A Basic Set of Mental Models for Understanding and Dealing with the CyberSecurity Challenges of Today

van den Berg, J.

Publication date
2020

Document Version
Final published version

Published in
Journal of Information Warfare

Citation (APA)
van den Berg, J. (2020). A Basic Set of Mental Models for Understanding and Dealing with the CyberSecurity Challenges of Today. *Journal of Information Warfare*, *19*(1). https://www.jinfowar.com/journal/volume-19-issue-1/basic-set-mental-models-understanding-dealing-cybersecurity-challenges-today

Important note
To cite this publication, please use the final published version (if applicable).
Please check the document version above.

A Basic Set of Mental Models for Understanding and Dealing with the Cyber-Security Challenges of Today

Jan van den Berg

Cybersecurity Chair
Delft University of Technology & Leiden University
Delft & Leiden, The Netherlands

E-mail: j.vandenberg@tudelft.nl

Abstract: *For most people, cybersecurity is a difficult notion to grasp. Traditionally, cybersecurity has been considered a technical challenge, and still many specialists understand it as information security, with the notions of confidentiality, integrity, and availability as its foundation. Although many have searched for different and broader perspectives, the complexity and ambiguity of the notion still thwarts a common understanding. While the author was developing and executing a MSc cybersecurity program for professionals with a wide variety of backgrounds and widely differing views on cybersecurity, the lack of a common understanding of cybersecurity was clearly evident. Based on these observations, the author began seeking and defining a new, transdisciplinary conceptualization of cybersecurity that can be widely agreed upon. It resulted in the publication of three scientific papers. This paper is an amalgam of the contents of the three supplemented with some extensions. It turned out that the previously introduced description of two key notions, cyberspace and cybersecurity, is still an adequate starting point. Described here is a set of additional mental models elaborating on these key notions and providing more detail on their meanings. The research suggests that this set of mental models strongly supports the description and analysis of current cybersecurity challenges and helps people understand how everyone, in his or her various roles, can contribute to reducing the related cyber risks. These claims are supported by presenting the modeling and analysis approaches of various MSc-thesis research projects executed by students when working on practical cybersecurity problems both within and outside their organisations. The author further discovered that, for a limited set of cybersecurity challenges, it was not yet possible to identify adequate mental models; this defines the agenda for future research.*

Keywords: *Cyberspace, Cyber Activities, Cybersecurity, Cyber Risk Management, Mental Models, Transdisciplinary View, Cyber Situational Awareness, Cyber Risk Mitigation*

Introduction

For most people, cybersecurity is a hard-to-grasp notion. Traditionally, cybersecurity has been considered as a technical challenge and still many specialists view it equivalent with Information

or Information Technology (IT) security, with the notions of 'Confidentiality, Integrity and Availability' (CIA) as starting points. Also, in information security standards such as the widely-acknowledged International Organization for Standardization/International Electrotechnical Commission (ISO/IEC) 27000-series (ISO/IEC JTC1 2018, 2005), the key asset chosen is usu-ally 'information', and the 'preservation of the confidentiality, integrity, and availability' of information is defined as the key information security challenge.

This rather abstract conceptualisation may be rather clear to IT specialists (like those working in hardware and software R&D where computers are general data and information processing devices), for policy makers, strategic managers, and end users; among others, this (classical) cybersecurity conceptualisation is difficult to grasp. Consequently, many actors in cyberspace have difficulties in defining what their role can and should be in securing the digital environment. In addition, these actors have often great difficulties in understanding each other's framing of their cybersecurity problems and the ways they should be tackled. This leads to the situation, in many cyber (sub-)domains, in which a coherent cybersecurity approach is missing. Based on this, there is arguably a need for a re-conceptualisation of what the cybersecurity challenge entails. More precisely the author asserts that there is a need for a broad transdisciplinary view on cybersecurity that everyone can grasp and that enables everybody to understand how he or she can contribute to securing cyberspace, in each of his or her cyber activity roles.

While developing and executing an executive MSc Program Cybersecurity for professionals (Cyber Security Academy 2014), the author worked on the creation of a holistic view of cybersecurity and discovered that mental models turn out to be very useful to create a common conceptualisation, understanding, and language about what cybersecurity essentiality is. This work resulted in two papers: van den Berg *et al*. (2014) and van den Berg (2018), in which two key notions of cybersecurity were brought forward (1) a clear conceptualisation of cyberspace, and, (2) a basic definition of what cybersecurity (that is, the security of cyberspace) is.

During further cybersecurity research as well as continued execution of the MSc program, these ideas were elaborated on by collecting all kinds of additional models and best practices in attempts to deepen the new conceptualisation. This resulted in a third paper—van den Berg (2019)—containing a first sketch of additional mental models that are thought to be most essential. The current paper is essentially an extension of the last paper with one extra mental model and a lot more details such as additional argumentation, examples, and references. As a result, this paper describes how far the research has come by sketching a basic set of mental models that are thought to be most essential. In addition, the author describes which cybersecurity basic mental models are still lacking. In order to validate the proposed set of essential mental models, the paper also describes some examples of cybersecurity research in which these models have been applied.

The remainder of this paper is structured as follows. The second section discusses a basic model of 'cyberspace', consisting of three layers and three supportive mental models, one corresponding to each layer. This creates the basis for describing in the third section what the 'cybersecurity challenge' essentially is using another basic model related to a cyber-risk management cycle, supplemented by a series of supportive mental models. Also identified are some gaps in the research body of cybersecurity knowledge, identifying three key topics of cyber in

the research needed in the future. The fourth section attempts to validate the proposed set of mental models the results from recent research in which these models have been applied. Finally, the fifth section draws conclusions and summarises future research topics.

Cyberspace and Its Security Concerns

Three-layer model of cyberspace

In an attempt to update and broaden the view of what modern cybersecurity entails as put forward in the ISO/IEC 27000 series, the ISO/IEC standard 27032 (ISO/IEC JTC1 2012) defines cyberspace as "the complex environment resulting from the interaction of people, software and services on the Internet, supported by worldwide distributed physical Information and Communications Technology (ICT) devices and connected networks". This framing is somewhat related to the Enterprise Resource Planning (Jacobs & Weston 2007) type of thinking where, in a layered approach, business processes, as executed by people and machines, are enabled by supportive IT services. The framing also relates to the purport of the well-known 'People, Process, Technology' triangle, where—in the context of software applications—people are split into end users (of applications) and application creators (that is, IT specialists), where (business) processes relate (and should be aligned) to the strategic business goals of the organisation (to be fixed by the strategic management), and where the supportive software applications enable better business decisions by relevant decision makers (Halo Business Intelligence 2009).

Inspired by the above-given frameworks of thinking that make an explicit distinction between technology (the IT) and people (using it or making decisions about it), the author has designed a new conceptualisation of cyberspace (van den Berg *et al*. 2014; van den Berg 2018; van den Berg 2019) consisting of a three-layer model, shown in **Figure 1** (below, left side) describing the basic cyberspace model. The middle layer describes the key assets of cyberspace and concerns the socio-technical layer of cyber activities as being executed by people in an attempt to reach their personal, business, or societal goals. This conceptualisation considers the activities performed using IT as the key assets, not the IT itself. Examples of cyber activity behavior include communication using one of the many available apps, searching on the World Wide Web, executing financial and other transactions, fundraising and crowdfunding on the WWW, manufacturing goods and products using design software and robots, controlling critical infrastructures (such as energy supply, water supply, and transport of goods and people), up until criminal cyber activities of all kinds (hacking into computer systems, stealing intellectual property, and selling illegal products on the dark web), law enforcement pursuits (related to the mentioned criminal cyber activities), and cyber warfare operations (such as cyber intelligence, defense, and attack) proliferated. Note that many of these activities (for example, financial transactions and control of critical infrastructures) are being executed by people in strong interaction with 'intelligent IT services' that (also) make all kinds of decisions. These intelligent IT services can also be considered part of the socio-technical layer.

The inner layer of the shown cyberspace model concerns all IT that enables cyber activities by providing all kinds of underlying supportive services, especially related to world-wide communication facilities. Putting the IT layer under the socio-technical layer shows that cyber activities are basically IT-enabled activities.

