

TRUST TRUMPS SECURITY

BEER SIJPESTEIJN

Why trust is key to well-being in cyberspace and security a means to its
achievement

Master Thesis
MSc Philosophy of Science, Technology & Society
Department of Philosophy
Faculty of Behavioural, Management & Social Sciences
University of Twente

June 5th, 2015

Beer Sijpesteijn: *Trust Trumps Security*, Why trust is key to well-being in cyberspace and security a means to its achievement, June 5th, 2015


SUPERVISORS:

Johnny Søraker
Michael Nagenborg
Toon Segers
Tim van Essen

LOCATION:

Enschede, The Netherlands

LICENSE:

 Creative Commons Attribution-ShareAlike 4.0 International

FOR GREAT JUSTICE

— Zero Wing

ABSTRACT

Cyberspace holds the promise of increasing the well-being of billions, improving their lives through such affordances as access to information, contact with like-minded people, and providing commercial opportunities. This thesis contends that trust is key to cyberspace's success and of paramount importance to well-being in cyberspace. Trust in cyberspace is not trivially present though. Our increasing dependence on cyberspace and the raising stakes in cyber-conflicts have governments ramp up their cyber-capabilities. If not thought out well, such efforts threaten to be counter-productive and undermine the principles that promote well-being in cyberspace.

This thesis forms an inquiry into trust in cyberspace. A characterization of cyberspace is developed, extending technical network design principles to include the virtual and cognitive aspects of cyberspace, whilst acknowledging its evolutionary nature. Cyberspace is characterized as a grand decentralized network, conjoining myriads of circuits that represent different meanings that cyberspace has for different users at different times; circuits are made up of nodes of a content, logical or physical nature. Next is an analysis of the conception of trust as security, which turns out to be problematic with regard to well-being in cyberspace. Although security is an important condition for many affordances, it is not a suitable alternative to the fundamental role of trust to well-being in cyberspace. Security interests of different actors clash, while well-being in cyberspace requires an agnostic core. A key element of trust is the leap of faith that provides room for individual interpretation and decision making about the trustworthiness of, in casu, the agnostic core of cyberspace.

Combining these insights, a novel conception of trust in cyberspace is presented. Trust in cyberspace is conceived as a succession of leaps of faith from node to node, each time assessing the trustworthiness of the next node. If all nodes in a circuit are deemed trustworthy, trust in that particular meaning at that instance of cyberspace exists. Security proper enhances the trustworthiness of nodes and hence is a means to achieving trust and thus well-being in cyberspace. Consequently, what those who are committed to well-being in cyberspace should do, is finding ways to promote the trustworthiness, as perceived by as many people as possible, of the nodes that they can influence. If adopted by governments, this would align the responsibilities they feel with the promises of cyberspace.

“Cyberspace. A consensual hallucination experienced daily by billions of legitimate operators, in every nation, by children being taught mathematical concepts... A graphic representation of data abstracted from the banks of every computer in the human system. Unthinkable complexity. Lines of light ranged in the nonspace of the mind, clusters and constellations of data. Like city lights, receding.”

— William Gibson in *Neuromancer*

ACKNOWLEDGMENTS

With great delight I present this thesis that I wrote in order to obtain the degree of Master of Science in the Philosophy of Science, Technology & Society at the University of Twente. I started with the first preparations for my thesis in the spring of 2014. The ensuing summer I developed a rough idea about the subject (what I then tentatively called *ethics of cyber warfare*), found supervisors, and landed an internship during which I could write my thesis. Monday the 1st of September I officially started my six-month internship at Deloitte Cyber Risk Services. Now, nine months later, I am almost done.

I want to express my gratitude to my supervisors, Tim and Toon at Deloitte, Johnny and Michael at the University of Twente, for all the effort they put in helping me write this thesis. I also want to thank the other Michael, my initial primary supervisor who seized a great job opportunity outside academia, for helping me find direction in the first months of my thesis. Furthermore, I want to thank everybody inside and outside Deloitte whom I talked, sparred and discussed with to further develop my ideas. I also want to express my appreciation for the patience of so many around me when I would respond with little enthusiasm when they showed interest in my thesis — at those times I just wanted to have something else on my mind.

What I am really happy about is that for both Deloitte and the University of Twente it was no problem at all that my thesis took more than the official time. The extra months I got allowed me to develop and improve my ideas, until I got to something that I believe is a novel approach to an actual problem with a result that I really support. Having been able to conclude my student days in such a way is tremendously satisfying and very valuable.

As a final remark: Thank you, Alan Turing, for being a founding father of computer science. Thank you, Socrates, for being a founding father of philosophy. Not thank you, postwar Britain and city-state of Athens respectively, for forcing your geniuses to commit suicide.

CONTENTS

1	INTRODUCTION	1
1.1	Problem statement	2
1.2	Research question	3
1.3	Structure of thesis	4
2	CHARACTERIZING CYBERSPACE	7
2.1	Promises of cyberspace	7
2.2	The problem of defining cyberspace	9
2.3	Characterizing cyberspace using design principles	11
2.3.1	A network of networks	11
2.3.2	Layers as nodes	13
2.3.3	The end-to-end principle	16
3	TRUST AS SECURITY	21
3.1	The need for security	22
3.1.1	Attributes of information security	22
3.1.2	Governing security in cyberspace	23
3.2	Securitization of cyberspace	25
3.3	Whose security?	28
3.4	The counter-productivity of trust as security	31
4	TRUST AND WELL-BEING	33
4.1	Functions of trust	34
4.1.1	The intrinsic value of trust	35
4.1.2	The instrumental value of trust	36
4.1.3	Virtuous and vicious circles	38
4.2	Bases of trust	41
4.2.1	The possibility of trust in cyberspace	41
4.2.2	Accounts of trust in cyberspace	43
4.3	The leap of faith	44
5	REGAINING TRUST IN CYBERSPACE	47
5.1	Recapitulation	47
5.2	Leaping from node to node	49
5.3	How leaping works	50
5.4	Security proper	53
6	CONCLUSION	57
6.1	Conclusion	57
6.2	Normative implications & future work	58
	BIBLIOGRAPHY	61

INTRODUCTION

The title I have given to my thesis is: *Trust trumps security*. This might seem a bit bold, but it really is the essence of the argument I present. In this thesis I am concerned with the fate of cyberspace and all the people who (want to) use it to enrich their lives. In the last decades cyberspace has pervaded our lives and changed the way our society works. I would contend that most of these developments are for the better, although at the same it seems that some of the magic that surrounded cyberspace in the early days with its promises to vastly change and improve the future evaporated. Perhaps such events as the dotcom-crash around the turn of the century or the barrage of articles about people, companies and governments being hacked that flood the news these days have served as a cyber reality check for many. Or maybe expectations were set too high all along and cyberspace is not as disruptive as people thought it. It is difficult to say and provide a conclusive explanation, and in addition to that we should not forget that we are still in the midst of the digital revolution. Many more innovations are predicted, such as 3D printing, wearables and the internet of things as major developments expected to impact our lives in the near future.

Even though future developments are notoriously difficult to predict, we can still try to steer them a bit and discuss what we do and what we do not consider to be good developments for our society. The argument I present in this thesis does just that. I will argue that *trust* is a condition sine qua non for cyberspace to improve people's lives. After having provided a conceptual analysis of both cyberspace and the phenomenon of trust I will introduce a conception of how trust works in cyberspace, how it enables users to benefit from what cyberspace has to offer to improve their lives. I will pay special attention to the relation between trust and security. As I will argue, they are not the same thing. Although cyber security receives much attention, it is trust that is key to improve peoples lives in cyberspace. Security is still important, but in a facilitating role to trust. This is why I make that claim in the title of my thesis: trust trumps security.

I will use the rest of this chapter to set the stages for my thesis. First I will elaborate the problem I am addressing in this thesis. In this problem statement I will set the scene in a more informal sense. In the second section I will be more formal and formulate the research question that I will try to answer with this thesis. In the last

section will I will explain how I structured this thesis in order to get the buildings blocks that I need to answer my research question.

1.1 PROBLEM STATEMENT

Taken to the extreme, the problem that is central to this thesis seems to take the form of a classic motif: Collapsing under the weight of your own success. The idea behind this motif is that something (and that something in this thesis would be cyberspace) is so successful, that it causes powers to be set in motion that have unexpected and undesired effects that ultimately undermine that something and causes it to collapse. Before being cast an alarmist: I do not think that the collapse of cyberspace is nigh. Far from it actually, I think, or at a minimum hope, that cyberspace is still in its infancy.

But I do believe that if we want to be able to enjoy the benefits of cyberspace for long times to come, we have to take precious care of it. Therefore, in this thesis, I want to raise attention to the possibility that because of its success and the ever-increasing role cyberspace plays in our lives, developments may occur that have the potential to undermine the bases of what made cyberspace so successful. However, if we identify such developments in time and address them properly, which means both from the perspective of causes of such developments as from the perspective of the long-term development of cyberspace, undermining can be prevented. On the contrary, it could even lead to reinforcement of cyberspace.

Obviously, I am not just talking talking about theoretical possibilities or hypothetical dangers. There are real developments going on in cyberspace that I consider to be problematic. Addressing those in the way I just described (so with respect to both sides) is what this thesis is about. The problem that I see is that because cyberspace is so successful, the stakes are getting higher in clashes of interests and other cyber conflicts, which could be the cause of effects that undermine the original success factors of cyberspace. The success factors of cyberspace are its open nature that has no bias towards any activity and the low barrier of entry. This has and still is enabling billions of people from around the globe to have access to information, be in touch with other cultures and like-minded people and be entrepreneurial in an unprecedented way. The cyber conflicts I refer to are a range of activities that take place in the world of organized cybercrime, state-sponsored hackers, hacktivists et cetera. Their activities and subsequent security countermeasures exploit or (potentially) undermine the success factors of cyberspace.

The problem is that such developments can not simply be undone. It is not an option to go back to the cyberspace of say twenty years ago. That would also nullify the large beneficial majority of cyberspace developments of the past twenty years, which does not seem desirable to me either. In addition to this, security concerns are real and valid concerns that must be addressed. For example, without secure environments people will be hesitant to trade online or maybe even exchange ideas online, and these are definitely desired activities in cyberspace. Also, the fact that the stakes are getting higher in cyberspace is unsurprisingly: the challenge lays in how to handle the ensuing clashes and conflicts.

As will be clear by now, my claim is that trust is a key concept in addressing the described problem. I will now turn to explaining how I aim to that, by formulating a research question and after that explaining the structure of this thesis.

1.2 RESEARCH QUESTION

Up until now I have informally described what the problem is that I am concerned with in this thesis. But that is not sufficient to start my analysis. In order to be able to do that I have formulated a research question. An answer to that question should automatically address the problem raised in the previous section. The research question that I will try to answer in this thesis is as follows:

RESEARCH QUESTION: Why, where, and how should trust be established in cyberspace?

It is clear that my research question in essence is an inquiry into the establishment of trust in cyberspace from three different perspectives, as represented by the three different interrogative words. In order to answer the research question, I will split it in the three subquestions that are entailed in it. Besides contributing to clarification, these subquestions (or rather the answers to them) correspond with three main contributions that I will present in my thesis. The subquestions are the following:

SUBQUESTION 1: Where should trust be established in cyberspace?

SUBQUESTION 2: Why should trust be established in cyberspace?

SUBQUESTION 3: How should trust be established in cyberspace?

The observing reader might have noticed that I changed the order of the subquestions in comparison with the order of the interrogative

words in the research question, by switching why and where. Not too much should be sought behind that. Aesthetically, the order in the research question is most appealing to me and seems like a logical order for answering the question as well, but for the flow of the thesis it made sense to answer the where-subquestion before the why-subquestion.

One concern that might be raised at this point is that I might overlook an elephant in the room, namely the question of what cyberspace is. This concern is not warranted though, for I actually dedicate a whole chapter to it. I made the decision however to not include it as a separate subquestion, because I believe that would divert attention from what this thesis is really about, namely trust in cyberspace. The question of what cyberspace is is indeed relevant for this, but more as a precondition. That is why I do address it, but not as a subquestion.

1.3 STRUCTURE OF THESIS

Before I will describe how I structured the contents of my thesis in chapters, I first want to create some structure by clarifying some of the terminology I will use throughout the thesis.

I use *trust question* as a term denoting the point when trust becomes an issue. In other words, the trust question is asked (implicitly) when a user has to judge whether or not he trusts (some part of) cyberspace. I will also speak of *affordances* of (trust in) cyberspace. An affordance is the opportunity that an object creates for an actor to achieve something. A doorknob affords a person to open a door, trust affords a person to for example empower their autonomy. Furthermore I will use the terms *prudential value* and *well-being*, that are closely related, but not exactly the same (Taylor, 2013, p.10). Both are about what is good for a person and therefore serve as explanans to why something should be strived after. Well-being is what someone has if his life is going well for him, so it expresses something absolute. Something has prudential value (which is a different kind of value than moral or aesthetic value) if it contributes to making someone's life go well, but it says nothing about the overall evaluation of someone's life. To illustrate this, giving a condemned person his favourite last meal can have prudential value for him, even though his level of well-being is low. Prudential value can be positive or negative, actual or potential and intrinsic or instrumental. Most of the time I will not explicate these modalities, but in such situations I will mean positive prudential value of some sort. Prudential value and well-being are connected, because they can be translated into each other. If something has positive prudential value for someone, than that something adds to (or

increases) the well-being of that someone. This also works the other way around and in a negative sense.

In this thesis I will often argue that something has prudential value or promotes well-being in cyberspace. This is used as an ultimate reason of why something is good and should be interpreted as being good for an (undefined) user of cyberspace. What already begs to be addressed now is the connection between on the one hand trust and on the other hand prudential value and well-being. The short answer is that the existence of trust has prudential value and thus promotes well-being. How this works will be explained in Chapter 4.

In total there are six chapters in this thesis, the first of which is this introduction. This will be followed by four chapters in which I will present my contributions and finally a concluding chapter where I will answer my research question. In Chapter 2 I will meet the aforementioned precondition of answering what I mean with cyberspace. I will begin with discussing the societal promises of cyberspace, followed by a characterization of cyberspace that I will base on three principles of technical network design, but that I will adapt to include more than only the technological side of cyberspace. This characterization conceives cyberspace as a decentralized network of many circuits that in turn are composed of different nodes. The next concept that I will discuss in Chapter 3 is security. This chapter will serve as a foundation for answering the where-subquestion. Here I will explain why security is not the same as trust and why confusing security for trust can have negative effects on cyberspace. This subquestion will actually be answered in a negative sense, by arguing where it should not be established. What this implies will become clear further along, when addressing the how-subquestion. But first, in Chapter 4, I will answer the why-subquestion through an extensive review of literature which also deals with definition questions regarding trust. Here the connection between trust and prudential value and well-being is made. At the end of this chapter I will introduce a notion of trust that revolves around a leap of faith, that will serve as a framework for the subsequent chapter. There, in Chapter 5, I will use the knowledge gained in the previous chapters to answer the how-subquestion. Once we know where the trust question arises in cyberspace and have a working conception of trust, I will fuse the two together in what I will call my conception of trust in cyberspace. The core idea of this conception is that trust questions are connected to the leaping from node to node. Finally, in the concluding Chapter 6 I will articulate an answer to the research question, describe the normative implications and suggest future work.

CHARACTERIZING CYBERSPACE

Before addressing any of the questions about the establishment of trust in cyberspace, I will dedicate this chapter to outlining what I mean with cyberspace. This is needed to prevent that question from turning attention away from answering the research question.

