digitalization makes us proud!* Such a small country, but nevertheless, it has intelligent people who make such a meaningful contribution to the whole world.

In this quote, it is apparent that the practice of technomoral care provided by the government to their citizens in digitalization and the human-centered values pinpoint towards other values and pride of belonging to the Western liberal democracies.

However, the other extreme of that would be vice - the excess of care that results in disinterest, or in oppression and control. This can be apparent because the value of care is not that straightforward in different techno-social environments. Thus, the virtue of care that is informed by the shared value of care in a particular context (including technological) can indicate how the virtue of care is promoted and expressed within a virtue-vice continuity (Aristotle, 1934/1996). Take, for instance, *North Korea where* *technomoral care is expressed in a vicious way as totalitarian control, surveillance, and violation of basic human rights; where interests of the state is all that matters and needs of the people are disregarded as minor.*

Considering all the above mentioned analysis, following the aim to depict and unpack missing values in the debate around cyberwarfare*, I claim that a combination of the frameworks within the domain of philosophy of technology: technomoral virtue ethics (culturally sensitive) and techno-moral change (historical) is a comprehensive integrative approach that encompasses the spatial and temporal dynamic in values and, practices informed by them resulting in a habituation of the virtues.* In other words, virtues that are informed by values, which are constantly affected by techno-moral change, should also be a matter of deliberate development in the light of new types of affordances provided by technologies.

By bringing the Estonian case study, it was demonstrated why *the value of care should be understood through the prism of techno-moral change and how this relates to the development of corresponding virtues and vices.* Oppositely to perceiving care as a mere value, *care as a virtue would require some considerations on how to distinguish responsible care from caring too little or too much.* Particularly, because the value that informs care as a practice has significantly changed its connotations through time, at the same time, despite being universally relevant, *the practices informed by the value of care have heterogeneous forms of expressions in the diverse cultural contexts according to its level of technological advancement.* By unpacking the value of care through the caring practices and forms of expression of care, such analysis becomes a good indicator of how the value of care is shared, practiced, and interpreted in each particular cultural and technological context. So, in order to depict how technomoral practice would look like if it was informed by evolving values, it is necessary to gain empirical insights about how these practices are manifested in the contemporary world.

Furthermore, following the discussion in the sections 3.1 and 3.2, *I consider care to be a particularly relevant practice that needs to be cultivated for military services and linked with the governmental decision-making apparatus about cybersecurity services.* Because in the cyber realm, due to the specifics of technologies and systems, cyber warriors cannot sufficiently care within a single act, and express care similarly in the different techno-social contexts. The abilities to resolve particular situations are limited when the security of the system was not granted in advance by providers, governmental authorities, and cybersecurity agencies.

That is the main difference with regular warfare, where a soldier could take care of the civilians, and, for instance, move them to a different location. *In cyberwarfare, the expression of care, without provided security service, would be problematic.* At the same time, the very nature of cyberwarfare allows different connotations of care, as the distinction between military and civil context, that is crucial for Aristotelian virtue ethics, is no longer an obstacle. Indeed, my take is that, in the international law that regulates cyberspace, reexamination of the limits of care as a virtue, in the light of new and emerging technologies, constitutes an essential step for the promotion of the individual and communal flourishing in that context.

Let me elaborate on the latter point. In comparison with countries that are not digitized and cyber attacks do not necessarily threaten their government, as Haataja (2017) notices, Estonia aback in 2007 had already integrated internet-based solutions for the vast majority of aspects of public life, and thus was particularly vulnerable to the threats of cyber attacks. During these attacks, Estonian citizens (and residents) could not accomplish their basic and administrative needs online, for instance, getting a prescription from a doctor, getting access to the school cabinet and filling in social security and tax returns. Instead of merely threatening the state, or affecting the information system, these attacks brought disruption to the social order, since *each citizen could feel the effect on their daily banking or other personal and individual practices and tanks*.

From this view, a merely state-centered perspective on this case, as well as claims in favor of information integrity and sovereignty as the values to be protected, are definitely not coherent with the full picture. The attack, within its complex cultural and historical context, could reveal to the citizens the incapability of the governmental authorities to deal with challenges and thus, express their inability to fulfill the promises of care by technological means in form of provision and maintenance of resilient system of innovation that the government promotes.

Reflecting upon the “information integrity,” the relevant values instead (e.g. justice, trust, care, and informed by them practices informed by them) are rather traditional and familiar to everyone (although interpreted differently). Were these attacks aimed to threaten an “information integrity,” or were they rather aimed to target the microclimate among diverse national groups who live in the same country? Or, was it perhaps aimed to show that the digitalization agenda in Estonia was particularly weak, since it could not protect its own citizens from such security risks? These *destabilizing attacks could prove the Estonian government technologically helpless, careless to their citizens of both cultural inherencies, and* *thus, incapable to provide maintenance and resilience of the security system.* This point is especially interesting within the scope of the rule 114 (point 2) in TM 2.0.

Yet, *in the cyber domain, each citizen should be secured enough from the cyber breaches beforehand, and not afterwards.* Thus, when one reads - awareness is needed not just during preparatory stage, but at all times - this pushes the boundary further and signifies care as one of the practices that need to be deliberately unpacked. Take, for instance, the use of blockchain technology to securely store data, different security measures for access to citizens' data ranging from regular passwords to USB reader devices (like in Estonia), and many other examples of how care in the cyber realm fractured and new type of practices of care, unfamiliar before, were afforded. At the same time, in the rule 114, even more than in the other two rules, it is unclear who is obliged to provide constant care: how? when? and from which perspective? This and other points will be further elaborated in the next section.

3.3 Discussions and Objections: Moving Cyberwarfare Beyond State Level

While being inspired by the bottom-up insights about care in the Estonian case, I made an attempt to contribute to the reading of TM 2.0 with these insights to fill in the value gap with practical observations. In this discussion section, I aim to reflect upon the dynamic in virtues and vices that correspond to the specific value of care, which should be encompassed when aiming on the development of regulatory framework. Since it is assumed that virtues and vices are interconnected and interdependent in the context of achieving *eudaimonia*, or utterly human flourishing, care is the ultimate value in "selfless service." As it was suggested in the previous section, the value of care should not be perceived in a reductionist manner, as it appears now in TM 2.0 as something opaque, vague, and abstract. Thus, to narrow the focus in this thesis, I suggest to localize the discussion just around the rule 114 point 2, as it seems to create more contradictions and raise more questions than the other two rules.

A remark about the limitation of this research is important here. Even though I indicate care as an intrinsic value that is crucial for *eudaimonia*, each component that constitutes *eudaimonia,* like justice, honesty, trust, and others, deserve to be explored separately. To provide a full account of *eudaimonia* in this sense, I do ultimately claim that substantial research shall be done with regard to each component of *eudaimonia*, highly relevant issues that provide transparent security. These are virtues that correspond with the values that have shared meaning in each particular context but are vulnerable in the context of cyber threats.

However, detailed elaboration on each of these values and virtues goes beyond the scope of this master thesis. Due to that, I excluded potential investigations on other values and virtues like justice, trust, and honesty, while being aware that the empirical data contains numerous insights on their matter, as well as on the matter of concern of this thesis. Thus, by endowing the lacking explanations about care in TM 2.0, I make a preliminary step to motivate further research where other components of *eudaimonia* which are necessary to achieve it will be extracted.

As it was suggested in the section 3.1.1, following the Golden Mean rule, care can be perceived as a mean in continuity between excess and deficiency. It has a flexible nature that can be affected by the different circumstances that occur in people's lives to shift it to one edge (constant surveillance) or another (indifference). As the R2 pointed out in the interview: "In a hybrid type of warfare, people [are] used as a means to achieve certain goals. We could see this happening recently during the US 2016 Presidential elections, and [something] similar was happening in Estonia back in 2007."

Following the notes that were made during the interviews with R3 and R4, in the case of Estonia, the citizens trust in the government's ability to skillfully take care in the context of innovation that they deploy was targeted. From this view, the discussion should start with an underlying assumption: to ethically tackle cyberwarfare one needs to reveal a broader contextual picture, rather than merely focusing on the cyber attack itself and its direct consequences.

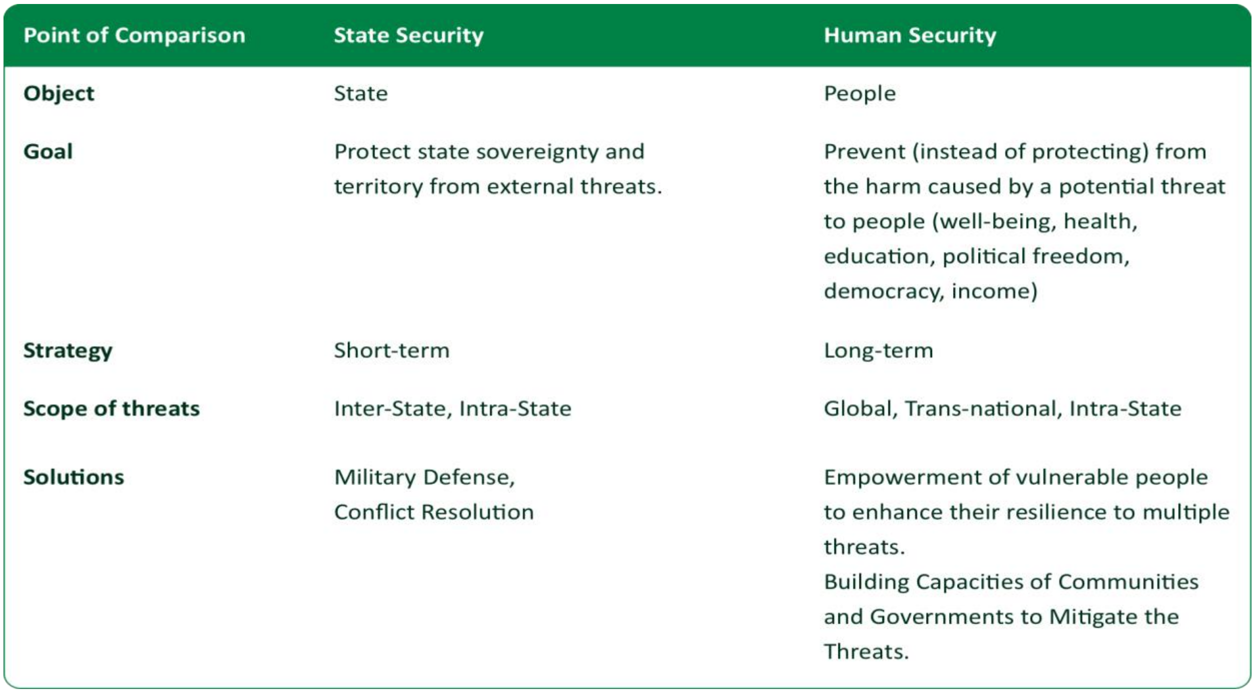
The latter idea is not sufficiently covered in the ethics of cyberwarfare, but is very often articulated in International Relations and the securitization discourse around the strategies of hybrid wars to pick an environment where value tension appears to be dynamic. In other words, cyber attacks in Estonia in 2007 were mainly aimed to sabotage and challenge the Estonian government on their tendency towards the intensive digitalization of the country (Haataja; 2017), by targeting the weak areas - the existing value conflicts. In this strategy, every failure of the government to protect their citizens would result in broken promises, unjustified expectations, and the discrimination of one ethnic group or the other. Gradually increasing the conflict between Estonian and Russian inhabitants could even potentially turn into real civil war.

