Less is more

A theory driven research project to determine what collaboration design best supports the sharing of pragmatic cyber security related information between organisations

by L.L.J. (Bart) Spijkervet - 1222503
Systems Engineering, Policy Analysis and Management (SEPAM)
Faculty of Technology Policy and Management (TPM), section ICT
Delft University of Technology

Graduation committee:
Prof. dr. Y. (Yao-Hua) Tan ----Yao-Hua Tan
Dr. M.L.C. (Mark) de Bruijne
Prof. dr.ir. J. (Jan) van den Berg

February 2014

Summary

In the last few years many cases of espionage, copying or stealing of information, and destruction of data and (possibly even) physical items of all sorts of organisations by in- and outsiders using computers hit the news. The stories typically represent cases of successful attacks which were detected. However, not all attacks are detected, or only long after the attack actually took place. As a result of that, the number of attacks which actually took place is higher. These situations are particularly undesired as because of the attacks some person or organisation incurs a loss. They are confronted with a security accident, being a negative consequence at their expense. This is what cyber security is supposed to prevent. As conveyed by Von Solms and Van Niekerk (2013), cyber security is about securing individuals, organisations or nations that function in the digital environment.

Unfortunately there is no real prospect for the yearly amount of cyber security accidents to suddenly drop considering the current state of cyber security. Generalized the attackers are considered to be ahead of the defenders in terms of their capabilities. As a result, the attacks are allegedly becoming more insidious. Furthermore, they are also allegedly becoming more targeted. The net result is that fewer organisations get attacked by the same, yet more powerful attacks. Unfortunately, again generalized, the defensive side is not really keeping up. This starts with the problem that defenders already have a natural disadvantage. But given the current complexity of cyber security resulting from interconnectivity, dependencies and changing use of technology, no organisation can oversee what *has to be done* to keep the environment secure. Add to this that organisations are considered to underinvest in defensive measures and the net result is that less information *can* be kept secure and that even less information *is* actually stored in secure manner. Finally, also the cyber security accidents, if detected in the first place, are not communicated productively. As a result thereof other individuals, organisations or nations can be affected by the initial breach too. Examples are a breach of security at a power plant, potentially allowing for a catastrophic interference of power plants by some adversary.

To improve the defensibility of organisations numerous parties pose that organisations have to share information on cyber security -incidents and -accidents with each other. Sharing information on incidents might help avoid an actual accident at another organisation. For example by informing that adversaries were able to bypass security using a backdoor. With accidents the organisation could inform others not only of the cause, but also of the consequences of an attack in order to minimize further damages. An example thereof is to warn the users and other organisations if user credentials were acquired by adversaries in an attack.

At the moment there are quite a few collaborations in different countries. The collaborations vary quite a bit in ways such as:

- how they came about (such as being specifically oriented on cyber security or that cyber security is added to the agenda of a pre-existing collaboration),
- who is part of the collaboration (such as anybody in a specific industrial sector or larger organisations irrespective of the sector they are in), and
- how they are organized (ad-hoc in small groups or in a large collaboration involving some central bodies for analyses).

Given these vast differences between the collaborations, but also given that the participants are confronted with different attacks the question is what ‘the best’ cyber security collaboration design is like.

The goal of this research project is to:

> *determine what the default collaboration design should be, for the purpose of sharing pragmatic, cyber security related information, and to identify critical factors in further shaping the design of the collaboration*

The factors to decide upon are presented in a roadmap. Using the roadmap, participants are guided in what they have to consider in the development of their specific information sharing collaboration. The scope is limited to a collaboration that wishes to address the in this research defined 'unknown-unknowns'. An unknown-unknown represents cases in which an attacker is able to cause an incident (or worse), without the defender being aware of the possibility thereof. The other options are depicted in Table 1.

Security state	Resulting in
Unknown-unknown	Undetected incidents or even accidents
Known-unknown	Unavoidable incidents, avoidable accidents
Known	Avoidable incidents

Table 1: Categorization of security states, as defined and used in this research project

The goal is to quickly discover why the incident is possible. Such as finding out that the adversary was able to bypass the security of a specific application. Once this is known, the organisation enters the 'known-unknown', meaning it is known that something is wrong with the application, but not why this is possible. The final state is the 'known', such as that it is known that the default password was never changed.

Identification of critical factors leading up to the design of collaboration

Methodology of identifying critical factors

In order to identify the factors affecting the ultimate design of the information sharing collaboration two models are used. The first is a model by Kowtha et al. (2012), hereafter referred to as the 'Kowtha model'. The model is developed on the basis of observations of actual collaboration centres. The model is originally used to uniformly describe collaboration centres. Because of that, it is reasoned that the model also presents a portion of the factors that parties will have to consider in the design of a new collaboration centre. To structure, rename and complement the derived factors the second model is used: the Bow Tie model. The purpose of that model is to systematically analyse risks in an organisation. For that the model provides a way of thinking and structuring of the chain of events leading up to accidents. Using the Bow Tie model the, from the Kowtha model derived, factors can be put into perspective.

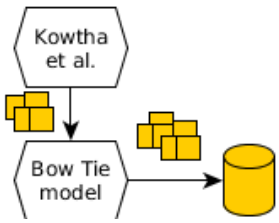


Figure 1: Extraction and collection of the initial set of factors from two models

To guide the parties in developing an information sharing collaboration the factors had to be positioned onto a roadmap, with each factor being a step. A moment at which participants have to agree upon some aspect of the collaboration. For that the first round was about defining the method of filling the roadmap with the factors. This round and the subsequent steps are illustrated in Figure 2. Given the lack of a context from what the factors are identified it is difficult to assess completeness of the list of factors which would have to be added to defined roadmap. For that purpose three simplified collaboration scenarios were defined that limit the complexities of normal collaborations. For each scenario the relevant factors were, in the second round, selected depending on their useful in defining the collaboration for that scenario. This resulted in an initial set of factors which are added to the roadmap. To complement the set of factors a theory relevant to the scenario at hand was employed. The respective theory is discussed hereafter in this summary. In the third round, using the theory, readily defined steps were put into perspective and additional factors were added to the roadmap. These second and third were repeated three times, for each of the three scenarios.

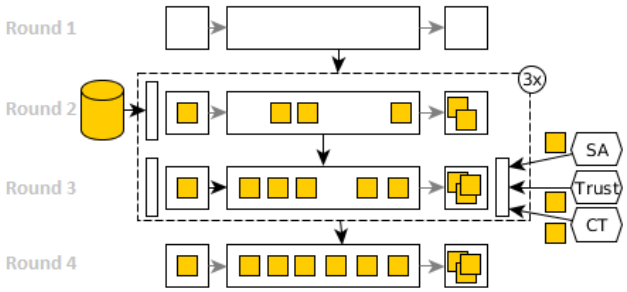


Figure 2: The four rounds to end up with the roadmap consisting of factors that help define the design of the information sharing collaboration.

The fourth round was about the outcome and discussing its implications, what lead up to the definition of the suggested default collaboration.

First scenario: static autonomous collaboration of equal and fully committed participants
The first scenario is primarily oriented on the information exchange and its presumptions, hereby intentionally ignoring many of the additional complexities of collaborations. The main assumption underlying this scenario is that of the collaboration being static in nature. The configuration of participants will not change and the participants are not affected in any way by the outside world. The main goal of the first scenario is to define the actual goal of the collaboration. Participants will have to agree upon what type of attack they will focus on, with an attack being defined as coming from some adversary, directed on some organisation, bearing some level of risk.
The goal of information sharing is in this research translated into the presumed goal of improving Situation Awareness (SA) in some respect, such as the security of an application or whether there are security breaches. Given the importance of SA to this research the theories on SA are employed for the purpose of identifying additional factors. According to the broadly accepted explanation by Endsley (1995), SA can reach three levels:

- the first level is about knowing which elements (such as physical systems, software applications and users of the systems) are present in some environment,
- the second level is about the comprehension of the current state of those elements (such as noticing that the server is sending spam messages because the server never sends emails in the weekend and now it does to unknown recipients),
- the third level is about foreseeing what those states mean for the future situation (such as foreseeing that, should the spamming continue, this will result in a blacklisting of the IP address of the server preventing it from successfully delivering email to recipients).

The level of SA individuals or organisations is considered to affect the decision making capability of that party. For example, limited SA of the security of applications might result in non-decisions to decommission the use of the application. And the lack of such a decision might allow an adversary to take benefit of a weakness to acquire information. The proposed solutions of information sharing collaborations come down to the attempt to improve the SA of a party by means of the SA of somebody else.
There are multiple views of individual SA and SA in a collaborative setting. Of the collaborative options, Distributed Situation Awareness (DSA) is considered to be particularly useful to information sharing on cyber security. It takes benefit of the fact that parties have slightly different perspective. Even if those parties would look at the same information, they would interpret it differently based on their different, unique yet *compatible* perspectives. With DSA the focus is on such compatibility. This presupposes that the parties are compatible in that they are able to collaborate and that the SA of that participant is of use to the others (referred to as compatible SA). A collaboration entails being able to provide some other party with information to improve the SA of that party. And for that the party has to know who knows what, who needs to know what and how those parties can be reached.

Second scenario: static collaboration of equal participants
The second scenario is similar to the first scenario but no longer presupposes the collaboration to be impartial to its environment. The environment of a collaboration is made up of those parties, artefacts and topics that are relevant to the collaboration. Relevance is the resultant of the ambition of the collaboration. Depending on the ambition certain aspects of the environment will be considered and other aspects will not. Participants of the collaboration would have to agree upon whether they will jointly focus on future attacks, current attacks or the aftermath of some attacks. These attacks take place in the environment of the collaboration. Therefore the collaboration is affected by and supposed to respond to actions and activities in its environment.
Whereas the focus in the first scenario was on the exchanged information in a collaboration, the focus of the second scenario is on the interdependence of decisions regarding the goals, the method of collaboration and what the goals and methods effectively demand from the participants in terms

required levels of trust amongst the participants. And depending on the sensitivity of the exchanged information different levels of trust are required. To understand the constituents of these different levels of trust the book on trust by Nooteboom (2002) is used.

Third scenario: dynamic collaborations of heterogeneous participants
The final scenario depicts the most complete and most 'realistic' picture in this research. It no longer presupposes participants to be identical. The focus in the third roadmap is on the development of configurations of organisations in a collaboration. Organisations have different levels of awareness on cyber security related matters, have different needs and might have a vastly different background (aside from the backgrounds of the person representing the organisations). As a result, individual participants of the collaboration will look at the cyber security problem differently. Together the participants define their view of reality, the situation which they will consider. And the collaboration is about achieving the highest possible awareness of that situation.
To explain the impact and importance of having different participants in a collaboration, the configuration theory is employed. The main notion is that participants in a configuration define their view of reality together and that this reality in turn affects the participants. But, participants do not necessarily find themselves entirely in the definition by a single configuration. They might be involved in multiple collaborations, they are *multiple included*. It is this multiple inclusion that allows parties to come into contact with different definitions of reality. And these different definitions are the source of conflict which allow for redefinitions. Introduction of other views on reality allows configurations to reconsider their view of reality. An example of all this is that a configuration of smaller organisations might think the reality is that they are safe from attacks. But some of the participants might have recently been confronted with a different reality in another configuration. The other configuration might have redefined its view based on some recent event in which a smaller organisation was attacked.

Resulting roadmap with steps that have to be taken in the development of a collaboration
The final roadmap consists of the steps that participants have to take in the development of an information sharing collaboration. Each step poses a decision moment on some aspect. The roadmap is divided into three phases: the current state, the development phase and the desired end state. The ultimate goal is to close the gap between the current state and the desired state by means of some sort of collaboration. For that, in the current state participants have to define the goal of the collaboration. The definition of the goal comes down to:

- defining what type of attack the participants of the collaboration will focus on, with the type of attack being enclosed by (i) the adversary that launched the attack, on (ii) some target, with the attack bearing some (iii) risk in terms of the expected consequence of the attack,
- defining how the collaboration is positioned relative in time to the considered attack, meaning the participants have to decide whether they wish to avoid that type of attack, thwart those types of attacks or learn from those types of attacks, and
- defining the scope of influence, being whether the collaboration only considers attacks affecting organisations in a specific country and intends to influence organisations in that country or whether the collaboration considers influencing international organisations too.

Such a definition of the goal takes place in five steps, as depicted in the right part of Figure 3.

Current state ↓	Development of the collaboration model ↓						Desired end state ↓		
formulation goal of collaboration	roles	collabo-ration structure	topic of shared information	level of information sharing	method of sharing information	type of response	adversary	target	risk-value
	timeliness of response	environ-ment	external interaction	scale	maturity		time dimension	scope of influence	

Figure 3: A depiction of the steps representing factors which the participants of the collaboration have to decide upon in the development of an information sharing collaboration, for the purpose of improving the cyber security. Such a decision comes down to defining what the desired end state should be like and this has to be backed up with a suitable collaboration design.

In order to be able to reach the desired end state, as in being able to avoid, thwart or learn from some type of attack, the participants have to collaborate in some way. The participants have to agree upon the design of the collaboration in terms of organizational structure, collaboration process and shared information. To guide such a design effort several steps were identified on which participants necessarily have to agree. These steps are depicted in the middle part of Figure 3.
Importantly, the decisions taken in each step have consequences for other choices that need to be made along the way. Although sharing all kinds of information on weaknesses in defences (an option of 'topic of shared information') at a raw data level (an option of the 'level of information sharing') is possible, this possibly affects the volume of shared information and requires participants to trust each other with that type information. Such mechanisms, which are derived from the three theories, make that there is some sort of default design for information sharing collaborations. This will default will be discussed in the next section.

Default design for information sharing collaborations

Based on the interdependence of the steps the main finding is that it is unproductive to opt for a single large collaboration. However, a default design can be discerned, which was the main goal of this research. The default design is to use *multiple, small, focused collaborations of similar organisations, with the organisations being part of multiple collaborations*. Such a design is considered to be the most effective and most efficient way in achieving the goal of addressing the unknown-unknowns. Herein similarity of organisations plays a crucial, double role. If organisations in a collaboration differ too much:

- their views of reality will be different meaning they consider different things to be important,
- the situation they consider is different and with that their awareness is different, and
- they have a harder time to empathize with the other organisations.

As a result, such a collaboration tends to remain at the strategic or abstract level of discussions. It is at that level that organisation can still find each other. But for uncovering unknown-unknowns participants have to reach a more practical level of discussion and exchange practical information. Without this, the effectiveness of the collaboration in addressing the unknown-unknowns is compromised. However, exchanging more practical information on unknown-unknowns implies that organisations have to exchange more raw information, yet relevant information with each other. The information cannot be processed too much as in the process subtle details might be lost. But at the same time this means the amount of information shared will increase, but has to be limited based on relevance. With such a balancing it is important that organisations understand what the other needs, but also whether the other can be trusted with the information. With that, at the other end increasing differences between organisations challenges the efficiency of the collaboration. Information has to be relevant. But for that, first organisations have to be willing to share the information. And for that, they have to assess the trustworthiness of the other organisations, which is more difficult if the organisations differ too much. The tendency will be to assure that the other organisations can be trusted, such as by using extensive contracts. But these contracts hurt the efficiency of the collaboration in the long run. For these reasons, the organisations in a collaboration

have to be similar, in order for them to understand what has to be done, what the other needs in terms of information and whether the other can be trusted with the information.
But similarity is at the same time also the villain of effectiveness because organisations can become blind to what is actually needed. They are not contested with whether what they think is reality, what has to be done and how it has to be done is actually true. For that reason it is deemed necessary that organisations have to be part of multiple collaborations. With that, they will come into contact with different views of reality. And with that, at times, organisations (and the collaborations they are in) are forced to reconsider their view of reality.