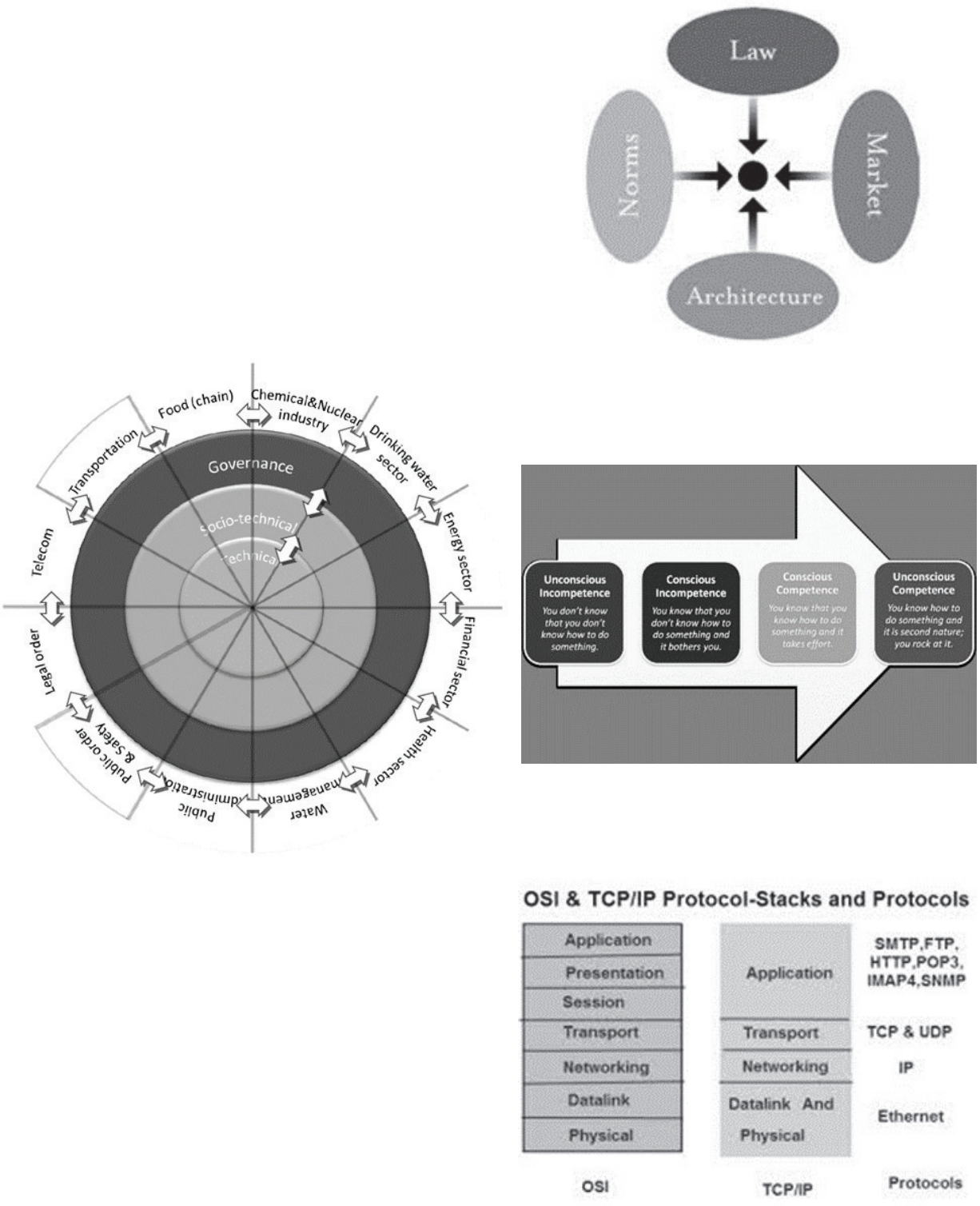


Figure 1: The 3-layer model of cyberspace (left) (van den Berg *et al*. 2014), and three models describing the key cyberspace issues per layer (right) (Lessig 1999/Surisetti 2017; Adams 2011/Steinberg 2020; Tanenbaum 1996/Cope 2019)

The outer layer of the model concerns the governance layer of rules and regulations that should be put in place to properly organise the two other layers, including their security. This relates, for example, to Internet governance issues next to rules and regulations that influence human behavior in cyberspace.

The three-layer model visualisation also shows a subdivision of cyberspace in cyber sub-domains to emphasise that cyber activities in different domains often have different characteristics and, consequently, different security requirements. Here, it is further asserted that the model can be applied at several levels of organisational aggregation, like the private environment at home, the small and medium enterprise level, the urban environment, state level, and multi-national company level, up to the continental and global levels.

Additional mental models for conceptualising cyberspace

Having defined the fundamental cyberspace model, it is helpful to formulate three additional mental models that define the basic cyberspace challenge per layer: see **Figure 1** (above, right side). The socio-technical layer is looked at first. Due to the continuous process of ubiquitous digitisation in all domains of society, the amount and variety of cyber activities people currently execute at home, when traveling, at work, and beyond, is enormous and is still growing. For many, it is quite challenging to adequately cope with the rapid digitisation developments and to stay competent as 'homo digitalis'. The basic challenge for adequate cyber behavior, which includes secure cyber behavior, may therefore be formulated as becoming and staying 'unconscious cyber competent'. This is visualised in the center, right side of **Figure 1**: the basic challenge is that every cyberspace actor, regarding all of his or her cyber activities, takes the path from the state of being 'unconscious incompetent', via 'conscious incompetent' and 'conscious incompetent', to the final state of being 'unconscious competent' (Adams 2011, Steinberg 2020). Since intelligent IT services are also part of the socio-technical layer, as previously noted, IT workers should also behave competently, meaning they can also be expected to make sound decisions. For example, a transaction server who is processing financial transaction is expected to execute legitimate financial transactions and to refuse to execute fraudulent ones.

To clarify the key issues of the governance layer, the author chose a mental model shown at the top, right side of **Figure 1**. It concerns the 'four modalities of regulation in cyberspace', as proposed by Lawrence Lessig (Lessig 1999, Surisetti 2017), as being 'laws' (next to rules, policies, and regulations), 'norms' (informal societal rules), 'markets' (to create the right incentives for stake-holders), and 'architecture' (which concerns physical or technical constraints on cyber activities). It should be clear that this framing of the four modalities of regulation is precisely in line with the three-layer model of cyberspace: the modalities (laws, norms, and markets) steer cyber activities (in layer 2) from a governance perspective (in layer 3), while the modality architecture (in layer 1) puts constraints on the cyber activities one executes using a technical approach. An example of the latter can be found in role-based access control: depending on a person's role(s) in an organisation, he or she is granted access to a specific set of services and applications while others are made inaccessible to him or her.

For the technical (IT) layer, the key issues relate to the two protocol stacks used to describe computer networks, namely the OSI and TCP/IP protocol stacks. The working of the TCP/UDP and IP

protocols at, respectively, the transport and network layer are especially essential to understand the ways worldwide communication is enabled. In addition, the application layer is interesting to know about with networking applications such as electronic mail, file transfer, and the World Wide Web services (Tanenbaum 1996; Cope 2019). These networking applications are often embedded in the real applications end users execute by means of Applications Program Interfaces (APIs). As a final observation related to understanding fundamentals of current digital technology, the frequently made distinction between Information Technology (IT) and Operational Technology (OT) needs to be mentioned. Here, IT relates to 'traditional' information technology (related to data and information processing, mostly in an office environment); while in OT, computers are used to control (often complex) physical processes in a factory environment (for example, "power plants, oil rigs, manufacturing assembly lines, and inventory management processes") with technologies like ICSs, SCADA, and PLCs (Coolfire, 2019). However, due to new technological developments, such as the Internet of Things (IoT), the distinction between IT and OT is becoming less strict but is still important to understand since the related cybersecurity challenges are quite different. This will be further discussed below.

The mentioned protocol stacks and IT/OT distinctions are at the core of what every cyberspace actor should understand, to a certain extent, to become 'unconscious competent' in his or her cyber activity role, or to be able to make sound cyber governance decisions.