Every misconduct and careless behavior from the government authorities' side would immediately show their incapability to provide the level of security they promised, the level of care they pledged for the protection of the national heterogeneity of Estonian society. Thus, the target of this event was apparently not just digital infrastructure, not just the government, not the 'informational integrity,' but the citizen's trust in innovation and the ability of the government to handle the protection, in other words, to take care about their citizens and prevent reoccurrence of cyber attacks in the future.

Importantly, the Estonian case also indicates *how relevant and skillful practices of care have changed over time.* In the past, the Estonian government would not need to advance security measures circularly: for each individual citizen and for the whole country network in general. However, in the contemporary world, the empowerment of vulnerable people to enhance their resistance to multiple cyber threats and the development of the capacity for communities and governments to mitigate cyber threats are a key component of the skillful technomoral care about the security of the system. In particular, what this represents, as it was also illustrated in the case of Estonia, is the human-centered values of human security.

Let me elaborate upon these observations. It is notable that human-centered security concerns are, if not more, then equally relevant as state-security concerns because, as was revealed in the case of Estonia, empowering individual citizens by considering their diverse needs pinpoints towards the human-centered values that should be incorporated. To make this claim more concrete, in the Table 3 I contrasted state-centered security with a human security paradigm.

Table 3. State-centered security versus Human security



Point of Comparison	State Security	Human Security
Object	State	People
Goal	Protect state sovereignty and territory from external threats.	Prevent (instead of protecting) from the harm caused by a potential threat to people (well-being, health, education, political freedom, democracy, income)
Strategy	Short-term	Long-term
Scope of threats	Inter-State, Intra-State	Global, Trans-national, Intra-State
Solutions	Military Defense, Conflict Resolution	Empowerment of vulnerable people to enhance their resilience to multiple threats. Building Capacities of Communities and Governments to Mitigate the Threats.

The debate between human security and national security reveals that, when it comes to human intervention and others types of potential threats, the question that is relevant here is: why is it that we tend to focus on state stability, and not on more concrete elements such as access to food, water, jobs, education, health, and others? The Estonian case is a good example where some of these concerns are applicable. *Due to the fact that cyberwarfare is not just about cyber attack, but rather about cyber peace as a mean to achieve certain goals, human-centered long-term security concerns are highly relevant. Yet such skillful practices of technomoral care as empowerment of vulnerable people to resist cyber threats and the development of capacities to mitigate cyber threats should be promoted not just on the local level, but on the level of legal proceeding like TM 2.0.*

Going back to the case study, Estonian governmental authorities and not cyber warriors or soldiers fulfilled the requirement of the rule 114 and employed the value of care by equally distributing information security in the country in inclusive, engaging, and educating ways that encountered the needs of different social groups. So, referring to the Rule 114, in this case, it is clear that the Estonian government took a *long-term solution by empowering their citizens' security (bottom-up micro level of individuals), not merely securing the country (top-down macro level of state)*. By doing so, they *straightened the trustworthy relationship with their citizens and expressed what can be called technomoral care responsibly, skillfully and empathetically by technological advancement, that goes hand-in-hand with individual security advancement.* The Estonian government largely invested *not just on the digitalization of the state, but also on the digital education of every citizen.* This type of response indicates two important points:

- The intrinsic value of care was an overarching 'pillow' that was expressed and promoted in all those actions that were taken by the Estonian government since the cyber attack. It is important in itself, but also to grant trustworthy relationships, to express respect to human dignity and to resolve the situation in a just and inclusive way. *By expressing care in a virtuous way, without discrimination of the Russian population, or elderly people, the Estonian government created educational programs in order to include different social groups equally (Russian population, Estonia population, elderly, kids).* Care about all social groups in Estonia was promoted by technological means. Examples of that would be the accessibility to digital platforms in both the Estonian and Russian languages, promotion of instructors, and other support services on both languages.

- Constant upgrades of security measures for every user. In particular, it is evident that the Estonian government took a long-term path in preventing cyberwarfare, instead of merely defending from it. Thus, their agenda in response to the events in 2007 aimed to enhance their resilience to multiple threats and building capacities of the communities and governments to mitigate the threats, measures all aligned with a human security perspective. In other words, *the empowerment of vulnerable people to enhance their resistance to multiple cyber threats together with the development of the capacities to mitigate those threats illustrates circularity between human-centered and state-centered security, where the second can be better granted by the enhancement of the first.*

Indeed, building upon the observation that citizens do not necessarily depict cyber attacks from the context in which they occur, meaningful insights could be extracted from the citizen's experience of the context where this attack manifests. In the case of cyber attacks on Estonia in 2007, it was demonstrated that the government took a proactive role in empowering their citizens and building capacities to mitigate different cyber threats that may occur in the context of cyberwarfare. *Such 'prevention' policy is the most effective way in which care towards civilians and their objects can be skillfully expressed.* This sense of "care" seems to be very different from the sense meant in TM2.0 and in JCW. Hence, by recalling all the aspects of the theoretical and empirical analysis, I argue that while TM2.0 and JCW leave out the significance of care as virtue by design, such sense of care is insufficient for the proper development of regulations for cyberwarfare in a sense broader than the act of war that aligns with the principle of JWT.

However, I want to emphasize that since this thesis is mainly dealing with an ethical question, instead of making a strong claim in favor of human security, I propose to highlight the relevance of human-centered values as an important preliminary step. In particular, because there are two objections posed to the idea of shifting towards a human security paradigm exploration which goes beyond the objective of this thesis: "on the one hand, the claim that less inter-states warned more casualties in intra-state does not necessarily prove that national security has become useless; on the other hand, public policy requires prioritizing certain aspects over others, it cannot just give same attention to everything concomitantly" (Mahmud et al., 2008, p. 69). Nevertheless, this analysis attempted to indicate the advantages of the shift towards a human-centered perspective on security that will potentially move the understanding of the regulations of cyberwarfare beyond a state level.

3.4 Objections, Limitations and Future Trajectories

I want to conclude by recalling some objections. The first would be: up to what extent the critique of state-centrism suggested by this thesis is different from the critiques to traditional warfare? This critique of state-centrism is unique for cyberwarfare because it would be problematic to consider the possibility to sufficiently care in the case of a cyber attack, due to a significant dependence from the existing level of security advancement, technological development, and digitalization (DiMaggio & Hargittai, 2001). In other words, it would be difficult for a cyber warrior to take care during operation when basic digital security means are not met (consider digital divide issues). The one can never protect enough, the one always needs to maintain, or, alternatively create a resilient system. Hence, an aretaic reading of care, in this case, would pinpoint towards the need to perceive care as the maintenance of a resilient digital system, rather than an act in a particular situation according to the principle of discrimination, as if it would be in traditional warfare.

Then, one might object: how would the findings of the current analysis be applicable to the context of larger countries than Estonia? Considering that people can rarely care sufficiently about their own digital security themselves to the extent of cyber threats that exist nowadays, the party which is responsible for security (government, military, cyber security agencies, private companies) is obliged to provide "constant care" on a daily basis. However, it is unclear who should care when we talk about larger countries, for instance, the U.S. Because the threats are mostly implied in (almost all) technologies and third parties provide security by default, such that no user can sufficiently affect such security on her own, certain requirements should motivate these providers of security to somehow consider more seriously the idea of a human-centered focus. That could be considered as a minimal first step to promote human-centered values in cyberspace in countries with diverse technological, social, and political landscapes.

Another objection would be that care ethics can similarly support the core idea of this thesis: that the care needs to be reconsidered within the scope of cyber justice and digital divide (DiMaggio & Hargittai, 2001). Since some countries are more digitally advanced than others, risk and benefits are distributed in an unequal and unjust way (DiMaggio & Hargittai, 2001), empowerment of the vulnerable seem to fit the objective of care ethics. Moreover, cyber operations raises multiple challenges regarding the digital divide and the currently existing asymmetry in cyber justice. Technomoral care, similarly to the care ethics, would integrate these justice concerns in order to inform a long-term solution for building capacities to mitigate cyber threats and empower vulnerable people to resist those threats.

However, I consider the idea of heterogeneous techno-social traditions to be important. As it was pointed out in the case of Estonia, care is a traditional value that evolved together with the digitalization of the country. Nevertheless, future research is needed to adjust the analysis to the heterogeneous techno-social context of different countries and not just Estonia. In further research, the aspects to be considered are: i) different capacities to mitigate cyber threats should be build according to particular needs of each technological and cultural context, ii) these needs should be acknowledged, signified, and encompassed, which should result ideally in iii) the empowerment of vulnerable people that enhances their ability to resist cyber threats by providing advanced security measures based on their current situation.

Significantly, one might object aretaic ethics in the first place, as a theory that, similarly to deontological theories, still aims on universal maxims about what virtues should be considered as relevant. However, this would be a very limited representation of aretaic ethics and its capacities. Instead, aretaic ethics, in particular thechnomoral virtue ethics, can potentially serve as a fruitful foundation for legal theories and enable them to tackle the issues raised by diverse cyber threats based on the promotion of human-centered values and human flourishing. Currently, civilian population requires empowerment since, as statistics shows (Figure 4, *see in the Appendix D*), literacy about cyber threats and self-awareness of preventive measures is dramatically low (even though this data is limited to the U.S., it is still a strong point to consider when analyzing the situation in Europe).

Hence, this objection is rather weak because deontological theories in comparison to aretaic ones are voiceless; whereas the voice of aretaic theory speaks about the rules in TM 2.0 as lacking an emphasis on a skillful and empathetic expression of care by technological advancements together with investments in the promotion of self-awareness of up-to-date security measures.

Last, but not least, since this thesis is a preliminary step for a serious consideration about the shift from deontic legal theory to aretaic legal theory, care was distinguished as one of the aspects that is vividly relevant for *eudaimonia*, the ultimate goal of aretiac legal theories. However, there are also other virtues like trust, justice, civility, and many more, that deserve further exploration under the overarching idea of the “legal turn” in the context of cyberwarfare and the regulations for cyber operations at large.

Moreover, more data is required to make data samples more diverse and encompassing, and not bounded to a single country, as this current research is. Hence, this can be a future direction for a doctoral research aimed at the highly impactful outcome for the benefit of digital and non-digital societies. Even though this point is only a preliminary step towards such direction, it is an important first step that integrates philosophy of technology and that takes cyberwarfare beyond the state level in accordance with existing legal challenges and considerations about actual cases.