I will first turn to the promises of cyberspace. Especially techno-utopians, but also mainstream society conceived cyberspace as a liberating technology that would be of huge prudential value to mankind. The last decade however has shown that cyberspace has its dystopian sides as well. Fear of cybercrime and cyber warfare and power plays for control over cyberspace negatively affect trust in cyberspace, which has negative prudential value to users as well.

If we want to abide to well-being in cyberspace, I will argue, we will need to recognize the importance of trust. In Chapter 5 I will present my conception of how trust in cyberspace works. That too requires having a conception of what cyberspace is, next to earlier mentioned reason. Rather than providing a definition of cyberspace I will develop a characterization of it, that better caters to its evolving nature. This characterization will cover both the cognitive and technical aspects of cyberspace, although I will remain loyal to such fundamental notions of cyberspace as a network of networks, layers, and the end-to-end principle.

2.1 PROMISES OF CYBERSPACE

The advent of cyberspace is closely connected with techno-utopianism. When in the 1990s access to personal computers and connections to the Internet for households started to gain traction, this gave rise to a narrative that combined 1960s counterculture with techno-libertarianism. As [Turner \(2008\)](#) explains, the promises of cyberspace were that of an ideal society based on decentralization, egalitarianism, harmony and freedom. Such ideals are reflected in declarations and manifestos published and spread around that time. [Barlow \(1996\)](#) laments against governments and borders, emphasises the equality and freedom of expression that cyberspaces promotes, [May \(1992\)](#) praises that governments will have less power to regulate, impose taxes and control economic interactions, and [Hughes \(1993\)](#) stresses the importance of

cryptography to the open society in the electronic age.

A key ideological concept is that computers are seen as personal tools of liberation (Barnes, 2008, p.825). This is striking, for in the view of the 1960s counterculture, computers loomed as technologies of de-humanization, tools of the military-industrial complex. So there is not something that inherently connects computers to techno-libertarianism (Turner, 2008, p.3). That connection is part of what is called the Californian Ideology, for it was in there were this ideology “emerged from a bizarre fusion of the cultural bohemianism of San Francisco with the high-tech industries of Silicon Valley” (Barbrook and Cameron, 1996).

Despite the appeal that the Californian Ideology had (and continues to have) to many, it has also received a lot of criticism. Jacobs (2001) argues that in this ideology, the concept of utopia has been thoroughly degraded and commercialized. Already around the year 2000 scholars got the idea that cyberspace as an anarchic environment was living on borrowed time, due to the growing commercialization and concerns of anxious governments (Spinello, 2001, p.137). Although one of the premisses of the Californian Ideology was the absence of governments in cyberspace, Lessig (2006) began arguing that government regulation might become necessary to protect the affordances of cyberspace that have prudential value. Lessig identifies four modalities of regulation whose implication all need to be considered: code, law, markets and norms. All four can contribute to well-being in cyberspace, but can just as well undermine it.

The notion that governments have interests in cyberspace and develop cyber policies by now has become commonplace. However, activities by governments in cyberspace have not solely had the purpose of providing prudential value to the well-being of as many citizens as possible. Morozov (2012) has exposed the naiveté of cyber-utopians who thought cyberspace is inherently liberating. As became clear in the aftermath of the Arab Spring, authoritarian regimes used digital technologies for repression just as much as protesters used them for their agenda. Subsequently we learned from the classified documents leaked by Edward Snowden that its not only authoritarian governments that monitor their civilians in cyberspace through mass surveillance, but Western democratic governments as well (Greenwald, 2014).

It has become clear by now that cyberspace has not developed as envisaged by the cyber-utopians of the 1990s. If you support the arguments of Lawrence Lessig and others who feared for the privatization of cyberspace, there may even be upside to the fact that the ideas of the Californian Ideology have not come to full effect. But only as long as the promises of cyberspace to promote well-being

are not disregarded at the same time. The role that governments will play in subsequent developments of cyberspace seems crucial here. In Chapter 3 I will explain how even benign intended governmental activities in cyberspace can become counter-productive. However, that does not mean that I think governments should stay away from cyberspace. I believe governments do have a role, even a responsibility, in promoting human well-being in cyberspace. I think it is a modest role though, were governmental policies should be aimed at the promotion of trust in those parts of cyberspace that are under their influence. It should still be the individuals that use cyberspace that should decide how and for what they want to use it and governments should abstain from prescribing or nudging towards desired behaviour outside of the confines of what is conceived as criminal offline as well. In order to elaborate on this later, I will turn towards providing my characterization of cyberspace now.

2.2 THE PROBLEM OF DEFINING CYBERSPACE

Providing a satisfying definition of cyberspace is no sinecure¹. Military strategists and lawmakers who have been tasked with developing cyber policies have been struggling with it for years. [Singer and Friedman \(2014\)](#) provide the example of the Pentagon, that has issued at least twelve different definitions of what it thinks of as cyberspace over the years. Every time reasons came up to reject the prevailing definition. Sometimes because the definition was too narrow and did not include aspects that it should, other times because it was too wide and encompassed everything from computers and missiles to the light from the sun. At its essence, they proceed, cyberspace is the realm of computer networks (and the users behind them) in which information is stored, shared, and communicated online.

But rather than making the next attempt at finding perfectly worded definition of cyberspace, it is more useful to look at what exactly people try to capture in definitions of cyberspace. According to Singer and Friedman, cyberspace is an information environment handling digitized data in different ways (copying, moving, storing et cetera), but at the same time it is the computers and networking technologies on which this happens. So on the one hand, cyberspace is something virtual, but not exclusively, for on the other hand it has physical aspects. Furthermore, cyberspace is a man-made domain (in opposition to the other four domains of warfare: land, sea, air & space) and people behind their computers and in charge of managing the in-

¹ Previously I have written an essay covering the main idea of this section (that cyberspace can better be characterized or described than defined) for a university course. This section is a digested version of that essay.

frastructure are also an important part of cyberspace. As they put it: “cyberspace is defined as much by the cognitive realm as by the physical or digital” (Singer and Friedman, 2014, p.14).

Another important observation is that cyberspace is constantly evolving, more precisely in three different ways. Firstly in size and scale, secondly qua the technologies that form it (think about the billions of smartphones that currently are part of cyberspace, but did not exist ten years ago) and thirdly the politics surrounding cyberspace. It is this evolving on multiple axes that in my view clashes with the desire to define it. Cyberspace is far from definite and something definite is what a definition tries to capture. Alternatively, I suggest to use the word *characterize* instead. I oppose *defining*, on the basis that its connotation contains a certain element of permanence. By using ‘characterizing’ I want to make room for more dynamic approach in capturing the notion of cyberspace, which I think is more suitable to the phenomenon.

Singer and Friedman distinguished between the cognitive, physical and virtual realm, and each of these has its own ontology. Consider what a certain byte (as a unit of information) is in these ontologies. Cognitively, it might represent a letter I typed. Physically, it might be stored on a hard disk inside a computer somewhere, where it is stored by configuring a magnetic field in a certain way. And virtually, ‘inside the computer’, it is a set of 8 bits that are stored at a certain data address. None of these alone fully addresses what that byte constitutes in cyberspace. Cyberspace is, in the terminology of Pickering (2001), at the zone of intersection of these realms.

Instead of sticking to the ontology of one fixed realm, I think Andrew Pickering’s notion of an ontology of becoming (Pickering, 2008) is more suitable for characterizing cyberspace and its evolving constitution. Whatever cyberspace’s form or shape is at this moment, it is not permanent so. New technologies will change what we do in cyberspace, our changing behaviour in turn may spur the development or change the usage of other technologies. This is the way cyberspace has evolved so far and most likely will continue to evolve. This is what Pickering calls a dance of agency between humans and nonhumans (Pickering, 2001, p.6).

Having said this it is time to turn towards my characterization of cyberspace. I want to base this on three main principles that are normally used to describe the technological design of the Internet. In the next section I will explain how I want to expand these principles to cover the whole of cyberspace.

2.3 CHARACTERIZING CYBERSPACE USING DESIGN PRINCIPLES

For my characterization of cyberspace, I will stick to three principles that are used in the technological design of the Internet, but broaden their scope to include the cognitive aspects of cyberspace as well. The first principle is to describe cyberspace as *a network of networks*, the second is the concept of *layers* and the third principle is the *end-to-end principle* that prescribes to keep complexity at the edges of the network. I will now turn to introducing each principle and right away explain how I adapt it to my characterization of cyberspace.

2.3.1 *A network of networks*

The characterization of the Internet as a network of networks stems from its technological architecture. It can be found formulated like this in RFC 1122² by Braden (1989). What it aims to capture is that the Internet is not one big network, but an interconnection between many publicly or privately operated networks mediated by TCP/IP protocols. Such networks can be governmental networks, regional networks, institutional networks and so forth, that are connected at internet exchanges such as AMS-IX in Amsterdam. This exchanging of data between different providers is called peering and explains the *inter*-part in Internet. This is a very rudimentary description of the internet, Krol and Hoffman (1993) already mentioned over two decades ago that the Internet quickly became much more complex, for example through the interfacing with non-IP networks (that are based on other protocols), but the principle of the Internet as a network of networks remains.

For my characterization of cyberspace, I want to take this notion and extend it to include users, or the cognitive realm in the terminology of Singer and Friedman, too. So cyberspace is much more than the technical infrastructure and the devices it connects. It also includes the uses to which it is put (DiMaggio et al., 2001, p.308) and the users who are connected to it. Therefore, viewing cyberspace as a network of networks is not limited to a technological interpretation of networks. People and their interpersonal relations can be viewed as networks as well. Hannerz (1992) applies the concept of a network of networks to the global ecumene in an article about the development of meaning in the context of globalization. According to him, "it [are] dispersed institutions and communities, groupings of people regularly coming together and moving apart, short-term relationships or patterns of fleeting encounter, which offer the contexts in which glob-

² RFC stands for Request For Comments. They are publications that describe the technical developments and standards for protocols used to run the Internet. They are maintained by the Internet Engineering Task Force.

alization occurs as the personal experience of a great many people in networks where extremely varied meanings flow" (Hannerz, 1992, p.46).

This is very relevant for a characterization of cyberspace, because the world-spanning reach of cyberspace makes it ecumenic. I think the description above is very applicable to what happens in cyberspace, where users from around the world with different backgrounds connect. Hannerz argues that a network of networks is a good metaphor to describe in a "reasonably orderly way (without necessarily aiming at rigour of measurement) about some of the heterogeneous sets of often long-distance relationships which organize culture in the world now—in terms of cumulative change or enduring diversity." (Hannerz, 1992, p.51). Multiple times in the article he pays attention to viewing the global ecumene not just a singular network, but a network of networks. This is to allow his theory to take a less totalistic, more pluralistic direction, open to decentralized, mutable ideas of structure than that a metaphor of a singular network would allow.

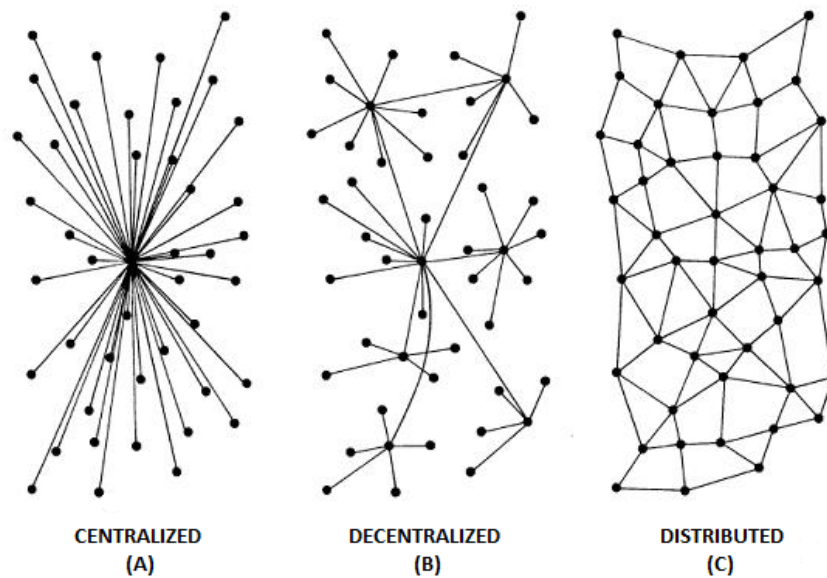


Figure 1: Three different network topologies.

There are different ways that networks can be structured. Following Galloway (2004) I will distinguish between three topologies: centralized, decentralized and distributed networks. Their structures are respectively illustrated in Figure 1. Centralized networks are hierarchical in structure. Each hierarchical level has one central node that serves as a hub that connects all others. Those other nodes in turn can be the central node of a lower hierarchical level, but power is always wielded from top over bottom. An example of a centralized network is the judicial system. Decentralized networks have hubs too,

but those interconnected without a central point that exercises control over all others. The airline system is an example of this, with minor airports with limited destinations and major airports that serve as a hub, but without one central airport that must necessarily be passed on a trip. Distributed networks finally move away from central coordination and vertical hierarchies towards autonomous nodes that can be connected to any other node, without any necessary intermediary hubs. This topology is the ideal topology of the Internet. However, as network of networks already suggests, some hierarchy remains in its design with internet exchanges such as AMS-IX as hubs. Since these hubs are interconnected but without one central point between them, the structure that best describes the topology is that of a decentralized network.

To summarize, Hannerz describes the global ecumene as a grand network connecting many sub networks of people with extremely varied meanings. It is such a metaphor of a network of networks that I want to use to characterize cyberspace. It has a decentralized structure, for the grand network connects many sub networks. For clarification, I will refer to those sub networks as *circuits* from now on, for reasons that will become clear soon. First now I want pay attention to a crucial component of every network: the nodes it is composed of.

2.3.2 *Layers as nodes*

Nodes are the connection points, redistribution points and endpoints that are connected to each other through a network. A connection between two endpoints in a network usually goes via (or ‘hops over’) a range of nodes that are interconnected. Primarily I see these nodes as parts of the circuits, but indirectly that makes them nodes of the grand network (e.g. the whole cyberspace) too.

For my definition of the nodes of cyberspace I again want to revert back to a technical concept from the Internet. This time to the concept of layers, a fundamental architectural feature of the Internet. In this architecture the design of computer networks and the way they communicate is modelled using different protocol layers that are stacked on top of each other. Data flows through the layers, each of which has its own role in the communication process. For more information see for example [Kurose and Ross \(2008\)](#) who describe the five-layer Internet protocol stack consisting of the application, transport, network, link & physical layers.

Just like with a network of networks, I am not the first to expand the notion of layers to outside the technical domain. Multiple au-

thors have used a conception of layeredness in articles on internet governance. Benkler (2000) makes a distinction between “three levels of the information environment”, firstly the physical infrastructure layer that includes wires, cables and the radio spectrum, secondly the logical infrastructure layer that covers software programs and applications, and finally the content layer that represents the data and activities of users. This distinction in three layers is the point of departure for Lessig (2001), who discusses the commons (publicly held goods) in cyberspace. He argues that in communication systems each of these three layers (although he uses the term code layer instead of logical layer) are either privately controlled or free (in which case the layer is a common). The Speakers’ Corner in Hyde Park in London is an example of a communication system where all layers are free, cable television is an example of the opposite where all layers are controlled. With regards to cyberspace, Lessig warns for developments that aim to exercise control over all layers in cyberspace, whilst the innovative powers of cyberspace result from layers being free — especially the code and content layer.