Security concerns of cyberspace layers

Having conceptualised cyberspace in three layers, it is possible to determine the security concerns per cyberspace layer and per cyber subdomain. First, if one reconsiders the security concerns of the key assets in cyberspace being the cyber activities, it can be said that that cybersecurity essentially concerns the security of cyber activities, which is about the security of cyber behavior. It is evident that the security requirements of a cyber activity strongly depend on the type of the activity and its context (that is, the cyber subdomain in which it is executed). For example, the requirements related to the execution of a financial transaction in a public environment (primarily in the IT environment) relate to secure payment behavior: careful use of debit/credit card, checking the amount before paying, shielding the keyboard of the payment equipment while typing the pin number, and inspecting the receipt for errors. When considering the automatic control of a critical infrastructure, such as water supply (a clear OT example), the cybersecurity requirements are very different and primarily focus on guaranteeing continuous automatic supply of clean water to recipients, monitoring this process, and committing necessary interventions "through SCADA systems attached to Distributed Control Systems (DCS), Programmable Logic Controllers (PLCs), Remote Terminal Units (RTUs) and field devices" (ISACA 2016). It can be stated in general terms that issues of secure cyber activity behavior also relate to prudent behavior in social networks: for example never clicking on a URL in an email; minimising or discontinuing the use of USB drives; always choosing strong passwords; protecting and storing passwords adequately; limiting or stopping the downloading of files from the Internet; making backups on regular basis; protecting electronic devices properly; and carefully monitoring transactions, supply, production, and delivery processes; and activities on the dark web, among others.

For the technical layer, the classical requirements of information security are Confidentiality, Integrity, and Availability (CIA) (ISO/IEC JTC1 2018). For the financial payment example mentioned above, all three CIA requirements are relevant, while more refined technical requirements might be added here like secure Identification, Authentication and Access (IAA) control; the use of anti-virus software; automatic patching; and real-time monitoring of transactions and of automated control activities. More details on this are given below

As a consequence of separating cyberspace into the three layers previously mentioned, the author's approach discriminates between cybersecurity (being the security of cyber activities/ behavior in layer 2) and information security (being the security of IT/OT in layer 1), a distinction usually not made in current practice.

Continuing this line of thought, it is further observed that incidents involving the technical layer (often termed information security breaches) are actually cyber threats to the cyber activities executed in the socio-technical layer. If such cyber threats, emerging as information security incidents in the technical layer, also result in incidents in the socio-technical layer, this can be termed cybersecurity incidents or cybersecurity breaches, which again show an important difference in meaning of 'cyber' and 'information'. In short, within this conceptualisation of cyberspace and cybersecurity, information security (with a focus on IT) is truly different from cybersecurity (with a focus on behavior).

Finally, the security concern of the governance layer encompasses the establishment of rules and regulations for both the socio-technical and the technical layers; this is accomplished using the governance modalities discussed previously and in accordance with the chosen 'risk appetite', which is discussed below. So once again, governance rules and regulations should be related to both secure cyber activity behavior (cybersecurity in layer 2) and secure IT/OT (information security in layer 1). Principal stakeholders in this instance tend to be board members of organisations and legislative entities in governmental institutions of nation states.

Modelling the Cyber Security Challenge

Before diving into the cybersecurity challenge, which arguably concerns a risk management challenge, it is relevant to examine the modern notion of 'risk'. According to modern standards, risk is the 'potential for gaining or losing something of value', or according to ISO/TC 262 (2018), the positive or negative "effect of uncertainty on objectives". In the financial world, this phenomenon is well known since investments can result in an actual return that is higher (opportunity) or lower (loss) than the expected return. In cyberspace, similar symptoms are observed since digitisation usually offers expected opportunities (such as efficiency, cost reduction, and convenience, for example). At the same time, it also enhances the 'cyberattack surface', which creates higher cyber risks. For cyber-security decision makers it is always wise to take this two-sided view of cyber risk into account (which helps to properly balance expected digitisation opportunities and losses); the focus of the discussion in the remainder of this paper is mainly on the negative part of cyber risks.

Bowtie model and cyber risk management cycle

The bowtie model is a well-known basic model used in safety and security science. Going from left to right in the figure below, the model considers (intentional and unintentional) threats, incidents,

and finally the impact of the latter. Threats may result in incidents. Incidents occur with a certain probability or likelihood, and the risk of an incident is defined as the expected impact of this incident, that is, risk = likelihood x impact. In cyberspace, the bowtie model can be used to model cyber threats, cyber incidents, and their impact. To avoid cyber incidents, 'preventive' measures can be taken to reduce the probability of their occurrence. To reduce impacts of a given incident, 'repressive' measures (like measures related to detection and recovery) can be taken. For more details on (the use of) the bowtie model, see United Nations (2015). **Figure 2** (left side), provides a visualisation of the bowtie.

Repeat 'forever'
(in all 'relevant' cyber sub-domains)
1. Identify the critical cyber activities (sometimes termed the 'crown jewels')
2. Identify & assess their cyber risks (potential gains & losses)
3. Define acceptable cyber risk levels
4. Decide way(s) of dealing with the risks
5. Design & Implement cyber risk measures
6. Monitor effectiveness of measures taken

Figure 2: The bowtie model (left) (van den Berg et al. 2014), and the basic cyber risk management cycle (right) (ISO/TC 262 2018)

Many cyber activities are intentionally threatened by a variety of actors. These range from incompetent end users to script kiddies, ethical and unethical attackers, organised criminals acting on the dark web, and various agents of nation states. There are also unintentional menaces (for example, technical failures, human errors, and natural disasters) and related cyber incidents that may occur, sometimes with high impact. This explains why cybersecurity is actually a risk management challenge. Here, again, standards like those of the ISO (ISO/TC 262 2018) can help. It follows that, for proper risk management, a risk management process should be implemented. This risk management cycle is depicted on the right in **Figure 2**, in the context of securing cyberspace. Once again, within this framework, cybersecurity primarily entails risk management of cyber activities. Since the characteristics of cyber activities in different cyber subdomains often vary substantially, and therefore the related cyber risks vary as well, cyber risk management processes should be adapted to the context in which they are executed.

Additional models for conceptualising cyber risk management

Having defined and visualised the basic mental cyber security model and related cyber risk management cycle, it is now possible to sketch a set of additional models that provide background details that deepen these basic ideas. This is done by considering each of the six steps of the basic cyber risk management cycle. The first step concerns the identification of the critical cyber activities as executed by a person, organisation, or society. The critical cyber activities are the IT-enabled activities most frequently depended on and are, if disrupted, expected to have the highest impact. In society, critical cyber activities are related to critical infrastructures, like transportation (of goods and people), supply of water and energy (electricity, oil, gas), as well as financial, healthcare, and first responder services. In a digitalised corporate environment, data and infor-

mation are often considered as critical and are sometimes termed the "crown jewels" (Fredriksen 2018). Within the conceptualisation of cyberspace introduced in this paper, not their data per se, but the critical cyber activities of a state, corporation, or individual are considered as "crown jewels" (which are usually related to critical infrastructures, critical business processes, or important personal cyber activities); these need to be cyber secured with the highest priority. For a visualisation of the crown jewels, refer to **Figure 3**, below, left side (Royal Exhibitions 2020). However, before thinking about their security, they should be first identified and defined, which is still not a common practice in many organizations or for most individuals, as has often been observed.