Validation of the findings
The identified steps and the resulting default design were not validated empirically. However, the credibility was demonstrated using numerous examples from practice. Furthermore, in a thought experiment the main finding that a single collaboration is not a productive approach was demonstrated. Even if a technological platform could be developed that is able to extract the relevant information at the right level of abstraction, this still leaves trust as a concern. The extracted information still ends up at some organisation. Depending on the implementation the organisation has to still trust that organisation or trust the system. Either way, the organisation will still desire some limitation in the exchanged information, not knowing where it might end up and not knowing what they might get in return. As a result, it is reasoned that organisations will always want the collaboration to be limited, if not for usefulness than at least because they want it to.

Limitations
This research has two main limitations, one being that the findings themselves were not validated empirically. Just that using empiricism the credibility of the findings was demonstrated and further demonstrated using a thought experiment. Furthermore, the research had a rather binary approach towards security for the sake of clarity. Collaborations were considered to focus on secure or insecure matters, not on the conditions contributing to insecurity. With that, some additional difficulties are omitted. However, it is considered that these additional difficulties would only make the case for opting for multiple, smaller collaborations even stronger. A single, large collaboration would be forced to highly simplify conditions, failing to address the actual situations at hand.

1 INTRODUCTION

This chapter gives an introduction to the cyber security challenges that many organisations are confronted with. The first part of the introduction discusses the attacks, the impact thereof to- and the importance of- cyber security, and the collaborations intended to improve the cyber security of organisations. The second part of discusses the goal and methodology of the research project to improve the collaborations for improvement of cyber security. Finally the structure of the remaining part of the report is presented.

1.1 CYBER ATTACKS

In the last few years many organisations were attacked by so called cyber-attacks. Of cyber-attacks various definitions and conceptions exist (Hathaway and Crootof, 2012, p. 823). In this research cyber-attacks are defined as being:

> "A hostile act using computer or related networks or systems, and hereby affecting and/or disrupting and/or destroying an organisations' cyber systems, assets, or functions. The intended effects of cyber-attack are not necessarily limited to the targeted computer systems or data themselves. The activation or effect of a cyber-attack may be widely separated temporally and geographically from the delivery."

This definition is an altered version of the definition by the Joint Chiefs of Staff (Hathaway and Crootof, 2012, p. 824). The original definition and the motivation for the various changes are discussed in A1.

Classification of cyber- attacks

Cyber-attacks have resulted in various cases of espionage, copied and destruction of data, and (possibly even) destruction of physical items of all sorts of organisations. But the actual attack method leading up to those events vary quite a bit. They range from the attacks using highly advanced techniques on single facilities of an organisation to opportunistic attacks using common techniques on any organisation.

Examples of (high profile) cyber-attacks

Possibly the most sophisticated series of cyber-attacks was specifically targeted on the Iranian uranium enrichment centrifuges at Natanz (Symantec, 2013a). The attacks are considered to be made possible using a combination of the malicious software (malware for short) named Flame and the 'cyber weapon' named Stuxnet (Goodin, 2012a). The pieces of malware made (amongst others) use of an at the time unknown cryptographic collision attack (Flame)(Goodin, 2012b) and twenty vulnerabilities in software which were at the time unknown (Stuxnet)(Goodin, 2013a). Given the sophistication and the specific target of the pieces of malware it was believed that the development was sponsored by the USA, which was 'confirmed' later on (including Israeli involvement) (Anderson, 2012).

Stuxnet and Flame are highly advanced attacks, however they are exceptionally targeted. They focused at one specific facility. An example of another category of attack which was less targeted is Red October. It is a highly advanced piece of espionage malware, which was able to go unnoticed for five years (Goodin, 2013b). It gathered information from many organisations (such as governments, aerospace industry and the military) in at least 39 countries. (Goodin, 2013c)

A final example of yet another category of attacks are the high profile cases of attacks on energy companies using less sophisticated (or even blunt) attacks to (attempt to) acquire information (Kirk, 2011), wipe data (Goodin, 2012c) or allow access to (remote) control systems (Goodin, 2012d).

Besides all these cases of external attacks the insider attack is also a serious concern. With an insider attack, the insider could even just make use of 'normal' access rights. The only reason for labelling it an attack is that it could cause harm to the organisation. From an industry report, based on surveys, follows that about twenty percent of all discovered cyber-attacks were in 2011 the work of an insider or partner (Jeffries, 2013). Similarly, insider attacks account for 19% of all attacks at the FBI. But more

importantly, according to the FBI “those incidents were about twice as costly as *all* the attacks by outsiders” (italics added).(Jeffries, 2013) Finally, Verizon concludes, based on their findings from forensic investigations, that internal actors were involved in 14% of the cases and in 1% of the cases partners were involved (Verizon, 2013, p. 19).

1.2 CYBER SECURITY

The goal of cyber security, as expressed by Von Solms and Van Niekerk (von Solms and van Niekerk, 2013, p. 5), is to secure parties and the interest of parties (tangible such as properties and infrastructures, and intangible such as societal values) who function in the digital environment called cyber space (whether they are individuals, organisations or nations).

Stuxnet as a breach of cyber security

Stuxnet also drew much attention as it concretely highlighted another area in the more encompassing whole of cyber security. It indicated yet again the vulnerability of humans and their interest in the case of cyber-attacks. Cyber security is more than information security as indicated in figure 4 the topic of Information Security already received a lot of attention over the years, specifically in warranting the confidentiality, integrity and availability of information. With information security the asset which has to be protected is the information itself (von Solms and van Niekerk, 2013, p. 3). Attacks as Flame and Stuxnet were able to make use of vulnerabilities in hardware, software and information. The espionage software Flame could spread malware by being able to pose as being a legitimate update coming from Microsoft via the Windows Update mechanism. Stuxnet was able to affect Programmable Logic Controllers (PLCs). These devices are used to control all sorts of machines, such as elevators but also centrifuges, such as those in Natanz.

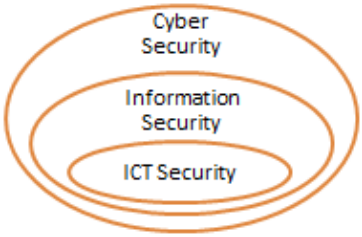


Figure 4: Euler diagram that visualizes the relationship of cyber-, information(using ICT)- and ICT security, adapted from (von Solms and van Niekerk , p. 5).

But crucially, the goal of Stuxnet was not really about affecting ICT or information stored on ICT systems. It was about what that information is used for. By changing the information which was fed to PLCs the adversary would be able to destroy the centrifuges. In all this, the information is the vulnerability to the functioning of the centrifuges. And that functioning affects humans and their interest.

A shared concern

Cyber security breaches can indirectly (loss of productivity) or directly (financial reimbursement) cost organisations money. But cyber security can affect more individuals and organisations than the organisation whose digital environment got attacked initially. An example is that one organisation got attacked and that via this organisation the partner organisation can be affected too. For example because there is a secure, yet open connection between the two organisations. And in extreme cases breaches in one sector can even directly affect other sectors. Think of the role of the power grid plays in supplying a stable supply of energy to various sectors. A blackout caused by some adversary will affect all kinds of other organisations.

To protect humans and their interests, the challenge of cyber security is therefore about reducing the risk humans are exposed to via cyber space. And with information being stored on digital systems in various sectors cyber security is a concern to society, not just to some (part of a) sector.

1.3 COLLABORATION INITIATIVES TO IMPROVE CYBER SECURITY

To combat the sheer increase in manifestations of security breaches due to all sorts of attacks information sharing is considered to be necessary. As expressed by Francis Maude at the launch of the Cyber Security Information Sharing Partnership (CISP) it is ‘abundantly clear’ that the complexity of cyber security is such that not a single organisation has ‘anything like’ complete overview of what is going on (Maude, 2013). Art Coviello of RSA too thinks that information sharing can help and

herein envisions an important role for governments. Furthermore, he considers that the defenders are behind on attackers in terms of capabilities. (Schellevis, 2013) Vice-president of the European Commission Neelie Kroes said governments have to provide incentives to organisations to invest in security and to share information on threats and attacks. But governments would also have to lead by example. (Kroes, 2012)
Currently there are indeed collaborations between organisations in the United States (National Infrastructure Advisory Council, 2012), the United Kingdom (Cabinet Office, 2013) and the Netherlands (NCSC, 2013, p. 39). However, in January 2012 the National Infrastructure Advisory Council (NIAC) of the United States labelled many public-private partnerships as being "relatively immature, leaving a large gap between current practices and an optimal system of effective public-private intelligence information sharing". (National Infrastructure Advisory Council, 2012, p. ES–1) Similarly in the Netherlands there is no continuous sharing of information amongst public and private companies yet (NCSC, 2013, p. 39).
Apart from the fact that the collaborations are not to satisfaction, they also vary quite a bit in terms of the scope. Some collaborations are rather small, others span almost entire sectors and yet other collaborations even consist of organisations that are in different sectors.

1.4 RESEARCH

The vast differences in designs of collaborations, without a clear indication of what the best design is, forms the motivation for this research project. The main goal of this research project is to:

> *determine what the default collaboration design should be, for the purpose of sharing pragmatic, cyber security related information, and to identify critical factors in further shaping the design of the collaboration*

Hereto first critical factors would have to be identified which define the design of the collaboration. These factors refer to the collaboration itself, but also to the information or the process of the collaboration. Each of the identified factors will be illustrated as steps on a roadmap. The roadmap should be interpreted as a method, in the way Hevner et al. described it. Methods have the purpose of defining processes, to "...provide guidance on how to solve problems, that is, how to search the solution space.". (Hevner et al., 2004, p. 79) The research will not result in an instance of a roadmap. Such an instance would describe what the organisations would have to do at some moment in time to setup a collaboration. Instead, the roadmap guides the organisations on how to approach the development of information sharing collaborations. For that, each identified step on the roadmap depicts a moment at which participants of the collaborations have to decide upon some aspect of affecting the collaboration.
The scope of this research is limited to the design of collaborations on:

- sharing information on cyber security related matters,
- about wilful/intentional activities by some adversary (internal, pseudo-internal or external),
- that can cause harm to other organisation.

1.5 RESEARCH METHOD

This research is largely based on theories with situation awareness at the level of collaborations as the focal point. The starting point of the research project was to find literature by means of databases like Scopus but also the web search of Google. The main purpose was to determine the state of research and implementations of information sharing collaborations. Based on this state, the research is being positioned on the actual collaboration, with special attention to trust, which commonly was ignored. The initial search terms were combinations of 'collaborations', 'information sharing' and 'cyber (security)'. In Scopus the searches were limited from 2011 till 2013. Later on the search entries were changed into forms of situation awareness in general and in conjunction with 'cyber'[1]. This change was inspired based on what the motivation behind information sharing could

[1] Cyber situation(al) awareness is a somewhat popular term, but commonly referred to as individual situation awareness on cyber-security.

be. On the basis of the resulting initial set of publications, studies and pilot projects the main demarcation of the research was realised. To assess the completeness, to complement and to structure the findings additional, (more) generic, theories and concepts on collaboration were added. Herein the information sharing in a collaborative setting took priority over generic collaboration theories.

Ultimately two models, a descriptive-detailed one by Kowtha et al. (2012) and a prescriptive-abstract one being the Bow Tie model, formed the basis of the remainder of this research, as displayed in the left side of Figure 5. The model by Kowtha et al. was developed with the intention of being able to characterise actual collaboration centres.

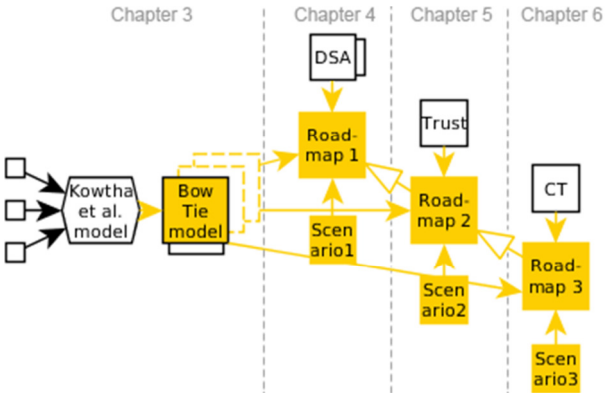


Figure 5: A simplified overview of the contribution of two models (Kowtha et al. and the Bow Tie) and some theories (Distributed Situation Awareness as DSA, Trust and Configuration Theory as CT) to the development of three roadmaps (1,2,3), which are based on three scenarios (1,2,3).
Black bordered (structure) and white solid fillings (contents) of the shapes are things I cannot take credit for. (The two arrows with no fillings are simply to reflect a specialization relationship, otherwise they would have been solids too.)

In this research the various ways in which Kowtha et al. characterize collaborations centres are used as a starting point for the development of a design of a new collaboration. The model is discussed in more detail in 3.1.1. The Bow Tie model is used to restructure the findings from the model by Kowtha et al. The usefulness of the Bow Tie as a method for (re)structuring the findings is discussed in 0. A coarse evaluation of the alternatives to the Bow Tie is presented in A2.1. Additionally, based on initial findings from the publications, such as the model by Kowtha et al., additional, more specific publications on three specific aspects were retrieved. These aspects – distributed situation awareness (DSA), trust and the configuration theory – are still a prominent part of this research. Their usefulness is discussed in the report (DSA in 4.1 && A3.2.1.2.5, trust in 5.1 && A3.2.2.1 and the configuration theory in 6.1.3 && A3.2.3.1).

Given the lack of a renowned method regarding both the information and collaborations (the Bow Tie is useful to some extent), three collaboration scenarios with increasing levels of complexity were constructed. These scenarios are used to be able to focus on specific complexities of information sharing collaborations and to be able to assess whether certain aspects therein are missing. Each of these three scenarios is actually about a set of assumptions per scenario. These scenarios, and the more tacit assumptions, formed the starting point of the intended roadmaps, which identify what steps parties have to consider in the development of a collaboration. As there are three scenarios envisioned, there are also three roadmaps. These are not distinct roadmaps, but rather extensions of each other, with the third being the most comprehensive one. For the development of each roadmap the initial focus was on defining the goal of the information sharing collaboration. The definition of the goal is the result of a set of decisions, which take place in a number of steps. This was followed by and identification of the steps that have to be taken to define on how to meet that goal. The initial inspiration for the different decisions which have to be taken came from the findings from the model by Kowtha et al. and the Bow Tie model. These decisions, represented as steps on the roadmap, were complemented and redefined on the additional theory.

1.6 STRUCTURE

The upcoming second chapter covers the challenge of cyber security in more detail. It discussed the challenges of cyber security, the need for collaborations and the motivation why readily available collaboration models do not seem to suffice.