3.5 Conclusion

In the Chapter 3, a principally different reading of TM 2.0 was proposed. It was argued that a broader perspective on the values, more specifically, human-centered values that are at stake in cyberwarfare is needed for the development of a solid legal framework. By going beyond the state level and focusing on the value of care as referred in TM 2.0 in the Rules 59, 92, and 114, two different ways of reading the manual were suggested based on the ethical theory of technomoral virtues (Vallor, 2016). On one hand (Rule 59, 92), care is understood in the sense of the principle of discrimination, and on the other hand (Rule 114), care is understood in the sense of technomoral virtue. *It was concluded that the value of care does not reveal in what way care as a trait and as a practice is developing to grant civil rights to the societies who face dramatic expansion of interconnectedness of technologies with a minimal level of security. At the same time, the value of care does not have a homogeneous meaning around the world, since technological risks and benefits are distributed unequally. This requires an adjustment to the types of practices of care at the level of technological advancement of the particular context.*

This analysis was supported by the empirical insights from several interviews and a case study of the cyber attacks on Estonia in 2007. As a result, *the combination of the theoretical and empirical analysis proposed in this chapter resulted in a preliminary step in favor of the shift towards a human-centered perspective on security.* Subsequently, it was pointed out that the understanding of the regulations of cyberwarfare should move beyond the state-centered security level. *It was concluded that the shift towards human security might be a fruitful option for further development of the regulations for cyberwarfare, not solely for the cases that are classified as an act of war, but also for those that fall beneath the threshold.*

General Concluding Remarks

A year after the Tallinn Manual 2.0, a comprehensive academic study on how International Law applies to cyber operations, was published, many tensions and urgent challenges are waiting to be explored. In particular, these issues are urgent considering the rapid technological development in the Internet of Things, Robotics, Biomedical technologies and Geoengineering, and the relatively slow advancement of the legal frameworks. Although TM 2.0 is a good starting point, it appear to be tied with a state-centered security paradigm and bounded by IL and JWT. This thesis offered a critical appraisal of TM 2.0 and more general ethical debate around the regulations for cyberwarfare.

Due to the international nature of the manual, it does not represent the point of view of a particular state. As the Dutch Minister of Foreign Affairs Stef Blok highlights: “the Netherlands doesn't necessarily agree with everything in it [TM 2.0],” he continues, “but this is the only way we can impose a price on bad behavior.”[21] However, considering how widely cyber threats are spread around the world and how vulnerable new and emerging technologies that enter every level of individual, social, economic, and political being are, it remains arguable if the imposition of the price on behavior is the only objective to address. This thesis, with the help of philosophy of technology, made an attempt to show that instead of merely imposing the price on certain behaviors, it is particularly important to shift the focus and promote the right type of behaviors.

In order to proceed to my argument, I started the first chapter with an introduction and the methodological scope. Then, in Chapter 2, a critical examination of the conceptual and regulatory debate around TM 2.0 and cyberwarfare at large was done. The key issues on conceptual and regulatory gaps were indicated. Since the TM 2.0 is a set of guidelines on how IL applies to cyber conflict that predominantly relies on JWT, there are certain limitations in its applicability which were framed as the regulatory gap. These challenges were introduced within the scope of the tension that appears between JWT, TM 2.0, and actual case studies. It was concluded that, apart from the conceptual and regulatory gaps, the value gap is limiting the analysis when cyberwarfare is at stake. The value gap that is caused by a state-centered security paradigm, which underlies IL and constitutes the basis of TM 2.0, was made apparent here.

Furthermore, in order to reveal the relevance of such observation, I contrasted the state-centered security with the human security paradigm and pinpointed the human-centered values that

[21] *See report (p. 9)*: http://puc.overheid.nl/doc/PUC_248137_11

are somehow dismissed. In this respect, in Chapter 2, three main positions in the ethics of cyberwarfare were further examined: 1) the 'moderates,' who are the proponents of JWT and whose foundational account is hold in TM 2.0; 2) the 'skeptics,' who are largely critical about the necessity of the military context and JWT and implicitly object to the state-centered values brought by this context; and 3) the 'radicals,' who proposed the discontinuity approach and the JCW framework based on the onto-centric genuine value of 'integrity of information system.'

At the same time, only the 'radicals' did suggest a constructive solution to re-evaluate the basics of IL - JWT in the light of new and emerging technologies. Yet 'information integrity,' the genuine value suggested by them, appears to be an instrumental value and, thus, is misleading because of its insensitivity to the different cultural and technological contexts in regards to the distribution of risks and benefits of technologies. It was shown that due to the relevance of the risks and benefits distribution of new and emerging technologies and the correspondence with cultural heterogeneity in the regulation of cyberwarfare, JCW appears to be a context-insensitive framework that suggest instrumental values which in the macro level rests on the state-security paradigm and, ultimately, does not fill in the existing value gap. Thus, an alternative view on the relevant values was required in order to fill in the value gap currently existing in the debate around TM 2.0 and the regulation of cyberwarfare at large.

In Chapter 3, the principally different reading of TM 2.0 and respectively, JWT, was demonstrated. This was tailored with the explanations of technomoral virtue ethics (Vallor, 2016). Under this light, the technomoral care was provided with temporal and spatial dimensions of concern and was subsequently proposed as an encompassing practice of skillful and empathetic nature, which is based on the intrinsic value of care that is commonly recognized as important. By taking cyberwarfare beyond the state-centered security level and focusing on the value of care as referred in TM 2.0 in the Rules 59, 92, 114, two different ways of reading it were suggested. It was indicated that TM 2.0 implies two different connotations of care.

In the Rule 59, the notion of care was understood in the sense of the principle of discrimination, and thus was limited to be a requirement for the military personnel. However, in the cyber domain it is not an entirely right statement that military personnel grants cyber security. Often, those are third parties, private companies, governmental securitization programs, or cyber security agencies. Thus, certain virtues that were overlooked in JWT and TM 2.0 but however appear to be relevant, in particular, care, were revealed and critically re-examined in the light of new and emerging technologies.

This became a link to the second way of reading the notion of care in TM 2.0 which appears to be significantly relevant, but nevertheless, dismissed in the debate. In the Rule 114, the notion of care was understood in the sense of technomoral care, as a character trait, as a practice that should grant a sufficient level of cybersecurity to civil societies who face dramatic expansion of interconnectedness of technologies. Indeed, since the value of care does not entail a homogeneous meaning around the world, the technological risks and benefits are distributed unequally and that is why the practices of care should be adjusted to the level of technological advancement of the particular context in order to promote basic literacy in cybersecurity.

Such conclusion was based not merely on the theoretical analysis, but also supported by the empirical insights from the interviews and elaborations upon the case study of cyber attacks on Estonia in 2007. Thus, by shifting from the prioritization of the values of state-centered security to the human-centered values of human security, this challenge can be addressed by a regulatory framework that will address 'grey zones' of IL by foundationally relying on the human-centered values of human security.

The combination of all these components of the analysis led to the answer on the main research question: *In what way could an ethical theory of technomoral virtues fill in the value gap that currently exists in the Tallinn Manual 2.0 and, more broadly, in the debate surrounding the development of regulations for cyberwarfare?*

As a response to this question, it was suggested that the combination of the theoretical and empirical analyses based on the ethical theory of technomoral virtues had served as a preliminary step of the shift towards a human-centered perspective on the security and aretaic legal theory. Based on this theoretical underpinning, *it was argued that a broader perspective on values, more specifically, human-centered values that are at stake in cyberwarfare, is needed for the development of a solid legal framework that will regulate cyberwarfare.* From this view, *the ethical theory of technomoral virtues is able to fill in the value gap, by taking the regulations of cyberwarfare beyond state level. This way, the focus on a) prevention instead of merely protection, and b) promotion, instead of prescriptive regulation may lead to empowerment of vulnerable people to enhance their ability to resist cyber threats and the deliberate building of capacities to mitigate cyber threats on the international level.*

While this idea requires further investigation and research within the domain of human security and ethical foundations of legal theory, this master thesis made a strong preliminary claim towards this direction. By suggesting this critical interdisciplinary analysis, I aimed to point towards the 'legal turn' to aretaic legal theories and the significance of virtues that are practiced through technological means. The shift from deontological ethics to an aretaic ethical theory is particularly fruitful for the development of a legal framework because it provides more possibilities to capture different culturally-technological contexts with respect to the dynamic nature of human-centered values informed by these diverse technological practices.

Bibliography

Alford, L. D. (2000). *Cyber Warfare: Protecting Military Systems*. United States: Air force materiel command wright-patterson AFB OH.

Anscombe, G. E. M. (1958). *Mr. Truman's Degree*. Pamphlet published by author, Oxford.

Aristotle. (1934/1996). *The Nicomachean ethics* (H. Rackham, Trans.; S. Watt, Eds.), Herts: Wordsworth Editions.

Arquilla, J., & Ronfeldt, D. (1993). *Cyberwar is Coming!*. Santa Monica, CA: RAND Corporation. Retrieved from*:* https://www.rand.org/pubs/reprints/RP223.html.

Barrett, E. T. (2015). Reliable Old Wineskins: The Applicability of the Just War Tradition to Military Cyber Operations. *Philosophy & Technology*, *28*(3), 387–405. https://doi.org/10.1007/s13347-014-0185-4

Barrett, E. T. (2013). Warfare in a New Domain: The Ethics of Military Cyber-Operations, *Journal of Military Ethics*, 12(1), 4-17. DOI: 10.1080/15027570.2013.782633

Beard, M. T. (2014). *War rights and military virtues: A philosophical re-appraisal of Just War Theory* (Doctor of Philosophy (PhD)). University of Notre Dame Australia. Retrieved on July 3, 2018 from https://researchonline.nd.edu.au/theses/96

Bentham, J. (1890). *Utilitarianism*. London: Progressive Pub.

Boenink, M., Swierstra, T., & Stemerding, D. (2009). Exploring Techno-Moral Change: The Case of the Obesity Pill. *The International Library of Ethics, Law and Technology Evaluating New Technologies,* 119-138. doi:10.1007/978-90-481-2229-5_9

Brekke, T. (2005). The Ethics of War and the Concept of War in India and Europe. *Numen*, 52(1), 59-86.

Brey, P. (2008). Do we have moral duties towards information objects? *Ethics and Information Technology, 10*(2-3), 109-114. doi:10.1007/s10676-008-9170-x

Brey, P. (2007). Is Information Ethics Culture-Relative? *International Journal of Technology and Human Interaction*, *3*(3), 2–24.

Chappell, T. (2012). Aristotle. In T. Angier (Eds.), *Key Ethical Thinkers* (pp. 33-55). London: Bloomsbury (Continuum).

Chappell, T. (2013). Virtue ethics in the twentieth century. In D. Russell (Eds.), *Cambridge Companion to Virtue Ethics* (pp. 149–171). Cambridge Companions to Philosophy. Cambridge: Cambridge University Press.