Solum and Chung (2003) take this notion further and develop the *layers principle* that in short amounts to the maxim that Internet regulators should respect the integrity of the layers. By this they mean that if they want to regulate something that affects a certain layer, then that regulation should also be a measure that takes place in that layer. For example, if a country wants to block certain content in cyberspace, they should not seek to achieve that by severing the physical connections, because content and physical are different layers. Subsequently, Solum and Chung argue that layer-violating regulations are ineffective and are hurtful for innovation in cyberspace. A similar case is made by Whitt (2004) who argues that instead of imposing outdated policy frameworks on cyberspace, regulators and legislators should adapt their policies to the constitution of cyberspace. That means adapting a layers-based framework. Whitt proposes a model framework of four layers that is very similar to (and inspired by) the three layers of Benkler. He maintains the physical layer and content layer as described by Benkler, but splits an applications layer off of the logical layers. The logical layer then covers the software that runs in the back to make networking possible, the application layer covers the software the user interacts with.

For the adaptation of layers as nodes of the networks I want to stick to the tripartite made by Benkler. I do think a distinction between a logical layer and an application layer has practical use, but it is not fundamental. This actually holds for other models of layers too. Where they distinguish more layers it is basically a subdivision of the logical layer. This can be illustrated using the model proposed

by Solum and Chung that consists of six layers, effectively those of Kurose and Ross's Internet protocol stack with Benkler's content layer stacked on top (Solum and Chung, 2003, p.3). Translated to Benkler's model, this top content layer is of course its own layer. Similarly, the bottom layer of the Internet protocol stack is the physical layer, which for obvious reasons identifies with Benkler's physical layer. The four layers that are left – application, transport, network & link – all still have their own distinct roles, but they are all logical — software-based that is. When translating layers to nodes it is not necessary any more to call them separate layers. They are just different nodes with different roles, albeit of the same category.

The three layers of Benkler are transformed into three categories of nodes. This means that a node can either be a content node, a logical node or a physical node. These three categories identify with the three realms of cyberspace mentioned before by Singer and Friedman. Content nodes represent the cognitive realm, so the users of cyberspace and all their digital representations inside it, identities and manifestations such as all sorts of user accounts, e-mail addresses and video game avatars. Logical nodes represent the virtual realm and as I have already explained there is a wide ranges of instantiations of such nodes. This includes social media sites and office tools, but also the firmware of the GSM chip inside a smartphone or the user interface of an industrial control system. Lastly, physical nodes represent the physical realm, covering such nodes as glass fibre cables and mobile 4G networks.

In models of layers a layer receives data, transforms it in some way and then passes it on to the next layer. Similarly, nodes are the points in a network that exchange data with adjacent nodes and (in a very broad sense) perform an operation on the data. Whilst layers only interact with other layers directly on top or under them, nodes can be connected to a theoretically infinite amount of other nodes, which allows them to distribute data amongst them. Moreover, connections between nodes are flexible. New connections can be made and older ones can be detached. Such connections can be diverse in nature and both inter and intra node-categories. Entering text in a command-line interpreter is an example of a connection between a user (a content node) and some computer program (a logical node), A java-program that makes use of the java virtual machine is an example of a connection between two logical nodes and the Ethernet standard is an example how a logical node connects to a physical node. That nodes perform an operation on data should be interpreted in a broad sense, example operations are translation between protocols, the guaranteeing of properties such as authentication, but also the merely passing on of data. So the transferring of data by a coaxial cable counts as

an operation as well. In this sense it can be said that nodes have an inside and an outside. The inside is where the performing of one or more operation takes place, the outside is composed of the fringes at which it is connected to other nodes.

I started this subsection by saying that nodes are the connection points, redistribution points and endpoints of a network. The core of the network is where many nodes connect and data is distributed, fulfilling the hub function in decentralized networks. Most of this is done to move data from one endpoint to another. Usually endpoints are users, such as a group of people who are teleconferencing, but certain computer systems (especially when artificial intelligence technologies become more advanced in the future) could be considered endpoints as well. Whereas the hub function is part of the core of cyberspace, these endpoints form the edges of cyberspace. The last technical design principle I want include in my characterization of cyberspace as well is concerned with these endpoints or edges of the network and will be introduced now.

2.3.3 *The end-to-end principle*

The last principle I want to introduce is called the *end-to-end principle*. It states that in a network complex operation should be performed at the edge of the network. In other words, it is the endpoints where (through policy or technology) complex though desired properties or functionalities of that network or specific connections should be implemented. The core of the network should remain stupid in that sense and mostly concerned with merely connecting nodes and distributing data. In other words, it should remain agnostic to who uses the network for what. In short, the end-to-end principle can be summarized as follows: smart applications, stupid network.

As we have gotten used to now, this principle has a technical origin. [Saltzer et al. \(1984\)](#) argue that implementing functions such as bit-error recovery, security using encryption, duplicate message suppression, recovery from system crashes and delivery acknowledgment yield little value compared to their cost. Such costs can be in performance, but also ease of use or interoperability. The canonical example is TCP/IP, where IP is a dumb protocol for moving datagrams between networks and TCP is protocol on top of it that is responsible for setting up and maintaining dependable connections between hosts.

This principle has been taken by Lawrence Lessig who considers it “one of the most important reasons that the Internet produced the in-

novation and growth that it has enjoyed” (Lessig, 2006, p.44). In terms of this thesis that means that it is of great prudential value. Lessig proceeds to list three important consequences the principle has for innovation, thereby showing why this principle is important beyond its technical rationale and consequences. Firstly, new innovators can easily connect their innovation to the existing network, which itself does not need to be altered. Secondly, because the network is not optimized for any particular existing application, the network is open to innovation not originally imagined. Thirdly, since the network is a neutral platform, it does not discriminate against certain data in favour of other, which works against domination by incumbent applications (this is what discussion about *netneutrality* are about) (Lessig, 2001, p.37).

In their article on the layers principle, Solum and Chung praise the end-to-end principle, but contend that it actually is an articulation and abstraction of implicit ideas inherent in the layers model. It is their claim that “the normative content of the layers principle is a superset of the normative content of the end-to-end principle” (Solum and Chung, 2003, p.7), for it provides “guidance for regulators where the end-to-end principle is silent or indeterminate”. However, since I moved from layers to nodes, the layers principle and its interdiction on the violation of the separation between layers is not applicable. Therefore I will remain with the end-to-end principle, for it is more suitable to my characterization of cyberspace as a network of networks with nodes of which some are endpoints. As I will explain in more detail later, the end-to-end principle is not as much about the constitution of cyberspace as it is about how that constitution promotes well-being in cyberspace.

Now to bring these three design principles together, I suggest to characterize *cyberspace as a decentralized network of networks consisting of nodes that are either of a content, logical or physical nature, where complex particular functions should be implemented near the endpoints.*

With a network of networks I mean that cyberspace is the union of myriads of sub networks that I call circuits. The term circuit incorporates the idea of a closed network with a starting point that is also the finish. These circuits are how individual users experience cyberspace, this individual experiencing makes a circuit closed from others. For each user cyberspace has a different meaning, an amalgamate of the different ways a user experiences it, or in other words, the use cases it offers, such as mobile banking, contacting friends or gaining knowledge. Cyberspace does not necessarily have one meaning at a time and with the evolution of cyberspace the meanings can evolve too. This is because it has many parallel use cases and those change. Each

circuit represents a use case and since many use cases (such as following the news) are shared by users these circuits overlap. This brings us back to cyberspace as a whole, the grand network that spans all circuits. This characterization of cyberspace has several benefits. For once, through showing that in spite of it being problematic to define cyberspace, it is still possible to talk about cyberspace in general, because of all the overlap in the meaning it has for users. Secondly, because it leaves room for future developments of cyberspace, thereby accounting for its evolutionary nature. And thirdly because this intentional non-committing allows users to let cyberspace be what they want it to be, which is best for their well-being.

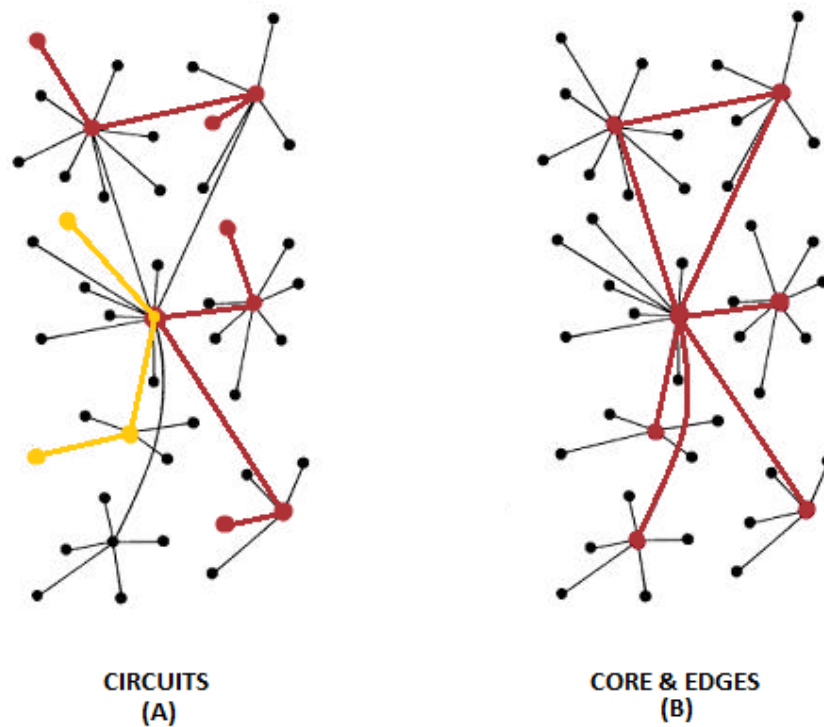


Figure 2: Circuits and the distinction between core and edges.

If users execute a certain use case, a connection is made through the corresponding circuit between the user and his destination. This connection goes through a set of nodes, that as I have explained are of a content, logical or physical nature. Nodes are not inherently connected to one specific circuit. They can be a part of many circuits connected to different users. Big ocean spanning fibre glass cables are a clear example of this, for they form nodes in millions if not billions connections every day. Previously I have mentioned that circuits overlap, the sharing of one or multiple nodes is how this happens and it is visualized in part A in Figure 2. In the figure dots represent nodes and lines between nodes mean that those nodes are connected. Three example circuits are highlighted and it can be seen how circuits over-

lap through the sharing of nodes.

The end-to-end principle in cyberspace, at last, prescribes that measures that are taken to stimulate certain behaviour or ensure certain properties ought to be implemented at the edges of networks close to endpoints. This is desirable, because it allows user to implement the functions they wish for in their usages of cyberspace, which contributes to their well-being, but does not force those upon other users, thereby not affecting their well-being. When the core of the network — those nodes that are part of many circuits, are left unaffected by particular needs, so stays agnostic, they remain useful to as many users as possible which maximizes its utility towards general well-being in cyberspace. The difference between the core and edges of cyberspace are illustrated in part B of Figure 2 where I highlighted the core.

— ∞ —

In this chapter I provided my characterization of cyberspace that is maximally attuned to promote well-being. But this conception alone will not make people use it. They must still trust cyberspace before its potential positive prudential value can be realised. In the next chapter I will turn to analysing the role of security in cyberspace, which is often confounded with trust, but as I will show actually is not the same.

TRUST AS SECURITY

Now that my characterization of cyberspace is established, it is time to turn to addressing the research question. In this chapter I will do that by addressing the where-subquestion:

Where should trust be established in cyberspace?

The challenge is to determine where in the constitution of cyberspace trust can be made possible. A common response to this challenge is to come up with ways to ensure that trust is present. The rationale behind this is that (with well-being as ultimate goal in the back of our mind) trust in cyberspace must be secured. Actors then try to achieve this through a mixture of security technologies and policies. The idea is that if these technologies and policies are implemented and executed correctly, a certain set of properties (such as authentication and identification) can be ensured. This rules out undesired behaviour, making cyberspace trustworthy to people and hence promotes well-being. Since this conception of trust heavily depends on the security technologies and policies that are put in place to ensure trust, [Nissenbaum \(2001\)](#) has called it a vision of *trust as security*.

At first sight this line of reasoning appears logical and it contains many good points, but it contains many pitfalls too and ultimately proves counter-productive to well-being in cyberspace. Still, because it is such a dominant conception of trust and contains some useful parts, I will spend this chapter to the analysis of trust as security. Firstly, because it helps understanding what is partially the status quo and partially an undesired (to those in favour of promoting well-being) realistic future scenario of the development of cyberspace. Secondly, because it provides something to juxtapose the alternative conception of trust I will present later to. In that conception there is still a big role for security technologies and policies, but in a more modest way.

This chapter is made up of four sections through which I will provide my full analysis of trust as security. The first section is about why there is a need for security in cyberspace and what the role of governments is in that. In the next section I will introduce the securitization of cyberspace, which is at the core of why trust as security is bad for well-being in cyberspace. The third section is about conflicting security interests in cyberspace that ultimately make that trust as

security is an insufficient conception of trust in cyberspace. Finally, in the fourth section I will combine the insights from the previous sections and argue why trust as security is counter-productive to well-being in cyberspace. In addition, I will point out what lessons we can learn from this for my alternative conception of trust.

3.1 THE NEED FOR SECURITY

This section is split in two parts. In the first part I will describe how security is commonly conceived in cyberspace and how this relates to trust. This is from a perspective of information system design. The second part is concerned with the role that governments play with regards to security in cyberspace.

3.1.1 *Attributes of information security*

The type of security we talk about in the context of cyberspace is called information security. Usually it is defined as the composite of a set of attributes that provide certain properties to a system. Classic is the CIA-triad, which distinguishes three attributes important in information security: Confidentiality, integrity & availability. For an information system to be secure its properties must meet “the combination of confidentiality, the prevention of the unauthorized disclosure of information, integrity, the prevention of the unauthorized amendment or deletion of information, and availability, the prevention of the unauthorized withholding of information” (Pfleeger, 2000). Online banking is a good example to illustrate this. Confidentiality is required because banking involves sensitive personal data, integrity is needed because both the bank and the user must be certain that they are in fact talking to the other (so only the right users can transfer money from the right accounts) and of course the online banking system must be available for usage. If these three properties are met, online banking offers more prudential value to users than traditional banking, which imposes limitations in both time and location through openings hours and limited fixed physical presences respectively.

Information security is not limited to these three attributes though. Multiple extensions to the CIA-triad have been proposed. For example, Cherdantseva and Hilton (2013) developed an information assurance and security octave which next to confidentiality, integrity and availability includes accountability, auditability, authenticity, non-repudiation and privacy. My point is not here to debate what properties combine to a proper definition of information security or if some should be removed or added. It is merely to show that it is through

such attributes that information systems can have that security is conceived in cyberspace. The attributes provide certain assurances that make all sorts of interactions in cyberspace possible and of added value. This way, security can promote well-being. How security is linked to trust through breaking it down in attributes is illustrated by [Avižienis et al. \(2004\)](#). They link security to dependability, through their sharing of attributes, and trust in their conception is defined (without further arguing why) as *accepted dependence*. This shows how from a technical perspective it makes sense to conceive trust as security: they both depend on a set of attributes whose properties must be met in order for the information system to be used for its purpose.

3.1.2 *Governing security in cyberspace*

At least in western society, but actually in most places of the world, governments are responsible for the security of their citizens. Providing security, against both fellow citizens and foreign threats, can be considered the core duty of a state. This relation between state and citizen is known as the social contract in political theory. Historically, in the writings of Thomas Hobbes and John Locke, the social contract describes the giving up of some of one's natural rights to a sovereign government in exchange for security offered by that sovereign government. John Rawls revived the concept in the 1970s and made the principles of justice and the design of basic social institutions part of the social contract ([Cudd, 2013](#)).