The methodology is discussed in the third chapter. In chapters four until six, as displayed in Figure 5, the three roadmaps with their accompanying scenarios and theories are discussed:

- The first roadmap focuses on a collaboration in a crisis situation which is in no contact with the environment and with all participants being exclusively available to the collaboration. The chapter concludes with an example of what a simplified collaboration along the lines of a followed clean roadmap would be like.
- The second roadmap builds on the first. Contrary to the first, the collaboration is now under bidirectional influence of the environment. The environment consists (at least) of the organisations of the participants, but also relevant adversaries and target organisations. With the bidirectional influence, opportunistic behaviour of participants is a concern. Furthermore, the participants now have to decide upon how they position their collaboration relative to attacks which take place in that environment. Some might decide to collaborate on ongoing attacks, others on trends in attacks and others on learning from how attacks were handled. The final paragraph of the chapter lays out a simplified example of what collaboration could look like and the challenges which will come up.
- The third roadmap again builds on its preceding roadmap and now no longer considers all participants to be identical. The participants differ in terms of their definition of the situation at hand, their awareness and their capabilities. With all that, this roadmap focuses more on the challenges of the collaboration itself and the courses of possible development.

In the seventh chapter follows a rather extensive discussion about:

- how the main the roadmaps and the three theories relate to each other, and together affect
- the courses of possible development of collaborations, leading to
- the suggested default design of the collaboration.

The eighth chapter focuses on the (difficulty of) validation of the research findings within the turnaround of this research project. In chapter nine the conclusions, research limitations and the suggestions for future research are presented. In the appendices first the definitions of terms, as explained and presented in the report, are presented in A1. In A2 a, in the report used, risk analysis method is discussed. Finally in A3 the method of identifying the set of factors determining the design of a collaboration (A3.1), the resulting set of factors (A3.2) and an extensive discussion of the factors is presented (from A3.2.1 onwards).

2 THE NEED FOR COLLABORATION ON CYBER SECURITY

In this chapter a coarse overview of the current challenges in the cyber security domain is presented. The chapter starts with a brief explanation of some of the terminology that has been used in this research. This is followed by a section on the increase of- and direction of- witnessed attacks on organisations. In the third section one of the intended and used methods (i.e. information sharing by means of collaboration) to stop or at least slow down the number of attacks is presented. The fourth and final section discusses why the intended method is amidst a promise and a delivered promise. The underlying reason is the actual motivation and further subject of study in this research.

2.1 DECOMPOSING AND DEFINING ASPECTS OF CYBER ATTACKS

In the introduction a few cases of cyber-attacks on some type of target and its consequences were presented. In this research:

- an attack is defined as a threat launched by some adversary on the assets of some target, resulting in an incident or even an accident,
- a threat is defined as the technique(s) adversaries use which affects the target in some way,
- a technique refers to an undefined method such as a specific type of malware which harvests credentials or to log on to a system using the harvested credentials which potentially could cause harm to the organisation,
- a vulnerability is a weakness or gap in the protection efforts which the technique is able to exploit,
- the asset is the thing (such as devices, data or other valuables) one tries to protect, and
- an incident is an occurrence which does not have negative consequences to the organisation (such as a scanning attack for weaknesses by the adversary), whereas with an accident there is a negative consequence (such as a loss of data).

For convenience these (and some other) definitions are also presented in A1 on page 73.

2.2 THE CHALLENGES OF CYBER SECURITY

In brief, what we can notice is an increase in *witnessed* attacks and in particular an increase in *witnessed* accidents. However, we're not sure whether the total amount of attacks and accidents is also increasing, let alone know the rate of changes in the amount of attacks. The assumption is that there is an increase in attacks. But more importantly, there is an increase in the amount of accidents. After all, what is noticed is that the attacks are increasingly targeting smaller companies (Symantec, 2013b, p. 4). Such organisations have fewer resources to protect themselves against the attacks and with that they are an easier target.

In this section the attacks are discussed from the result of attacks back to the initial activities. First the known consequences which became *public* are discussed, followed by brief discussions of targets, attacks, treats and ultimately the adversaries.

2.2.1 UNKNOWN TOTAL COSTS OF CYBER-ATTACKS

As discussed in the introduction attacks can involve different types of techniques (ranging from well-known to unknown) on different selections of (types of) organisations, resulting in altered, copied or stolen data. Those events might result in consequences as missed opportunities for revenue (such as with copied intellectual property), financial repercussions or destruction of physical items. According to Ponemon Institute the 'average total organizational cost of data breach' as incurred by American organisations was 5.4 million US dollars in 2012 (Ponemon Institute, 2013, p. 5). McAfee estimated that the costs of cybercrime worldwide are 300 billion US dollars, although the CTO of McAfee admitted that it is very difficult to estimate the true costs (Hutchinson, 2013). This was also pointed out by quite a few authors (Florêncio and Herley, 2011, p. 8). The estimates of the costs of cyber-attacks are typically the result of flawed surveys or even "random guesses" in reports (Anderson et al., 2012, p. 6).

Florêncio and Herley demonstrate cyber-*crime* surveys suffer from:

(i) the difficulty of achieving a representative sample from of a heavy tail distribution, which appear to be case with cyber-crime data (Florêncio and Herley, 2011, pp. 3–5). A heavy tail distribution implies that *a small portion of the population* greatly affects the cyber-crime figures because *their crime figures are relative extreme*, their over- or under representation in the sample of the survey therefore greatly affects the findings,

(ii) the inability to verify the accuracy of the retrieved survey input (mistakes, misunderstandings[2] or incompleteness) and to remove outliers, which becomes particularly troublesome with heavy tail distributions (Florêncio and Herley, 2011, pp. 5–7), and

(iii) surveying a rare phenomenon has an impact to the findings because the majority of the population "will have nothing useful to say" which (1) makes getting a representative sample a concern, especially with regard to the representation of affected parties, the parties that have something useful to say, because they might be over- or underrepresented in the sample, it also (2) reduces the effective sample size because a large portion of the survey is unaffected by the rare phenomenon and therefore not be of interest, and (3) a portion of respondents lies and because the aforementioned difficulties the contributions of the liars carries more weight in the resulting findings. (Florêncio and Herley, 2011, pp. 7–8).

These three sources together, with the third reinforcing the effect of the former two, makes it hard to come up with proper, *consistent*, results representing the populations (Florêncio and Herley, 2011, p. 2). All this results in often inherently flawed findings on cyber-crime figures, over *or* understating the actual costs of *known* cyber-crime. All this does not even factor in the fact that not all companies even know whether they have been (successfully) attacked, what the direct accompanying costs are and what the indirect accompanying costs are. Regarding the latter, the consequences of an attack could be that another target gets 'attacked'. Knowing that many people re-use their user credentials on websites, with the Adobe hack, other websites and users could be affected too. And quite possibly it could be that this second attack was the final goal all along. As discussed by Mandiant, the aim could be to attack some target with the intention of gaining 'normal' access to partner organisations (Mandiant, 2013, p. 4)i.e. the unknown attacks and their accompanying costs. With such a scheme, only in the first attack rather sophisticated techniques might be used to merely acquire credentials, not to use those at that location too, to minimize suspicion or chances of detection, should the organisation be on alert based on the attack itself.

As a result of all this, we do not really know what the actual scale of the problem is. All we notice is an increase in accidents. Instead presenting an estimate of the overall, global costs of cyber-crime by attributing *some*[3] price as being the cost of an attack by *selected* parties and generalizing this to a population we'll have to abide with the known attacks, at least for now[4]. Even those values are rather capricious over time. But the very fact the number of known attacks increases should suffice for now. Add to this that increasingly organisations with fewer resources are targeted and it is likely to expect that the amount of accidents will increase even further.

The known consequence of these attacks in 2012 was an increase in witnessed levels of industrial espionage and data theft (Symantec, 2013b, p. 14). And although the overall number of data

[2] Florêncio and Herley refer to situation in which respondents answer what they think is the value of what the attacker (adversary) took, not what the total costs were to the target. Anderson et al. went through great lengths to avoid these issues by first decomposing the cost of cybercrime, distinguishing the categories of criminal revenue, direct losses (like damages), indirect costs (opportunity costs) and defence costs (Anderson et al., 2012, p. 12). Next they searched for all sorts of publications to ultimately, allegedly provide "the first systematic and comprehensive examination of cybercrime costs" on UK level as well as a global estimate of the costs (Anderson et al., 2012, p. 33).

[3] How do you value the costs of an attack on a start-up, which was about to release their innovative concept, yet has its invention stolen overnight and released by some other party?

[4] Although Anderson et al. came up with a detailed set of carefully disaggregated figures on cyber-crime their first publication still contained quite some important caveats (incompleteness, rough estimates and varying levels uncertainty) (Anderson et al., 2012, p. 30).

breaches was down by 26%, the median of the number of stolen identities per breach had increased by 3.5 times[5] (Symantec, 2013b, p. 17).

2.2.2 ATTACK

An attack was defined as a threat (2.2.4) launched by some adversary (2.2.5) directed on some (possibly random) target (2.2.3). The attack can be distinguished into two parts. First, somehow, the target has to come into (indirect) contact with the adversary. In their Internet Security Threat Report of 2013 Symantec states that this part of the attack is becoming increasingly 'insidious' (Symantec, 2013b, p. 14). An example of such an insidious first phase is a *watering hole* attack. With these attacks adversaries observe or anticipate what kind of websites the target organisations frequently visit. The adversary than inspects these websites for vulnerabilities that would allow the adversary to redirect the visitor of the website to another location that in turn infects the computer of the target. That other location thus contains the actual threat. (Symantec, 2013b, p. 21) But the adversary could also send a *spear phishing* email to the target. In that email there commonly is an email which poses to go to a 'normal' website, but instead goes to another location. A location which, just like the watering hole attack, contains the actual threat. Another way would be to approach the target directly. The latter could be by trying to login at a system (using readily available credentials from a prior leak or non-revoked credentials after a termination of a contract).
After the adversary and the target came into contact the adversary confronts the target with the threat that could cause actual damage to the attacked organisation. This is referred to as an 'accident' in this research. If there is an attack, but without consequences, it is an incident. The threats resulting to incidents or accidents are discussed in 2.2.4.

2.2.3 TARGETS

The recipient's end of the attack is increasingly targeted. Hereby social engineering techniques are used to impersonate respectable organisations. But also the aforementioned 'watering hole attack' is an example of increasingly targeted attacks. (Symantec, 2013b, pp. 20–21)
In 2012 Symantec *noticed* a 42% increase in targeted attacks (Symantec, 2013b, p. 10) resulting in a global average of 116 attacks per day (Symantec, 2013b, p. 14). Furthermore, these attacks are increasingly aimed at the smaller organisations. Of all known targeted attacks, 50% was aimed at organisations with less than 2.500 employees and 31% was aimed those with less than 250 employees (Symantec, 2013b, p. 4). A motivation for targeting smaller organisations is that their defences are weaker, there are more of those and they have valuable data. But they could also be "spring board" into larger organisation. (Symantec, 2013b, pp. 20–21) For that, acquired data from a smaller organisation can be highly valuable. With the data, the adversary might be able to compose an email containing specific details allowing it to successfully pose as being a trustworthy email.

2.2.4 THREATS

In the end attacks are typically about acquiring information, disrupting operations or forcing payment by the victim. Malware is a capable method for this and is "most frequently encountered' threat by organisations. Of the surveyed organisations, 66% were attacked by malware. (Kaspersky, 2013). Malware is short for an undefined piece of malicious software. It is an umbrella term for types of software like:

- ransomware (software that restricts access to computers or data for the purpose of forcing payment by the owner of the computer or the data),
- spyware (to acquire information) or
- Trojans (which opens the system up to all sorts threats, such as by providing a backdoor to a system or installing spyware that logs keystrokes on a keyboard). (Emisoft, 2012)

According to security firm F-secure, Trojans were most of the times involved in infections. However, in 2012 the company noticed of all detections in their protection environment in 28% of the cases it

[5] Two sides: possibly more reports in 2011 and on the other side some less known performance compared to 2011 in 2012: less attacks, more silent or less detections (p17).

involved an exploit-based attack. (F-Secure, 2013a, p. 7) Such an exploit is the actual malicious part in some pieces of malware. An exploit attacks a vulnerability of (or the surrounding) an assets for the purpose of installing malware into a system (F-Secure, 2013b, p. 4). The most sophisticated attacks make (amongst others) use of a 0-day vulnerability (or even multiple). Those vulnerabilities were not publicly known, even to the producer of the asset containing the vulnerability. In 2012 Symantec witnessed 14 *new zero-day* vulnerabilities, which could be exploited (Symantec, 2013b, p. 24). Those zero-day vulnerabilities are worth quite a bit of money, making them only available to the more serious adversaries. Such zero-day vulnerabilities don't lose value overnight as they first have to be detected. After that, it will take a while until a patch is issued to address the vulnerability. But in end zero-day vulnerabilities are not the sole reason of exploitation of systems. Symantec mentioned in its report that most exploitations of vulnerabilities are not the newest at all. A lack of patching by organisations and consumers makes their systems an easy target. (Symantec, 2013b, p. 26) Actually, exploitation of vulnerabilities for which solutions are readily available is "typical occurrence" according to SophosLabs (Baccas, 2012, p. 2). Along these lines, Solutionary Research discovered that of the vulnerabilities that are exploited by the most popular exploit kits 58% are well known for over two years (Solutionary, 2013).
Regarding the near future Symantec claims the highly advanced attacks will be reverse engineered. And the "open sourcing" available malware toolkits will make it easier to develop new malware as well. (Symantec, 2013b, p. 54) With that, also the second part of the attack, the threat itself, becomes more sophisticated. And all these new and more advanced threats pile upon the readily existing threats that cause harm to organisations.

2.2.5 ADVERSARIES

Roughly speaking adversaries can be classified into criminals, protesters and governments (A3.2.1.1.3). But knowing with certainty who is behind an attack is very difficult. The level of sophistication of an attack might suggest involvement of some adversary, but as mentioned, reverse engineering or accidental collateral damage might obfuscate the real type of adversary. (Symantec, 2013b, p. 19) As an example, Symantec suggests that the technique that was used to wipe hard drives on a Saudi oil firm might be copied and used for other purposes by another party. Or it might, unintentionally, end up at another organisation and cause harm, suggesting a far less targeted attack from some type of adversary. (Symantec, 2013b, p. 19) In both cases, the sophistication of the attack or the (lack of a highly) specific target might obfuscate who is actually behind the attack.
Looking ahead, Symantec expects an increased amount of state-sponsored attacks, characterised by high levels of sophistication. Although the target is quite possibly also fairly targeted, collateral damage from these attacks is considered to be possible and this necessitates preparation from organisations. (Symantec, 2013b, p. 54) Should the amount of these types of attacks increase and the expectation of reverse engineering come true, it reinforces the sophistication of attacks in general.

2.3 CLASSIFICATION OF VULNERABILITIES

Summarized, many accidents are the result of impaired awareness of different organisations, at different of levels. Many organisations are at their level confronted with opponents that are ahead. They merely act based on detected accidents and sometimes do not even know that accidents took place, such as that information was copied by adversaries.

Some organisations are targeted by highly determined adversaries. Those adversaries are able to hide their tracks, and use the most advanced attacks exposing unknown vulnerabilities. And hereby possibly even the mode of delivering the malware is sophisticated and unheard of. With that the scenario behind the malware called Flame is described. Other organisations fall prey to far less advanced attacks. Attacks that exploit vulnerabilities that were discovered a long time ago. But for all sorts of reasons the organisation might not been able to patch the system. Or they did not anticipate that users often reuse their passwords, passwords that were leaked in a recent attack.

To be able to distinguish between different types of security matters, different types of unknowns are distinguished. These are displayed in Table 2.