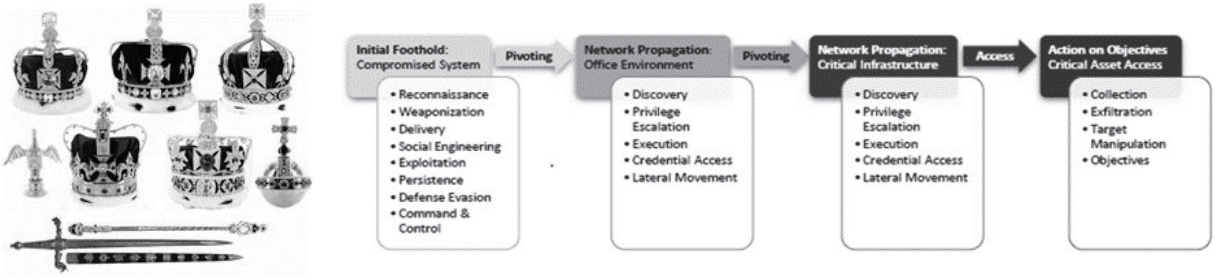


Figure 3: Two additional mental models related to the first two steps of the basic cyber risk management cycle: crown jewels (left) (Fredriksen 2018; Royal Exhibitions 2020) and the unified kill chain (right) (Pols 2018)

Once a person, organisation, or society has identified his/her/its critical cyber activities, the second risk management cycle step, identifying and assessing cyber risks, can be addressed. For unintentional disruptive threats, it is important to analyse the consequences of all possible technical failures, human errors, and natural disasters. When considering intentional attacks, the unified kill chain model can be applied for analysing and defending against possible attack behavior: this model (visualised in **Figure 3**, right) describes in detail all possible steps an attacker can choose in attempts to disrupt critical cyber activities and, therefore, can be used to identify, understand, and classify existing and upcoming attack strategies (Pols 2018).

Such analyses are only possible if it is precisely understood how the cyber activities are taking place on an IT system connected to the Internet, for example, in terms of normal and abnormal behavior and in the ways the activities are threatened. In more formal terms, it can be said that there is an urgent need to increase cyber situational awareness. The general notion of situational awareness was introduced in 1995 by Micah Endsley (Endsley & Jones 2016) and applies here in the context of securing cyberspace (**Figure 4**, below)(Lankton 2007). With respect to cyber situational awareness, some steps have been taken in multiple domains. For example, it is quite common for the execution of financial transactions to be monitored in real time, and many organisations have already established or plan to establish (Information) Security Operations Centres (SOCs). These centres "monitor and analyse activity on networks, servers, end-points, databases, applications, websites, and other systems, looking for anomalous activity that could be indicative of a security incident or compromise. An SOC is responsible for ensuring that potential security incidents are correctly identified, analysed, defended, investigated, and reported" (Lord 2019). In addition, law enforcement organisations, like the police, actively monitor operations on the dark web (and are sometimes successful in finding the criminals behind malicious activities), and nation-state intelligence organisations are constantly looking at

what threatening cyber activities other states are executing, like distribution of fake news to influence people's behavior (for example, to disrupt democratic elections), and include cyber intelligence as well as defensive and offensive operations in the military domain. However, at other places in cyberspace, especially within many OT environments, the creation of sufficient cyber situational awareness is often still in its infancy.

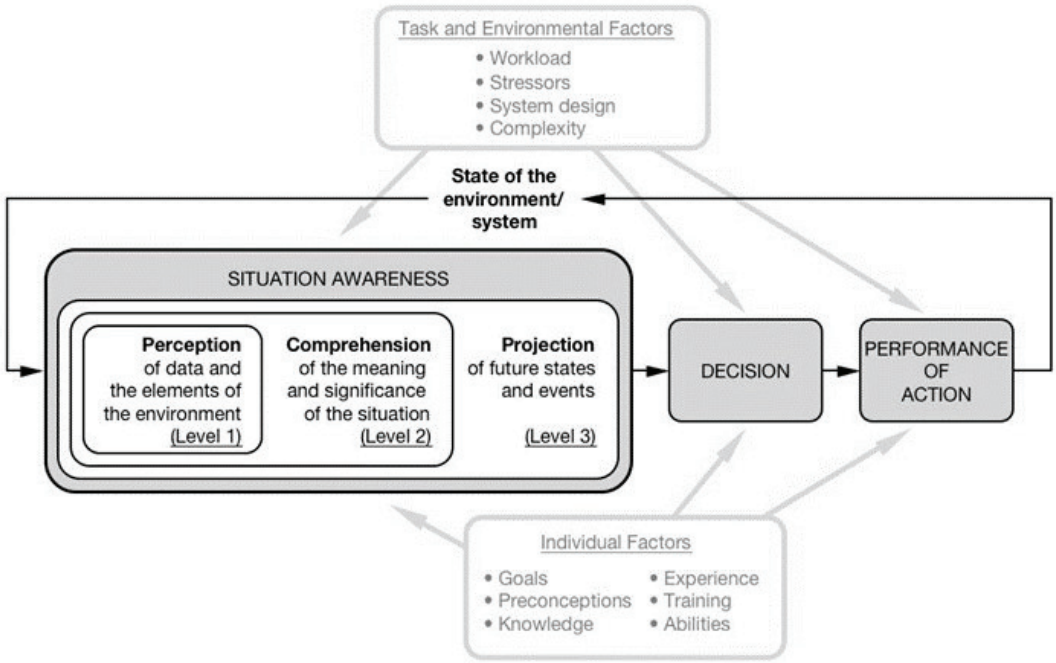
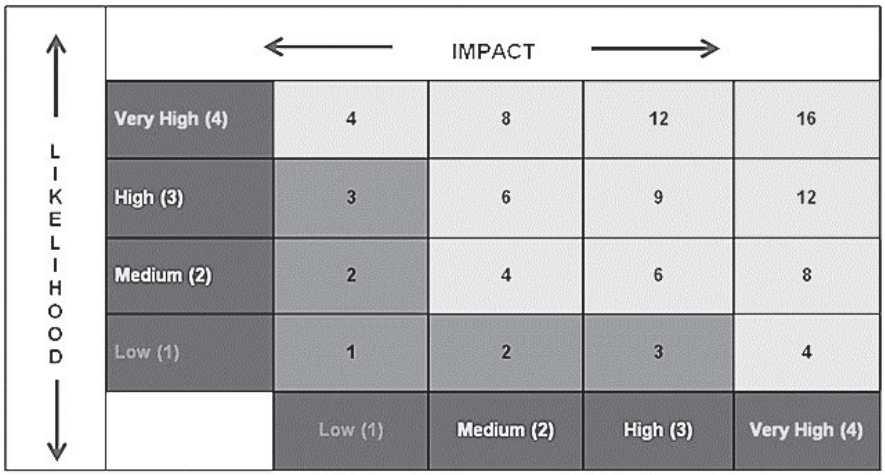


Figure 4: An additional mental model related to the second step of the basic cyber risk management cycle: (cyber) situational awareness (Endsley & Jones 2016; Lankton 2007)

Having created adequate cyber situational awareness, one can try to assess the risk related to possible cyber activity incidents, usually, in terms of Likelihood x Impact. There are numerous (both qualitative and quantitative) methods available to make such assessments (for an overview refer to ISO/IEC TC 262 [2009]); however, in practice it often turns out to be a difficult task, due to, for example, lack of relevant data. **Figure 5**, below, depicts a framework following this basic idea—the 'risk matrix' (Caldes 2016). The Risk Scores shown in the matrix cells (1, 2, 4, 6, …, 16) have been calculated here as the product of LIKELIHOOD (ranging from low (1), via medium (2), high (3) to very high (4)) and IMPACT (using the same range numbers 1 to 4).This figure completes the set of four additional mental models related to the first two steps of the basic cyber risk management cycle.



Risk Score	Rating
0 – 3	Low
4 – 6	Medium
6 – 9	High
10 – 16	Very High

Figure 5: Yet another mental model related to the second step of the basic cyber risk management cycle: cyber risk assessment using the risk matrix model (ISO/IEC TC 262 2009; Caldes 2016)

The next step of the basic cyber risk management cycle concerns determining 'acceptable cyber risk levels' for each of the critical cyber activities. This relates the so-called 'risk appetite' of an individual, organisation, or society. Defining the risk appetite falls, for the most part, outside the scope of science since it concerns a choice and is often based on personal judgments and insight. However, some remarks are relevant here. For critical infrastructures, governments often determine the required risk levels, as is common practice in the worlds of finance, energy supply, flood defense, and IT systems, for example. Being non-compliant with the related rules and regulations can result in severe penalties, which of course influences the risk appetite of an organization. Similarly, shareholders have their own ideas about managing (cyber) risk and the related risk appetite, so their voices are often decisive in the risk appetite choice of an organisation.