There is much more political philosophy to the social contract, for example whether it is a justification or explanation of the role of governments in the security of their citizens, but I will leave that discussion for what it is. The reason I invoked the concept is because it relates to the question of with whom in the trust as security discourse the responsibilities for security in cyberspace lay. As I explained already in Chapter 2, the Californian Ideology is very libertarian and sought to diminish if not abolish the influence of governments in cyberspace. I also mentioned [Lessig \(2006\)](#) who later argued that government regulation might be needed to protect the affordances of cyberspace. A further argument is made by [Taddeo \(2014\)](#) who argues that authorities actually have a duty to positively promote security in cyberspace. She claims that citizens have a claim-right towards a secure cyber-sphere (the environment in which their online persona interacts, this seems to be close to what I would call the union of the circuits of a user or that individual's experience of cyberspace). A claim-right is a right that exists as the duty of another than the claim owner to perform a given action. In casu the claim-right exists in the duty of a government towards providing a secure cyber-sphere

to its citizens. Taddeo grounds this claim-right in three functionings (access to information, shaping of personal narrative and enjoying a high level of connectivity) that the online persona has according to her. The capability to achieve such functionings amounts to individual well-being in this theory. However, Taddeo continues to argue, in response to the claim-right of citizens, governments have the right to interfere (within regulated boundaries) in cyberspace. This is a cyberspace version of the social contract.

Whether one follows and agrees with Taddeo's argument, endorses Lessig's point or from a realist perspective accepts the status quo: governments are involved with security in cyberspace and are active and powerful presences. In his *National Cyber Security Framework Manual Klimburg (2012)* distinguished five different mandates of national cyber security. These are five different issues that could be addressed by different government departments. The mandates are the following:

1. Military Cyber
2. Counter Cyber Crime
3. Intelligence and Counter-Intelligence
4. Critical Infrastructure Protection & National Crisis Management
5. 'Cyber Diplomacy' & Internet Governance.

I am mostly concerned the third mandate, but will first shortly clarify the other four. The first, military cyber, is commonly referred to as cyber warfare and has received much attention in recent years in both media and academia. Although it is a very actual subject where new developments follow in rapid succession, consensus seems to be forming among experts that a real cyber war has not taken place yet. Famous incidents such as the Stuxnet worm or the denial of service attacks on Estonia in 2007 that are sometimes called acts of cyber warfare are better grouped under the third mandate. The second mandate, counter cyber crime, is concerned with criminal activities in cyberspace that impact both citizens and businesses. A distinction can be made between two types of cyber crime. Firstly, cyber crime in a wide sense, which includes traditional criminal activities with a digital component, such as swindling on online market places. Secondly, cyber crime in a narrow sense, where the criminal schemes depend on properties or activities specific to cyberspace. The fourth mandate is concerned with the physical protection of infrastructure that is critical to the functioning of society, such as public utilities, finance or telecommunications. Lastly, the fifth mandate of 'Cyber Diplomacy' & Internet Governance is about interstate efforts such as the development of norms and standards for cyber behaviour and confidence building measures between nations in cyberspace. An example of how this takes shape is the Global Conference on Cyberspace 2015

that was held in The Hague on the 16th and 17th of April 2015¹.

As I said, I want to focus on the third mandate, which is about intelligence and counter-intelligence. This is because it is within this mandate where most state activities related to cyber security that have been in the news the past years can be positioned. Intelligence and counter-intelligence covers a level of cyber-conflict below warfare and above crime, although on both sides the borders can be blurry. The main motives of actors within this mandate are (geo)political power and influence, and the interest of different actors can clash. This also includes such activities as theft of intellectual property, especially when it concerns advanced or trade regulated technologies. From the side of states it are intelligence agencies and secret services that are mandated to operate in this area. Other typical actors are hacktivists, (sometimes state-sponsored) hacker collectives and other non-state actors with certain agendas. Because many operations are of a secretive nature and in general attribution is difficult in cyberspace, it is often difficult to irrefutably determine who is behind certain activities. However, through publicly published reports and leaked information such as the Snowden files, it has become clear that many governments invest heavily in offensive cyber capabilities for their intelligence and secret services in order to strengthen their grip on cyberspace.

This observation will serve as a point of departure for the next section on the securitization of cyberspace. An arms race in the development of cyber capabilities between states endangers the positive affordances to well-being in cyberspace, even if those states' initial motivation is the protection of its citizens.

3.2 SECURITIZATION OF CYBERSPACE

In the previous section I have established that in cyberspace the interest of different actors from around the globe clash. The academic discipline pre-eminently concerned with such clashes is the field of international relations. Traditionally, there have been three competing paradigms in the study of international relations: realism, liberalism and constructivism. The main theoretical proposition of realism is that self-interested states constantly compete for power or security. Liberalism in contrast upholds the idea that the concern for power is overridden by economical and political considerations. In their desire for prosperity, the interdependence of states prevents conflicts between them. Realism and liberalism have in common that the main units of their analyses are states. This is different with constructivism, that focuses on individuals and especially elites. The main theoretic-

¹ For more information see: <https://www.gccs2015.com/>

cal proposition of constructivism is that state behaviour is shaped by elite beliefs, collective norms and social identities (Walt, 1998, p.38).

Barry Buzan in his book *People, States, and Fear* argued that security as a concept equals in rank with power (as the focus point of realism) and peace (as the focus point of liberalism). His point was that up until then security, if addressed at all, had a heavy military emphasis (Buzan, 1983, p.3). This is what is called a traditionalist conceptualization of security. Buzan objects to such a narrowly focused notion of security, amongst other reasons because it can be abused “as a justification for actions and policies which would otherwise have to be explained”, which makes it “a political tool of immense convenience for a large variety of sectional interests” to political and military elites (Buzan, 1983, p.9). He argues for a more holistic approach of security analysis in a wider sense. Since, as I have shown in the previous section, cyber-conflicts and accompanying security concerns in cyberspace also manifest in a wider sense than (in Klimburg’s words) military cyber, I will follow Buzan’s way of analysing security.

The school of thought that Buzan co-developed and belongs to is called the Copenhagen School. Because the Copenhagen School treats security as the outcome of a specific social process it can be considered a constructivist theory, although it has realist roots too (Williams, 2003). The key concept in the analysis of the Copenhagen School is called *securitization* — this is the specific social process I just mentioned. The seminal work in this field, *Security: A New Framework For Analysis* (Buzan et al., 1998), describes the process of securitization. The central idea is that security is not a subjective or objective condition, but an intersubjective activity. They invoke speech act theory, arguing that when an actor talks about security in the context of a certain referent object, it does something to that object, it securitizes that object. At least, the actor tries to securitize that object to a certain audience. Only if the audience is persuaded to accept the issue as a security threat we can talk about successful securitization. The result of securitization is the effect Buzan objected to, that the allotting of disproportional attention and resources to solving the securitized problem is justified. Additionally, addressing the problem is put in the hand of experts only. Buzan et al. (1998) define a spectrum of how public issues are addressed, from non-politicized to politicized to securitized. In the first case an issue is not considered a state issue and is addressed through technical or consensual means. In the second case the issue is considered part of public policy and through political debate about resource allocation the issue is addressed. As I said, in the third case due to a perceived threat-urgency the issue is left to experts and scrutiny by media and politicians is considered taboo. This is what enables the disproportionate allocation of resources and

legitimises the application of extraordinary means.

The Copenhagen School treats security in a wider sense and in *Security: A New Framework For Analysis* five sectors in which securitization took place were identified: military, political, economic, society & environment. For each of these sectors they explained how the securitization discourse works. At the time of writing, cyberspace and cyber-conflicts were not as prominent as these days, but by now cyberspace has been securitized too. This has been clearly explained in an article by Lene Hansen and Helen Nissenbaum. They argue that through this discourse, security in cyberspace has changed from a technical issue (which they call computer security) to a societal issue (which they call cyber security) (Hansen and Nissenbaum, 2009, p.1160). In their terminology cyber security equals to computer security plus securitization. Subsequently, they identify three modalities in which cyber security manifests (Hansen and Nissenbaum, 2009, p.1163-p.1168). The first modality they call hypersecuritization, which hinges on multidimensional cyber disaster scenario's that predict massive catastrophes ("electronic Pearl Harbor's") even though there is no precedent whatsoever that justifies this. Secondly, they identify cyber securitizations of everyday life that mobilize individuals by emphasizing cyber threats to their daily lives. This modality makes extensive use of medical metaphors, such as viruses. The third modality they call technification. This discourse is focused on the privileged role of technical experts in cyberspace, by stressing how difficult security in cyberspace is and arguing that it surely should not be dealt with by non-experts. This clogs what the actual level of threats is and facilitates the coupling of such threats with technoutopian solutions, as if they are neutral technical solutions.

The securitization of cyberspace is what thrives the contemporary cyber security arms race and has, as Harris (2014) calls it, led to the rise of a military-internet complex, the cyber equivalent of the cold war military-industrial complex. However, as is part of the Copenhagen School argument, securitization and (potential) accompanying arms races have negative effects and might not even lead to proper security. Therefore, in the long run desecuritization, the moving back in the realm of the politicized (e.g. the ordinary public sphere) of security issues is preferred (Buzan et al., 1998, p.29). In the case of the securitization of cyberspace, the core problem is what in international relation is called *the security dilemma*, which I will elaborate upon in the next section.

3.3 WHOSE SECURITY?

If somebody would ask if it is important to secure cyberspace it sounds like a no-brainer: of course this is important. Securitization thrives on this reflex, if only because the opposite (no security) sounds like an even worse alternative. But what such a question leaves in the middle is *whose* security we are talking about. Increasing security is not some neutral operation that would instantly benefit all actors involved to the same extent. Even more so, it could even have negative impact on the security of some actors. For example, if I were to acquire a gun, this could be good for my security, but bad for the security of people around me. This is what is called *the security dilemma*, which one of its main promoters Robert Jervis summarizes as follows: “many of the means by which a state tries to increase its security decrease the security of others” (Jervis, 1978, p.169). The security dilemma has its origins in the realist school of international relations, but Buzan (1983) invokes the concept too and as I have said the Copenhagen School has realist roots as well.

	OFFENSE HAS THE ADVANTAGE	DEFENSE HAS THE ADVANTAGE
OFFENSIVE POSTURE NOT DISTINGUISHABLE FROM DEFENSIVE ONE	<p>1</p> <p>Doubly dangerous</p>	<p>2</p> <p>Security dilemma, but security requirements may be compatible.</p>
OFFENSIVE POSTURE DISTINGUISHABLE FROM DEFENSIVE ONE	<p>3</p> <p>No security dilemma, but aggression possible. Status-quo states can follow different policy than aggressors. Warning given.</p>	<p>4</p> <p>Doubly stable</p>

Figure 3: Four scenario's in offense-defense theory by Jervis (1978).

Crucial to the security dilemma is the balance between offensive and defensive measures. Jervis developed the so called *offense-defense theory*, which looks at the relation between offensive and defensive measures from two perspectives: Firstly, if the two are distinguishable from each other, and secondly if one has an advantage over the other (Glaser and Kaufmann, 1998, p.47). These two perspectives combined leave for four scenario's to which the security dilemma can unfold as illustrated by Figure 3.

The worst combination is if offense has the advantage and offensive posture is not distinguishable from defensive, a scenario that Jervis called *doubly dangerous*. In this scenario arms races are likely and incentives to strike first are high. In the second scenario defense has

the advantage but offensive and defensive posture are still indistinguishable. The security dilemma is still present and intense, though less explosive, because pre-emptive action is less attractive. If offense has the advantage, but offensive and defensive measures are distinguishable, which is the third scenario, the dilemma is less intense, although there are still security threats. This is because the dominance of offense might seem seductive to some actors who feel like there is something to gain for them. The last scenario, where defense has the advantage and offensive posture is distinguishable from defensive posture, is the opposite of the first and therefore called *doubly safe* (Jervis, 1978, p.211-214). This final scenario permits a way out of the security dilemma. There would be no incentives for a first strike and the intentions (offensive or defensive) of an actor who takes security measures would be clearly signalled. Hence, this scenario would be worthwhile to strive for.

The security dilemma as a concept stems from the Cold War, but it is very much applicable to cyberspace and helps understand current developments in cyber security. We can speak of a cyber security dilemma “when efforts by one state to enhance the security of its digital infrastructure, either through the development of offensive or defensive cyberwarfare capabilities, decrease the cybersecurity of others” (Rueter, 2011, p.35). Unfortunately, it seems that at present offensive activities trump defensive ones. Nicholas Rueter names three main reasons for this: Cyber attacks can be carried out almost instantaneously; cyber warfare has extremely low costs of entry with potentially high returns on investment; and it provides a means of enhancing security and coercing others without causing loss of life to either side (Rueter, 2011, p.38). Furthermore, distinction between offensive and defensive measures is a hard challenge in cyberspace. Not per se because viruses are difficult to distinguish from firewalls, but because it is difficult to determine what are cyber security activities in the first place. In addition, the actors involved (from the side of states) often have both offensive and defensive responsibilities, and as we have seen operate on the level of intelligence and counter-intelligence, which adds to the opaqueness and makes distinction even harder.

The combination of dominance of offense over defense in cyberspace and it being hard to distinguish between offensive and defensive measures, suggests that the securitization of cyberspace leads the cyber security dilemma to unfold into the doubly dangerous scenario. There are no winners in such a scenario, at most actors that have an upper hand over others for a short while, before being caught up with in a cyber arms race. No nett security is being gained in such a situation, despite all the money and effort put in it (Dunn Cavelty, 2014, p.710).

This on its own is already a reason why trust cannot be secured in cyberspace and why trust as security is a flawed conception of trust.

But there is more to it. So far I have only reasoned from the perspective of state security and its desire to secure its digital infrastructure. However, as I argued in Chapter 2, cyberspace also has a cognitive side (e.g. its users), next to its technological and virtual constitutions. The security of users is at stake too and if not acknowledged will suffer under the securitization of cyberspace by nation states. Security of users means that they can use cyberspace the way they wish to use it in a secure way. This amounts to the protection of their civil liberties and rights in cyberspace. Users need not only be protected against cyber-criminals or foreign powers (which was the primary task of a state according to social contract theory), but also against their own state.

“The state always seeks to limit liberty” argues Waldron (2003), who scrutinizes the metaphor of balancing between security and liberty that is often invoked in securitizing discourses. The idea behind the metaphor is that in exceptional times, when there allegedly is a major threat to security, it is justified to violate some civil liberties. In cyberspace this materializes for example in discussions about data encryption. Encryption allows people to communicate with each other without being snooped upon. This is a great technology for political dissidents in countries under authoritarian rule to share their ideas with others inside and outside the country. But the same technology is available to criminals or terrorists who plot their misdeeds. According to some, in order to provide security against such criminals or terrorists, encryption should be forbidden or technologically weakened, so their plots can be detected by security agencies, even though this infringes everybody’s privacy. Waldron provides several reasons why scrutiny is at place with this metaphor, the most important for us is the one about unintended effects of security measures: “We need to consider the possibility that diminishing liberty might also diminish security against the state, even as it enhances security against terrorism” (Waldron, 2003, p.195). Security of citizens against their state is a legit security concern in cyberspace as well and sometimes it even competes with national security (Saco, 1999, p.270).

In this section I have tried to show that there is not one universal conception of security in cyberspace that is beneficial to everybody. There are security clashes between states through securitization, and the security dilemma can lead to doubly dangerous situations where nobody’s security is improved. Citizens also need protection against their own states when those try to infringe civil liberties in the name

of national security. Security clashes are omnipresent in cyberspace.