Charles, A. (2010). Cyber attacks widespread, says report. *The Guardian,* Retrieved on February 22, 2018 from https://www.theguardian.com/technology/2010/jan/28/cyber-attacks-hacking

Chloe, A., (2009). Russian Group's Claims Reopen Debate On Estonian Cyberattacks. *Radio Free Europe / Radio Liberty*, Retrieved June 13, 2018 from http://www.rferl.org/content/Russian_Groups_Claims_Reopen_Debate_On_Estonian_Cyberattacks_/1564694.html

Clarke, R. A., & Knake, R. K. (2011). *Cyber War*. HarperCollins.

Clausewitz, C. V. (1997). *On War*. In J.J. Graham (Eds.), Hertfordshire: Wordsworth Editions Limited.

DiMaggio, P., & Hargittai E. (2001). From the 'Digital Divide' to 'Digital Inequality': Studying Internet Use as Penetration Increases. *Center for Arts and Cultural Policy Studies*, Princeton University.

Eberle, C. J. (2013). Just War and Cyberwar. *Journal of Military Ethics,* 12(1), 54-67. DOI: 10.1080/15027570.2013.782638

Efrony, D., & Shany, Y. (2018). A Rule Book on the Shelf? Tallinn Manual 2.0 on Cyberoperations and Subsequent State Practice. *American Journal of International Law, 112*(4), 583-657. doi:10.1017/ajil.2018.86

Ess, C. (2009). Floridi's Philosophy of Information and Information Ethics: Current Perspectives, Future Directions. *The Information Society, 25*(3), 159-168. doi:10.1080/01972240902848708

Eye from the sky. (n.d.). In Wikipedia. Retrieved on October 2, 2018, from https://en.wikipedia.org/wiki/Eye_in_the_Sky_(2015_film)

Fildes, J. (2010). Stuxnet worm 'targeted high-value Iranian assets'. *BBC NEWS*, Retrieved April 19, 2018, from https://www.bbc.com/news/technology-11388018

Floridi, L. (2008d). Information ethics: A reappraisal. *Ethics and Information Technology*, 10(2), 189–204.

Floridi, L. (1999). Information ethics: On the philosophical foundation of computer ethics. *Ethics and Information Technology,* 1(1), 37–56.

Gertz, N. (2014). *The Philosophy of War and Exile: from the Humanity of War to the Inhumanity of Peace*. In T. Brooks (Eds.), New York: Palgrave Macmillan.

Glorioso, L. (2015). Cyber Conflicts: Addressing the Regulatory Gap. *Philosophy & Technology*, *28*(3), 333–338. https://doi.org/10.1007/s13347-015-0197-8

Haataja, S. (2017). The 2007 cyber attacks against Estonia and international law on the use of force: an informational approach. *Law, Innovation and Technology*, 9(2), 159-189. DOI: 10.1080/17579961.2017.1377914

von Heinegg W.H. (2014). The Tallinn Manual and International Cyber Security Law. In T. Gill, R. Geiß, R. Heinsch, T. McCormack, C. Paulussen, and J. Dorsey (Eds.), *Yearbook of International Humanitarian Law, vol 15.* T.M.C. Asser Press, The Hague

Held, V. (2011). Morality, care, and international law. *Ethics & Global Politics,* 4(3), 173-194. DOI: 10.3402/egp.v4i3.8405

Held, V. (2006). *The ethics of care: Personal, political, and global*. New York: Oxford University Press.

Herzog, S. (2011). Revisiting the Estonian Cyber Attacks: Digital Threats and Multinational Responses. *Journal of Strategic Security,* 4(2), 49-60. DOI: http://dx.doi.org/10.5038/1944-0472.4.2.3

Howard, D. (2014). Virtue in Cyber Conflict. In M. Taddeo & L. Floridi (Eds.), *The Ethics of Information Warfare* (pp. 155–168). Cham: Springer.

Hughes, D., & Colarik, A. (2017). The Hierarchy of Cyber War Definitions. *Intelligence and Security Informatics Lecture Notes in Computer Science,* 15-33. doi:10.1007/978-3-319-57463-9_2

Hutchby, I. (2001). Technologies, Texts and Affordances. *Sociology,* 35(2), 441-456. doi:10.1017/s0038038501000219

Ihde, D. (2010). *Heidegger's Technologies: postphenomenological perspectives*. In J. Caputo (Eds.), New York: Fordham University Press.

Ilves, T. H. (2018). If Liberal Democracies are to Survive in the Digital Era, they Must Create a Defence Organization. Keynote speech on 10th annual cyber security conference CyCon 2018 in Tallinn. Retrieved on June 28, 2018, from https://www.youtube.com/watch?v=j_Fx5TPwQ4E

John, S. (2007). How to take deontological concerns seriously in risk–cost–benefit analysis: a re-interpretation of the precautionary principle. *Journal of Medical Ethics*, *33*(4), 221–224.

Kant, I. (1998). *Groundwork of the metaphysics of morals*. In M.L. Gregor (Trans., & Eds.), Cambridge: Cambridge Univ. Press.

Kisilyov, E. (2017). Personal interview. Retrieved on February 13, 2018, from https://m.youtube.com/watch?v=CqfvEU3O0_c

Kvale, S. (1983). The qualitative research interview: A phenomenological and a hermeneutical mode of understanding. *Journal of phenomenological psychology,* 14(2), 243-266.

Latour, B. (2005). *Reassembling the social: an introduction to actor-network-theory*. Oxford New York: Oxford University Press.

Lazar, S. (2017). War. In *The Stanford Encyclopedia of Philosophy* (Spring Edition), Edward N. Zalta (ed.), Retrieved from: https://plato.stanford.edu/archives/spr2017/entries/war/

Leder, F., Werner, T., & Martini, M. (2009). Proactive Botnet Countermeasures: An Offensive Approach. In C. Czosseck, & K. Geers (Eds.), *The Virtual Battlefield: Perspectives on Cyber Warfare*. Washington, DC: IOS Press.

Leetaru, A. (2017). What Tallinn Manual 2.0 Teaches Us About The New Cyber Order. *Forbes.* Retrieved on February 25, 2018, from https://www.forbes.com/sites/kalevleetaru/2017/02/09/what-tallinn-manual-2-0-teaches-us-about-the-new-cyber-order/#11b2f6d9928b

Lemos, N.M. (2015). Value. In A. Robert (Eds.), *The Cambridge Dictionary of Philosophy* (pp.1100-1101). Third Edition. Cambridge. Cambridge University Press.

Lin, P., Allhoff, F., & Abney, K. (2014). Is Warfare the Right Frame for the Cyber Debate? In M. Taddeo & L. Floridi (Eds.), *The Ethics of Information Warfare* (pp. 39–60). Cham: Springer.

Maathuis, C., Pieters, W., & Berg, J. V. (2016). Cyber weapons: A profiling framework. *2016 International Conference on Cyber Conflict (CyCon U.S.)*. doi:10.1109/cyconus.2016.7836621

Mahmud, H., Quaisar, M. M., Sabur, M. A., & Tamanna, S. (2008). Human Security or National Security: The Problems and Prospects of the Norm of Human Security. *Journal of Politics and Law, 1*(4). doi:10.5539/jpl.v1n4p67

McGuinness, D. (2017). How a cyber attack transformed Estonia. *BBC NEWS*. Retrieve April 21, 2018, from https://www.google.ru/amp/s/www.bbc.com/news/amp/39655415

McMahan, J. (2010). The Just Distribution of Harm Between Combatants and Noncombatants. *Philosophy & Public Affairs,* 38, 342-379. doi:10.1111/j.1088-4963.2010.01196.x

Meredith, C., & Christou, T. A. (2009). *Not in my front yard: Security and resistance to responsibility for extraterritorial state conduct*. Nijmegen: Wolf Legal (WLP).

Militair Rechtelijk Tijdschrift Jaargang 111 - 2018 - 3 Cyber Special. (2018). Den Hague. Retrieved on August 30, 2018, from http://puc.overheid.nl/doc/PUC_248137_11

Moreno, V. C., Reniers, G., Salzano, E., & Cozzani, V. (2018). Analysis of physical and cyber security-related events in the chemical and process industry. *Process Safety and Environmental Protection,* 116, 621-631. doi:10.1016/j.psep.2018.03.026

Noor, K. B. M. (2008). Case Study: A Strategic Research Methodology. *American Journal of Applied Sciences*, 5(11), 1602-1604.

Pattison, J. (2016). The Case for the Nonideal Morality of War: Beyond Revisionism versus Traditionalism in Just War Theory. *Political Theory, 46*(2). 242-268. doi:10.1177/0090591716669394

Pew Research Center. (n.d.). Knowledge of various cybersecurity topics among adults in the United States as of June 2016. In *Statista - The Statistics Portal*. Retrieved on November 1, 2018, from https://www.statista.com/statistics/807380/united-states-cybersecurity-topics-awareness/.

Pinch, T., & Bijker, W. (1982). The Social Construction of Facts and Artefacts: or How the Sociology of Science and the Sociology of Technology Might Benefit Each Other. *Social Studies of Science,* 14, 399-441.

Ponemon Institute, & Raytheon. (n.d.). Greatest cyber threat risks over the next three years according to senior information technology practitioners worldwide as of 2017. In *Statista - The Statistics Portal*. Retrieved on November 1, 2018, from https://www.statista.com/statistics/883716/greatest-cyber-threat-risks-expected-by-senior-it-practitioners-threat-type/.

Porter, B. F. (2017). *The Good Life: Options in Ethics*. Rowman & Littlefield Publishers.

Proofpoint. (2017). *The Human Factor Report.* Retrieved from*:* https://seclab.stanford.edu/courses/cs203/lectures/humanfactor.pdf

Radin, A. (2017). *Hybrid Warfare in the Baltics: Threats and Potential Responses.* Santa Monica, CA: RAND Corporation. Retrieved from: https://www.rand.org/pubs/research_reports/RR1577.html.

Reid, K., & Flowers, P., & Larkin, M. (2005). Exploring lived Experience. *The Psychologist,* 18, 18-23.

Reijers, W., & Coeckelbergh, M. (2018). Narrative technologies meets virtue ethics in alternate reality. *ACM SIGCAS Computers and Society, 47*(4), 96-106. doi:10.1145/3243141.3243152

Rid, T. (2012). Cyber War Will Not Take Place, *Journal of Strategic Studies,* 35(1), 5-32. DOI: 10.1080/01402390.2011.608939

Robinson, P. (2006). *Military Honour and the Conduct of War.* New York: Routledge.