Having assessed the relevant cyber risks and having chosen the acceptable cyber risk levels, the fourth step concerns the decision of how to deal with the assessed risks. A well-known principle from safety and security science is that there exists basically four response strategies to negative risk (Dorfman 2007). The first is 'avoidance'—stopping the risky cyber activity at stake. The second one is 'transfer'— making another party responsible for the risk through insurance or out-sourcing. The third option is simply 'accepting' the risk in case it falls within the designated risk appetite. The final response strategy is 'mitigating' the risk to the defined acceptable risk level by reducing the probability and/or impact of a cyber threat. The four possible 'risk response strategies' are shown in **Figure 6**, below (Rowley 2014).

Step five of the basic cyber risk management cycle concerns the design and implementation of cyber risk measures, which are relevant in case organisations have adopted the strategy of risk mitigation in the previous step. This concerns a complex challenge since an abundance of preventive and repressive mitigation measures exist. Within the author's conceptualisation of cyberspace, the challenge boils down to designing and implementing a balanced set of cyber risk mitigation measures in all three cyberspace layers. A simple example may illustrate the basic idea. Consider the case of using USB drives, which is often a risky cyber activity since

malware can be easily and quickly transferred if these drives are used. Measures to mitigate such a malware infection risk at the socio-technical layer concern measures related to cyber activity behavior—someone who feels himself or herself a potential target for a malware infection attack via a USB drive might decide neither to use such a device, nor to allow anyone else to use it on his or her PC or laptop. The identified USB infection threat might also be mitigated at the technical layer by disabling all USB ports in the IT environment at stake or, in a less restrictive approach, by monitoring USB drive connections and scanning for infections before allowing data retrieval from such drives. Finally, at the governance layer, rules might be made official that USB drives are not allowed inside a certain IT environment and, in case of a cyber incident occurrence due to a violation of this rule, the person who violated the rule (financial or other) will be responsible for any consequences derived as a result of breaking this rule.

Figure 6: A mental model related to the fourth step of the basic cyber risk management cycle: risk response strategies (Dorfman 2007; Rowley 2014)

In practice, cyber risk mitigation is usually a much more complex challenge than the simple example shows. Once again considering critical infrastructures in a digitised society, it is immediately observed that usually many stakeholders (often a combination of public and private actors) are involved, each one with specific responsibilities for the risk mitigation challenge. This thwarts the design and implementation of balanced risk mitigation approaches that are both effective and efficient. Here the author identified a vast research topic for the near future since little attention has been given thus far to the cyber risk mitigation challenge of designing and implementing balanced sets of cyber risk mitigation measures.

However, one is not left completely empty-handed with respect to filling this knowledge gap. For example, several best practices have existed for years in the technical domain. To effectively monitor and manage data traffic in complex computer networks, an important prerequisite is a transparent IT/OT architecture, with clear compartments or, in other words, "when managing security domains, the IT environment should be classified into discrete, logical entities that ease man-

agement activities (granularity) and minimise negative impact (compartmentalisation)" (Gibbs 2007). In line with this, a common practice in networking environments has been to separate the IT and OT environments, although this approach is currently undermined due to further digitisation of OT, its integration with IT, and the emergence of IoT. Therefore, best practices are currently being developed to unify IT and OT network security solutions. For a detailed discussion on this, see Skybox Security (2018).

Finally, in his book on computer networks, Andy Tanenbaum asserts that "every layer (in the OSI and TCP/IP-model) has something to contribute" (1996). This view is strongly related to the more general 'Swiss Cheese model' created by James Reason in 1997 (Skybrary 2016). The model has its origin in safety science (Ly & Thimbleby 2014), which emphasizes that a layered defensive approach is needed in order to deal successfully with existing vulnerabilities. In **Figure 7,** vulnerabilities in the layers of defense are visualized by holes, and attack trajectories by straight arrows. It is further shown that one attack trajectory (lover arrow) is not, since it is blocked by the second layer of defense. Due to its clearness and simplicity, the author included it in the basic set of mental models described in this paper.

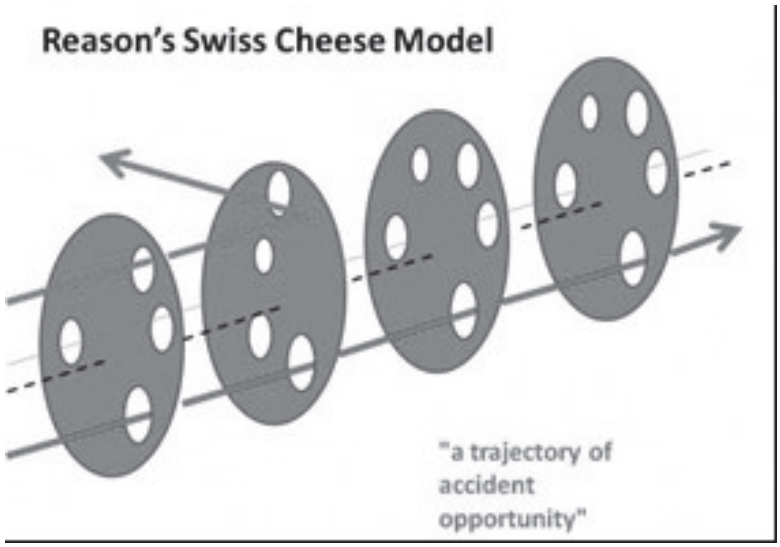


Figure 7: A general mental model related to the fifth step of the basic cyber risk management cycle: the Swiss cheese model of accident causation (Skybrary 2016)

At the socio-technical layer, cyber risk measures seek to understand and promote behavior change of cyberspace actors based on an analysis of what security behavior these actors are expected to perform, their behavioral determinants, and what behavioral change approaches can be applied in order to be successful (Blythe 2013). This is rather new field of study, where the psychology perspective is relevant. For example, it has been observed that cyber security awareness training campaigns are often not effective because "changing behavior requires more than providing information about risks and reactive behaviors—firstly, people must be able to understand and apply the advice, and, secondly, they must be motivated and willing to do so—and the latter requires changes to attitudes and intentions" (Bada, Sasse & Nurse 2015). The time has arrived to further take up these research challenges using a psychology perspective in order to better understand how cyberspace actors can be trained to become sufficiently 'unconsciously competent'.

At the governance layer, cyber security rules and regulations, among others, should be put in place within organisations, within society, and in the personal lives of cyberspace actors. It is worth noting again that, due to the rapid digitisation of society, actors such as corporate board members, politicians, and educators/parents often have great difficulties in defining effective rules of the cybersecurity 'game'. Fortunately, some additional models are available that can be applied to designing solutions. For example, to implement am often-stated need for Public-Private Partnership (PPP) in cyberspace, models from situational economics are relevant. The first, a visualisation of which is given in **Figure 8** (left), relates to the "institutional design for complex technological systems" in such a way that "socially desired objectives are realized" (Koppenjan & Groenewegen 2005). This paper discusses how "arrangements between actors that regulate their relations, that is tasks, responsibilities, allocation of costs, benefits and risks, can be designed" (Koppenjan & Groenewegen 2005). So this theory is also applicable for institutional design of solutions for secure cyberspace. In addition, the author observed that, in practice, PPPs also emerge spontaneously, especially in cases of tough cybersecurity challenges related to potentially high cyber risks. For example, in order to cope with large-scale Distributed Denial of Service (DDOS) attacks that threatened the Logius identity management platform used by crucial government agencies in the Netherlands, multiple governmental organisations and Internet Service Providers (ISPs) collaboratively created a relatively simple technical solution called the 'Quality Peering Platform' (Santana 2019; Grutter 2019). Additionally, in other instances of significant cybersecurity incidents occurring, the sudden emergence of PPPs in cyberspace was observed, for example, during the DigiNotar affair (Hoogstraaten 2012) and the Hansa case (Greenberg 2018). However, despite these positive examples of spontaneous PPPs emerging during cyber incidents, examples of structural, institutionalized PPPs are still relatively rare.