3.4 THE COUNTER-PRODUCTIVITY OF TRUST AS SECURITY

The question now at hand is what we can learn from the insights we have gained so far and what their ramifications are for trust in cyberspace. The first insight is that it is not sufficient to consider security in cyberspace a non-politicized issue that can be addressed through technological means such as the CIA-triad. Because of the importance of cyberspace to the functioning of society and the accompanying rise of stakes in cyber-conflicts, governments feel the urge to have presence in cyberspace and the cyber capabilities to promote their interests. Even if that would be undesirable from a (techno-utopian) ideological point of view, this is not going to change any time soon. Besides, governmental presence in cyberspace is very defensible from a social contract point of view. The challenge however, is to find ways in which cyber security efforts made by governments do not defeat their purpose. Their activities should have positive prudential value to their citizens and preferably other users too. They should definitely not have negative impact on non-citizens, for that would invoke counter measures from other governments. If due care is not taken, governments risk setting developments in motion that spiral to the wrong scenario of the cyber security dilemma, the doubly dangerous scenario.

The desired scenario, doubly safe, which is in line with well-being in cyberspace, is where defensive measures can readily be distinguished from and have advantage over offensive measures. Achieving this requires substantive effort and dedication from involved actors. Before anything else this requires that cyber security issues are addressed as politicized rather than securitized issues. Moving away from considering security a holy grail that justifies extraordinary means allows us to see the incongruence between security and cyberspace on a fundamental level. The fundamental problem here is that where security is always about some actor's security against some threat (such as states versus other states or civilians against their state) and thus inherently incorporates biases, it is crucial to the success of cyberspace that the core of the decentralized network remains agnostic, and thus unbiased, about who uses it and for what. If people would not trust the core to be agnostic, this would discourage them to use cyberspace to the extent that they would have otherwise, which has negative prudential value for them. This is why the conception of trust as security is flawed and why even benign attempts to achieve this will be counter-productive to well-being in cyberspace. In other words, when the core of cyberspace becomes part of clashing security

discourses, trust and well-being in cyberspace are the first victims.

This is also recognized by The Netherlands Scientific Council for Government Policy (WRR) who in a policy report argue that the public core of the Internet should be considered a global public good that should be safeguarded against unwarranted intervention by governments and other parties that erode trust in cyberspace (WRR, 2015). However, it is important too to note that the conclusion that it is counter-productive to replace trust with security is something markedly different than a full dismissal of security. What I will henceforth call *security proper*, which includes information security and its sets of attributes that provide certain properties to a system such as the CIA-triad, retains its importance. Security proper is security in a more modest sense that promotes well-being in cyberspace by enabling services that are otherwise not possible, and by serving as a measure to enhance trust in cyberspace. I will elaborate on this vindication of security in a dedicated section in Chapter 5, where I will show that at the edges of cyberspace (satisfying the aforementioned end-to-end principle) security proper is still important and of prudential value. The crux here is that because security proper is a quality that is implemented at the edges of cyberspace, it has no impact on the neutrality of the core of the network. When for example the property of confidentiality is implemented through encryption, the data is encrypted near one end-point, the source of the data, and decrypted near the other, the data's destination. It has no interest in what goes on in the core of the network, except for the encrypted data to be conveyed. That is how it differs from the trust as security discourse that was rejected because it was incompatible with an agnostic core of cyberspace.

— ∞ —

In this chapter I have explained why the conception of trust as security is counter-productive to well-being in cyberspace. This has its ramifications for answering the where-subquestion. Trust in cyberspace requires a core of cyberspace that is agnostic to clashing interests of different actors. Because security inherently favours some actor over another, it is not a suitable mechanism to promote trust in cyberspace, therefore security proper should be implemented in the edges of cyberspace. Regarding the core of cyberspace, this leaves the question for a working conception of trust open. In the next chapter I will address that question, by reviewing and discussing literature on trust.

The next part of my research question that I want to address is the why-subquestion. In essence it questions what the values of trust are that make it worthwhile striving for (in the context of cyberspace):

Why should trust be established in cyberspace?

For reasons that will become clear in the final section of this chapter I will refer to these values as the functions of trust. Answering the subquestion is only part of this chapter though. This chapter provides a literature-based analysis of trust in cyberspace that covers much more than its functions.

Researchers who study the role of trust in our society often consider [Luhmann et al. \(1979\)](#) to be the seminal work that established trust research as a field¹. According to Luhmann trust is a very effective mechanism to reduce social complexity ([Luhmann et al., 1979](#), p.8). Trust turns uncertainty into risk, something that people are much better at dealing with. This reduction of complexity is applicable both to trust in cyberspace and trust in society (which is a dubitable distinction to begin with, cyberspace can better be considered part of society). In this section I will describe the state of art in research on trust in cyberspace. Through literature review I will establish what the main issues are in this area and what answers to them have been suggested by philosophers and sociologists alike. In the final part of I will introduce the Simmelian notion of trust developed by [Möllering \(2001\)](#) as a framework that is useful firstly for ordering the existing discussions on trust in cyberspace and secondly for serving as a basis for the notion of trust in cyberspace that I will develop in Chapter 5.

Academic interest in trust in cyberspace developed alongside the increase of the importance of cyberspace in society. So roughly speaking it can be said that it started gaining momentum from the turn of the millennium, even though some earlier work dates from the 1990's and before. So far I have used the term *trust in cyberspace*, but different authors have used different term to describe this, such as 'online trust' ([Turilli et al., 2010](#); [Ess, 2010](#)) and 'e-trust' ([Taddeo, 2009](#)). Since the philosophical discussions argued for in those papers are not about what precise terminology to use, I consider those terms to be interchangeable. I will therefore stick to using 'trust in cyberspace' or

¹ See for example [Möllering \(2001\)](#); [Nissenbaum \(2001\)](#); [Taddeo \(2009\)](#)

sometimes even shorter to merely ‘trust’ in this chapter. Furthermore, it is important to distinguish between the trustor and the trustee. If in a trust relation it is said that person a trusts person b, then person a is the trustor, the one who trusts, and person b is the trustee, the one who is being trusted.

In order to answer the question what trust is, I will break the question down in several parts. I will discuss each of these parts in a separate section. Firstly, I will argue why we need trust in the first place, thus addressing the why-subquestion. This will be dealt with in a section about the functions of trust, where I will make a distinction between intrinsic and instrumental values of trust, and how they relate to prudential values of trust. Next will be a section on the bases of trust. Here I will begin with discussing the issue if trust in cyberspace can exist at all. After arguing that trust in cyberspace is possible, I will discuss on what grounds trust can be based. In the final section of this chapter I will summarize the discussion on trust in cyberspace. Here I will argue that an important aspect of trust has received too little attention: The leap of faith. This is where I will introduce the Simmelian notion of trust, which does properly pay its dues to the leap of faith, making it a suitable framework for my notion of trust in cyberspace.

4.1 FUNCTIONS OF TRUST

Trust, what is it good for? In Chapter 1 I have already provided a short answer to the question of why trust is important. As I have argued, trust is an important condition to well-being in cyberspace. At that point I have not gone into further detail of the workings of trust. In staying away from providing a conceptual analysis of trust, a danger exists in confounding the concept of trust with its functional properties. Much research in trust is focused mostly on these functional properties only (Möllering, 2001, p.404). Even though this is only part of the full picture of trust, it is worthwhile to perform further analysis. This section is focused on the functions of trust, by which I mean the different ways trust can be valuable, and I will split this in three parts. The first part is on the intrinsic value of trust, why it would be worthwhile to pursue trust as an end in itself. The second part is on the instrumental value of trust, or the goods of trust for individuals and society. Both are forms of prudential value (Taylor, 2013, p.11). In the third part I will discuss virtuous and vicious circles, of whom it can be argued that they are functions of trust that can conceptually be placed as hybrids between being intrinsic and instrumental valuable.

4.1.1 *The intrinsic value of trust*

Theories of trust that emphasize the intrinsic value of trust see trust not merely as a means to achieve something else, but stress that it is end in itself. To be more precise, such theories talk about the virtuousness of trust. Either of trustworthiness as a virtue itself (McLeod, 2014) or rooted in virtues such as patience, perseverance & empathy (Ess, 2010) or transparency & honesty (Turilli et al., 2010). I will proceed with discussing each of these theories separately.

McLeod (2014) discusses trustworthiness as a virtue. Referring to Potter (2002), she makes a distinction between specific trustworthiness and full trustworthiness. The former refers the trustworthiness of a trustee towards a specific trustor that is not generalizable. The example given is that of a convicted felon and his mother. The felon might be trustworthy towards his mother, but in general we would be hesitant in considering this person to be trustworthy. The latter, full trustworthiness, refers to the moral disposition of a trustee to be trustworthy towards everyone. It is in this sense that trustworthiness can be a virtue. This however does not answer yet what makes a person trustworthy, what this moral disposition entails. Nancy Potter provides an Aristotelian conception of trustworthiness, which means that striving to be virtuous is striving to find a mean between extremes. According to her a trustworthy person is “one who can be counted on, as a matter of the sort of person he or she is, to take care of those things that others entrust to one and (*following the Doctrine of the Mean*) whose ways of caring are neither excessive nor deficient” (McLeod, 2014; Potter, 2002, p.16).

Charles Ess is concerned specifically with trust in computer-mediated communications (Ess, 2010), so trust in cyberspace with explicit emphasis on the fact that the trustor and trustee are not in direct contact. He argues that trust and virtue are interconnected and relevant for several important reasons. Trust is central in the human condition, by which he means that it is a necessary condition to achieve harmonious and efficient communities. Those in turn are required to deal with the vulnerability and need to be able to depend on others that marks the human condition (Ess, 2010, p.293). The connection with virtue ethics makes trust a primary component of our moral character. In order to develop trust, a person needs to combine the virtues patience and perseverance. Patience stimulates the development of trust, because it showcases one’s commitment to a relationship. Perseverance is the virtue that shows “the willingness to push through conflict or misunderstanding to reconnect with one’s partner on the other side of the breach” (Ess, 2010; Vallor, 2010, p.9). Related as well

is the virtue empathy, the ability to feel with others.

A concern however that Ess raises is whether it is possible to cultivate such virtues through computer-mediated communications. Theoretical accounts on how patience, perseverance and empathy develop stress the importance of embodied co-presence (i.e. being physical present to the other), for example because of the importance of non-verbal communication. The question is, in other words, if this conception allows for trust in cyberspace. Charles Ess believes so, arguing that a blurring between online and offline (what Floridi (2015) calls *onlife*) might have as a positive consequence that the problematic aspects of disembodiment in cyberspace could be overcome. An important motivation for Ess' focus on virtue ethics is that it is the Western world's school of ethics that is closest to Buddhism, Confucianism and Hinduism. His hope is that these can serve as a source for a pluralistic global information and computer ethics (Ess, 2008). Other virtues connected to trust are transparency and honesty. Transparency and honesty are two main parameters in assessing a potential trustee's trustworthiness (Turilli et al., 2010, p.14). They can stand at the base of a virtuous circle, which I will discuss later on.

Summarizing this subsection, several authors have pointed out that trust can have an intrinsic value. As an end in itself, trustworthiness could either be seen a virtue on its own, or as a combination of several virtues. The main takeaway here is that trust is not necessarily something that is achieved through efforts by the trustor, but also a moral disposition worthwhile to strive for by the trustee.

4.1.2 *The instrumental value of trust*

The instrumental value of trust describes trust as a means. If trust exists between a trustor and a trustee it might serve as a foundation for many goods. These goods of trust can be divided in two types, namely individual goods for the trustor or trustee and social goods, that have value for society as a whole. As is so often in such distinctions, a strict demarcation between these two does not exist, for the goods of society affect individuals and vice versa. I will begin with discussing how trust is good for the autonomy, knowledge and moral maturity of individuals and subsequently turn to the social goods of cooperation and social capital.

Both for trustors and trustees trust can be an affordance for individual development. John Weckert emphasizes how trust contributes to the autonomy of the trustee (Weckert, 2005, p.98). If an employer trusts an employee rather than monitor all her activities, it affords

the employee to take the responsibility to do her work right. This can contribute to her self-esteem and overall well-being, thus empowers her autonomy. Note that this does require autonomy to be seen as a socially constituted property (McLeod, 2014). Knowledge is another individual good that depends on trust. In order for a person to gain knowledge, he must trust the testimony of others. This is a very practical requirement, for it is simply impossible for one person to fully check all facts about the world presented to him (think about the roundness of the world, or a person's day of birth) (McLeod, 2014). This is of course not to say that people should easily accept all claims that others make towards them. Mechanisms are created to help people decide when trust is justified. For scientific knowledge, peer-reviewing is such a mechanism that serves as a base of trust. I will say more about that later on.

Moral maturity of the individual is a good that benefits both trustors and trustees. Being trusted, as we have seen when describing autonomy, affords people to behave responsibly. It reflects the belief of the trustor that the trustee has a moral and mature character, which can engender the trustee's self-respect. This also benefits trustors, for putting trust in others also positively reflects their moral maturity.

The thought that trust is good for the moral maturity of trustors and trustees at the same time already hints at the goods of trust for society. In the introduction of this chapter I already referred to Luhmann who studied the role of trust in society and stated that trust reduces social complexity. As is clear by now that is not the only function of trust, but it remains very important as it explains how trust contributes to well-being in cyberspace. Therefore, I will elaborate a bit further on what Luhmann meant with reduction of social complexity, and on how it promotes cooperation and social capital.

We live our lives in a complex society where many forces are at work. According to Luhmann we need mechanisms to act successfully in such a social system. Trust is such a mechanism, for it reduces complexity and uncertainty to a system based on three parameters: familiarity, expectation and risk (Taddeo, 2009, p.4). Familiarity describes the acquaintance that the trustor has with the trustee and the context in which the trust question takes place. Expectation is what the trustor believes to result from trusting. There is always a risk that comes with the choice to trust, it is up to the trustor to decide what risk she is willing to take. Now if a person is in a complex social situation where she is required to act, to make a decision, she could freeze up because there are so many parameters she must determine in order to make a decision, such as possible intentions of others, contextual variables or other contingent factors. Trust then serves as

a model that allows her to bring this back to these three parameters and make a decision fast. In other words: it reduces social complexity.

If there is a high level of trust in a society, opportunities for cooperation are vastly higher, which is to everybody's benefit (Gambetta, 1988). Trust promotes cooperation because it lowers the barriers to engage with others, for example by removing the incentive to check the work performed by others. This translates into more time for other activities for both the trustor and trustee. In a society with a high level of trust things go smoother and more efficient (Weckert, 2005, p.97). The advantages of trust for cooperation are actually a manifestation of how trust is a form of social capital (Weckert, 2005; Nissenbaum, 2001, p.107). The idea of trust as social capital has been developed by Putnam (1994), who argued that it enabled people to work together for common purposes.

A high amount of social capital allows people to take collaborative action and reach outcomes that otherwise could not be achieved. This is beneficial for both the participating persons and society as a whole. A low level of social capital on the other hand leads to societal inefficiency. In that sense a lack of trust can be said to be expensive. When there is less trust, other mechanisms need to be set up to govern societies. Formal rules and regulations need to be developed and a whole bureaucratic apparatus must be created and maintained to enforce them: police, lawyers, auditors et cetera (Weckert, 2005, p.97). To summarize this point in the words of Fukuyama (1995): "Widespread distrust in a society (...) imposes a kind of tax on all forms of economic activity, a tax that high-trust societies do not have to pay".