Security matter	Manifestation of security matter	Example in the case of vulnerabilities
Unknowable-unknown	No harm (yet)	A vulnerability that is present in (e.g.) software, but not known to *anybody* yet
Unknown-unknown	Unavoidable incidents or accidents	A vulnerability that is exploited by an adversary, but so far unknown to other parties (e.g. the producer of the software)
Known-unknown	Avoidable accidents	A vulnerability that is now known to defending parties, allowing the defending party to (e.g.) (temporarily) discontinue the use of the vulnerable (e.g.) software
Known	Avoidable incidents	A vulnerability for which a solution is available, such as a patch removes the vulnerability

Table 2: A classification of security matters in types of unknowns, such as unknown vulnerabilities. As will be discussed in 2.4.1, the focus of this research is on information sharing on 'unknown-unknowns'

The four types of security matters represent consecutive security stages. For illustrative purposes, the table depicts the view on vulnerabilities. Herein an unknowable-unknown represents a so-called 'zero day vulnerability' which is present in a piece of software, but *nobody* knows of. The unknown-unknown represents a case in which the 'zero day vulnerability' is used, but the vulnerability is not known. Stuxnet exploited many zero day vulnerabilities, possibly resulting in accidents with the centrifuges. The known-unknown represents the case in which the affected party notices (or is made aware) is aware something is wrong with a piece of software. The final stage is that the vulnerability is discovered and that a solution is available. This solution could range from not using the affected asset (which is a trade-off of the risks of continuing to use the asset), disabling/protecting the vulnerable part of the asset or possibly even patching the vulnerability.

Importantly, these unknowns could be about everything. For example, accidents could also be defined as unknowns. Some organisations had to be alerted that their systems were compromised by an adversary and that their data was copied. And that in turn could result in another type of unknown to another organisation. For example that with the copied data user credentials were acquired, which leave systems of other organisations vulnerable.

2.4 STATE OF COLLABORATION

The need for collaboration

In order to improve the defensibility of organisations numerous persons expressed that organisations have to (be allowed to) share information with each other. Examples are:

- Neelie Kroes, vice-president of the European Commission (Kroes, 2012),
- Francis Maude, minister for the Cabinet Office (Maude, 2013),
- Keith Alexander, director of the U.S. National Security Agency (Finkle and Menn, 2012).

The motivations for information sharing collaborations boil down to:

- pooling information about new threats, new vulnerabilities and insights on new defences against all kinds of attacks (Maude, 2013), all this to prevent as many accidents as possible,
- sharing information with each other in the case accidents took place, to make sure other organisations are on alert for the same attacks (Maude, 2013), but also that they can precautionary measures against consequences of the attack (such as revoking user credentials like Facebook and Diapers did recently after the attack on Adobe (Goodin, 2013d)), and possibly also to,
- discover unknown incidents or accidents by detecting patterns in the activity on systems of different organisations (Zhao and White, 2012, p. 458).

Current state of current collaborations

Although there are many collaborations in development in many nation states, with the United States appearing to be the frontrunner, the setup, directions and developments differ quite a bit.

In the United States the collaborations are organised per sector (National Infrastructure Advisory Council, 2012). In the United Kingdom a collaboration (CISP) is launched in which organisations from different sectors are part of the same collaboration (Cabinet Office, 2013).

But also the way the collaborations develop differ. With some of the collaborations in the United States cyber security is merely added to the agenda (National Infrastructure Advisory Council, 2012), hereby benefiting from prior arrangements (Rashid, 2013). Other collaborations, such as CISP in the United Kingdom are specifically developed for cyber security (Maude, 2013).

Finally, also the developments vary quite a bit. According to a study on public/private information sharing partnerships in the Unites States by NSS Labs the progress is 'uneven' (Rashid, 2013). NSS Labs particularly noted the limited progress in critical infrastructures. And explanation being: "We are still struggling to find and enable the right level of public/private cooperation and responsibility assignment to protect the nation's critical infrastructure" (Rashid, 2013). Still, overall the information sharing arrangements tend to discuss topics at a strategic level (Rashid, 2013). After all, big picture discussions are easier in large groups[6]. It also limits issues with clearance on confidential issues (particularly of private parties), such as with critical infrastructure topics. Limited concrete threat related data is shared, which is the type of information organizations need. NSS Labs hereby specifically singles out private sector participants that need information that is "specific, timely, and actionable". (Rashid, 2013) Finally, in their its report NSS Labs warns that information sharing programs have to think seriously about civil liberties and privacy, also as legislative attempts have met strong opposition (Rashid, 2013). In contrast, in the United Kingdom the public and private organisations, from a wide variety of sectors are about to *voluntary* share information with each other in one new collaboration (Maude, 2013). However, at first the collaboration is solely open to larger organisations representing the critical infrastructures (Maude, 2013).

2.4.1 Approach of this research to guide in setting up a pragmatic information sharing collaboration

Given that the designs, scopes and success of the various information sharing collaborations vary quite a bit, the question is if there some recommendable default design can be distinguished. The identification of such a design is the main research goal of this research. Specifically this research focuses on the development of a collaboration:

- that has the purpose of sharing information allowing to detect unknown-unknowns,
  These unknown-unknowns are the real challenge because they require participants to share possibly confidential information. Furthermore, with these unknown-unknowns different organisations might have different roles at different times. At some moment in time the organisation can be affected, whereas at another they could be necessary to make sense of

[6] Such as with the most mature and most successful one, the financial service centre with 4.400 organisations.

an attack at another organisation, and at yet another they might be able to point to a solution.

- that forces organisations to clearly articulate the intentions of the information sharing collaboration and to back this up with decisions regarding the design of the collaboration, and

  Such an articulation of the intentions is considered to be a contributing factor to the motivation of organisations to actually share the information. Organisations should know why they would have to collaborate, for what purpose. Articulation of the purpose can make the collaboration interesting to organisations.
- that motivates organisations to actually be willing to share information because they trust the other participating organisations.

  Whereas the preceding bullet focused on the being of value, this one is about whether organisations are also inclined to actively share information. They should be willing to share information because in some way they trust the other organisations up to the point that there is a net value to them to join.

With this focus on willingness of information sharing this research is supposed to complement prior studies and projects that focus on the facilitations of the needs in information sharing. Those other studies and projects focus:

- on the collaboration architecture (such as CAIS (Skopik et al., 2012)) or information sharing platforms (such as MISP (ENISA, 2013, p. 23)),
- on the identification of roles, competencies and tasks in information sharing systems (such as CAIS (Skopik et al., 2012)) or the organisational structure in communities under different alert levels (Zhao and White, 2012, p. 460), and
- on various ways of standardizing indicators of compromise (CybOX and IODEF) and securely exchanging such information (RID), representation of threat related information (STIX, OpenIOC) and automated delivery thereof (TAXII)(ENISA, 2013, pp. 12–21)

But for this research to focus on information sharing on unknown-unknowns it has to consider the supportive collaborative environment. It is about identifying the required steps in setting up such a collaboration, about discussing extreme options for these steps and discussing the impact of decisions on these steps for the sharing of information. All in all, the challenges this research has to deal with are:

- challenges on the shared information such as the topic of shared information (such as on vulnerabilities, security incidents or on how to best defend against some specific threats), the type of shared information (data, information or knowledge) and the type intended response following from the collaboration (such a recommendation to the participating organisation or a coordinated response to an external body on behalf of participating organisations).
- challenges on how the collaboration should be organized to be able to serve the intended information sharing purpose (such as on who can join the collaboration, what the organizational structure is like and how the information is exchanged).

Crucially, it is the main contention of this research that the information sharing needs supportive structure. After all, the collaboration is not a goal and is not the actual issue. The organisation or design of the collaboration should be such that insights on unknown-unknowns can productively be shared. And hereto the organization of the collaboration should be supportive, not leading

Summarized, and to sharpen the preceding formulations of the goal, the goal of this research is to:

- identify *critical decision steps* relevant to the development of a collaboration, that will be presented on a roadmap allowing organizations to setup,
- a custom collaboration that supports the sharing of pragmatic information on unknown-unknowns, for which the organizations are *willing* to provide the information

The methodology to reach this goal is discussed in the third chapter, the results are discussed from chapter four onwards.

3 DEVELOPMENT OF METHODOLOGICAL ASPECTS

In this chapter the in this research used models and roadmaps are discussed. The goal of this research is to come up with a roadmap identifying the required steps to be taken to setup a collaboration. This is achieved in three stages, resulting in a base roadmap and two stacked refinements thereof. The required steps are identified using two models, three theories and three scenarios. The two models are discussed in 3.1, the roadmap in 3.2. The three theories and three scenarios are discussed along with the three roadmaps in the chapters four until six.

3.1 IDENTIFICATION OF INITIAL SET OF STEPS USING TWO MODELS

To identify an initial set of steps which have to be considered in the development of the roadmap, two models are used: a descriptive-empirical one by Kowtha et al. (2012) and a prescriptive-theoretical one called the Bow Tie model. The motivation for the selections of these models and their purpose is discussed in the sections 3.1.1 and 3.1.2.

3.1.1 EMPIRICAL KNOWLEDGE (KOWTHA)

The foundation of the contents of the roadmap is provided by the descriptive model of cyber defensive collaboration centres, which was developed by the Applied Physics Laboratory of the John Hopkins University. Hereafter the model will be referred to as the model by Kowtha et al.. The purpose of the model is to be able to *characterize* operation centres uniformly. For that, the developers identified several factors and additional subdivisions that highlight all kinds of aspects of an operation centre. Using the model it is possible to quickly identify similar operations centres to find "opportunities for collaboration or complementary activities". (Kowtha et al., 2012, pp. 3–4)

The initial model is developed on the basis of interviews, observations, document reviews and questionnaires on "about half-dozen" centres from defence/intelligence, federal/civilian and commercial sectors. The revised version which is used in this research was improved on the basis of actual data from four of such centres. (Kowtha et al., 2012, pp. 3–4) The revised model now covers the content (as in topics discussed) and development of operations centre related aspects (Kowtha et al., 2012, p. 3).

Figure 6: Visualization of the scope of a centre (Kowtha et al., 2012, p. 32)

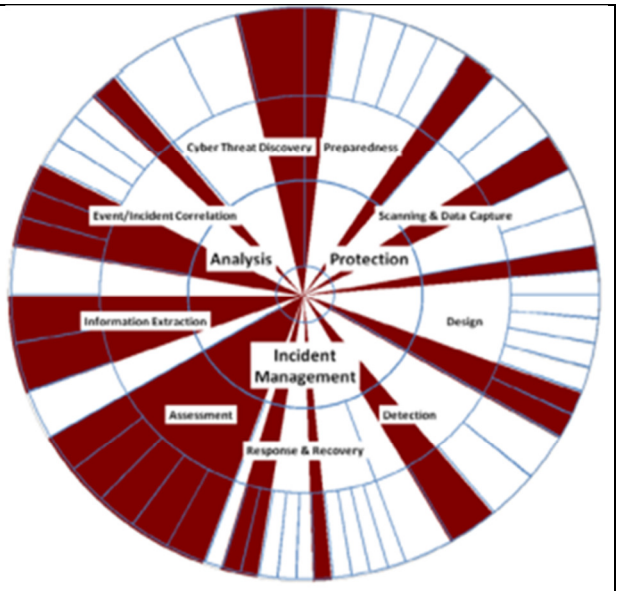


Figure 7: Visualization of the activities as performed by a centre (Kowtha et al., 2012, p. 35)

The model is actually a set of visualizations of centres onto some dimensions, which are described by factors, its attributes and the respective values. Two of such visualizations are

Figure 6 and Figure 7. The factors are used to differentiate the operation centres in terms of goals and means of operation. (Kowtha et al., 2012) The dimensions and attributes are in this research used as steps of the roadmap model. For a complete coverage of all dimensions, factors and attributes of the model by Kowtha et al. (and its representation in the roadmap model), please refer to A3 on page 81.

Being the resultant of an analysis of 'completed' centres, some of the preconditions or elements which actually make the centre function properly might have been overlooked and not become part of the model. For example, the level of trust between participants in the collaboration. As a result, the model cannot be considered to be a complete model. Neither can it be considered a 'correct' model as it is based on studies of 'random' actual centres. It is quite possible that those centres were not per se the best centres relative speaking and more importantly, not per se actually performing entirely as desired. The model was never intended to assess the quality of a centre anyway (Kowtha et al., 2012, p. 3). But it was the best descriptive model that was encountered in the literature study phase of this research that both studied multiple centres and uniformly presented the results.
Not only was it the best, it is also considered to be of use to this research. Given the fact that the model is supposed to be able to describe actual collaboration centres it is considered that it can also be useful as a starting point of defining new ones. The descriptive model provides the topics of discussion (the dimensions), the options (factors) and evaluation aspects (the attributes) which in the start-up of a centre should *at least* be discussed.

3.1.2 THEORETICAL KNOWLEDGE (BOW TIE MODEL)

Whereas the model by Kowtha primarily focuses on (characterising) actual operation/collaboration centres, it is less pronounced about what information is shared. To structure findings in that regard and to complement the findings to get a better coverage of what parties can discuss in a collaboration the Bow Tie model is used. The Bow Tie model (illustrated in Figure 8) is a theoretical model which focuses on actually improving the safety of an environment. *It is a qualitative, event based, risk analysis method to structure and systematically analyse the risks and measures to minimize the risk in an organisation using protective barriers*. (Nordgård, 2008)(RPS, 2012)
The Bow Tie is selected because it is both useful in conveying possible topics of discussions for collaborations, but it is also useful in the actual collaboration itself. The reasons for that are that it:

- the approach as conveyed, in thinking about cause-effect and introducing barriers is relevant and *applicable* to the challenge of cyber security,
- is rather *intuitive* to use to systematically analyse events leading up to accidents because of its orientation on causes and effects and containment of blocking events and effects using barriers, and
- is *useful* in provides insights on what kind of aspects information can actually be shared between organisation, such as that some barriers are ineffective in stopping some threats.

A more extensive discussion of the considered alternative types of models can be found in A2.1.