A second governance model that is helpful in designing cybersecurity solutions relies upon 'social contract theory' (Bierens, Klievink & van den Berg 2017). This paper argues that for cyberspace (as for the physical domains of land, water, and air), appropriate social contracts must be established. In their approach, a distinction has been made between a direct social contract (between citizens and the government) and an indirect social contract (between citizens and the government but via private organisations), a visualisation of which is given on the right side of **Figure 8**, below. The new cyber social contracts should explain which rights and freedoms and under which conditions individuals consent to surrender to which governmental bodies, corporations, and other organisations, in exchange for protection of their remaining rights and freedoms. This entails, for example, under which specific conditions (related to suspicious negative cyber activities such as fraud, distribution of child pornography, and extortion), personal privacy should be waived to allow the police to analyse personal data related to these activities, to eavesdrop on personal smart phones, or (even) to break into laptops and servers to search for evidence of those activities.

Step six of the basic cyber risk management cycle must also be considered; in this instance, it is essential to monitor the effectiveness of the measures taken. Here the author asserts that this strongly relates to the creation of cyber situation awareness as discussed previously. Measuring the effectiveness of measures taken is challenging, but through smart monitoring of cyber activities in the socio-technical layer and in the IT processes of the technical layer, it is possible to gain significant insight. Many of the examples previously discussed are relevant in this instance.

For example, it has become commonplace to monitor and analyse, in real time, a wide variety of financial transactions in attempts to prevent fraud. In addition, national intelligence services are actively monitoring cyber activities of foreign states and state-sponsored actors related to, for example, digital espionage and attempts to hack global international institutions such as the Organisation for the Prohibition of Chemical Weapons (OPCW) (GOV-NL 2018). Meanwhile, police and other law enforcement organisations throughout the world are actively monitoring the dark web in search of a variety of illegal activities (recall, for example, the Silk Road [US Department of Justice 2014] and Hansa [Greenberg 2018] cases). Finally, computer network researchers are monitoring and analysing a wide variety of Internet traffic to, for example, discover new threats based on new emerging attack tools. Since such monitoring activities are closely tied to the creation of cyber situation awareness as previously discussed, there is no need to introduce an additional mental model.

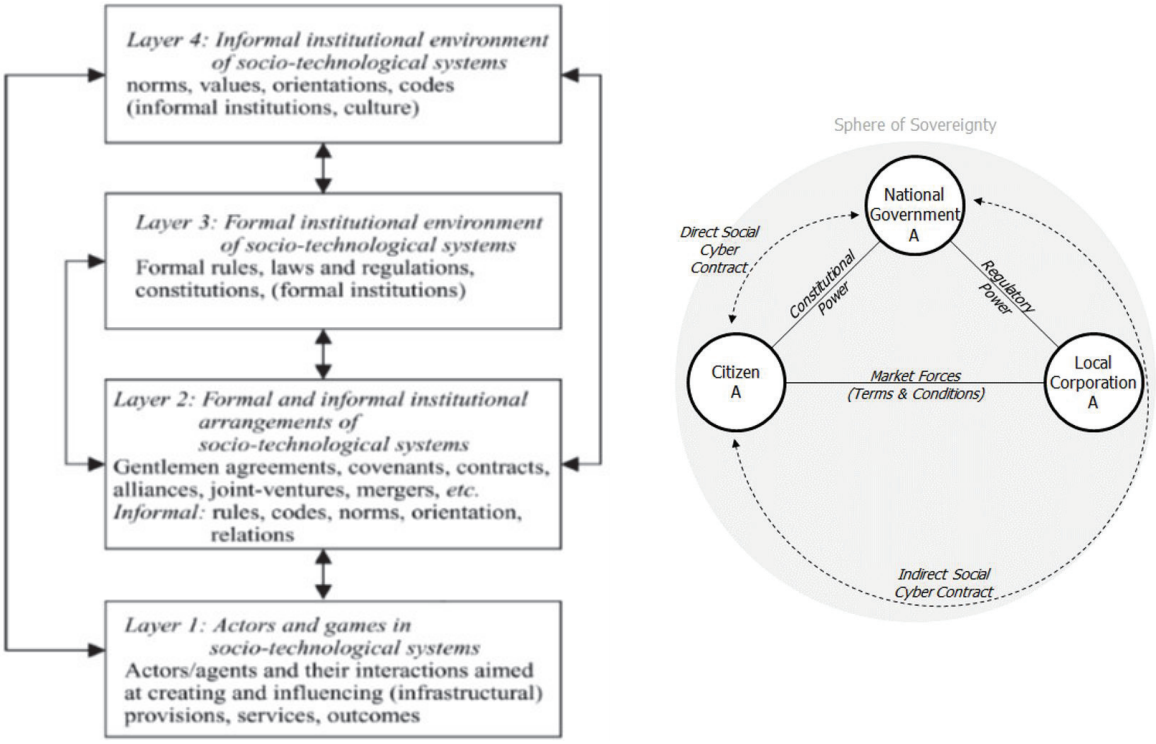


Figure 8: Two additional mental models related to the fifth step of the basic cyber risk management cycle: the four-layer diagram of institutionalisation (left) (Koppenjan & Groenewegen 2005) and the direct and indirect social contract model (right) (Bierens, Klievink & van den Berg 2017).

Key Challenges for the Near Future

Reconsidering the analysis results provided above, this section concludes by describing three key challenges for the near future to arrive at adequate cyber security levels.

1. Creation of Cyber Situation Awareness: although some progress has been made, the state of the art of understanding what happens in cyberspace is still insufficient. It is often said that states, organisations, and individuals have limited insights on the (in)correctness of relevant cyber activities and, consequently, the related cyber risks. Cyber situation awareness is crucial for understanding cyber risks, for measuring the effectiveness of cyber risk mitigation measures, as well as for dealing with fundamental dilemmas like those between cyber opportunities and negative cyber risks (discussed above), and between privacy and cyber security. With respect to the latter, if cyber risks are high (for example, socially disruptive), people tend to be willing to accept privacy limitations to help law enforcement agencies catch the perpetrators.

2. Methodologies for Arriving at a Balanced Set of Cyber Risk Mitigation Measures: being aware of the possibility of taking or applying various cyber risk mitigation measures in the technical, socio-technical, and governance layer of cyberspace, efficient and effective methodologies and best practices for selecting a balanced set of those measures do not yet exist. This is considered an important research challenge for the near future.

3. Implementing Structural Public-Private Partnerships (PPPs) for Securing Cyberspace: although there is a growing number of emerging initiatives of increased cooperation in various cyber subdomains, one might say that—as compared to PPP implementations in physical world domains (land, water, air, and space)—the development of those for securing cyberspace is still in its infancy. This may be caused by the enormous complexity of cyberspace with almost 4 billion people connected via the World Wide Web on the one hand, and a few big players on the other (for example, the 'Big Five' tech companies). However, governments cannot escape their responsibilities, as part of their social contracts with their citizens, to mutually strive to build PPPs for achieving a more secure fifth domain of cyberspace. Researchers can support this development by recommending suitable relationships between relevant cyberspace stakeholders.

Use of Mental Models

As previously discussed, the set of mental models introduced in this paper has been collected based on lecture discussions and research collaboration with cybersecurity professionals while they pursued an executive master's degree (Cyber Security Academy 2014). In an attempt to validate the choice of models presented, the use of several mental models during the execution of the research of some students when composing their final thesis was discussed.

To begin an elucidation, it is worth noting that the cyberspace model of three layers has been often applied, for example, to structure the results of an analysis of the security of eHealth services of Dutch General Practitioners (Willems 2017) by mapping the communication between patients and general practitioners as the core cyber activities in the socio-technical layer, enabled by Transport Layer Security (TLS) services in the IT layer, and by positioning the relevant governance actors as well as the relevant health rules and regulations in the governance layer.

In other research, the three layer cyberspace model was used to structure a large set of requirements for designing a multi-stakeholder roadmap to implement ISP-based consumer vulnerability management in the in-home domain (Bastiaanse 2018). To do so, Bastiaanse analysed modern in-home environments in terms of cyber activities as being executed at home by adults and children in the socio-technical layer, by describing technical equipment (like network termination points, modems and connected devices such as laptops, [game] computers, tablets, solar panels, and smart phones) in the IT layer, and by paying attention to the laws relevant to the in-home domain (like the telecommunications laws and laws related to computer criminality, among others) in the governance layer.