Ruus, K. (2008). Cyber War I: Estonia Attacked from Russia. *European Affairs, 9(1-2), Columbia International Affairs Online.*

Schmitt, M. (1999). Computer Network Attack and the Use of Force in International Law: Thoughts on a Normative Framework. *Columbia Journal of Transnational Law, 37*, 1998-1999. Retrieved from: https://ssrn.com/abstract=1603800

Schmitt, M. (2017). Grey Zones in the International Law of Cyberspace. *Yale Journal of International Law Online 1,* 42(2), 1-21. Retrieved from: https://ssrn.com/abstract=3180687

Schmitt., M. (2017). Peacetime Cyber Responses and Wartime Cyber Operations Under International Law: An Analytical Vade Mecum. *Harvard National Security Journal,* 8, pp. 239-282.

Schmitt, M. (2017). *Tallinn Manual 2.0 on the International Law of Cyber Operations: What It Is and Isn't.* Webcast from the launch of the Tallinn Manual 2.0 on the International Law Applicable to Cyber Operations at the Atlantic Council in Washington, D.C. Retrieved on August 2, 2018 from https://www.justsecurity.org/37559/tallinn-manual-2-0-international-law-cyber-operations/

Schmitt, M. et al. (2017). *Tallinn Manual 2.0 on the International Law Applicable to Cyber Operations*. Cambridge: Cambridge University Press. doi:10.1017/9781316822524

Shackelford, S. J. (2010). Estonia three years later: A progress report on combating cyber attacks. *Journal of Internet Law*, *138*, 22–29.

Shachtman, N. (2009). Kremlin Kids: We Launched The Estonian Cyber War. *Wired*. Retrieved June 4, 2018 from https://www.wired.com/2009/03/pro-kremlin-gro/

Sleat, M. (2017). Just cyber war?: Casus belli, information ethics, and the human perspective. *Review of International Studies,* 44(2), 324-342. doi:10.1017/s026021051700047x

Smith, J. A. (2011) Evaluating the contribution of interpretative phenomenological analysis, *Health Psychology Review,* 5(1), 9-27, DOI: 10.1080/17437199.2010.510659

Smith, J. A., Larkin, M., & Flowers, P. (2009). *Interpretative phenomenological analysis: Theory, method and research*. London: SAGE.

Smith, J. A., & Osborn, M. (2015). Interpretative phenomenological analysis as a useful methodology for research on the lived experience of pain. *British Journal of Pain, 9*(1), 41-42. doi:10.1177/2049463714541642

Smith, P. T. (2018). Cyberattacks as Casus Belli: A Sovereignty-Based Account. *Journal of Applied Philosophy, 35*(2), 222-241. doi:10.1111/japp.12169

Smith P. T. (2017). Towards a Richer Account of Cyberharm: The Value of Self-Determination in the Context of Cyberwarfare. In M. Taddeo, L. Glorioso (Eds). *Ethics and Policies for Cyber Operations. Philosophical Studies Series*, 124. Springer, Cham.

Spradley, J. P. (1979). *The ethnographic interview.* New York: Holt, Rinehart and Winston.

Solum, L. B. (2018). Virtue as the end of law: an aretaic theory of legislation. *Jurisprudence: An International Journal of Legal and Political Thought*, 9(1), 6-18, DOI: 10.1080/20403313.2017.1369725

Stone, J. (2013). Cyber War *Will* Take Place! *Journal of Strategic Studies,* 36(1), 101-108. DOI: 10.1080/01402390.2012.730485

Swierstra, T. (2016). Introduction to the Ethics of New and Emerging Science and Technology. In R. Nakatsu, M. Rauterberg, & P. Ciancarini (Eds.), *Handbook of Digital Games and Entertainment Technologi*es (pp. 1271-1295). Springer Singapore.

Swierstra, T., & Waelbers, K. (2012). Designing a Good Life : A Matrix for the Technological Mediation of Morality. *Sci Eng Ethics*, *18*, 157–172. https://doi.org/10.1007/s11948-010-9251-1

Taddeo, M. (2012). An Analysis For A Just Cyber Warfare. In C. Czosseck, R. Ottis, & K. Ziolkowski (Eds.), *2012 4th International Conference on Cyber Conflict* (pp. 209–218). Tallinn: NATO CCD COE Publications.

Taddeo, M. (2014). Information Warfare and Just War Theory. In M. Taddeo & L. Floridi (Eds.), *The Ethics of Information Warfare* (pp. 123–139). Cham: Springer.

Taddeo, M. (2016). Just Information Warfare. *Topoi*, *35*, 213–224. https://doi.org/10.1007/s11245-014-9245-8

Tamkin, E. (2017). 10 Years After the Landmark Attack on Estonia, Is the World Better Prepared for Cyber Threats? *Foreign Policy*. Retrieved on May 15, 2018 from https://www.google.ru/amp/foreignpolicy.com/2017/04/27/10-years-after-the-landmark-attack-on-estonia-is-the-world-better-prepared-for-cyber-threats/amp/

Tavani, H. T. (2013). *Ethics and Technology: controversies, questions, and strategies for ethical computing*. In B. L. Golub & J. Ling (Eds.), Nashua, NH: River University.

Tikk, E., Kaska, K., & Vihul, L. (2010). *International cyber incidents: Legal considerations*. Tallinn, Estonia: Cooperative Cyber Defence of Excellence (CCD COE).

Vallor, S. (2015). Moral Deskilling and Upskilling in a New Machine Age: Reflections on the Ambiguous Future of Character. *Philosophy & Technology,* 28(1), 107-124. doi:10.1007/s13347-014-0156-9

Vallor, S. (2016). *Technology and the Virtues: a Philosophical Guide to a Future Worth Wanting*. New York: Oxford University Press.

Volkman, R. (2011). Why Information Ethics must begin with Virtue Ethics. In A. T. Marsoobian, B. J. Huschle, E. Cavallero and P. Allo (Eds.), *Putting Information First: Luciano Floridi and the philosophy of information.* Oxford: Wiley-Blackwell.

Walzer, M. (1977). *Just and unjust wars: A moral argument with historical illustrations*. New York: Basic Books.

Wong, J.K., & Solon, O. (2017). Massive ransomware cyber-attack hits nearly 100 countries around the world. *The Guardian,* Retrieved April 11, 2018, from https://www.theguardian.com/technology/2017/may/12/global-cyber-attack-ransomware-nsa-uk-nhs

Zimmerman, M. (2015). Intrinsic vs. Extrinsic Value. *The Stanford Encyclopedia of Philosophy.* Metaphysics Research Lab, Stanford University.

Appendix 1. The Conceptual Justifications: the Definition of Cyberwarfare

While cyberwarfare often pops up as a name for the different cyber attack cases on the media, the cross-disciplinary environment around the definition of “cyberwarfare” is still rather problematic. In their paper “The Hierarchy of Cyberwar Definitions” (2017), Daniel Hughes and Andrew Colarik provided a comprehensive definitional analysis of 159 publications that contained notions like ‘cyber war’ and ‘cyber warfare’ and different types of spelling of such a concept in their titles and abstract keywords. Following this search criteria, their discourse analysis included all publications that were published before 2016. This resulted in an academic output that provides an extensive conceptual background for the further ethical analysis of cyberwarfare. Hughes and Colarik deduced their textual and disciplinary analysis to several conclusions. I will list some of their findings in order to pick a definition and ‘map’ the cross-disciplinary conceptual environment around cyberwarfare.

First, although the notion of war is classically attributed to the act of war, and is different of what warfare implies, studies showed that there is no basis to definitely distinguish between ‘cyber war’ and ‘cyber warfare.’ Their quantitative analysis revealed that out of 159 publications, 39 only used ’cyber war,’ 43 use ’cyber warfare’ and the rest 75 used both terms without any distinction (Hughes & Colarik, 2017, p. 17). Thus, to strictly follow the aim of this thesis in unpacking the values in TM 2.0, I will consistently stick to the term ‘cyberwarfare.’ However, I will use it as synonymous to ‘cyber war,’ ‘cyberwar,’ and ‘cyber warfare,’ with awareness of the fact that conceptual clarity around cyberwarfare is an urgent inquiry for philosophers and ethicists, that unfortunately, goes beyond the scope of this thesis.

The second conclusion the authors made which is relevant for further conceptual clarity, is that “despite location in a domain ostensibly concerned with the explication and implications of newly emerged technologies and modalities, a majority of articles do not offer explicit definitions of either cyber war or cyber warfare from which to base their analysis” (Hughes & Colarik, 2017, p. 29). The authors pointed out the distinction between an explicit and an implicit definition. The former being the one that is particularly formulated and referred in the text, and the latter that has, without specific articulation, an implicit reference to the debate. More specifically, out of 159 publications, only 56 used explicit definitions. This indicates a raising conceptual gap fueled by the stipulative use of the ‘cyberwarfare’ definition and a dis-convergence upon particular terminological formulas. That is why, for the sake of conceptual clarity, I will further select the most comprehensive single definition.

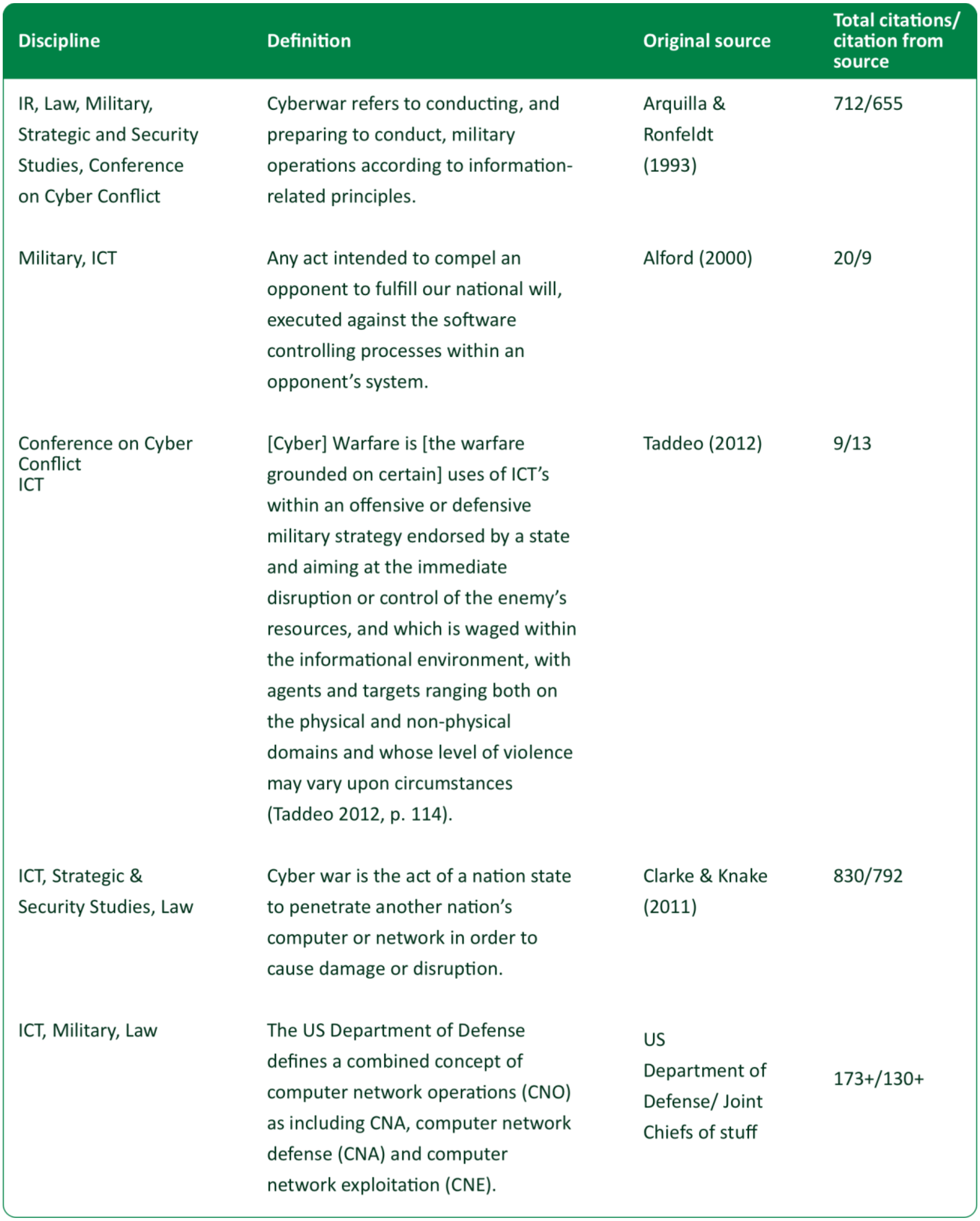
Considering all the above-mentioned points, it is clear that the discourse is inherently inter-disciplinary, and as Hughes & Colarik pointed out, the definitions that arise from different disciplines migrate across the articles. Strategic and Security Studies, Military Studies, International Relations, Law, ICT, Ethics are dominant disciplines in the debate around cyberwarfare, thus the latter is a multidisciplinary issue. However, the conceptual lack of clarity around cyberwarfare is a significant challenge for the flow and accuracy in the cross-disciplinary debate. As R1 commented on the difficulties to find an agreement upon the definition of cyberwarfare in the interview:

> Difficulties with the definition of cyberwarfare are in essence similar to the ambiguities and complexities in a conceptual agreement on the definition of such phenomena as terrorism. It became a politically loaded term, with no universal definition.

A similar point was made by R2 - the idea of politicization of the security issue - which is commonly discussed in International Relations Studies.

In this view, I suggest to critically examine several definitions in order to get an understanding of the conceptual complexities surrounding cyberwarfare and evaluate which one is the most comprehensive. Following Hughes & Colarik (2017), it is possible to extract five core definitions from the cross-disciplinary discourse around cyberwarfare. These are distinguished in their hierarchical order according to the widespread academic usage and quantitative analysis of the number of citations (Hughes & Colarik, 2017). I propose to take a close look at the conceptual differences of these definitions. They are introduced in Table 4.

Table 4. Breakdown of cross-disciplinary definitions (Hughes & Colarik, 2017, p. 24)



Discipline	Definition	Original source	Total citations/ citation from source
IR, Law, Military, Strategic and Security Studies, Conference on Cyber Conflict	Cyberwar refers to conducting, and preparing to conduct, military operations according to information-related principles.	Arquilla & Ronfeldt (1993)	712/655
Military, ICT	Any act intended to compel an opponent to fulfill our national will, executed against the software controlling processes within an opponent's system.	Alford (2000)	20/9
Conference on Cyber Conflict ICT	[Cyber] Warfare is [the warfare grounded on certain] uses of ICT's within an offensive or defensive military strategy endorsed by a state and aiming at the immediate disruption or control of the enemy's resources, and which is waged within the informational environment, with agents and targets ranging both on the physical and non-physical domains and whose level of violence may vary upon circumstances (Taddeo 2012, p. 114).	Taddeo (2012)	9/13
ICT, Strategic & Security Studies, Law	Cyber war is the act of a nation state to penetrate another nation's computer or network in order to cause damage or disruption.	Clarke & Knake (2011)	830/792
ICT, Military, Law	The US Department of Defense defines a combined concept of computer network operations (CNO) as including CNA, computer network defense (CNA) and computer network exploitation (CNE).	US Department of Defense/ Joint Chiefs of stuff	173+/130+

From these five definitions it becomes clear that some scholars (Clarke & Knake, 2011; Taddeo, 2012) explicitly specify a nation-state as the main actor, and others do not. Nevertheless, they all implicitly presuppose it by inserting loaded concepts like 'national will' and 'military operation.' At the same time, while Alford's (2000) definition refers to software as the target, other four definitions provide more space for potential targets referring to informational environments and computer networks. The most encompassing definition is the one offered by the ethicist Taddeo (2012). I do not claim it is a universal definition. However, for the operational purposes of this thesis, Taddeo's definition provides substantial clarity on the nature and components of cyberwarfare.

In what follows, I elaborate a bit further on some details of Taddeo's definition and insert some clarifications. According to Hughes's and Colarik's (2017) observation, Taddeo's definition is conceptually rooted in two often cited and influential definitions proposed by Clarke & Knake (2011) and Arquilla & Ronfeldt (1993, p. 29). Both of these definitions presuppose military context. At the same time, in her definition, Taddeo specifies the informational environment as a core location, which extends from Arquilla & Ronfeldt's (1993) reference to information-related principles.[22] However, what is considered under these principles and why are they relevant in a conceptual analysis of cyberwarfare? In order to answer this question, I turn to the discussion of cyberwarfare as compared to the classical conceptualization of warfare.

Arquilla and Ronfeldt (1993) claim that the idea of cyberwarfare goes far beyond merely improving and defending the state's own systems while attacking the enemies. They refer to the context of classic war theory suggested by Carl von Clausewitz, framing the idea that cyberwar "is characterized by the effort to turn knowledge into capability" (Arquilla and Ronfeldt, 1993, p. 32). The authors illustrate this claim with the strategies used by the Mongols in the 13th-century,[23] who succeeded partly due to their dominance on the informational battlefield. To phrase it differently, when one sees a bigger picture of the battle and calculates different strategies according to an extensive knowledge about this picture, the chances of winning become relatively high, in spite of potential advantages of the adversary.

[22] The authors formulated a clear division between the typology 'netwar' and 'cyberwar.' They defined 'netwar' as "a new entry on the spectrum of conflict that spans economic, political, and social as well as military forms of "war" (Arquilla and Ronfeldt, 1993, p. 28). While 'cyberwar,' may instead "raise broad issues of military organization and doctrine, as well as strategy, tactics, and weapons design" (Arquilla and Ronfeldt, 1993, p. 32). In other words, the guiding assumption is that the role of the organization and technology is significant as such, because compared to 'netwar,' 'cyberwar' "may be applicable in low- and high-intensity conflicts, in conventional and non-conventional environments, and for defensive or offensive purposes" (Arquilla and Ronfeldt, 1993, p. 31).

[23] Quoting the note about 13 century Mongols example directly from the source: "We might appear to be extrapolating from the U.S. victory in the Gulf War against Iraq. But our vision is inspired more by the example of the Mongols of the 13th Century. Their "hordes" were almost always outnumbered by their opponents. Yet they conquered, and held for over a century, the largest continental empire ever seen. The key to Mongol success was their absolute dominance of battlefield information. They struck when and where they deemed appropriate; and their "Arrow Riders" kept field commanders, often separated by hundreds of miles, in daily communication" (Arquilla and Ronfeldt, 1993, p. 24).

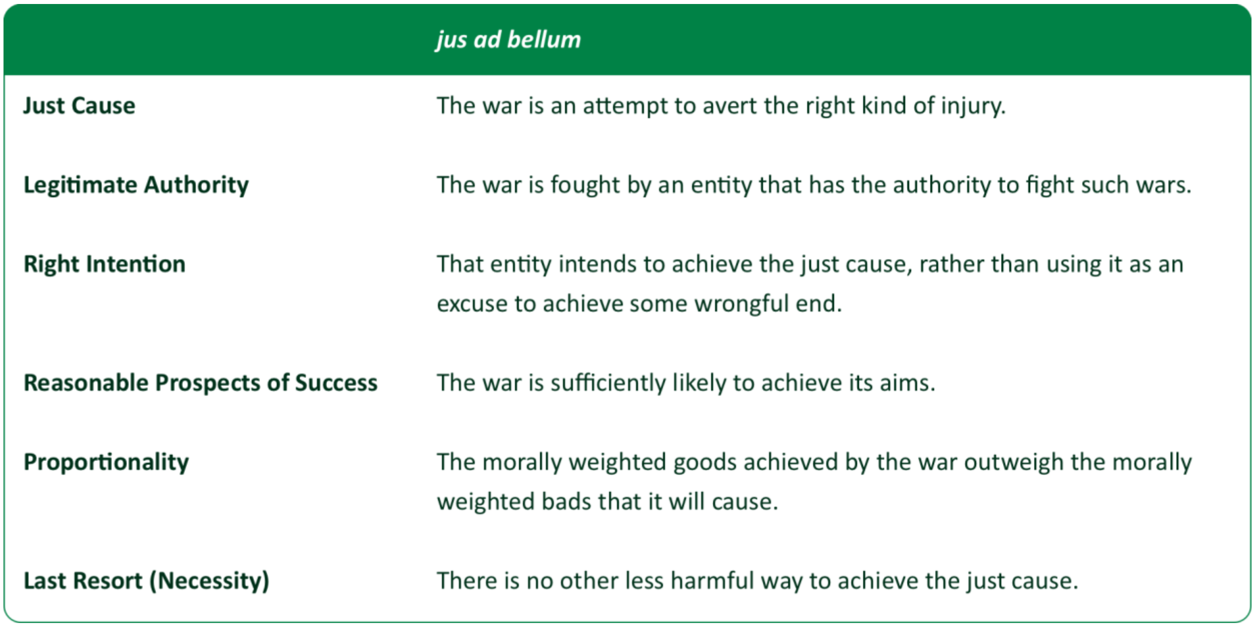
The information becomes a sort of “weapon,” that, in collaboration with other means, contributes beneficially in conventional and non-conventional environments of low- or high-intensity conflicts for defensive or offensive purposes.

In this view, Taddeo’s definition incorporates two points: 1) it adopts the idea of diversity of wars, in a Clausewitz sense, by explicitly distinguishing between ‘physical and non-physical domains’ of operation; 2) it highlights the informational nature rooted in the Information Revolution, and thus, affords the non-physical environment for strategic operations. Thus, considering the transversality included in the categorization of Taddeo’s definition, I consider this definition to be the most comprehensive from those listed in Table 4.

Appendix B. The Core Principles of Just War Theory

The six principles of *jus ad bellum* relate to one another (Lazar, 2017) and can be best described in a nutshell as represented in Table 4.