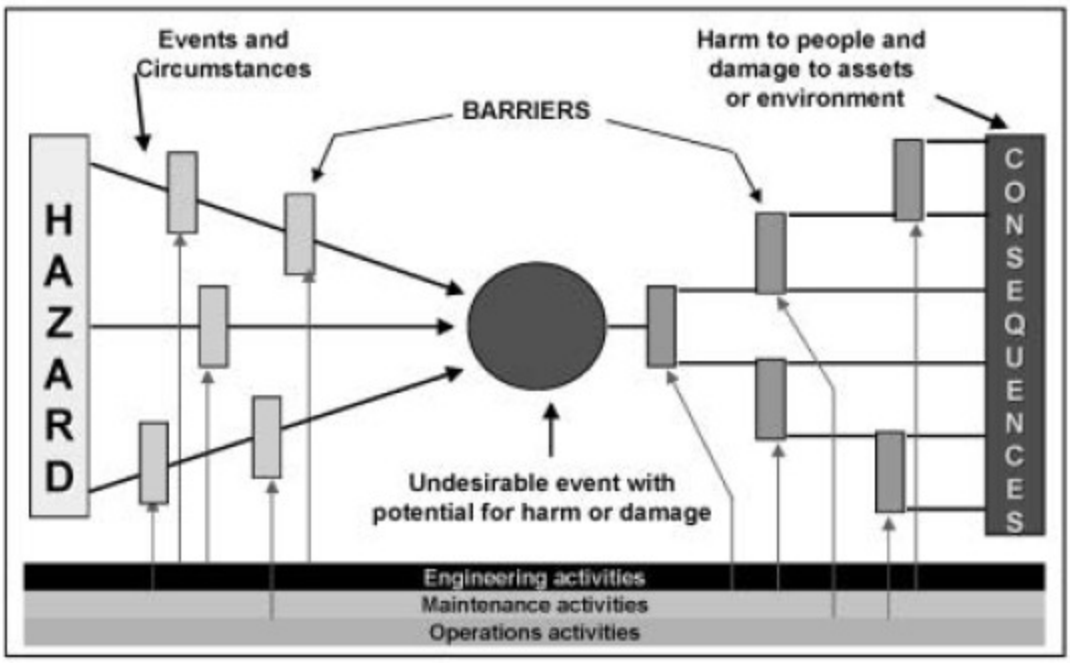


Figure 8: The Bow Tie model as visualised by Shell International Exploration & Production (Léger, 2008)

The Bow Tie model is used as a prescriptive model in this research to structure and verify the completeness of possible topics of shared information. The Bow Tie model covers all aspects of risks. Starting point of every Bow Tie model is the identification of a hazardous activity (such as storing confidential information) as displayed in the left part of Figure 8. Related to such hazardous activities are threats (in the figure represented by arrows of 'events and circumstances'), which can cause a loss of control over the asset (an adversary being able to access the data) and ultimately result in some negative consequences. With that the Bow Tie covers the threats, vulnerabilities and consequences; the quantitative components that make up risk. Adding quantities would allow for risk analyses. To minimize risks Bow Tie models are about the introduction of barriers and mitigation elements as defensive mechanisms to stop threats, as is characteristic to epidemiological models. A more extensive discussion of the Bow Tie model can be found in 0 at page 77.

In this research, given the open character of the cyber security problem and the focus on collaboration between organisations, the *adversary*, *target* and *risk value* are added 'to' the Bow Tie model. The *adversary* is an undefined source of the threats and the *target* is the organisation which is under attack. Finally the *risk value* (or actually risk level) becomes 'part of' the Bow Tie (as discussed and illustrated in section 4.1).

3.2 DEVELOPMENT OF THE ROADMAP

As discussed the goal of this research is to develop a roadmap that identifies what participants have to decide upon to setup an information sharing collaboration. It does not say what to do, merely what can be done, at abstract level and in a descriptive way. For that, first, this section covers the template of the roadmap, identifying and distinguishing the different aspects of the roadmap. Next it is discussed how the template is used and along with that it is discussed how the scenarios and theories come back in all of this, which in turn necessitate the development of three roadmaps.

Aspects of the roadmap

The combination of the model by Kowtha et al., restructured according to and complementary parts identified using the Bow Tie is captured in a roadmap. The roadmap consists of four states and displayed in Figure 9:

- the current state,
- the desired current state,
- the 'backcast'-state (later on referred to as the 'development of the collaboration model'-state),
- the end state (or the goal of the collaboration).

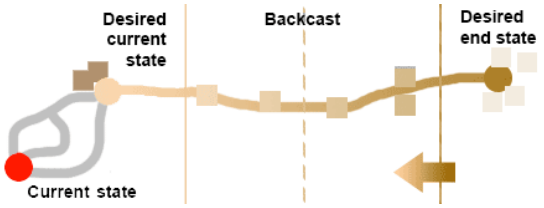


Figure 9: A roadmap-model with on the left the 'You are here'-dot, indicating the current state from which multiple possible routes are imaginable to next, the desired current state. The desired current state and the desired end state (all the way on the right) are separated by a backcast with some steps to be taken to come closer to the desired end state.

In this research this roadmap will be used from the end state back to the desired current state. This so called backcasting method is about defining a goal first and identifying the required preceding steps, as demonstrated in similar fashion by (Milan, 2008, pp. 29–33). With that, it is a counterpart of forecasting, which considers the elements currently available and forecasting their behaviour or values in the future. In this research this backcasting is taken to the extreme by ignoring the current state and (im)possibilities, as will be discussed at the end of this section. Every roadmap state contains one or more *roadmap steps*, which are to say topics organisations have to decide upon in the development of the collaboration. The options per topic are the roadmap step options.

Example of backcasting

> *An everyday example of all this is defining the end state as being checked in at the airport at one o'clock in the afternoon. Next would be to identify the necessary steps prior to this goal, such as the transportation from the home residence to the airport. For that specific step some options are possible, such as public transportation or a cab. And a decision on this step might affect prior steps as for example public transportation might take longer. That would require an earlier departure from home and might affect the (scheduled) available time for packing.*

The motivation to opt for backcasting is to ignore for the moment what is available and possible right now in terms of developing collaborations. The main goal is to define the goal and to find out how to reach the goal.

The end state is in this research about identifying what the collaboration on cyber security is supposed to be about, what it is supposed to achieve. What kind of risk should be treated and thus become the topic of discussion in collaborations between organisations on cyber security? To reach this goal preceding decisions and steps have to be taken. These steps are part of the 'development of the collaboration model'. The starting point is the undefined *desired current state*. This state is about having all ingredients in place *to be able to start the setting up of a collaboration*. The state is undefined as there will be not one combination of ingredients which make organisations want to participate. It could be fuelled by a considered need, by being forced[7], a perceived added value of participating or some combination thereof. The final part is the actual current state. After all, not all parties have the same perception, there is no such thing as *the* problem. For example, according to an American survey, many small and medium sized businesses perceive they are safe from cyber threats (National Cyber Security Alliance and Symantec, 2012). The current state thus defines the current state of perception of the problem in society, rules & regulation regarding security and the collaborations in execution. The two forms of current state are thus related to actual and required (hereby intended) readiness for collaboration. The current state is depicted, yet is *not* the focus of

[7] This is not the preferred approach as that way the participants will in all likelihood be more inclined to hold back in their information sharing efforts, not to say they will do the absolute minimum required. Ultimately they might recognize some added value, but the chances thereof are limited if many, if not all, participants have this attitude.

this research. There are various ways to end up from the current state in the desired current state. Examples are to use regulation (such as the, proposal to an, European directive requiring organisations to share information in case of data breaches), but also by leading by example by demonstrating the value of information sharing or (convincingly) confronting organisations with the fact that they are not safe from security breaches. All this can contribute to the perception of organisations that they have to collaborate, which closes the distance or gap between the current and the desired current state. But given the fact that the size of the gap, the sensitivity to the influences and the importance vary per organisation the gap and the current state altogether is considered to be important, but omitted. In this research is assumed that the organisations are in principle willing to collaborate. The focus is on how such a collaboration could take shape.

Roadmaps and scenarios as requirements

The findings of this research are not captured in a single roadmap. Instead there will be a base roadmap and two subsequent specialized versions. The design space for a roadmap for an information sharing collaboration is considered to be comprehensive and too complex to solve and illustrate at once. To separate concerns of complexity three scenarios of collaborations were envisioned which are the bases of a developed roadmap. Each situation introduces another specific complexity of information sharing collaborations. Every roadmap is developed on the basis of a set of assumptions regarding the type of collaboration and the scenario of the collaboration. They are to say the requirements dictating what the collaboration has to be able to cope with. Such as whether organisations participating in the collaboration are free to act or whether they have to answer to their clients. Every subsequent specialization has one or more assumptions removed, making it a more realistic collaboration but also more complicated. The outcome per scenario is a roadmap which is built on the basis of the gained knowledge.

The first scenario is about setting up orderly collaborations in a secluded environment with participants having a single agenda(-item) of sharing information in a collaboration model. It is assumed there is no interference of outsiders possible, the configuration of participants in the collaboration will never change and there are no discernable differences between the participants. To say, they are identical, working together in the collaboration in a common and shared goal.

The second scenario is about a collaboration which is under influence of (and influences) its environment. And with that the question is raised why somebody would share the information. But assumptions of the lack of discernable differences between participants and a common, shared goal are still in place.

The third scenario does not longer presuppose a single non-evolving collaboration with the same participants throughout time. Neither do the participants have to have the same goals, same views and the same capabilities. With that, this scenario best meets the challenge of large collaborations that are open to anybody.

Identification of the steps on the roadmap

The identification of the steps that are added to the roadmap is a 'two round' approach that takes place three times. Two rounds (round two and three in the overall process) per roadmap. This illustrated in Figure 10.

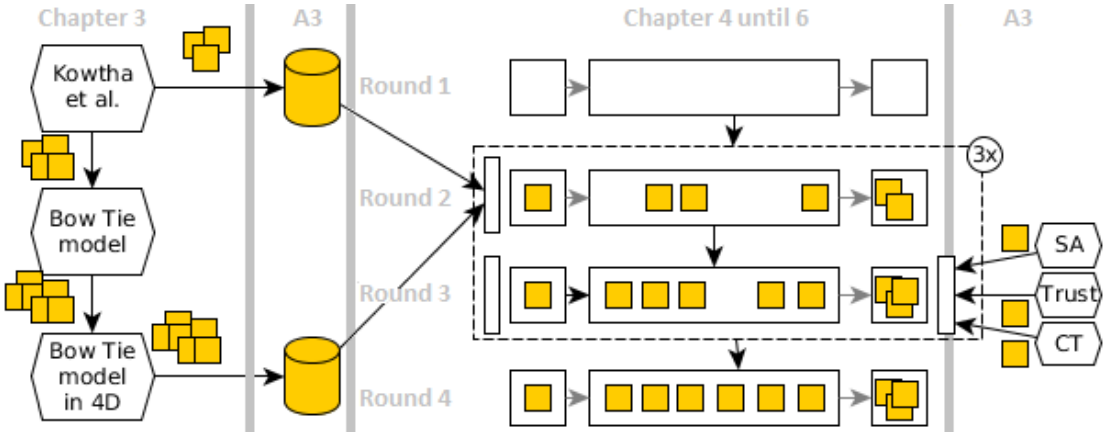


Figure 10: A visualization of the methodology. The figure depicts the two models (left), the process of adding steps to the final roadmap (right of the middle) and the three theories (utmost right) which provide some additional steps.

On the left it depicts the two models (Kowtha et al. and the Bow Tie) from which several steps were identified (the little squares). These steps were collected and served as the starting point of the filling of the roadmaps. (The full list of identified steps on the basis of Kowtha et al., and the occasional redefinition thereof, is depicted in Table 10 on page 82.)

For each roadmap relevant steps from the 'buckets' were selected and added to the roadmap, as depicted in the right part (in the black rectangle). Based on the scenarios (and the relevant, corresponding theories) some additional steps were identified and added. (The newly identified steps are depicted at the bottom of the aforementioned Table 10 on page 82.) This resulted in three roadmaps, with the third being most encompassing one, in being the result of three scenarios.

First round: Identification of the initial set of steps using the Kowtha- and the Bow Tie- model

The descriptive model by Kowtha et al. and the prescriptive Bow Tie model are used to identify what parties have to consider in the development of a collaboration. They serve as the source of the initial set of roadmap steps.

However, the models do not perfectly overlap nor complement each other. Both models miss some aspects and each will be complemented accordingly.

Being descriptive, the dimensions, factors and attributes identified by Kowtha et al. are in this report contrasted to the Bow Tie. Herein the Bow Tie is leading. The Bow Tie model is used in a prescriptive sense given its rich history of use and that it is focuses on the analysis of risk. The model by Kowtha et al. also extensively covers peripheral phenomena such as discussions of the design of the facilities. As a result thereof, some elements of the model by Kowtha et al. will become obsolete (such as the factors describing the facility in *number of desks* and the *layout type*), divided (such as *functional abstraction* into the *type of information sharing* and the *method of information sharing*) or repositioned (such as activity which takes a more prominent role as a dimension). Furthermore, some steps will be added, steps which were not in the characterization of operations centres by Kowtha et al.. An example is the topic of shared information which is far more extensively covered than in the model by Kowtha et al..

But what is missing in the Bow Tie model is the notion of time. The model by Kowtha et al. distinguishes activities of protection, incident management and analysis. This notion is added to the Bow Tie. Bow Tie models are about systematically analysing the risk of the hazardous activity, which might lead to a loss of control over some asset. But such a model can actually be used as:

- a forecast of such a scenario resulting in a loss of control,
- a depiction of an actual case of an organisation that got attacked, or
- a representation of the situation as a reconstruction, to learn what happened, how and why.

The entire list of identified steps and its origin is presented in the appendix in Table 10 on page 82. Along with the steps, the decisions on aspects of the collaborations, several options are identified. Some were already provided by Kowtha et al. These options are solely discussed in the appendix 0 with their respective step. In chapters four until six the roadmaps with the relevant steps are presented.

Second round: Identification of additional steps using additional theory

With each roadmap another specific theory comes into play, which deals with specific issues in a collaboration setting. The three theories are, in respective order, (distributed) situation awareness, trust and the configuration theory. The theories are used to explain behaviour of participants and point to crucial steps participants have to take or consider. For example, publications on situation awareness raised the awareness about the impact of 'organisational structure'. Depending on the organisational structure and the conditions would the collaborative performance differ. (Sorensen and Stanton, 2013) Because of that, 'organisational structure' is added as a step on the first roadmap. Given the position of the theories to explain behaviour and occasionally even dictate on what to do, the quality of the publications and the alternatives were evaluated. Such evaluations are discussed in the appendix (situation awareness in A3.2.1.2.5, trust in A3.2.2.1 and the configuration theory in A3.2.3.1).

In the end credit is taken for selecting, linking and restructuring the notions from the models and theories as steps in the three roadmaps. The actual roadmaps displayed in the right part of Figure 10, are discussed in chapters four until six. The motivation of the use of the three remaining theories (situation awareness, trust and the configuration theory) is covered in those chapters (4.1, 5.1 and 6.1).

Distinction of aspects of roadmaps

With all that the four recurring roadmap related terms are now introduced: the roadmap, roadmap states, roadmap steps and roadmap step options. Their relation with each other is illustrated in Figure 11.

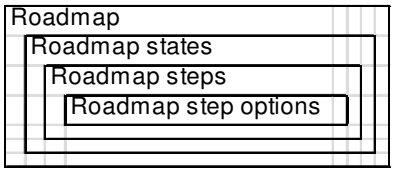


Figure 11: Distinction of aspects of roadmaps

In brief, there are three roadmaps, simply named 1,2 and 3. Each roadmap has three states: the current state, the 'development of the collaboration model'-state and the end state. Per state one or more *roadmap steps* are identified, representing decisions organisations will have to make about some aspect of the collaboration. And finally per roadmap step, or per decision, several options are identified. For convenience, the distinction is also presented in A1.

4 SCENARIO 1: A STATIC, AUTONOMOUS COLLABORATION OF EQUAL AND FULLY COMMITTED PARTICIPANTS

The base roadmap corresponds with scenario in which a collaboration of parties is situated unaffected by the environment for an undefined period of time to discuss some topics. Participants are under no influence of their own organisation, nor the environment of the collaboration. An environment that amongst others consists of attackers, public opinion and formal institutions such as laws. Due to this lack of pressure from the outside the collaboration can be considered static. There is no dynamic as there is no change in the configuration, the initial configuration of participants remains unchanged. And even the topic of discussion will remain the same.