In still another thesis, the cyber governance model of Lessig (1999) for cyber regulations modalities was applied to analyse the feasibility of three policy strategies for the government to make smartphone Virtual Private Network (VPN) services by default available for consumers and to select the most optimal policy strategy of these three policies (Ghaoui 2017).

As a final example, it is noted that a thesis has been written on the design of a model for cyber security supervision of 5G in the Netherlands to provide legal certainty and normative clarity in this upcoming complex digital environment. In this thesis, both the bowtie model and the three-layer model of cyberspace have been applied and integrated in one conceptual supervision model, which is shown in **Figure 9**, below:

The conceptual 'triple bowtie' supervision model for 5G cyber security in the Netherlands incorporates a risk-driven approach to supervision in the 5G 'building blocks' layer (in the technical layer), the 5G Provision Layer (also in the technical layer), and the 5G 'use case' layer (in the socio-technical layer) (on the right in the figure) and cooperation between relevant supervisory bodies (in the supervision layer) (on the left in the figure) to share information (for example, on norms, threats, risk assessments, cascading effects, interventions) and create cyber situational awareness in order to better ensure that NL can trust 5G (Wazir 2019).

For more details on these and other uses of the basic set of mental models mentioned in this paper, refer to the related master's degree theses, most of which are available online by clicking the 'Theses' link on the Cyber Security Academy website landing page (Cyber Security Academy 2014).

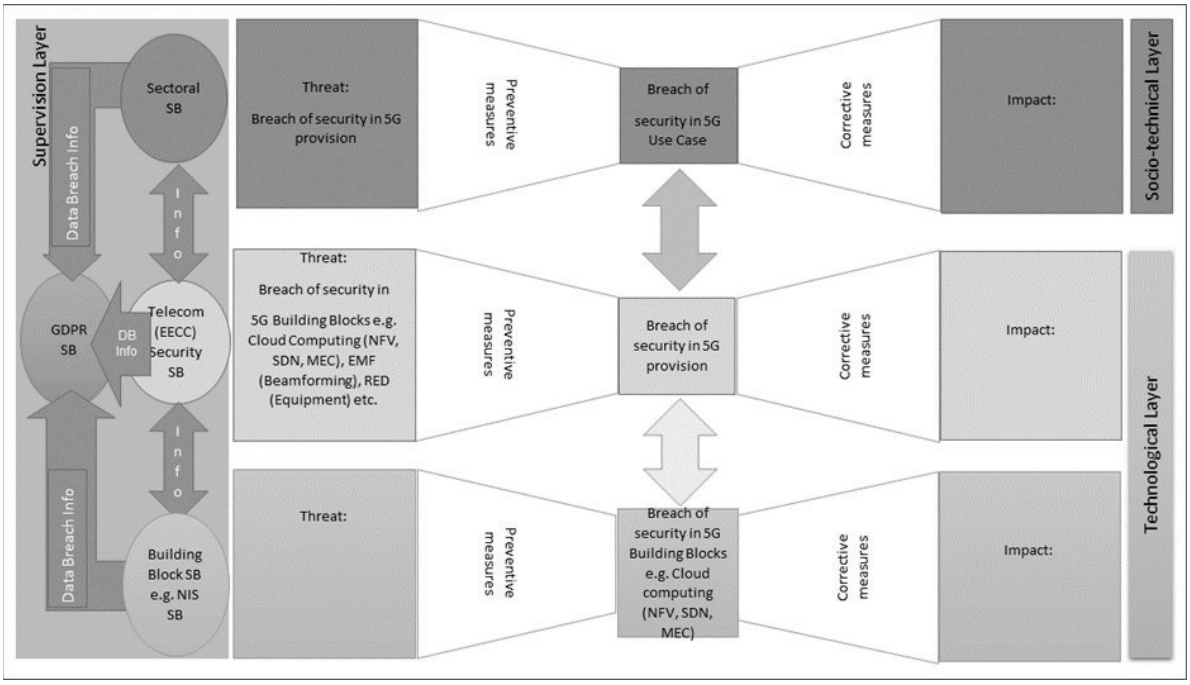


Figure 9: The triple bowtie cybersecurity supervision model with the 5G use case layer, the 5G provisioning layer, and the 5G building blocks layer (Wazir 2019)

Last but not least, it is important to note that the opposite has also occurred; namely, the related MSc-research resulted in a new fundamental mental model that can used by others; the 'unified kill chain' model (Pols 2018) mentioned above is an excellent example of this phenomenon (and, actually, is worthwhile to write a separate paper about).

Conclusions and Future Research

This paper has presented a set of essential mental models that, taken together, cover the key elements of cyberspace and can be used to secure cyberspace, that is, to implement cybersecurity. They provide an overview on what cyberspace basically entails and what the related cybersecurity challenges encompass. In day-to-day discussions between cybersecurity professionals and during the execution of cybersecurity research, these models and the related frameworks of thinking have proven useful and effective.

The choice of making an explicit distinction between IT and OT and underlying computer network services (together the enabling technologies of cyberspace layer one) on the one hand and the use of it by end users and intelligent applications (in terms of cyber activities in layer two) on the other is crucial for understanding the difference between classical information security (or IT security) and cybersecurity. Cybersecurity is the new notion and concerns the challenge of sufficiently securing our numerous cyber activities, all the things humans do with modern IT/OT. The consequence of this framing is that everyone can now understand what cybersecurity is about since cyber activities are started by people, and the conceptualisation avoids a one-sided focus on the security of IT/OT, which for many people is difficult to grasp

The second model with underlying mental models emphasizes that cyber security is essentially a cyber risk management challenge. Here, risk is considered a two-sided notion of being both an opportunity (positive risk) and a potential loss (negative risk). This opens the door for discussions on simultaneously assessing the potential advantages and disadvantages of (further) digitisation. In this way, cybersecurity is not just a cost factor but also a factor of potential profit, which encourages discussions on cybersecurity at the corporate board level. Finally, it is argued that, in one way or another, the design and implementation of the basic cyber risk management cycle is crucial for trying to create a sufficiently secure cyber environment in the cyber subdomain at stake.

During this research journey, the author encountered three key challenges related to the implementation of the cyber risk management cycle that need further investigation, namely, the creation of well-established cyber situation awareness in all cyber subdomains, the design of methodologies for arriving at a balanced set of cyber risk mitigation measures in these subdomains, and the further implementation of structural PPPs for cyberspace.

References

Adams, L 2011, *Learning a new skill is easier said than done*, Gordon Training International, viewed 13 November 2019, <https://www.gordontraining.com/free-workplace-articles/learning-a-new-skill-is-easier-said-than-done/>.

Bada, M, Sasse, AM & Nurse, JRC 2015, *Cyber security awareness campaigns: Why do they fail to change behaviour?*, *Proceedings of the International Conference on Cyber Security for Sustainable Society*, Cornell University ArXiv, viewed 2 Jan 2020, < https://arxiv.org/ftp/arxiv/papers/1901/1901.02672.pdf>.

Bastiaanse, H 2018, 'Multi-stakeholder roadmap for implementing consumer vulnerability management', Master's thesis, Cyber Security Academy, Leiden University, The Hague, NL.

Bierens, R, Klievink, B & van den Berg, J 2017, 'A social cyber contract theory model for understanding national cyber strategies', *Proceedings of IFIP EGOV-EPART 2017 Conference (EGOV-EPART2017)*, eds. M Janssen, K Axelsson, O Glassey, B Klievink, R Krimmer, I Lindgren, P Parycek, HJ Scholl, and D Trutnev, vol.10428, *Lecture Notes in Computer Science*, pp. 166-76.

Blythe, JM 2013, 'Cyber security in the workplace: Understanding and promoting behaviour change', *Proceedings of CHItaly 2013 Doctoral Consortium*, PaCT Lab, Department of Psychology, Northumbria University, Newcastle-upon-Tyne, UK, pp. 1-10.

Caldes, A 2016, *Residual risk scoring matrix example*, graphic retrieved 27 January 2020, <https://riskmanagementguru.com/residual-risk-scoring-matrix-example.html/>.