In this subsection I have provided an overview of the instrumental value of trust. There are individual goods and social goods of trust. Trust can be an affordance for the autonomy, knowledge and moral maturity of an individual. It is also a social good, since it reduces social complexity, promotes cooperation and can be considered a form of social capital. I have also argued why a lack of trust is expensive, societal as well as financial.

4.1.3 *Virtuous and vicious circles*

Several authors have paid attention to how the presence or absence of trust can lead to virtuous and vicious circles respectively. I will begin this subsection by introducing these specific accounts, although after that I will abstract from them in order to give a general account of virtuous and vicious circles.

The first virtuous circle builds on the aforementioned idea that the presence of trust makes interactions advantageous for the trustor. Interaction with others also increases the ability of individuals to appraise the trustworthiness of others, and therefore (Turilli et al., 2010, p.14) argue that these two combined initiate a virtuous circle that leads to a selection process. Through this process, trustworthy individuals are involved in a growing number of interactions, whereas conversely untrustworthy individuals become marginalized in the social system. The virtuousness of this circle lies therein that desired behaviour is positively reinforced and undesired behaviour discouraged. This virtuous circle is the core mechanism behind rating systems for customers and sellers such as used at eBay or Amazon.

In his review of Möllering's book on trust (which I will get to later when discussing the leap of faith), Nooteboom (2006) suggests that there may be a virtuous circle and a corresponding vicious circle connected to the "reflexive process of active interaction", e.g. that perspectives of trustors change along the process of trusting (Möllering, 2006, Chapter 4). The book mentions a "spiral reinforcement model of trust", Nooteboom argues that this spiral could work in two directions, positively virtuous or negatively vicious. The former would yield a cycle of trust, disclosure of information, acceptance of influence from others and relaxation of control, which would lead to more trust. The latter works in opposite direction, from misunderstanding to suspicion, increase of control and a breakdown of trust. These circles are at this point not more than concoctions of Nooteboom during reviewing, but they illustrate well that just as virtues have corresponding vices, for every virtuous circle a similar vicious circle can also exist.

This also emerges from Charles Ess' article that discusses one virtuous and three vicious cycles. The virtuous circle is rather curious in that it builds upon deception. If a trustee presents herself and acts better than she is (the deception part), she can be positively reinforced by the responses of others to this behaviour. This encourages and rewards acting better and hence encourages practising to actually become better (Ess, 2010, p.294). However, deception also lies at the heart of the third of the three vicious circles identified by Ess. Computer-mediated communication may negatively impact trust, through three vicious circles related with security, affordances and deception.

The vicious circle of security is set in motion when it is tried to overcome a lack of trust through certain intrusive forms of regulation and increased control. These can be counter-productive, for if such measures replace initial trust it drives out opportunities for trusting relationships and it will only become harder to regain trust (Thorseth

and Ess, 2009; Pettit, 1995; Ess, 2010, p.297). This vicious circle alludes to what trust as security, as discussed in Chapter 3. The second vicious circle is that of the affordances for trust-opposing behaviour that contemporary ICTs could stimulate. Previously, I have explained that Ess argues that the virtues of patience, perseverance and empathy are closely related to trust. Instead of stimulating such virtues, computer-mediated communications could stimulate escape, gaze and haste. Escape means that instead of being patient with others, people can through a click to close a screen or a short 'gotta run!' message back away from interaction much easier than in real life. Likewise, a lack of gaze or visual contact in online interactions (though not all, think about video conferencing) reduces the development of empathy. Similarly, the speed of these media can encourage a culture of haste, which hinders the cultivation of perseverance (Ess, 2010, p.296-298). These affordances may seem improvements in the short run, but since they counteract virtues that lead to trust, they are vicious in the long run. The third and final vicious circle is about deception. With the aforementioned merger of the online and offline world, the distinction of virtual and real fades as well. Ess argues that this "might afford greater temptation and occasions to practice deception and (...) a form of sexual infidelity" (Ess, 2010, p.300). The more online and offline life merges, the greater the chance that trust relations are harmed in this way.

In his conclusion, Ess paraphrases Vallor (2010) by articulating a challenge that we face: Recognizing that cyberspace offers various affordances, we should aim to design new technologies in such ways that stimulate virtuousness and discourage viciousness. To put this in the context of my thesis, this means that in order to promote well-being in cyberspace, we need to develop mechanisms and incentive behaviour that promotes trust. If done properly, this will on the one hand allow existing trust to reinforce itself and on the second to achieve the individual and social goods that are so valuable. The combination of these has great prudential value. Such mechanisms are of course virtuous circles and in essence their working is simple: People who decide to trust others should (trustors) and people who behave trustworthy (trustees) must be rewarded for this behaviour. Preferably not only through achieving what they hoped to achieve, but also through through an increase of trust, which would lower the barriers to trusting further. Such positive feedback loops are intrinsically as well as instrumentally valuable. Conversely, we need to be wary of (incidentally) creating vicious circles that have the opposite effects. Through these circles, I have already touched upon an aspect of trust that is equally important as the functions of trust and that is for what reasons people decide to trust. Such bases of trust are the

subject of the next section.

4.2 BASES OF TRUST

Up until now I have reviewed accounts of what trust is good for. However, this all amounts to nothing if people are not actually willing to place trust in others. In this section I will discuss what bases of trust are. Early discussion on trust in cyberspace were concerned with the question if it could exist at all. This discussion will be the first part of this section, which will end with arguments in favour of the possibility of trust in cyberspace. After that I will turn to discussing what different categories of accounts for trust in cyberspace have been suggested. Here I will distinguish between cognitive accounts, non-cognitive accounts and phenomenological accounts who are a little bit of both.

4.2.1 *The possibility of trust in cyberspace*

Early articles on trust in cyberspace put heavy emphasis on the question whether trust in cyberspace could be possible at all. The approach taken would be about as follows: first the author would look at theories of trust in offline society and identify what they thought were necessary conditions for the establishment of trust. Subsequently, they would take these conditions and discuss if they could be satisfied in cyberspace. Pettit (1995) and Nissenbaum (2001) are notable examples of such articles, both come to the conclusion that the conditions for trust they identified can inherently not be satisfied in cyberspace.

There are two main conditions that are identified. Different authors use slightly different names, but they amount to the same core issues. The first condition is about the absence of certainty about the identity of trustees in cyberspace. This includes concerns about the need for physical interactions with the other in the development of trust. The second condition is about the importance of a shared set of norms and values for trust, that is absent in cyberspace. Because I believe later reactions convincingly refute the objections against trust in cyberspace, I will now turn to explaining and directly rebutting each condition respectively.

The first condition revolves around online identities. In cyberspace it is possible to remain anonymous. Helen Nissenbaum argues that there is a double sidedness to this. On the one hand anonymity can have an empowering effect, for example for minorities and suppressed people. But on the other hand the lack of certainty of identity dis-

inclines trustors to trust. She stresses the importance of sustained identity, which we can see “as a thread upon which we string the history of interaction with others” (Nissenbaum, 2001, p.113). Such a diachronic identity of the trustee (i.e. an identity that does not change over time) would be difficult to establish and confirm by the trustor. However, later articles argue that this condition can in fact be satisfied. Although it might be true that establishing the physical identity of the trustee can be difficult, it is still possible to assess the trustworthiness of an online peer (Turilli et al., 2010, p.7). Rating systems such as mentioned in the subsection on virtuous circles allow trustors to make trust decisions based on the reputation of diachronic online identities. This way the problem of missing identities can be mitigated. The necessity of the condition is not challenged, but the condition can be satisfied.

A shared set of norms and values is widely regarded as an enabler for trust (Nissenbaum, 2001; Fukuyama, 1995, p.26). People are much more likely to trust others of which they know that they hold to a similar morality because they have a common background. Again the detractors of trust in cyberspace argue that this is a condition that cannot be satisfied. Cyberspace has an unstructured nature and there is broad heterogeneity in the background of users. Even in specific online communities that revolve around a certain interest participants belong to different cultures, religions, genders and nations. It is safe to say then that a shared set of norms and values is absent in cyberspace. However, the question in this case is if a shared set of norms and values really is a necessary condition for the emergence of trust. Turilli argues that although it definitely makes things easier, it is not a condition *sine qua non*. Following up on research by Yamagishi et al. (1999) they make a distinction between trust and assurance. Assurance is a mechanism to deal with uncertainty that requires a structured environment and thus might fail to emerge in cyberspace. Trust on the other hand depends on the proactive answer of individuals to the presence of environmental uncertainty (Turilli et al., 2010, p.6). Earlier I described how virtuous circles can stimulate the development of trust. Precisely this is why trust is a suitable mechanism to deal with uncertainty in cyberspace. So this condition is not an obstruction to trust in cyberspace, rather it is an argument for the importance of trust enhancing measures in cyberspace. In this case the necessity of the conditions has been refuted.

From this we can conclude that trust in cyberspace is possible. Differences remain between trust in the online and offline world, but the nature of these differences is topological (e.g. due to the differences in online and offline environment) rather than ontological (Turilli et al., 2010, p.13). Now that I have established the possibility of trust in cy-

berspace, I will turn to discussing accounts of trust.

4.2.2 *Accounts of trust in cyberspace*

There are different accounts of what trust in cyberspace is based upon. Following [Ess \(2010\)](#), I will describe three categories in which such accounts can be classified. The first category is that of rational or cognitive accounts, the second covers affective accounts and the third is that of phenomenological accounts.

The first account of trust in cyberspace defines trust as the result of rational or cognitive processes in the mind of the trustor. On this account, when trustor A trusts trustee B, then “A believes (expects) that B will do X in situation S” ([Weckert, 2005](#), p.101). The decision to trust in this case is purely rational and based on reasons that the trustor has to trust. An example of such an account is that of [Gambetta \(1988\)](#), who narrowed down trust to making decision based on calculating probabilities. Although such cognitive accounts of trust seems sensible at first glance, they have their limitations, for example because they do not account for how trust affects the behaviour of trustees. They also do not cover other experiences of trust such as that of children in their parents ([Ess, 2010](#), p.290).

Affective or non-cognitive accounts of trust acknowledge that there is more to trust than rational calculation. [Weckert \(2005\)](#) summarizes such accounts as “A’s attitude toward B is Y” or “A takes a certain stance, Y, toward B”. But accounts that are based purely on attitudes of the trustor are not satisfying either, for they do not explain why in some cases a trustor does decide to trust and in other cases does not.

This brings us to the category of accounts of trust in cyberspace that Charles Ess dubbed phenomenological accounts. Central to such accounts is a phenomenon to which cognitive and affective are inadequate classifications. Both are involved in trust. John Weckert provides such an account, which he himself calls a ‘seeing-as’ account of trust. A trusts B then means that “A sees B’s behaviour as trustworthy” ([Weckert, 2005](#), p.102). Weckert tries to capture both affective and cognitive aspects in this definition. Regarding the former Weckert argues that seeing-as is stronger than merely believing, although it still reflects a certain non-rational disposition of A towards B. Regarding the latter he argues that seeing-as still includes a certain reasonability for having this disposition.

Weckert’s account is not a final account in the sense that it for once and for all determines the correct bases for trust. [Taddeo \(2009\)](#) finds that his account fails to explain the emergence of trust in cyberspace

and its role in online interaction. In response to that I believe that Taddeo does not do right to Weckert because she largely ignores the cognitive side of his account, but that is a different discussion. What I want to take away from this subsection is that there are multiple categories of possible bases of trust. Based on (a mixture) of such accounts trustors decide to trust, to subsequently enjoy the prudential value of that. In the final section of this chapter I will discuss how finally the decision to trust is made.

4.3 THE LEAP OF FAITH

Up until this point I have discussed the functions of trust and the bases of trust. In the sections dedicated to these subjects I have referred to the same sources, which might make one wonder why I did not just discuss these articles one after another. The reason for this is that I find it important to emphasise the different aspects of trust. Most authors who develop a theory of trust (rightfully) spend time on discussing both functions and bases of trust. However, in these theories of trust in cyberspace no serious attention is being spent to how people get from having bases of trust to reaching the state of trust from which the functions of trust follow. Rather, this step is being dismissed as merely ‘deciding’ to trust. I have borrowed the terms functions of trust and bases of trust from Möllering (2001), who discusses precisely this issue, what links bases of trust to values of trust.

His main trouble with many prevailing theories of trust is that they presume a strong link between the bases of trust and the state of trusting, as if the one naturally flows over into the other. Sticking to the conception that trust turns uncertainty into risk, this would mean that that transition of states happens conveniently without a fuss. Möllering opposes such a course of events and develops a notion of trust that reverts back to writings by 19th-century philosopher Georg Simmel. This Simmelian notion of trust argues that bases of trust are connected with functions of trust through a leap of faith by the trustor, and describes the nature of that leap of faith. In this section I will introduce Guido Möllering’s Simmelian notion of trust, which will serve as a framework that I will use later on to argue how I think trust can be promoted in cyberspace. The framework exists of three steps, namely interpretation, suspension and expectation. These steps cover the bases of trust, leap of faith and functions of trust respectively. Möllering uses the metaphor of a valley to describe the trusting process, with a gorge that separates the land of interpretation from the land of expectation. Getting from the former to the latter requires a “mental process of leaping – enabled by suspension

– across the gorge of the unknowable” (Möllering, 2001, p.412).

The land of interpretation is where people develop their good reasons to trust. Möllering uses the term ‘good reasons’ to describe interpretative trust bases. Trustors start trusting trustees if they have good reasons to do so. What constitute good reasons depends on the person and the moment, and can comprise different types of bases. Simmel already identified what I have elaborated upon in the section on bases of trust, namely that there are multiple possible categories of bases of trust. These are not just solely rational, but can be affective as well and mostly are a combination of the two. In that sense bases of trust are (as Simmel would have put it) weak, because what makes bases of trust into good reasons cannot be hardly determined. Möllering asserts that the “quest for a single best way of mapping trust bases is ultimately futile” (Möllering, 2001, p.412). This is why he calls this approach to trust hermeneutical and why this step is called ‘interpretation’, for hermeneutics is the theory of interpretation and understanding of a person’s motives.

When trust has been established, the trustor is in “a state of favourable expectation regarding other people’s actions and intentions” (Möllering, 2001, p.404) from which the various functional consequences can follow that I discussed in the section on functions of trust. What is interesting to note is that according to this theory, after the state of favourable expectation has been reached, “the process continues and the land of expectation becomes the land of interpretation from which the gorge will soon need to be crossed again” (Möllering, 2001, p.414). I will be using this later, arguing that trust in a system can be achieved through step-by-step trusting its constitutive parts.

Finally, the crossing from the land of interpretation to the land of expectation is where the magic of trusting seems to take place. Möllering calls this step suspension and it is where the leap of faith is made. At this point the trustee has reached the point of having (personal) good reasons for trusting although the outcome is still uncertain. “Suspension can be defined as the mechanism that brackets out uncertainty and ignorance, thus making interpretative knowledge momentarily ‘certain’ and enabling the leap to favourable expectation” (Möllering, 2001, p.414). In the moment of suspension the trustor lets go of control and turns to what Simmel calls “quasi-religious faith” which he argues stands outside the categories of knowledge and ignorance.