4.1 METHOD

Besides selecting the relevant identified steps of the model by Kowtha et al. and the complementation using the Bow Tie model a complementary theory will be added. In this case Situation Awareness (SA) as a describing theory. The relevance of this theory is to provide a background in the presumption of the intended collaborations of participants having some level of awareness of cyber security related matters. The next step is for the participants to improve their awareness by 'sharing' their awareness. Such 'sharing' of SA at the collaborative level is also covered in this chapter, albeit in a more conceptual prescriptive manner. Of the high alternatives the Distributed Situation Awareness (DSA) applies best to this research. The main differentiating factor, as compared to the alternatives, is in the leveraging of different perspectives of DSA. A discussion of SA and DSA follows in section 4.3.

Based on the topic of collaboration and Kowtha et al. additions are made alongside the Bow Tie model. The underlying reasons are further explained in A2.2.2. One of the additions is the target, illustrated on the right side of the Bow Tie. Furthermore the risk value, although not really intended with the Bow Tie (CGE, 2013), is instated as the depth of the Bow Tie model. The counterpart of the target is added on the left side. The result is displayed in Figure 12.

Figure 12: The Bow Tie model with three dimensions. Attacks are launched by adversaries (left) and can have consequences to an organisation (right) if defensive measures fail. The risk of such a failure is illustrated with the depth of the Bow Tie model.

4.1.1 ASSUMPTIONS OF THE COLLABORATION MODEL

The entire described setting of the collaboration is based on quite a few strict assumptions. These assumptions are reflected in the roadmap as well, accompanied by some additional assumptions:

- An autonomous collaboration
  The assumed autonomy of a collaboration refers to the ability of parties in a collaboration to act freely. Neither the environment of the collaboration, nor the own organisation of the collaborator can influence the participants in any way.
- An unchanging collaboration over time
  Coupled with autonomy the degrees of unpredictability are limited to those generated by the participants of the collaboration. The collaboration can truly act autonomously if there is no change of coalitions in the collaboration which can result in new dynamics or a leakage of information.
- Identical quality levels of participants in the collaboration
  Quality in this assumption refers to the level of the skillset the participants bring into the collaboration. Preferably the participants are all top of their class *in their respective fields*, as in, they have full awareness of their field. The possibly vastly different backgrounds of the participants might result in troublesome collaboration. Those can be affected by whether one invites a participant at the start or not. Once the participant is part of the collaboration, there is no way out for the participant. The purpose of this assumption is to not have a discussion about whether some other participant *from the same field* would be more useful

in the collaboration. The relevance of this assumption becomes clear in the final roadmap, as discussed in chapter 6.

- Minimizing the amount of steps in the end state is preferred
  This assumption is about minimization of the amount of goals formulated of the collaboration related to those steps. A collaboration should have some clear goals, a clear purpose. However, defining too many goals will result in some means being defined as a goal in itself. And although defining goals on all sorts of aspects results in a focused discussion, it also limits the amounts of remaining options to achieve those goals. Finally, the amount of goals should not be confused by the clarity of those goals. In a sense the minimization comes down to actually defining the goals as end states and means as actual means. However, ultimately this distinction can become blurred. As an example, the goal of a collaboration can be to discuss how to thwart some threat (d-dos) on some target (banks) as the d-dos can be a smokescreen to some actual attack. But in this example the goal of thwarting a d-dos is a mean to thwarting the 'actual attack' on the banks.
- The roadmap steps are orthogonal
  Orthogonally is about the ability to select a different option in a step, without affecting some other step. It is actually more of a desired situation. With an orthogonal design changing (and interchanging) any variable does not affect any other variable in any way, but is also not constructed of (some of) the same constructs. Orthogonal variables are a special case of linear independent variables. As explained by Rodgers et al. (Rodgers et al., 1984) in geometric terms, vectors of independent variables do not fall along the same line. With orthogonal variables the axes are even at perfect right angle to each other. They share no linear components. An example would be to select different participants, yet share the same topic of information. The two have nothing to do with each other, one is about who shares the information, the other what information is shared. In some cases the assumption is far from plausible. Even with structure, in some scenarios this won't be likely. For example, if some party desires to share confidential information, but there is no trusted independent party, organisational structures intended for such an arrangement are no longer usable either.
- For each instance of a roadmap there can be only one option (roadmap step option) selected per roadmap step
  This assumption has to do with the ability to determine the success of the collaboration. If there exist multiple organisation structures in a single collaboration it is hard to assess the contribution of the organisation structure to the actual performance of the collaboration.

4.2 Definition of the desired end state

The definition of the desire end state is about defining *which attacks* will be considered in the collaboration with the intention of thwarting those attacks. The used definition of an attack provides the first two steps to consider, as an attack comes from some (1) adversary and has an effect on some (2) target. This results in a spanning of *types of attacks*. And as probably not all attacks can be covered in the collaboration the (3) risk value of a Bow Tie is another step to decide upon beforehand. The risk value per Bow Tie is necessary to focus in a collaboration on the *specific attacks* of a certain type with a specific sense of urgency. The three steps are discussed in this paragraph in more detail.

Type of attacks

The type of attack represents an attack coming from (1) some adversary on some (2) target. Attacks are becoming increasingly targeted and insidious (Symantec, 2013b, pp. 20–21). This combination makes it increasingly difficult to defend against these techniques because fewer organisations face the same attacks. A response would be to focus in a collaboration on those targets facing similar attacks. Similarity can be reach as far as identical attacks just on a different target, but also more vague in terms of similar patterns of attack or similarity of targeting the same vulnerability, albeit

using a different method. The troublesome bit is that 'targeted' does not necessarily imply individual organisations from *the same sector*. Attacks could also be limited to multiple sectors, but hereby being limited to (at the extreme) one organisation per sector. With that quite a few target options of attacks are possible, such as (a combination of) type of sector, size of the organisation or type of activity.
On the other side of the attack is some type of adversary, which has some motivations, resources and some capabilities for attacks. The focus on motivation (driving force of who is being attacked) and capability (driving force of the sophistication of the attack) is the actual purpose of definition of the adversary of interest in the collaboration. The adversary is simply the catchy proxy. A coarse distinction of the more serious adversaries is in criminals, hacktivists and governments. But far more specific distinctions are possible. (A3.2.1.1.3)

Attacks
Selection of attacks is about discrimination on the basis of the risk value. Some of the considered type of attacks might not be worthwhile to discuss in a collaboration. Depending on decisions regarding adversaries and targets, representing the types of attacks, many attacks apply. But the risk of those attacks can vary quite a bit. Risk refers to the probability times the impact. And with that some collaborations can focus on 'bandwidths' of probability, impact and/or risks of attacks. A bandwidth has a minimum and a maximum. An attack which has a very low probability might not be worth the discussion (possibly regardless of the impact). But similarly, a probability of near one hypothetically could also not be worth the effort of discussion. In the latter case organisations should just be notified of the attack, but they should act upon it, regardless. Whether participants in a collaboration use risk or its components instead and how stringent is a matter of choice.
The practical implementation is troublesome (nevertheless actively used) on a collaborative level. It is hard to objectively determine the risk-value(-component). The environments of organisations are different, making some attacks more successful and with more impact than in others. And the subjective part further complicates the selection criterion. Not all organisations will be equally capable of assessing the risk-value(-component) accurately. The result is a less accurate risk-value(-component), but also the meaning of that value will be interpreted differently by parties. Some will accept a higher risk as accepting that risk is *considered* to be worth it (assuming there is even such an evaluation of costs and benefits).

Decisions regarding these three aspects of an attack is a balancing act. For example, a rather comprehensive definition of considered targets of attacks will yield more attacks to consider. This *requires* a collaboration which is able *to fulfil* the more ambitious goal of thwarting more types of attacks. It might for example require more participants or different types of information. An inaccurate match will result in (more) accidental discussions, with the possibility of a backfire. Participants might not be interested in the collaboration (as it might seem to cover too much irrelevant aspects) or the collaboration might prove unproductive.
It might appear a bit vague to define a goal of thwarting some attacks, based on decisions regarding considered adversaries, targets and risk value. In the end defining steps as being part of the end state, and thus the goal of the collaboration is a balancing act. (A coarse discussion of considered high level alternatives is presented in A3.2.1 on page 83.) In the selection of the steps the main consideration was to leave the degrees of freedom to the participants. Defining a collaboration to thwart specific attacks, using a specific mean, under all sorts of stringent conditions will be focused, but has its drawbacks too. Such a focus might not appeal to organisations to join, might stimulate single mindedness and (as a result) might even be unproductive. The current focus is on leaving sufficient options to avoid attacks in some ways (by not adding too many steps to the end state) and yet still meaningfully focusing the discussion. It allows participants to think of the motivations of adversaries and goals, not focusing on specific techniques. There is an increasing number of attacks possible and some inherent weaknesses (such as humans which use weak passwords or act less careful in stressful times), protecting against all these is merely a reactive mode of response. Instead,

a collaboration focusing on thwarting an attack could focus on securing the targeted asset. Whether that is an option and the manner in which this could be done differs, but at least it should be possible to discuss this in a collaboration, and not ignoring it entirely due to the focus of the collaboration.

4.3 DEVELOPMENT OF THE COLLABORATION MODEL

The 'development of the collaboration model' is about the means to the end of thwarting the defined specific types of attacks which was identified in the steps in the previous section. In this research information sharing collaborations are the method of choice.
In the end, the purpose of information sharing comes down to improving the situation awareness of participants in a collaboration. Presumably the improved situation awareness will result in better decisions which are presumed to result in better performance. First a brief explanation on the concept of situation awareness is provided, a more elaborate version thereof is provided from A3.2.1 onwards. Situation awareness on the collaborative level, specifically using Distributed Situation Awareness is discussed next. More details about the specific approach and the motivation for that approach are provided in A3.2.1.2.5.

Situation awareness
As described by Endsley Situation Awareness refers to the level of knowledge about a situation in an environment. She distinguishes three cumulative levels of situation awareness. (M.R. Endsley, 1995, pp. 36–37) . An agent with level one awareness merely perceives the element in the current environment. With the higher level of awareness the agent is also able to comprehend the current situation on the basis of those elements. The final, third, level is about understanding what the current situation means for the future. An agent with level three awareness of activity on some server knows:

- the users of the systems, running services and data on the system (level 1),
- some user currently active on the system copying a file it isn't allowed to access (level 2), and
- that this file contains confidential data that can have consequences (such as financial repercussions) to the legitimate owner if it falls in hands of an adversary (level 3).

From this follows situation awareness is awareness of a certain level on a certain environment/topic/situation.

Given the fact that some organisations had to be informed about consequences taking place implies that they have limited situation awareness of their own cyber security (related) environment. And others only knew their data was copied or they were spied upon after the fact. Whether this should be attributed to an increasing, yet insufficient, awareness or an increase of the total amount of cases is speculation. In the end what matters is that many organisation appear to not have level two awareness in all respects, let alone three. They are increasingly confronted with consequences and have to analyse what occurred, indicating they do (or at least did) not comprehend the current situation. For that reason, implicitly it is considered organisations have to collaborate to improve their situation awareness.

Situation awareness on a collaborative level
The purpose of the information sharing on cyber security related matters is to improve the situation awareness of some agent by another agent which has a (slightly) different perspective on the issue. The most relevant theory for that purpose on the topic of this research is Distributed Situation Awareness (A3.2.1.2.5, p91). In brief, this theory does not presuppose similar perceptions, as its counterpart Shared Situation Awareness does. To the contrary, it leverages the different perspectives. Participants are supposed to in particular have compatible Situation Awareness. Awareness which can prove helpful to somebody else. The key of DSA is to get the right parties (having the right compatible SA) in a collaboration, and also make sure they high level meta SA. Meta SA is described as knowing where what information is available in a system (Stanton et al., 2006, p. 1291). Knowing how to reach those parties to offer or ask for information is of critical importance.

And with that, the links between parties determine the level or quality of DSA. (Stanton et al., 2006, p. 1308)
An example of all this would be an agent with level two or three awareness regarding access rights in some content management system. This agent might notice some newly registered users are able to high clearance level data. With DSA, the agent would not share the awareness of the elements itself (such as users and type of data), but the information of parties being able to register and acquire high clearance levels. Such an agent would have to know who in the collaboration has the same system, the same role and might not know the issue just yet. An alternative would be to simply inform everybody, but if everybody would do this, this would require more processing power by the recipients to filter out the relevant information.

With the preceding discussion on Distributed Situation Awareness two related steps to decide on are touched upon; structure and roles. Despite the high quality links between participants being considered essential, this does not presuppose all-connected structure. With such a structure anybody in the collaboration can reach anybody. Which structure is selected depends on the situation as it comes down to the most productive structure, which is the result of a balance of speed, effectively and efficiency. (More about this in A3.2.1.2.2).
The second part is the selection of roles. The roles refer to the position of the participants in the collaboration. Some are the consumers, possibly the targets of the attack. Others are the contributors, they provide the information on how to proceed given the situation the consumers might be confronted with. (More about this in A3.2.1.2.1.)
The structure, roles and method of sharing are all about the actual information transactions, or as defined in DSA, situation awareness transactions. Thus far the information was simply a blanket term to an undefined type of data on an undefined topic. The former is covered by the *level of information sharing* and refers to whether (raw) data is shared, information or knowledge. The topic is covered by the *topic of information sharing*. The topic can be on technical in the sense of covering specific portions of the Bow Tie, but it can also be more about the supporting environment. A social engineering attack to extract information by means of a custom website is less of a technical issue and more about the supporting environment (employees).
The final steps of the collaboration are about the definitions of the timeliness of a response in a collaboration and the type of response in a collaboration. The type of response is about the relationship of the collaboration to the affected parties. The collaboration can be really about improving situation awareness of the participating parties, allowing them to take benefit *individually* of their improved awareness. But the type of response could also be to have the collaborating unit to actually provide a response to the environment. With that, the improved awareness is merely a precondition to the collaboration being able to provide a clear answer or response in some matter. For example to advise, urge or command decommissioning of inherently unsafe systems, or to call for more stringent regulation to force the targets to patch faster.

4.4 THE CURRENT STATE

Backcasting implies the possibility of having a required start situation which is different from the current situation. In this case, the required start situation is about readiness of participants and necessary preconditions being in place to be able to *initiate* the setup of a collaboration. In the current situation parties might not see their role in the collaboration, or do, but have no desire to fulfil their role. Many smaller organisations are an example of the first, they think they are no target of attacks. But also failing to detect (un)successful attacks contributes to this. Other organisations do see their role, but feel like they won't get anything out of the collaboration. This appears to be the case with private companies in a public private collaborations. Private parties cannot get the practical insights from pubic bodies as discussions with them are on an abstract level as more concrete data is typically confidential.
The focus in this research is on intrinsic motivations to collaborate. Future participants should recognise the relevance of collaboration and also desire to productively participate in the

collaboration. The relevance could be initiated by demonstrating the relevance (by means of confrontation with attack figures) or using stimulating conditions. The latter refers to regulations regarding fines on data leaks. The actual desire to participate has to do with conveying a clear purpose of the collaboration. With ambiguous statements, such as information sharing collaborations, without expressing the level of shared information or the topic of information sharing it is hard to get commitment.