Coolfire 2019, *What is the difference between IT and OT?*, viewed 23 January 2020, <https://www.coolfiresolutions.com/blog/difference-between-it-ot/>.

Cope, S 2019, *The TCP/IP Model and Protocol Suite explained for beginners*, graphic retrieved 27 January 2020, <http://www.steves-internet-guide.com/internet-protocol-suite-explained/>.

Cyber Security Academy 2014, *About the CSA*, viewed 3 February 2019, <https://www.csacademy.nl/en/about-csa>.

Dorfman, MS 2007, *Introduction to risk management and insurance*, 9th ed., Prentice Hall, Englewood Cliffs, NJ, US.

Endsley, M & Jones, D 2016, *Designing for situation awareness*, 2nd ed., CRC Press, Boca Raton, FL, US.

Fredriksen, G 2018, 'Protecting the crown jewels', *Forbes*, viewed 17 February 2019, <https://www.forbes.com/sites/forbestechcouncil/2018/08/13/protecting-the-crown-jewels/#3eb2ba30a5a9>.

Ghaoui, N 2017, 'Policy strategies for VPN for consumers in the Netherlands', Master's thesis, Cyber Security Academy, Leiden University, The Hague, NL.

Gibbs, N 2007, 'Elements of a good security architecture', *Internal Auditor*, viewed 10 June 2019, <https://iaonline.theiia.org/elements-of-a-good-security-architecture>.

Government of the Netherlands (GOV-NL) 2018, 'Netherlands defence intelligence and security service disrupts Russian cyber operation targeting OPCW', NL Ministry of Defense, viewed 13 November 2019, <https://www.government.nl/latest/news/2018/10/04/netherlands-defence-intelligence-and-security-service-disrupts-russian-cyber-operation-targeting-opcw>.

Greenberg, A 2018, *Operation Bayonet: Inside the sting that hijacked an entire dark web drug market*, *Wired*, viewed 2 January 2020, < https://www.wired.com/story/hansa-dutch-police-sting-operation/>.

Grutter, H 2019, '*PoC rapportage kwaliteitspeering*', ('PoC Report Quality Peering'), GOV-NL Ministry of the Interior and Kingdom Relations, viewed 2 January 2020, <https://programmeringsraadlogius.pleio.nl/file/download/57979527/Bijlage%203c%20PoC%20Rapportage%20Kwaliteitspeering.pdf>.

Halo Business Intelligence 2009, *People process technology, the golden triangle explained*, viewed 3 February 2019, <https://halobi.com/blog/people-process-technology-the-golden-triangle-explained/>.

Hoogstraaten, H 2012, *Black Tulip Report of the investigation into the Diginotar Certificate Authority breach*, Fox-IT BV, viewed 4 April 2020, <https://www.researchgate.net/publication/269333601>.

ISACA 2016, *The merging of cyber security and operational technology*, viewed 2 January 2020, <https://cybersecurity.isaca.org/csx-resources/the-merging-of-cybersecurity-and-operational-technology>.

ISO/TC 262 2018, ISO 31000:2018, *Risk management, guidelines, International Organisation for Standardization* (ISO), Geneva, CH.

ISO/IEC JTC1 2005, ISO/IEC 27001:2005*, Information Technology - Security Techniques - Information Security Management Systems - Requirements*, ISO, Geneva, CH.

——2012, ISO/IEC 27032:2012, *Information Technology - Security Techniques - Guidelines for Cybersecurity*, ISO, Geneva, CH.

——2018, ISO/IEC 27000:2018, *Information Technology - Security Techniques - Information Security Management Systems - Overview and Vocabulary*, ISO, Geneva, CH.

ISO/IEC TC262 2009, ISO/IEC 31010:2009*, Risk Management - Risk Assessment Techniques, ISO*, Geneva, CH.

Jacobs, FR & Weston FC Jr. 2007, 'Enterprise Resource Planning (ERP) - A brief history', *Journal of Operations Management*, vol. 25, no. 7, pp 357-63.

Koppenjan, J & Groenewegen, J 2005, 'Institutional design for complex technological systems', *Technology, Policy and Management*, vol. 5, no. 3, pp 240-57.

Lankton, P 2007, *Endsley's model of SA*, graphic retrieved 27 January 2020, <https://en.wikipedia.org/wiki/Situation_awareness>.

Lessig, L 1999, *Code and other laws of cyberspace*, Basic Books, New York, NY, US.

Li, Y & Thimbleby, H 2014, 'Hot cheese: A processed Swiss cheese model', *Journal of the Royal College of Physicians of Edinburgh*, no. 44, pp. 116-21.

Lord, N 2019, *What is a Security Operations Center (SOC)?*, viewed 23 January 2020, <https://digitalguardian.com/blog/what-security-operations-center-soc>.

Pols, P 2018, 'The unified kill chain: Designing a unified kill chain for analyzing, comparing and defending against cyber attacks', Master's thesis, Cyber Security Academy, Leiden University, The Hague, NL.

Royal Exhibitions 2020, *The complete set of Replica Crown Jewels and Replica Queens Jewels,* graphic retrieved January 27 2020, <https://royalexhibitions.co.uk/>.

Rowley, J 2014, *Global risk reports 2014–Managing catastrophic risks*, graphic retrieved 27 January 2020, < https://4squareviews.com/2014/02/20/global-risk-reports-2014-managing-catastrophic-risks/>.

Santana, J 2019, *Simple and successful solution against DDoS attacks*, Dcypher, viewed 10 June 2019, <https://www.dcypher.nl/en/simple-and-successful-solution-against-ddos-attacks-ive-tested-and-attested>.

Skybox Security 2018, *Why a unified approach to IT and OT network security is critical*, viewed 10 June 2019, <https://www.skyboxsecurity.com/resources/search>.

Skybrary 2016, *James Reason HF model: Swiss cheese model*, graphic retrieved 2 January 2020, <https://www.skybrary.aero/index.php/James_Reason_HF_Model >.

Steinberg, N 2020, *Becoming Unconsciously Competent in your dating life*, graphic retrieved 27 January 2020, <http://thelovetrep.com/becoming-unconsciously-competent-in-your-dating-life/>.

Surisetti, P 2017, *De-Coding Indian intellectual property law*, graphic retrieved 27 January 2020, <https://spicyip.com/2017/07/code-2-0-chapter-review-technology-ip-and-regulation.html>.

Tanenbaum, AS 1996, *Computer Networks*, 3rd ed., Prentice Hall, NJ, US.

United Nations (UN) 2015, *BowTieXP, bowtie methodology manual*, UN International Civil Aviation Organization (ICAO), viewed 13 November 2019, <https://www.icao.int/safety/SafetyManagement/SMI/Documents/BowTieXP%20Methodology%20Manual%20v15.pdf>, pp. 1-64.

US Department of Justice 2014, *Ross William Ulbricht, Silk Road*, Criminal Complaint by US District Court.

van den Berg, J, van Zoggel, J, Snels, M, van Leeuwen, M, Boeke, S, van Koppen, L, van der Lubbe, J, van den Berg, B, & De Bos, T 2014, 'On (the emergence of) cyber security science and its challenges for cyber security education', *Proceedings of the NATO IST-122 Cyber Security Science and Engineering Symposium*, Tallinn, EE, 13-14 October 13–14. (Winner of the Best Paper Award).

van den Berg, J 2018, 'Cyber security for everyone', *Cyber security best practices*, eds. S Frey & M Bartsch, Springer Verlag, Wiesbaden, DE, pp 571-83.

——2019, 'Grasping cybersecurity: A set of essential mental models', *Proceedings of the 18th European Conference on Cyber Warfare and Security*, Coimbra, PT, pp. 534-43.

Wazir, F 2019, 'Can NL trust 5G? A conceptual model for cyber security supervision of 5G in the Netherlands', Master's thesis, Cyber Security Academy, Leiden University, The Hague, NL.

Willems, D 2017, 'Caring for security: An analysis of the security of eHealth services of Dutch general practitioners', Master's thesis, Cyber Security Academy, Leiden University, The Hague, NL.