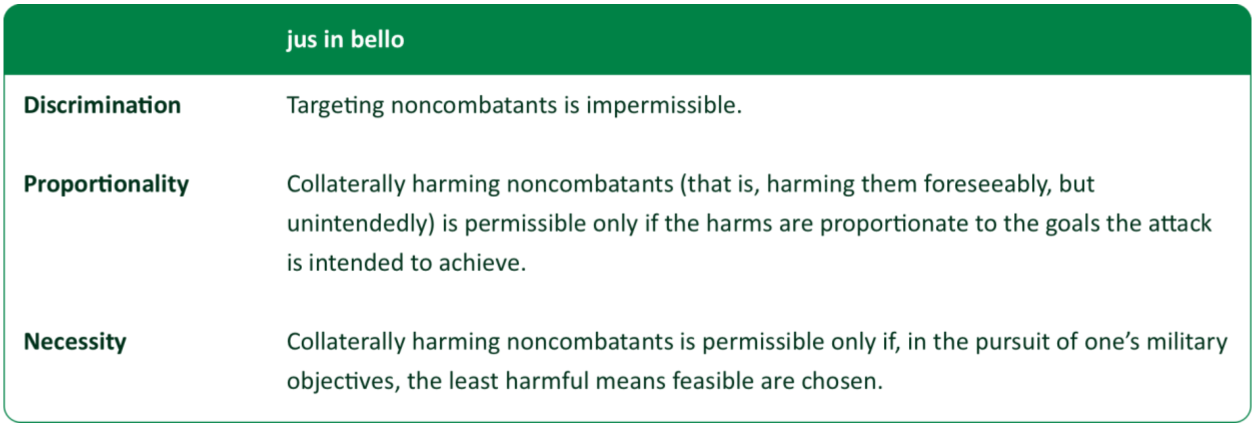
Table 4. *jus ad bellum* principles (Lazar, 2017).



	jus ad bellum
Just Cause	The war is an attempt to avert the right kind of injury.
Legitimate Authority	The war is fought by an entity that has the authority to fight such wars.
Right Intention	That entity intends to achieve the just cause, rather than using it as an excuse to achieve some wrongful end.
Reasonable Prospects of Success	The war is sufficiently likely to achieve its aims.
Proportionality	The morally weighted goods achieved by the war outweigh the morally weighted bads that it will cause.
Last Resort (Necessity)	There is no other less harmful way to achieve the just cause.

The principles of *jus in bello* hold that conduct in war must satisfy three following principles presented in Table 5.

Table 5. *jus in bello* principles (Lazar, 2017).

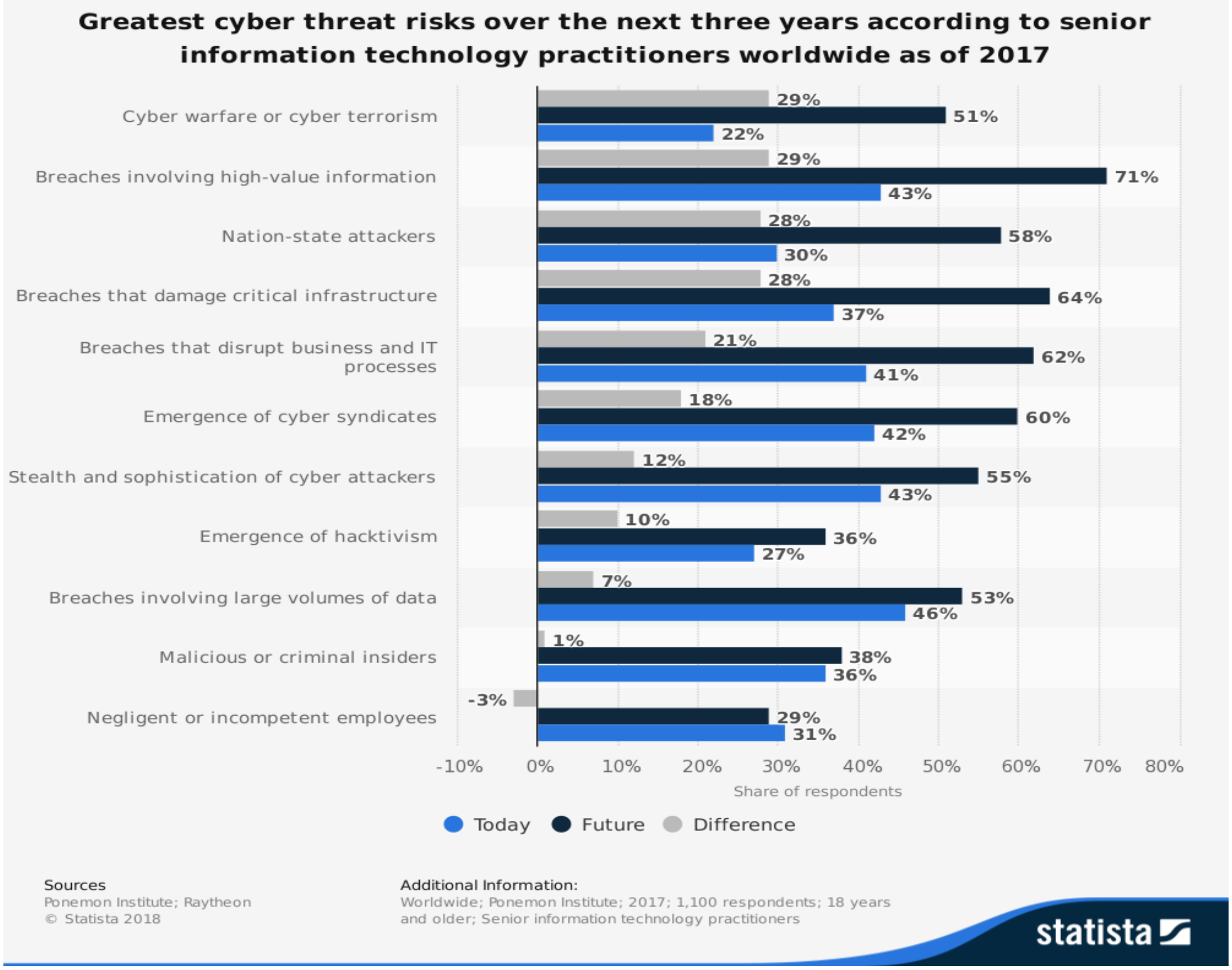


	jus in bello
Discrimination	Targeting noncombatants is impermissible.
Proportionality	Collaterally harming noncombatants (that is, harming them foreseeably, but unintendedly) is permissible only if the harms are proportionate to the goals the attack is intended to achieve.
Necessity	Collaterally harming noncombatants is permissible only if, in the pursuit of one's military objectives, the least harmful means feasible are chosen.

Appendix C. Quantitative Data About Future Cyber Threats

Considering quantitative data conducted by Statista about “Greatest cyber threat risks over the next three years according to senior information technology practitioners worldwide as of 2017,” those cases that are distinguished by TM 2.0 as an act of war are, according to 1100 respondents (senior information technology practitioners), just a part of the threat that appear to be relevant for security considerations (see picture 3).

Picture 3. Greatest cyber threat risks over the next three years according to senior information technology practitioners worldwide as of 2017[24].



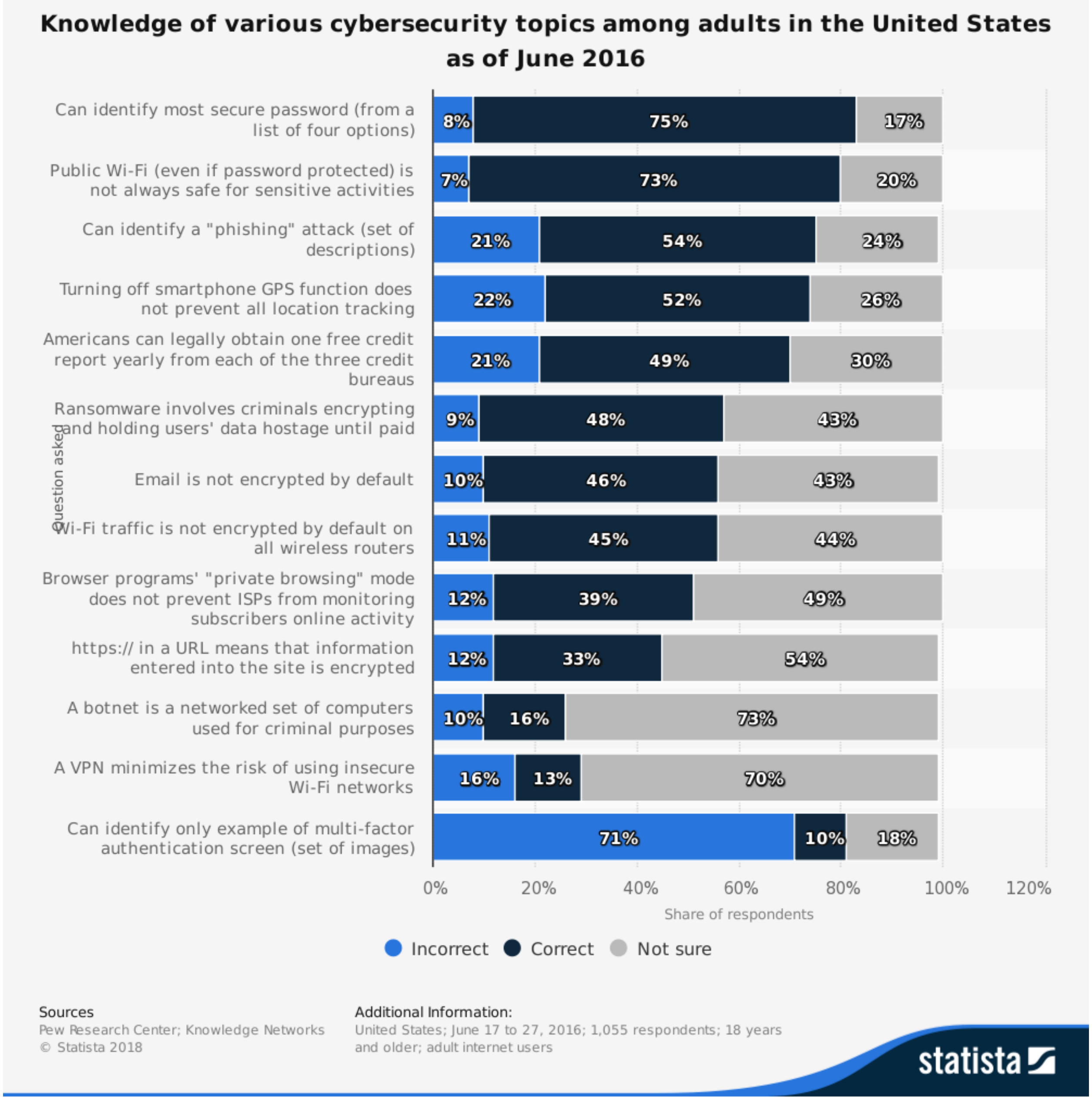
Sources
Ponemon Institute; Raytheon
© Statista 2018

Additional Information:
Worldwide; Ponemon Institute; 2017; 1,100 respondents; 18 years and older; Senior information technology practitioners

[24] *See* https://www.statista.com/statistics/807380/united-states-cybersecurity-topics-awareness/.