Empirical implementation

There are multiple ways by which the required start situation can be reached. Main routes are for example by institutions, increased awareness or by example. The institutional route refers to the entirety of (in)formal institutional environments and (in)formal institutional arrangements. It affects the perception of organisation about what they should do but also what they can do. Examples of what to do are provided by cyber security strategies setting out courses of action and executive orders/directives/laws stipulating what to (not) do. For example, a recent proposal to an European directive would *require* some types of organisations to report on incidents with significant impact [35, Art. 15]. But the institutions also affect what organisations can do. Herein the informal institutional environment, such as public opinion matters. Privacy and the discussion of espionage by governments might make it difficult to organisations to share information. The net result could be a limitation disclosure of incidents. And those could help set an example to what information sharing collaborations can be about. For example, Daley et al. suggest a collaboration piggybacking on coordinated incident handling. The authors suggest to share in earlier stages of incidents or accidents, not just after full recovery, with other organisations, which can be beneficial to the other organisations (Daley et al., 2011, p. 293). With that, a reinforcing effect also becomes apparent. The use of obliged reporting of consequences for the good (of others at that moment in time), not just to punish and shame the involved organisations.

4.5 SUMMARY

Information sharing collaborations are supposed to help improve the situation awareness of participating organisations. It is believed that *in general*

- increased awareness about some specific situation
- results in better decision making in that situation,
- which is supposed to result in better performance of the organisation in that situation.

In this research this is interpreted as that in collaborations organisations have to have focus on types of attacks they will consider and wish to be aware of. Not all attacks are equally useful to discuss and discussing all possible attacks requires a comprehensive collaboration.

For productive improvement of situation awareness to be able to thwart the attacks, the selection of participants, information exchanges in a supportive structure and timeliness are steps to consider. A more structured presentation of the aforementioned steps per roadmap state is presented in Figure 13.

Current state ↓	Development of the collaboration model ↓							Desired end state ↓		
formulation goal of col- laboration	roles	collabo- ration structure	topic of shared information	level of information sharing	method of sharing information	type of response	timeliness of response	target	risk-value	adversary

Figure 13: The first roadmap, including the relevant steps

The current state refers to the intended current state. Its corresponding step refers to the steps of the end state. Thwarting of specific attacks is the goal of a collaboration, a goal which should be clear prior to and take precedence to the (setup and use of) an information sharing collaboration. The usefulness of a collaboration depends on the compatibility of the situation awareness of present roles, sharing forms of information in some timely manner, ultimately allowing participants to have better situation awareness. And this improved situation awareness is believed to allow organisations to make (better) decisions and with that perform better (by actually thwarting attacks).

4.6 EXAMPLE

To explain the interplay of the discussed states and steps this chapter is concluded with an example.

4.6.1 ABUSEHUB-PLATFORM

Early in 2013 the association Abuse Information Exchange was set-up by seven Dutch Internet Service Providers. The purpose of the initiative is to collect and correlate data on botnets (and other forms of internet abuse) infections in one place called AbuseHUB. (Abuse Information Exchange, 2013)
The information can come from Trusted Complainers, such as SIDN (one of the initiating members), member Internet Service Providers and botnet monitoring parties as The Shadowserver Foundation. The result of the processing of all the information is sent to the respective ISP. (ISP Today, 2012) For the sake of net neutrality (which comes down to ignoring/not discrimination between types of traffic) ISPs limit themselves to analyses of spam, not what customers send and receive on the Internet (Security.nl, 2012a). Although with spam (nowadays) the actual sender is often able to hide himself, the machine which actually sent the message is not hidden. By virtue of the sent message (and for example it being part of a larger campaign) it is known something is wrong with that machine. It is one of the *possible consequences* of an infection, a loss of control over some machine. Other possible consequences could for example be involvement of the machine in a denial of service attack on some website or to spy on people who use the machine.
In case a machine turns out to be have sent spam due to an infection (and the type of infection is known), the ISP receives the required information (Abuse Information Exchange, 2013) which in turn contains the account of the subscriber (Security.nl, 2012a). Along with the containment the owner of the account gets contacted on how to proceed. These developments are already in full effect (Abuse Information Exchange, 2013). And hereby ISPs also warn competitors which turn out to have subscribers which have infected machines. This is not necessarily an indication of altruistic behaviour. Warning others about infections is beneficial as less spam and phishing emails will be sent. And receiving notifications from others is beneficial as it limits the chance of getting blocked by other ISPs for high spam volumes. (Security.nl, 2012b)

4.6.2 ANALYSING THE END STATE

Analysing the example, where possible, reveals there are two different stages. The first is the moment at which a computer becomes part of a botnet. The second is about the moment at which the computer that is now part of the botnet executes an attack on behalf of an adversary. With that there are two moments of avoiding the second attack, the attack that actually results in spam or worse. The first moment is about combating botnets, which is a common mean used by organised crime. The intent of the platform is to minimize the size of the botnet, which is a crucial factor for success in the first place. The more machines are part of a botnet, the more proxy adversaries (of a single type) organisations are confronted with. To combat a botnet one would thus have to limit the amount of zombies (infected machines). And for that one would have to stop the infections. Infections of machines owned by all sorts of parties, including regular civilians. An angle to achieve this is by protecting the civilian against the infection, yet this requires inspections of data. A troublesome bit from the point of view of privacy.
Rather, the platform is focusing on those machines which are already infected. In terms of the Bow Tie model the attack by some adversary was successful, a machine is part of the botnet and the consequence thereof is a fact: the machine is sending spam from the network of an ISP it is in. This is a thing the ISP might be *unaware* of, now and possibly in the foreseeable future. However, other parties such as botnet monitoring parties might become aware or might even be confronted with the consequences themselves. The (indirect) information exchange is thus all about improving the awareness of an ISP about an infection in the network of that ISP. On the basis of that awareness (and some information about the malware), the ISP can quarantine and inform the subscriber.
With all that the proxy adversary of a spam sending machine is identified. The target is less defined, it could be some other ISP or its customers, which receive the spam. And the impact focus is also unknown, quite possibly irrelevant. There will probably be some threshold to identify whether the

spam is actually sent due to an infection. If it is, the actual amount of spam sent is not relevant. Some machines are sending just a few messages and might get other duties next time. Another type of risk is whether the cost of disinfecting the machine (by means of instructing the owner or so) does not outweigh the benefit. But in all likelihood this will be more of an exception than an actual evaluation.

4.6.3 CHARACTERISATION OF THE COLLABORATION MODEL

In the collaboration the ISPs are one of the parties performing the role of trusted complainer along with some other trusted complainer. Furthermore, they are also a potential recipient of a complaint. Such a complaint would come from the analyses centre. With all that, ISPs do not really have a conflict of interest. Their complaints will typically be about others, and they indirectly receive complaints from others. In terms of structure, it appears to be a star network, with the centre being the central intelligence centre and the connected nodes being the ISPs and other Trusted Complainers.

Functional abstraction was about what part of the Bow Tie is discussed. In this case it is about removing the hazard in the first place. The hazard being defined as having an infected machine, which demonstrated to be capable of sending spam per instruction, be connected to the Internet.

To mitigate the situation the Trusted Complainer will have to share information, specifically the spam message (possibly, allowed per rules of the collaboration in part made anonymous), in some standardised form.

Possibly in terms of type of response there is a special role depicted for the central node, the central intelligence centre, being to dictate what the ISPs have to do.

In terms of timeliness 'yesterday' would be in order. After all, at the time the collaboration receives a notice of an infected machine it is on the basis of the consequences.

4.6.4 SUPPOSED INITIAL STATE PRIOR TO THE DEVELOPMENT

Not possible to tell with certainty, but given the goal of the collaboration and the method of collaboration it did not really require a lot of work to get the participants on board. It is relatively risk free and there is a relative clear benefit.

5 SCENARIO 2: A STATIC COLLABORATION OF EQUAL PARTICIPANTS

In contrast to the intended collaboration depicted in the preceding chapter, cyber security collaborations cannot be assumed as if they exist in some secluded environment. The environment from the collaboration point of view (consisting of adversaries, clients of participants and the remainder of the organisations of the participants) has an impact on the participants and the collaboration as a whole. The participants and the collaboration will be affected by developments in the environment. But the collaboration will also try to influence activities in that environment. The collaboration will try to stop attacks from happening in that environment in some way.

The factor of time in the interaction of collaboration and its environment was thus far ignored. The collaboration is related to a (set of) threat(s) and it is positioned relative in time to this threat. In the preceding chapter the attention was consistently framed on collaborations to *thwart* attacks of a certain risk level. The risk value of the attacks changes over time, making it at some point in time become a concern for the collaboration and at some point not anymore. At the collaborative level, in this research, this time period is referred to as the attack phase. Some organisation suffered an attack and others are on alert for that attack. The attack phase is the window in which the risk is such that the threat is (or should be) a concern to organisations. Collaboration on the attack phase appears to be the main driver of collaborations in the real world. But with an attack phase two other logical phases appear, prior and after this attack phase. Prior to the threat refers to identifying the threat before it becomes a threat. And after the fact refers to identifying whether there are any takeaways from the attack. And both provide options to collaborate on.

Importantly, with a collaboration taking place relative to the risk levels of a threat, it is also affected by it. Means and ends of the collaboration will change depending on whether it is about an actual threat taking place, what might occur or what has happened. Different participants will be involved, the motivation of participants will change and the environment will respond differently.

5.1 METHOD

With collaborations opened up to the environment, including participants' own organisations and public opinion, participants have not only to answer to each other. The environment can demand participants to perform or refrain from certain activities. This dynamic has an impact on trust. To better explain these complexities the constituents of trust are identified and discussed in a descriptive manner in section 5.2 (and A3.2.2.1). Trust is not discussed in a prescriptive manner as trust is not a goal. High levels of trust are considered to be desirable for productive collaborations, yet the collaboration could technically survive low levels. The intention is to understand the impact of certain decisions on trust. The insights on trust primarily follow from the book 'Trust: Forms, Foundations, Functions, Failures and Figures' (Nooteboom, 2002).

With the addition of the time dimension, which affects the topic of discussion, the assumption of orthogonality of the steps, as assumed in the previous roadmap model, had to be removed too. Orthogonality can no longer be considered if decisions in one step (time dimension) turn out to affect other steps (topic of information sharing). It is actually this lack of orthogonality that makes the prior decision of the positioning of the collaboration relative to the attack the real challenge, as will be discussed in this chapter.

5.1.1 ASSUMPTIONS

Contrary to the base roadmap autonomy of the collaboration is no longer assumed. With the lack of presumed autonomy it is also no longer necessary to assume an unchanged configuration of participants. A change in configuration poses another possible way to influence a collaboration. Either by actual change of participants or influence of the participants. Removal of autonomy also contributes to the necessity of removal of the orthogonality (including the weaker presupposition of independence) to avoid half-heartedness/ambiguity. With the collaboration being exposed to the environment, the environment can influence the collaboration, specifically affect steps or step options. But what defines the (relevant) environment, in turn is the result of decisions on who will collaborate. That way, indirectly decisions in one step affect the other. And with that the

orthogonality assumption that was posed in the preceding chapter no longer holds true. An example of an indirect influence would be to invite a party on privacy. Without involvement of the party, they might object to data sharing (containing information) between some participants. In case such a party is part of the collaboration the uncertainty might be taken away and data sharing, perhaps even raw data sharing, might not be opposed. With that, the involvement of a participant might affect the range of options available to a collaboration, in this case *via* the environment.
With the removal of three assumptions, this still leaves three assumptions, which were (also) already discussed extensively in the preceding chapter:

- For each instance of a roadmap there can be only one option selected per roadmap step
- Minimizing the amount of steps in the end state is preferred
- Identical quality levels of parties

5.1.2 ASSUMED RULES

Without the assumed orthogonality of roadmap steps the order of the roadmap steps is a concern. In this research it is assumed that both the order and the criticality of some steps are considered to be important:

- Some roadmap steps require a decision of all involved parties or even a sufficient level of maturity prior to discussing the next roadmap step. In a sense this result in a phase-gate roadmap model. Some steps are aggregated in one phase, concluded by a gate. At the gate participants have to agree upon continuation to the next phase involving additional roadmap steps.
  For example, in the base roadmap the decision regarding the roadmap step 'roles' has an impact on the possibilities of the type of response (or at least the timeliness of the response). If participants have a single role, multiple have the same role, and all the roles cover a wide spectrum, a rather powerful type of response could be possible. Importantly, it is mentioned that decisions regarding steps impact other steps, not that one takes priority over the other. Potentially the aim could be to get a powerful response on a highly specific matter in short time. That requires a more extensive phase in which parties are contacted to make that possible. With that, suddenly the roles are the dependent variable on the (decided) required type of response.
- With that the order of deciding upon roadmap steps and their division into phases becomes important and therefore the order and the phases have to be predetermined.

Based on the assumed importance of the order yet the variety of possibilities therein, the order and divisions in phases will not be discussed in this research.

5.2 NON-ORTHOGONALITY WITH TRUST AS AN INTERACTING LAYER

Changing selected options in some steps can affect selected options of other steps. It is an interrelation of on the one side the types of information and the awareness of parties and the other the collaboration of these parties on sharing the types of information. Thus far in the research the collaboration was considered to be the mean to the end of information sharing. For that purpose it was deemed necessary to define information sharing with its *supporting* elements of collaborations. But the collaboration is not merely supporting, it is also affected by the information. And important interacting layer in all this, in this research, is trust. You cannot simply select whatever you'd like. Sometimes you need something and that requires something, some trust.

Bit of background on trust

Thus far trust of participants in each other was ignored. However, with the focus on the sharing of pragmatic and potentially sensitive information such trust is an important presumption. Depending on the stakes (loosely: required information, intentions and the potential costs) and the setting (involved parties) the required level of trust can be low or high, but it is still presumed implicitly. A counterpart which relies less (or entirely not) on trust would be setup a collaboration with countermeasures to limit the opportunities for opportunistic behaviour. This would entail to initially

instate options to punish opportunistic behaviour, detecting such behaviour and actually meaningfully punishing for that.
In absence of a system and the opportunity of not participating in a collaboration we are, in this research, in essence interested in behavioural trust (A3.2.2.1). This type of trust is defined by Nooteboom as trust in actors (Nooteboom, 2002, p. 50). Behavioural trust involves two sides, the trusting party (trustor) and the trusted party (trustee). Trust is the result of the two components, sources of trust and reasons for trustworthiness. The two are related, not directly connected, in that the sources of trust to the trustor are supposed to help indicate the trustworthiness of the trustee[8]. (Nooteboom, 2002, p. 8) This connection is imperfect, it is in the end still a wager whether the trustworthiness was the case and the trust was warranted. At the very least trust is about knowing when a trustee can indeed be trusted and when not.
Trust and trustworthiness are not all or nothing entireties. There are different forms and they can reach different levels. Trust is the result of four aspects: "Someone (1) trusts someone (or something) (2) in some respect (3), depending on conditions such as the context of action (4) (Nooteboom, 2002, p. 38). In similar terms the same applies to trustworthiness. Nooteboom refers to this as the 'four place predicate of trust' and -trustworthiness (Nooteboom, 2002, p. 38). The result is a sort of demarcation, indicating the boundaries or the limits of trust and trustworthiness. Within those limits there is no calculation regarding trust. The trust 'in some respect' comes down to some form(s) of behavioural trust. These include forms as intentional-, competence- and dedication- trust (A3.2.2.1.3).