With bracketing out uncertainty and ignorance Möllering means acting as if the gaps of missing information and doubts of the trustor and the potential dangers of trusting are unproblematic. It is a logic of

'despite', 'although' and 'nevertheless', to deal with issues the trustor might be aware of, but cannot penetrate or resolve fully. Bracketing these vulnerabilities and acting as if they were resolved is the underlying idea of suspension (Möllering, 2006, p.115). It is arguable something we already start learning in infancy through dealing with the lack of control we have over the presence and absence of caretakers. We are not constantly aware that we take leaps of faith, but although they "may not be made consciously, they are not made unwillingly either" (Möllering, 2006, p.119), which means that suspension has a strong element of agency. This is to say that trusting is not an act of giving up or surrendering to fate, but an operation of the will of the trustor.

— ∞ —

In this chapter I have discussed the main points in academic research on trust in cyberspace. I have ordered the discussions in accordance with the three parts of Möllering's Simmelian notion of trust (although not in the same order). I began with discussing the functions of trust, which answers to the why-subquestion of this thesis. In this section I introduced intrinsic values of trust, instrumental values of trust and virtuous and vicious circles, listing many ways that trust adds to well-being in cyberspace. After this I turned to the bases of trust, where I showed how after initial scepticism by academics the possibility of trust in cyberspace has been embraced. Subsequently I introduced different accounts of trust in cyberspace and how they can be categorized as either rational, affective or phenomenological accounts. In the last section of this chapter I introduced the leap of faith that connects the bases with the function of trust and described how this happens through suspension.

The literature and concepts I introduced in this chapter will serve as a foundation to the notion of trust in cyberspace that I will develop in the next chapter.

REGAINING TRUST IN CYBERSPACE

After having put in place the necessary building blocks in the previous chapters, I am finally in position to write down my conception of trust in cyberspace. This is an answer to the how-subquestion of my thesis:

How should trust be established in cyberspace?

I will begin this chapter with a short recapitulation of the key concepts I have introduced earlier. These are my characterization of cyberspace, why trust as security is a misconception and counter-productive to well-being in cyberspace and the Simmelian notion of trust with its emphasis on the leap of faith. After that quick refreshment of key concepts I will fuse them together to form my notion of trust in cyberspace. The core idea of this notion is that overall trust in cyberspace is the amalgamation of a myriad leaps of faith from node to node. Explaining that view will be the next part of this chapter. This explanation will bring us to *trust enhancing measures* that strengthen bases of trust. Here I will return to security measures, arguing that their proper place is serve as trust enhancing measures at the edges of the network. But first, now, let me recapitulate the key concepts of the previous chapters.

5.1 RECAPITULATION

The three key concepts that will combine into my notion of trust in cyberspace have each been introduced in a dedicated chapter. In Chapter 2 I argued why formulating a definition of cyberspace is problematic and as an alternative provided a *characterization of cyberspace*. Subsequently, I turned to *trust as security* in Chapter 3 and explained why this conception of trust is counter-productive to well-being in cyberspace. After this, in Chapter 4 I turned to performing a philosophical analysis of trust. The take-away from this chapter was the *Simmelian notion of trust* that consisted of three steps: interpretation, suspension, and expectation. Let me now shortly refresh the ideas behind these three principles.

Cyberspace is not a fixed phenomenon, but an ever evolving and expanding environment. It combines cognitive, logical and physical aspects. Reverting back to three principles from technical network

design (the concept of a *network of networks*, the idea of *layers* that I transformed to nodes and the *end-to-end principle*), I characterized cyberspace as follows:

Cyberspace is a decentralized network of networks consisting of nodes that are either of a content, logical or physical nature, where complex particular functions should be implemented near the endpoints

The grand network is the union of all circuits. A circuit represent a certain use case that cyberspace has for a user and consists over interconnected nodes. Nodes come in three categories, they are either of a content, logical or physical nature. Nodes can be part of many different circuits, it is through such overlapping that all circuits combine in the grand network of cyberspace. If someone wants to implement specific properties (such as security), that should be done close to the endpoints, so the core of cyberspace stays agnostic to all the particular (and potentially conflicting) wants of users.

Trust as security is the idea that security mechanisms can provide the certainty that people wish to have before they using cyberspace. Although security or more specific attributes such as confidentiality, integrity and availability are important requirement for many activities in cyberspace, it is a wrong idea to substitute trust for security. Not only is security in this sense not a form of trust but a lack of trust, it also has the danger of leading to securitization of cyberspace. This is because of the security-dilemma: the security of one can be the insecurity of the other. Due to a lack of trust this can lead to a vicious circle that in the first place does not even increase the security of its participants, but more importantly is counter-productive towards the well-being of as many people as possible in cyberspace. Therefore, prudence is required with regards to how and what security measures are introduced in cyberspace.

Potentially, cyberspace is a great means that can have prudential value to many humans. However, in order for this potentiality to materialize, people need to feel like they can trust cyberspace. Following the Simmelian notion of trust, I split the phenomenon of trusting in three, from interpretation to suspension to expectation. More concrete, these phases are represented by the formation of bases of trust, the leap of faith, and execution of functions of trust. Trust is an ongoing process, after the land of expectation has been reached it becomes the new land of interpretation for the next trust question.

With this all fresh in mind again, I want to turn to taking the first step in formulating my notion of trust in cyberspace. This is the determination of where in the process of using cyberspace questions of

trust arise and how these are solved.

5.2 LEAPING FROM NODE TO NODE

The big question about trust in cyberspace is where the trust question takes place. What is it that must be trusted, before we say that we trust cyberspace? Right away it is clear that it makes little sense to see the trust question as a question about the system that cyberspace is as a whole. This would make trust in cyberspace a binary question, where you either trust it as a whole, or not at all. Both answers would have undesired consequences. Those who trust it as a whole could soon fall prey to all sorts of scams, those who do not trust it at all would throw out the baby with the bathwater and miss out on many affordances that have prudential value. Instead, what we need to find out in order to maximise the potential of cyberspace without falling victim to actors with bad intentions, is what the right level and place is in cyberspace to pose the trust question.

My claim is that the question of trust is best answered at the level of nodes, more precisely at the transition from node to node in a circuit. Further along in this section I will explain and illustrate this claim, but first I will shortly spend some time on the concept of *levels of abstraction*, of which the challenge of finding the right place to ask the trust question is an example.

The concept of levels of abstraction has its roots in formal methods which is a branch of computer science. Floridi (2011) has taken it and turned it into the main method of his philosophy of information. The method of levels of abstraction is a method to perform a conceptual analysis on a system. It is an epistemic method, by which is meant that it is a method to gain knowledge about the system, so it does not make ontological claims about what that system is (Van Leeuwen, 2014, p.16). Levels of abstraction refers to the idea that a system can be conceptualised on different levels in a hierarchy. The higher the level, the abstracter the description of the system. Details might seem to get lost when abstracting to a higher level, but properties of the system that did not make sense at a lower level might become visible. For example, it makes little sense to talk about the architectural style of individual bricks, but it does makes sense to talk about the architectural style of the building they form together.

Now to get back to the subject of my thesis, the question when viewing the system cyberspace is at what level of abstraction is the question of trust best placed. In my characterization of cyberspace I have introduced three levels, from top to bottom: the grand network,

the circuits and nodes. It is possible to go further down, although the make up of those levels then differ between the different types of nodes. For logical nodes for example one level of abstraction lower would be to look at the software code of the node. As I stated already, in my conception of trust in cyberspace the trust question is asked at the level of nodes. In other words, according to me the question: do I trust cyberspace? amounts to asking: what nodes do I trust? Per node then a leap of faith in the Simmelian sense needs to be made.

By placing trust at the level of nodes, I introduce a certain flexibility in trust in cyberspace that matches the flexibility of my characterization of cyberspace. This breaking down of the question of trust in cyberspace into many questions of trust in individual nodes has strong explanatory power. It can explain why we do trust cyberspace in some cases and do not in others, because in the case of lack of trust (where the leap of faith is not made) it allows to pinpoint the node(s) where it goes wrong. This would not be possible if I placed the trust question at the level of circuits. Although it would already provide a bit more leeway than placing it at the level of the grand network (I started his section by explaining why that would be worthless), it could quickly render cyberspace useless for those with an above average inclination to paranoia if they deem some circuit untrustworthy without further specification why so. It would also make it difficult to find out why some people do trust a certain circuits why others do not. Both cases require references to nodes that make up the circuit, which is another way of saying that the level of nodes is the better level to talk about trust. Putting trust in a level lower than nodes would be problematic for a different reason: as I already explained, the constitution of the layers under nodes differ between the different types of nodes. This would make it very difficult to make comparisons regarding the trustworthiness of nodes of different types. My fear is that this would complicate the trust question more than that it solves anything. Let us instead turn to the next section where I will try to make my conception of trust in cyberspace more clear. I will do this by articulating it in a shared vocabulary of my characterization of cyberspace and the Simmelian notion of trust, with an interwoven example of mobile banking.

5.3 HOW LEAPING WORKS

The question of trust in cyberspace is a question of trust in its constitutive parts. In the most abstract level, cyberspace is a grand network that is the union of a myriad of circuits. circuits represent the different use cases that users have of cyberspace. This level accommodates the diverse meanings that cyberspace has for individuals. One level

lower, circuits are made up of nodes, that form a path between the user and his destination in cyberspace. The nodes are the stepping stones of such paths. Each node can be part of many circuits, or in other words, each stepping stone can belong to many paths. When a person wants to do something in cyberspace, such as mobile banking, then there is a corresponding circuit connecting the user and his bank's online environment. Mobile banking is then one (of potentially many more) meaning(s) that cyberspace has for that user. Since there are some dangers to banking, the user is faced with the question whether he trusts cyberspace or not. In this case with this meaning of cyberspace (a place for mobile banking) that trust question comes down to: do I trust the nodes that form this circuit and connect me to my bank? With every transition from node to node in the circuit or every leap from stepping stone to stepping stone, the question is raised whether the user deems the next node trustworthy. Only if he trusts all nodes he is able to bank mobile, which brings a lot of convenience compared to the limitations of banking at a physical location with confined opening hours which has prudential value.

Seeing a transition from one node to another as a leap from stepping stone to stepping stone shows how I want to fit in the Simmelian notion of trust. A circuit starts with a content node, the digital representation of the user itself. Implicitly I take for granted that the user trusts itself. Then before leaping to the next node, the user is faced with the question whether his bases to trust the next node are strong enough. Remember that this is hermeneutical, e.g. it depends on the interpretation of the user. In case the user finds the bases to be strong enough he decides to make a leap of faith and trust the node. Now he is one step closer to his destination (the online banking environment) and the whole trusting exercise repeats, or as Möllering called it, the land of expectation becomes the land of interpretation. This process of trusting using leaps of faith continues for all nodes in the circuit until the leap to the last stepping stone is made. If indeed the last node is successfully reached and the trust question can be positively answered with the constrained that the answer is limited to this instance of this use case: yes, the user trusts cyberspace.

This is in essence my conception of trust in cyberspace. However, in the description above I assumed that every step went as desired. There remains more to say about leaping from node to node. Firstly, I already emphasised again that leaping is hermeneutical, which meant that it depends on the interpretation of the user if the bases of trust seem trustworthy enough to take the leap of faith. This means that different users that traverse circuits with a shared node can make a different decision about whether they trust that node, even if they have the same bases of trust. In fact, even the same user can make a

different judgement at a different moment for the same node. It is a good thing that my conception of trust in cyberspace has this property, for it does right to the importance of individual meaning giving that is central to my characterization of cyberspace.

A second remark is about what happens when somebody encounters a node he does not trust. At first instance then, it may seem that it is very unfortunate for the user, but his desired usage of cyberspace cannot go through. However, because it is possible to pinpoint what the untrusted node is, it might be possible to circumvent this node. This would mean that the user tries to find another route to the desired endpoint. Imagine the situation were someone is in a coffee bar and watching movies on his smartphone. Because movies can require quite an amount of data, the user makes use of the free WiFi offered by the coffee bar. This WiFi connection is then a node in his circuit for watching movies online. If later he wants to do mobile banking though, he might not trust the WiFi connection as a node in his mobile banking circuit, for somebody could be snooping on the connection (which is not really a problem in the case of funny cat movies). Here he could stop and decide not to mobile bank now, but he could also decide to turn off WiFi and connect using his mobile data connection. This circuit shares many nodes with the former, except the connection is a different node — one the user does trust. This too is a positive property of my conception of trust in cyberspace, one that demonstrates that it has a good fit with the dynamism of cyberspace.

Thirdly, what if a user decided to take a leap of faith that turned out to be unjust? For example when a user is connecting to his bank's online environment and it turns out someone was snooping on the connection and stole the user's login credentials. Unjustified trust is indeed problematic, for it is harmful to users and might make them wary in future situation which could limit their usage of cyberspace and thus indirectly have negative prudential value. Therefore it is important to at once minimize the impact of unjustified trust (such as through two-factor authentication in mobile banking) and at the same time minimize how often trust is misjudged by taking measures that enhance trust. This is where I see the proper role and vindication of security technologies and policies in cyberspace. I will elaborate on it in the next section, after a last remark about my conception of trust in cyberspace.

This fourth and last remark is a practical note to my conception. When people use cyberspace they are not consciously making leaps of faith all the time, but as Möllering already mentioned, not unwillingly either. Most of this leaping happens unconsciously, especially over paths that users have used before, such as when visiting a news

site. Still every now and then the question of trust arises. This can be in general cases, for example through the Snowden files or when there is news about privacy leaks. Such cases can prompt people to reevaluate the nodes they have considered trustworthy for a while. Also in specific cases the trust question can become one that must be answered consciously, for example when sensitive data is involved as in the banking over WiFi example, or when somebody tries something in cyberspace he has not tried before.

I have now presented my conception of trust in cyberspace and tried to clarify it by using examples. I have shown that in my characterization of cyberspace the level of nodes is the right level of abstraction to place the trust question. I have fused the Simmelian notion of trust to the transitioning from node to node in a circuit and visualized it as a leaping from stepping stone to stepping stone in a path from the user to its destination in cyberspace. Every next node is reached through a small leap of faith, if every node is deemed trustworthy then that circuit and its corresponding meaning of cyberspace is at that moment deemed trustworthy. Furthermore, I have provided four more remarks regarding my conception of trust in cyberspace. In the next section I will elaborate on one of those, the vindication of security as a trust enhancing measure at the edge of the network.

5.4 SECURITY PROPER

As I have hinted a few times already, my dismissal of trust as security is not a dismissal of security in total. In Chapter 3 I introduced the term security proper to refer to more modest role that I see for security in promoting well-being in cyberspace. Security proper promotes well-being in two ways that are not fully distinct, but worthwhile to explain separately. Firstly, security proper enables services in cyberspace that are otherwise not possible, secondly it serves as a measure to enhance trust in cyberspace.

Security proper includes information security, which was the technical approach to security. Through attributes made possible by such technologies, security proper enables many services to be implemented in cyberspace. In the previous section I already provided the example of mobile banking, which has prudential value over normal banking because it overcomes limitations regarding opening hours and the need to physically go to a specific location. Mobile banking is not possible if the properties of confidentiality, integrity and availability are not guaranteed. This illustrates why in order to reap some of the benefits that cyberspace can offer, such properties (there were more, such as those of the information assurance and security octave I re-

ferred to) must be assured. This is what security proper does. It is about security technologies and policies that enables users to use cyberspace for all sorts of things and protects them against criminality and other abuse. In this way, security has a positive impact on well-being in cyberspace and should be promoted. Governmental activities that realize the social contract in cyberspace in such a way, for example through computer emergency response teams (CERTs), to protect their citizens against malicious hackers, malware, swindlers et cetera are to be encouraged (because it is clear that they are of inescapably defensive nature). In this sense, trust also does not trump security: my thesis should not be interpreted as a calling on the removal of username and password based authentication systems or something of that nature. Such login systems allows people to know that they are really talking to their friends on social network sites, just as encrypted email allows interaction between political dissidents and auditable e-commerce systems enables people to judge what online stores are dependable. The list of how proper security has prudential value is endless.