Appendix D. Quantitative Data About Cybersecurity Literacy in the U.S (2016)

Picture 4. Knowledge of various cybersecurity topics among adults in the United States as of June 2016.[25]



Sources
Pew Research Center; Knowledge Networks
© Statista 2018

Additional Information:
United States; June 17 to 27, 2016; 1,055 respondents; 18 years and older; adult internet users

[25] *See*: https://www.statista.com/statistics/883716/greatest-cyber-threat-risks-expected-by-senior-it-practitioners-threat-type/.

Appendix E. Questions for Expert Interview (R1)

In May 2018 I conducted semi-structured interview with an expert in Computer Ethics. The areas of expertise of this researcher includes computer ethics, information ethics, military, security, political philosophy. The interview took approximately 60 minutes. The permission to record the interview was given, and name, gender and age are anonymized. Below in Table 6 I exemplify several question from the interview. Original raw data (recordings) was deleted, and anonymized material is stored on a separate USB as a transcript.

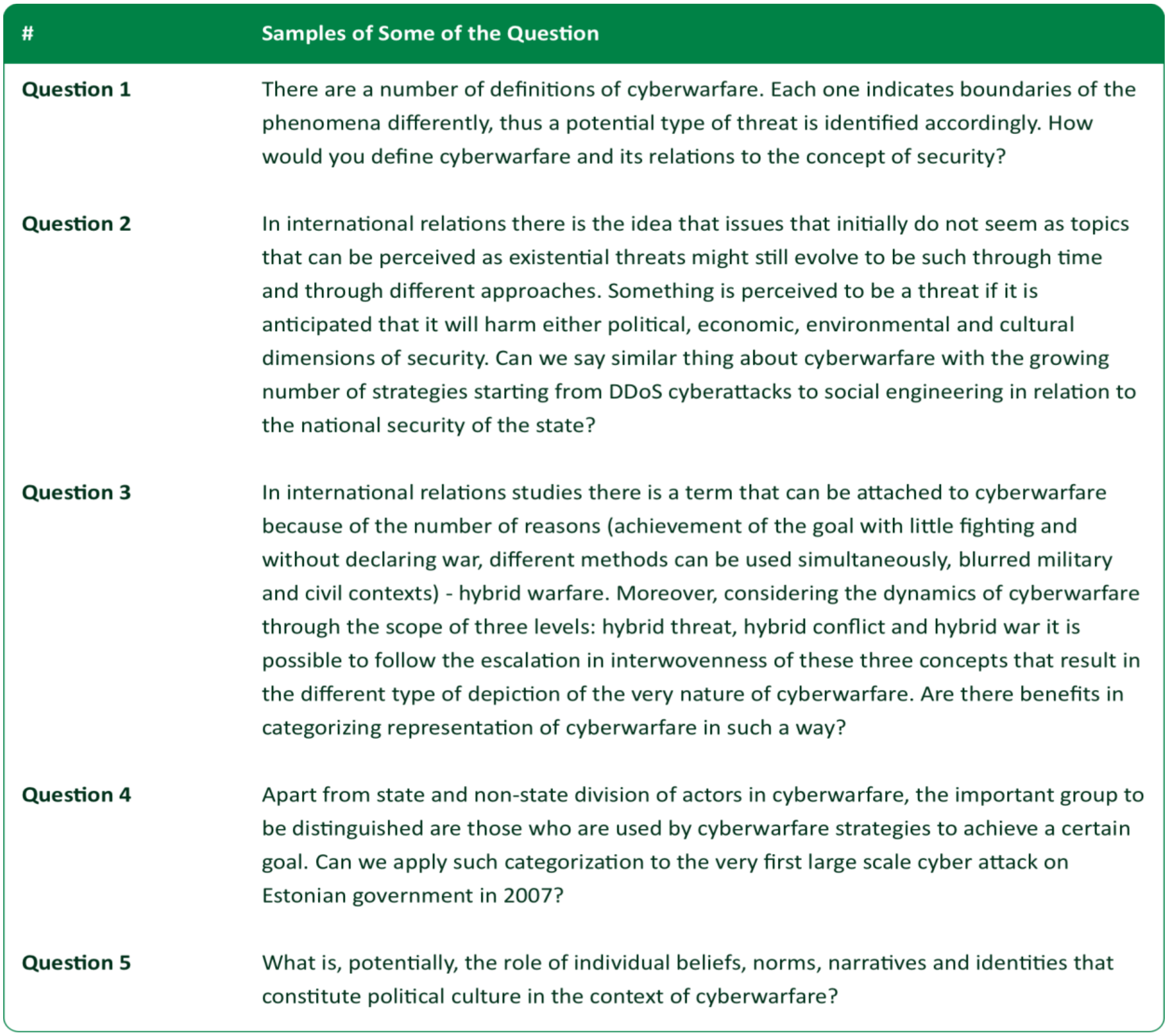
Table 6. Sample questions for R1.

#	Samples of Some of the Question
Question 1	There are many definitions of cyber warfare and its definition creates different boundaries of the phenomena which include or exclude certain aspects. For instance, some definitions exclude social engineering as the type of cyberwarfare. But other definitions include it. So, what do you think can be the definition of cyber warfare in relation to concept of security, within your domain for instance?
Question 2	So I wonder then if it is the problem with the way we frame cyberwarfare, or with the fact that we try to categorize it within those existing traditions?
Question 3	The most prominent scholars like Floridi and Taddeo are suggesting their JCW approach as the ethical basis for the analysis of cyberwarfare. However, JWT (theory) is criticized for a number of years as insufficient in the scope of new and emerging technologies. Why is it still so widely used and applied?
Question 4	Considering the nature of JCW as a foundational approach with extreme impartialism and universalism, is it really useful in regulation of cyberwarfare? Are there alternatives? What is in your opinion missing?
Question 5	Do you find values and virtues to be relevant in cyberwarfare? In this view, could the perspective offered by virtue ethics bring experiential aspect to make the analysis less abstract? Would that enhance the existing debate? What could be the possible objection and limitation?

Appendix F. Questions for Expert Interview (R2)

In April 2018 I conducted semi-structured interview with an expert in International Relations. The areas of expertise of this researcher includes international law, information warfare, military, security, political science. The interview took approximately 60 minutes. The permission to record the interview was given, and name, gender and age are anonymized. Below in Table 7 I exemplify several question from the interview. Original raw data (recordings) was deleted, and anonymized material is stored on a separate USB as a transcript.

Table 7. Sample questions for R2.



#	Samples of Some of the Question
Question 1	There are a number of definitions of cyberwarfare. Each one indicates boundaries of the phenomena differently, thus a potential type of threat is identified accordingly. How would you define cyberwarfare and its relations to the concept of security?
Question 2	In international relations there is the idea that issues that initially do not seem as topics that can be perceived as existential threats might still evolve to be such through time and through different approaches. Something is perceived to be a threat if it is anticipated that it will harm either political, economic, environmental and cultural dimensions of security. Can we say similar thing about cyberwarfare with the growing number of strategies starting from DDoS cyberattacks to social engineering in relation to the national security of the state?
Question 3	In international relations studies there is a term that can be attached to cyberwarfare because of the number of reasons (achievement of the goal with little fighting and without declaring war, different methods can be used simultaneously, blurred military and civil contexts) - hybrid warfare. Moreover, considering the dynamics of cyberwarfare through the scope of three levels: hybrid threat, hybrid conflict and hybrid war it is possible to follow the escalation in interwovenness of these three concepts that result in the different type of depiction of the very nature of cyberwarfare. Are there benefits in categorizing representation of cyberwarfare in such a way?
Question 4	Apart from state and non-state division of actors in cyberwarfare, the important group to be distinguished are those who are used by cyberwarfare strategies to achieve a certain goal. Can we apply such categorization to the very first large scale cyber attack on Estonian government in 2007?
Question 5	What is, potentially, the role of individual beliefs, norms, narratives and identities that constitute political culture in the context of cyberwarfare?

Appendix G. Questions for Estonian Citizen (R3)

In April 2018 I conducted semi-structured interview with an Estonian citizen from Russian ancestry (R3). The interview took approximately 60 minutes. The permission to record the interview was given, and name, gender and age are anonymized. Below in Table 8 I exemplify several questions from the interview in Russian language. Original raw data (recordings) was deleted, and anonymized material is stored on a separate USB as a transcript.

Table 8. Sample questions for R3.

#	Question	Translation from Russian
Question 1	В 2007 году, Эстония прогремела на весь мир новостями про так называемую Бронзовую Ночь, ряд протестов в в 2007 году сопровождаемый кибер атаками. Расскажите про ваши впечатления?	In 2007 there were a lot of news from Estonia about a so-called Bronze Night, a string of protests followed by cyber attacks. Can you please tell me about your impressions of it?
Question 2	Какая роль технологий в вашей ежедневной жизни? Какие плюсы и какие минусы?	What's the role of technologies in your everyday life? What are their upsides and downsides?
Question 3	Помните ли вы как происходил процесс внедрения технологий в обществе и какие сложности были в этом?	Do you remember how the process of implementation of technologies go and what types of difficulties you faced?
Question 4	Есть ли сложности в использовании технологий сейчас? Какая в целом ваша оценка таких изменений?	Are the any difficulties in using technologies right now? What do you think about these changes?
Question 5	А как граждане страны относятся к этим изменениям? Одобряют они эту динамику? Насколько люди открыты этим изменениям? Что бы вы могли выделить как ключевое для вас?	How do the citizens feel about these changes? Do they support this dynamic? How open are people towards these changes? What would you say in the key for you in all of it?

Appendix H. Questions for Estonian Citizen (R4)

In September 2018 I conducted semi-structured interview with an Estonian citizen from Estonian ancestry *(R4)*. The interview took approximately 60 minutes. The permission to record the interview was given, and name, gender and age are anonymized. Below in Table 9 I exemplify several question from the interview. Original raw data (recordings) was deleted, and anonymized material is stored on a separate USB as a transcript.

Table 9. Sample questions for R4.

#	Question
Question 1	Eleven years ago large scale cyber attacks took place in Estonia. Do you remember those events? What was it like? Could you share your experience? **Additional questions:** What services you could not use? How long the destructions continue? Were there any support provided?
Question 2	How technological environment changed in Estonia since then? Were there particular changes related to the events in 2007?
Question 3	How was digitalization accepted by the public? Were there any challenges during adaptation? Did you also face something like that? Was there any support provided to deal with those challenges? What in your opinion was important in all these processes?
Question 4	Can you remember other cyber attacks that you could actually feel (not read in news)? Does cybersecurity have any meaning in your life? How do you learn to protect yourself?
Question 5	What would you highlight as important during and after cyber attacks? Do any of these apply in Estonia? What would you say are the weaker and stronger sides of digital Estonia?