The second way in which security proper is of prudential value to well-being in cyberspace was already touched on in the third remark regarding my conception of trust in cyberspace. It is how (proper) security technologies and policies can work as trust enhancing measures. If a user is faced with the question if he trusts the next node in a certain circuit, he will evaluate the bases of trust he has and based on his interpretation of those bases decide to leap or not. Trust enhancing measures in this context are technologies or policies that can affect or themselves serve as bases of trust. Regarding mobile banking I provided the example where a user did trust his mobile data connection, but not the public WiFi connection. That is a technological example, but it can be based on policies too, such as online marketplaces where users have to be verified (using e.g. a phone number or credit card) before trading is possible. Such verification can be interpreted by many as a guarantee that enhances trust. If such a measure is decisive for trusting a node and through that a certain circuit, it is of great prudential value.

Implementations of information security attributes that are part of security proper are good examples of technologies that can serve as trust enhancing measures. As argued before, guaranteeing confidentiality through data encryption can be key in a political dissident's deciding to use cyberspace to be in touch with other dissidents. But there are many more measures that can be taken to enhance trust. Cyber diplomacy could lead to treaties that limit governmental activities in cyberspace such as snooping on emails, which people can interpret as good reasons to self-censor less than before. Businesses

can let independent parties audit their IT systems, to convince (potential) customers to use their services because the customers know the company protects their data well. There are endless possibilities, the challenge is to come up with convincing measures that as many people as possible will interpret as good reasons to trust, thus leading to having the most added prudential value as possible. The best way is to come up with measures that work as virtuous circles as described in Chapter 4. Since the interpretation of bases of trust by people can change, winning their trust is a never-ending process. Virtuous circles allows existing trust to reinforce itself through positive feedback loops. This results in the long-term establishing of trust, which has great prudential value.

— ∞ —

In this chapter I have addressed the how-subquestion, by introducing my notion of trust in cyberspace. I have introduced my notion that bases trust in cyberspace on leaping from node to node and subsequently proceeded explaining its workings and in particular discussing its implications, with special attention to the role of security proper. Having finished this means I am ready to conclude my thesis.

CONCLUSION

In the last four chapters I have elaborately detailed my analysis of trust in cyberspace. This makes that I am now ready to conclude this thesis in this final chapter. I will begin by answering the research question. Subsequently I will describe the normative implications of my work and connect this with an outlook at what could be next in future work.

6.1 CONCLUSION

In the introduction I began describing the problem statement that I was concerned with for this thesis in an informal way. I am concerned with the fate of cyberspace and want to prevent it from collapsing under the weight of its own success. Concretely put, the problem that I am concerned with is that because cyberspace is so successful, the stakes are getting higher in clashes of interests and other cyber conflicts, which could be the cause of effects that undermine the original success factors of cyberspace. The point of my thesis was not merely raising this concern, but addressing it by analysing what the sources of the described problem are and suggesting a way to overcome it. In order to do that I turned the problem statement into the following research question:

RESEARCH QUESTION: Why, where, and how should trust be established in cyberspace?

This research question consists of three questions regarding the establishment of trust in cyberspace in one. Those three questions were the three subquestions of this thesis and in the order that I discussed them were the following:

SUBQUESTION 1: Where should trust be established in cyberspace?

SUBQUESTION 2: Why should trust be established in cyberspace?

SUBQUESTION 3: How should trust be established in cyberspace?

After having first spent a chapter to describe my characterization of cyberspace, I spent a chapter per subquestion to describe what the issue behind that subquestion is and how to address it. In the recapitulation section of the previous chapter I already summarized the

contents of the chapters before, so that should all be fresh in mind. I will now turn to formulating an answer to the research question that combines all major insights of this thesis:

RESEARCH QUESTION: Why, where, and how should trust be established in cyberspace?

ANSWER: The establishment of trust in cyberspace is important because it has positive prudential value to the well-being of users in cyberspace. It is not to be confounded with security, for that have counter-productive effects due to the inherent bias in security. Within my characterization of cyberspace as a network of circuits comprised of nodes, the trust question materializes at the leaping from node to node, but it is not always answered consciously. The attributes of security proper are implemented at the edges of cyberspace and can serve as trust enhancing measures.

6.2 NORMATIVE IMPLICATIONS & FUTURE WORK

I have now provided an answer to my research question, but so far it is only a theoretical answer. What must happen next to actually achieve the establishment of (more) trust in cyberspace covers both the normative implications of my results and provides an outlook for future work. The final remarks of my thesis cover those.

For those who endorse my reasoning in this thesis and who are committed to promoting well-being in cyberspace, it follows that trust in cyberspace must be established as much as possible. Since cyberspace is a decentralized network, this implies a shared responsibility of all actors who have influence on some nodes in cyberspace. This actually includes every user of cyberspace, since their digital representations are content nodes, and indeed people are responsible for the trustworthiness of their digital representation. But the closer to the core a node is (e.g. the more circuits it is part of), the bigger the responsibility of the actor(s) that can influence it. This responsibility means that the actor(s) should take such trust enhancing measures, that in the interpretation of as many users as possible the node should be deemed trustworthy enough to make a leap of faith, for this maximizes the promotion of well-being in cyberspace.

Devising trust enhancing measures that make nodes trustworthy for as many people as possible is far from trivial. Not only do people interpret bases of trust differently, but as we have seen there are many clashes of interests in cyberspace. Coming up with trust enhancing measures that stay agnostic to them is quite intricate. This is what I

believe to be the main challenge of future work. I have no doubt that my theories can be refined, but what would be most beneficial is future work that takes a practical turn and investigates how the normative implications can be made concrete for governments, businesses, individual users and all other types of actors in cyberspace. For governments this could concern matters of cyber diplomacy, but also how to align their national security interests with the personal interests of users in cyberspace. For businesses this could entail amongst others how to design their systems so that they are deemed trustworthy by as many people as possible, which is good for both the business and (potential) customers. And for individual users, especially those who are not too technologically well-versed, it could be beneficial to learn how to make proper judgements about the trustworthiness of nodes. Step-by-step this would allow cyberspace to live up to its promises.

BIBLIOGRAPHY

- Avižienis, A., Laprie, J. C., Randell, B., and Landwehr, C. (2004). Basic concepts and taxonomy of dependable and secure computing. *IEEE Transactions on Dependable and Secure Computing*, 1(1):11–33.
- Barbrook, R. and Cameron, A. (1996). The californian ideology. *Science as Culture*, 6(1):44–72.
- Barlow, J. P. (1996). A declaration of the independence of cyberspace. In Ludlow, P., editor, *Crypto anarchy, cyberstates, and pirate utopias*. MIT Press.
- Barnes, S. B. (2008). From Counterculture to Cyberculture: Stewart Brand, the Whole Earth Network, and the Rise of Digital Utopianism (review). *Technology and Culture*, 49(3):824–826.
- Benkler, Y. (2000). From consumers to users: Shifting the deeper structures of regulation toward sustainable commons and user access. *Federal Communications Law Journal*, 52(3):561–579.
- Braden, R. (1989). RFC 1122: Requirements for Internet Hosts – Communication Layers.
- Buzan, B. (1983). *People, States and Fear: The National Security Problem in International Relations*. Wheatsheaf Books.
- Buzan, B., Wæver, O., and De Wilde, J. (1998). *Security: a new framework for analysis*. Lynne Rienner Publishers.
- Cherdantseva, Y. and Hilton, J. (2013). A Reference Model of Information Assurance & Security. *2013 International Conference on Availability, Reliability and Security*, pages 546–555.
- Cudd, A. (2013). Contractarianism. In *The Stanford Encyclopedia of Philosophy*. Edward N. Zalta, winter 2013 edition.
- DiMaggio, P., Hargittai, E., Neuman, W. R., and Robinson, J. P. (2001). Social Implications of the Internet. *Annual Review of Sociology*, 27(1):307–336.
- Dunn Caveltly, M. (2014). Breaking the Cyber-Security Dilemma: Aligning Security Needs and Removing Vulnerabilities. *Science and Engineering Ethics*, 20(3):701–715.
- Ess, C. M. (2008). Culture and global networks: hope for a global ethics. *Information technology and moral philosophy*, 195:225.

- Ess, C. M. (2010). Trust and New Communication Technologies: Vicious Circles, Virtuous Circles, Possible Futures. *Knowledge, Technology & Policy*, pages 287–305.
- Floridi, L. (2011). *The philosophy of information*. Oxford University Press.
- Floridi, L., editor (2015). *The Onlife Manifesto*. Springer International Publishing, Cham.
- Fukuyama, F. (1995). Trust: The social virtues and the creation of prosperity. Technical report, Free press New York.
- Galloway, A. (2004). *Protocol*. The MIT Press.
- Gambetta, D. (1988). *Trust: Making and Breaking Cooperative Relations*. Blackwell.
- Gibson, W. (1984). *Neuromancer*. Ace Books.
- Glaser, C. L. and Kaufmann, C. (1998). What is the Offense-Defense Balance and Can We Measure it? *International Security*, 22(4):44–82.
- Greenwald, G. (2014). *No place to hide: Edward Snowden, the NSA, and the US surveillance state*. Metropolitan Books.
- Hannerz, U. (1992). The global ecumene as a network of networks. In Kuper, A., editor, *Conceptualizing Society*, pages 34–56. Routledge London.
- Hansen, L. and Nissenbaum, H. (2009). Digital disaster, cyber security, and the copenhagen school. *International Studies Quarterly*, 53:1155–1175.
- Harris, S. (2014). *@War: The Rise of the Military-Internet Complex*. Eamon Dolan/Houghton Mifflin Harcourt.
- Hughes, E. (1993). A cypherpunk's manifesto. In Ludlow, P., editor, *Crypto anarchy, cyberstates, and pirate utopias*. MIT Press.
- Jacobs, K. (2001). Utopia redux. In Ludlow, P., editor, *Crypto anarchy, cyberstates, and pirate utopias*, pages 349–352. MIT Press.
- Jervis, R. (1978). Cooperation under the Security Dilemma. *World Politics*, 30(02):167–214.
- Klimburg, A. (2012). *National Cyber Security Framework Manual*. NATO CCD-COE.
- Krol, E. and Hoffman, E. (1993). RFC 1462: FYI on "What is the Internet?".
- Kurose, J. and Ross, K. (2008). *Computer Networking: A Top-Down Approach*. Pearson Education International, fourth edition.

- Lessig, L. (2001). *The future of ideas*. Random House, New York.
- Lessig, L. (2006). *Code 2.0*. Basic Books.
- Luhmann, N., Davis, H., Raffan, J., and Rooney, K. (1979). *Trust; and, Power: two works by Niklas Luhmann*. Wiley Chichester.
- May, T. C. (1992). The crypto anarchist manifesto. In Ludlow, P., editor, *Crypto anarchy, cyberstates, and pirate utopias*. MIT Press.
- McLeod, C. (2014). Trust. In *The Stanford Encyclopedia of Philosophy*. Edward N. Zalta, summer 2014 edition.
- Möllering, G. (2001). The Nature of Trust: From Georg Simmel to a Theory of Expectation, Interpretation and Suspension. *Sociology*, 35(2):403–420.
- Möllering, G. (2006). *Trust: Reason, Routine, Reflexivity*, volume 4. Elsevier.
- Morozov, E. (2012). *The net delusion: The dark side of Internet freedom*. PublicAffairs.
- Nissenbaum, H. (2001). Securing Trust Online: Wisdom or Oxymoron. *Boston University Law Review*, pages 101–131.
- Nooteboom, B. (2006). Book Review: Guido Mollering: Trust: Reason, Routine, Reflexivity. *Organization Studies*, 27(12):1907–1910.
- Pettit, P. (1995). The cunning of trust. *Philosophy & Public Affairs*, 24(3):202–225.
- Pfleeger, C. P. (2000). Data security. In Ralston, A., Reilly, E. D., and Hemmendinger, D., editors, *Encyclopedia of Computer Science*, pages 504–507. Nature Publishing Group.
- Pickering, A. (2001). In the thick of things. *Keynote paper given to the conference 'Taking Nature Seriously', University of Oregon.*, pages 1–18.
- Pickering, A. (2008). New ontologies. In Pickering, A. and Guzik, K., editors, *The Mangle in Practice: Science, Society, and Becoming*, pages 1–14. Duke University Press.
- Potter, N. N. (2002). *How can I be trusted?: a virtue theory of trustworthiness*. Rowman & Littlefield.
- Putnam, R. D. (1994). *Making democracy work: Civic traditions in modern Italy*. Princeton university press.
- Rueter, N. C. (2011). *The Cybersecurity Dilemma*. PhD thesis, Duke University.

- Saco, D. (1999). Colonizing cyberspace: 'national security' and the internet. *Cultures of Insecurity: States, Communities, and the Production of Danger*, 14:261.
- Saltzer, J. H., Reed, D. P., and Clark, D. D. (1984). End-to-end arguments in system design. *ACM Transactions on Computer Systems*, 2(4):277–288.
- Singer, P. W. and Friedman, A. (2014). *Cybersecurity and Cyberwar: What Everyone Needs to Know*. Oxford University Press.
- Solum, L. B. and Chung, M. (2003). Layers principle: Internet architecture and the the law. *Notre Dame L. Rev.*, 79:815.
- Spinello, R. A. (2001). Code and moral values in cyberspace. *Ethics and Information Technology*, 3:137–150.
- Taddeo, M. (2009). Defining Trust and E-Trust. *International Journal of Technology and Human Interaction*, 5(2):23–35.
- Taddeo, M. (2014). The Struggle Between Liberties and Authorities in the Information Age. *Science and Engineering Ethics*, 5(3):342.
- Taylor, T. E. (2013). Well-Being and Prudential Value. *Philosophy & Public Policy Quarterly*, 31(2):10–17.
- Thorseth, M. and Ess, C. M. (2009). *Technology in a Multicultural and Global Society: Worldwide Communication Online*. LAP Lambert Academic Publishing, Germany.
- Turilli, M., Vaccaro, A., and Taddeo, M. (2010). The Case of on-line trust. *Knowledge, Technology & Policy*, 23(3-4):333–345.
- Turner, F. (2008). *From Counterculture to Cyberculture*. University of Chicago Press.
- Vallor, S. (2010). Social networking technology and the virtues. *Ethics and Information Technology*, 12(2):157–170.
- Van Leeuwen, J. (2014). On Floridi's method of levels of abstraction. *Minds and Machines*, 24(1):5–17.
- Waldron, J. (2003). Security and liberty: The image of balance*. *Journal of Political Philosophy*, 11(2):191–210.
- Walt, S. M. (1998). International Relations: One World, Many Theories. *Foreign Policy*, pages 29–46.
- Weckert, J. (2005). Trust in Cyberspace. In Cavalier, R., editor, *The Impact of the Internet on Our Moral Lives*, pages 95–117. SUNY Press.
- Whitt, R. S. (2004). A Horizontal Leap Forward : Formulating a New Communications Public Policy Framework Based on the Network Layers Model. *Federal Communications Law Journal*, 56(3).

- Williams, M. C. (2003). Words, images, enemies: Securitization and international politics. *International Studies Quarterly*, 47(4):511–531.
- WRR (2015). De publieke kern van het internet.
- Yamagishi, T., Kikuchi, M., and Kosugi, M. (1999). Trust, Gullibility, and Social Intelligence. *Asian Journal of Social Psychology*, 2(1):145–161.