Trust and trustworthiness represent two (not mutually exclusive) angles of opportunity to make it possible for a trustor to actually trust a trustee. On the side of the trustee, this is about invoking potential sources stimulating cooperation (or at least discouraging opportunistic behaviour). On the side of the trustor, this is about improving the sources of trust to allow for an improved assessment of the trustworthiness.
Based on the sources of cooperation by Williams (1988) Nooteboom presented a division of sources ranging from egotistic to altruistic sources A3.2.2.1.4. In essence it comes down to whether the trustee is inclined to behave opportunistically, which is the result of incentives and opportunities to do so. Inclination can be affected. (Nooteboom, 2002, p. 203) This depends on the type of collaboration, such as the duration, oversight, and prospects of gain and loss. For that, parties could aim to take away as many opportunities away right from the start of the collaboration. But opting for this approach right from the start can:

- minimize the maximum potential of the collaboration(Nooteboom, 2002, p. 96), and
- less efficient in the long run, or
- turn out to be less effective in the long run than presumed (Nooteboom, 2002, p. 203).

The first is based on the consequence if participants turn out to be unable to break away from their initial social relationship(s). The presumed difficulty thereof is based on 'Crude law' by Deutsch indicating the typical type of relationship to reinforce itself (Nooteboom, 2002, p. 96). An attitude of lack of trust (or even distrust) would reinforce itself; participants would resort to the use of contracts and other protective measures against opportunistic behaviour. It leads Nooteboom to suggest to be careful at the start of the collaboration to not get stuck at the initial mode of collaboration (Nooteboom, 2002, p. 96). The other two consequences are related to the possibility that the extensive use of contracts might actually be less productive in the long run. The root cause of these two is uncertainty, which is particularly troublesome with contracts. Just like trust cannot be entirely calculative as not all possible future states and options of a collaboration can be known (Nooteboom, 2002), contracts cannot be drawn to accommodate all future possibilities, let alone in detail. The consequence is a required change of contracts envisioning all possibilities, or a more coarse contract

[8] A simplification, this, and entire section, actually refers to rational trust, but trust has also, less rational, sources. Trust is not entirely calculative (in terms of evaluation not probabilities), there is also an *assumption* of trustworthiness, a wager (Nooteboom, 2002, pp. 188–190).

from which participants can more easily escape. All in all, in a more practical sense regarding trustworthiness it is of importance to consider the circumstances of collaboration and (where necessary) complement this by means of some contracts for assurance.
On the side of the trustor it comes down to being able to 'assess' the actual trustworthiness of the trustee given the current circumstances. For that, trustors can employ some sources with different qualities. Sources of trust are distinguished by many researchers, including Nooteboom, as being knowledge based or cognition based, with the implicit claim of the two being extremes on a continuum (Nooteboom, 2002, pp. 12–13). Summarized knowledge based trust refers to the use of knowledge to *assess* the trustworthiness of the trustee (Nooteboom, 2002, p. 90). This knowledge can be based on stringent stimuli on the part of the trustee to collaborate (or not to defect), but also on the experience of vaunted behaviour in similar conditions. Cognition based trust has to do with the cognitive distance of the trustor and trustee. Smaller cognitive distances presents opportunities for empathy, possibly leading up to identification or friendship based trust (Nooteboom, 2002, p. 13). Crucially, (perfect) empathy allows a trustor to understand whether a current situation poses a realistic possibility for a trustee to become untrustworthy. It is about the ability to draw a conclusion from a trustworthiness in some situation and 'extrapolating'[9] it to an unknown other situation. In the end, the basis of trust is some mix of cognition and knowledge. Sufficiently large differences in cognition to yield novel insights because participants look at the world differently (but for that participants require more knowledge to assess trustworthiness), but not too large to preclude mutual understanding in terms of allowing some level of empathy (limiting the maximum cognitive distance).

The preceding is not supposed to indicate that there is no place for contracts at all, or that trust is the answer to all challenges in collaborations. Contracts and trust can be complementary and substitutes, with trust being a the preferred initial mode following the presumed 'Crude law' of initial interaction. For that Nooteboom suggests participants would have to start off with relationships with little risk (Nooteboom, 2002, p. 90). But the author also sees potential in certain *amounts* of control to get a relationship going. The next stage would to assess the trustworthiness of the and tolerance levels of trust. Finally, these levels the tolerance levels could be widened on cognitive grounds (such as empathy or even identification). (Nooteboom, 2002, p. 90) Trust can also in a way be the substitute to some detailed contracts, but this does not render contracts useless. Contracts turn out to be useful as an aid to memory and can actually still be detailed, it is just that the motivation for their existence can be more supportive instead of protective. (Nooteboom, 2002, pp. 122–123). In the end trust works in combination with different forms of governance, of which contracting is one of the options, as are the collaboration structure and reputation mechanisms (Nooteboom, 2002, p. 201).

Trust and its impact (per example of AbuseHUB)
In the example presented in the preceding chapter trust was present implicitly. The participants of the collaboration know each other and meet each other frequently. Furthermore, trust was not really challenged as there is not really much to lose and more to gain. After all, the example illustrated a case in which some party is better aware of negative consequences taking place, a machine sending spam, than the (indirect) responsible party. By sharing information about that, the responsible party becomes aware of the consequence (and the type of infection), allowing the party to take a decision on the course of action. To the recipient of the information such a collaboration is relatively

[9] Technically the trustworthiness is about trustworthiness of Bob in a specific circumstances 'X'. If the circumstances change into 'Y' (such as when trying to increase the collaboration), trustworthiness of Bob is no longer guaranteed. Should Bob indeed *be* untrustworthy in 'Y' than Alice should indeed not trust Bob in circumstances 'Y'. In the end that is the issue of trust. Alice has to *estimate* the trustworthiness of Bob in a circumstance and on that basis determine whether Bob is to be trusted. (Perfect) Empathy helps in all this as Alice is able to understand how Bob perceives this situation 'Y' and whether Bob is as a result thereof trustworthy or not. (But even if Bob is untrustworthy in 'Y', Bob is still trustworthy and in 'X'! In the end, scaling trust into unmarked territory is a wager with empathy helping out in the estimation of trustworthiness.)

comfortable. The information is about a consequence, revealing nothing more than usually with an ordinary email about the environment of the ISP. Additionally, it does not require any information from the ISP. The (consequence-related) information is provided by a known *trusted* party (Trusted Complainers), to an analyses centre with *known*, rather similar members (limited cognitive distance). As long as these providers and the analyses centre behave responsibly and without prejudice such collaboration can expect relative good support. Interestingly, due to the setup the main parties in the collaborations will meet each other time and time again. Reputations are therefore a factor. As a result, there is also an incentive for such responsible behaviour without prejudice. All this makes trust and trustworthiness easier, minimizing the need for highly detailed, protective, contracts. This does not necessarily mean there will never be bad times, but conditions are pretty favourable. Relatively limited sensitive *information* being exchanged, the parties know each other and they meet each other regularly in a variety of settings.

In more challenging real world situations the intention is to add cyber security related matters to agendas. Examples are the chemical sector and the financial sector in the United States. In a sense they try to piggyback on readily existing trust relationships and collaborations on security. Cyber security is 'simply' added. In the chemical sector the parties extensively make use of personal relationships for information sharing in general (National Infrastructure Advisory Council, 2012, pp. C–10). A specific case of cyber security information, on industrial control systems, is shared in a small community (National Infrastructure Advisory Council, 2012, pp. C–12). Cognition based trust is therefore possible, allowing to accompany the fact that the context is slightly different. It still leaves the forms of behaviour participants would have to trust in. But it appears the relationships are about information, requiring informational trust. This is not that much of a stretch from other topics, and actually more of a difference of context. It would be somewhat different if the relationship is about competence trust or material trust. Those scale more difficult to different topics.

Although the extensive focus in the chemical industry on small communities of inter personal relationships, smaller companies miss out on important threat information (National Infrastructure Advisory Council, 2012, pp. C–12). The banking and finance sector similarly has a rich history in risk intelligence sharing, albeit with very large numbers or organisations. Information security is yet another important topic, especially as the banking and financing sector too is considered to perform essential operations. (National Infrastructure Advisory Council, 2012, p. B–2) And due to the interconnectivity of the financial sector cyber security is 'widely recognized' to require information sharing, instead of a competitive aspect. To share this information 'strong and robust' channels are used, contrary to the use of personal relationships as with other sectors. This is the result of the developed clear roles and responsibilities regarding information sharing (National Infrastructure Advisory Council, 2012, p. B–17). With that, the banking and financing sector appears to have different balance of knowledge versus cognition based trust, possibly complemented with (more stringent) contracts.

Finally, the initiative in the United Kingdom called Cyber Security Information Sharing Partnership (CISP) is developed from the ground. The collaboration is consist of organisations from various sectors. Initially it focuses on large organisations, but later on Small and Medium sized Enterprises would join. (Maude, 2013) However, as expressed by John Colley on RSA Conference 2013 large enterprises 'mistrust' smaller organisations. They would not share information with organisations they do not know. (Stevenson, 2013) The Chief Technology Officer of FireEye, Greg Day, said that ultimately the purpose was to customize the information that organisations would receive. That way also smaller organisations would be able to understand the shared information. Larger organisations already have the capabilities to understand the information themselves. (Stevenson, 2013) With such processing possibly the request could be fulfilled to anonymize the data. However, as pointed out by Colley, in doing so valuable and necessary details might get lost actions of anonymizing data. (Stevenson, 2013)

5.3 DEFINITION OF THE DESIRED END STATE

5.3.1 ACTIVITY

As mentioned collaborations are considered in this research to be positioned relative in time to the attack. This is reasoned from a collaborative perspective, not from the perspective of some individual organisation. A collaboration addressing some attack phase gets initiated as soon as the first incident (possibly accident) occurs at some organisation. This organisation might have installed its countermeasures, but from a collaborative point of view the risk is full in effect. Other organisations will not have the necessary means installed to minimize the likelihood of a (successful) similar attack or to decrease the possible consequences.

In the end a collaboration can thus be about:

- identifying a future attack (what might happen and what can we do about it?),
- thwarting an attack (what happens and how can be recover?), and/or
- learning from an attack (what has happened?),

which makes collaborations come down to do the more liberal definition of collaborations being oriented relative in time to the attack. After all, nothing stops one from having a collaboration learn from attacks and also use these insights for identifying future attacks. It is suggested, yet considered to be troublesome as will be discussed later on this research. What for now it is important to remember is how the collaboration is positioned relative to the attack and being of relevance to the other roadmap phases and steps. To shorten the reference to the decision of whether the collaboration is about future attacks, thwarting attacks, learning from attacks or a combination thereof, the decision is regarding 'the fourth dimension' is referred to.

With the addition of the time dimension the trouble with risk levels becomes apparent. Risk levels are rather difficult to determine at times of an attack due to subjectivity. But they become even harder to assess prior to the attacks, due to levels of uncertainty in the estimation of probability of threats related to some consequences. And after the fact, the risk levels are not really the incentive to discuss attacks. Rather for example suggestions of common elements, purely the consequences of some attacks or (the resulting) responses of the environment of some attacks is a motivation to start discussions.

5.4 DEVELOPMENT OF THE COLLABORATION MODEL

This chapter is all about defining the activities of a collaboration which is open, responding to events in its relevant environment. The relevant environment is the entirety of organisations (some of which will be represented in the collaboration), institutions and events affecting those organisations and institutions.

Decisions regarding the setup of the collaborations affect what the relevant environment is like. The aforementioned inclusion of a participant is an example. But also the intentions of the collaborations affect and are affected by the environment. A collaboration intended to thwart an attack will have to respond to an attack. Such a collaboration might get more leeway, yet in a specific hands-on direction of stopping attacks. More elegant or comprehensive solutions to avoid attacks are interesting, but the main objective is to at least stop attacks. On the other hand, should the collaboration be about avoiding types of attacks the relevant environment will be different. The level of discussion will be at a higher (hierarchical) level, the acceptable means will be different and the overall sense of urgency will be different.

Aside from that, the collaboration might also try to directly influence the environment, not merely having some (relevant) environment as a resultant of decisions. This more active approach comes down to improve the usefulness of the environment such as in providing helpful information, getting acceptance for activities (such as sharing of confidential data) and to clearly position the collaboration in relation to other initiatives.

With all that, the environment is a factor and a step. Some decisions can be made to try to actively influence the environment, others are the result of decisions which affect what the relevant environment is like and that remaining relevant environment will affect (and be affected by) the collaboration. Finally, collaboration also have to consider what impact the environment can have on the collaboration and possibly to take precautionary measures for that. An example of that is that, as part of the laws on 'freedom of information' the general public might ask national government what information is shared. For that reason the Rabobank asked whether they could classify shared information a state secret (Lange, 2013). All this to avoid that for example vulnerabilities become public knowledge. (For a little more about this and the possible consequences, please refer to A3.2.2.3.1 on page 100.)

5.5 SUMMARY

The most important change which adds complexity to the base roadmap is the removal of the assumption of total autonomy of the collaboration. This was supported using the assumptions of orthogonality, the non-existence of the environment and the stability of the configuration of participants of the collaboration.

The resulting readily identified steps are presented in Table 3. The steps are still separated in two separate roadmaps for the sake of overview in which roadmap/chapter they were discussed. The steps are actually added to the total list of steps per roadmap phase.

Current state ↓	**Development of the collaboration model ↓**							**Desired end state ↓**		
formulation goal of colllaboration	roles	collabo-ration structure	topic of shared information	level of information sharing	method of sharing information	type of response	timeliness of response	target	risk-value	adversary
	environ-ment							time dimension		

Table 3: A visualization of the second roadmap. The steps that are identified specifically for the second roadmap are depicted on the second row. The readily identified steps of the first roadmap are depicted on the first row.

The main difference of the second roadmap is reflected in the addition of the time dimension (*the fourth dimension*) and the environment to the overall set of steps per roadmap phase. With the removal of orthogonality the steps are no longer considered to *not* influence each other, directly or indirectly. An interplay is possible. The interplay can manifest itself through the environment (which can only partially be influenced) and the implicit trust (due to the effect of uncertainty from within and of the environment). The environment became a factor as the parties in the collaboration no longer have to just answer to the others in the collaboration. But the environment is not entirely a given, it is about the relevant environment, which depends on the intentions of the collaboration. By adding a participant to the collaboration the corresponding organisation will have a different position in the environment. But also with a different ambition, such as preventing types of attacks instead of thwarting specific attacks, the environment will be different. And with that change of the environment, some of the steps will be affected. The questions are what participants are willing, able and allowed to share in a collaboration. This depends on the environment and the participants of the collaboration. Depending on the configuration of participants and the information that might be present in the collaborations the environment might grow concerned. If large organisations would work together with organisations like Google and Facebook other parties might become concerned. They might oppose such information sharing. But there are other ways in which the environment can affect the collaboration. Think of regulation that dictates whether information can actually be shared. Perhaps the organisations are not allowed to keep some log files long enough.

5.6 EXAMPLE

Contrary to the example of the preceding chapter, this is not a fully elaborate example nor even an example of an actual collaboration matching the joint roadmaps. The intention is to demonstrate the principle as it is more speculative with many possible variations. A fully elaborate example would distract from the principle and is therefore considered not to be worth the effort. Furthermore, the

impact of trust and the topic of information exchanged is discussed to a limited extent as it was already covered extensively.
In this research the decision on the collaboration relative to the incident is considered to have a profound effect on the steps in the 'development of the collaboration model' phase. As a result, in this example solely the three distinct time options are discussed, 1) those predating the attack, 2) during the attack and 3) after the attack. This is still to be considered at the collaborative level, yet the first option is about predating an attack to anybody in the collaboration. Nobody who is part of the collaboration even knows about the attack, let alone encountered it.

5.6.1 PREDATING THE ATTACK

Collaborations which are predating the attack are about identifying the future (type of) attacks. The actual attack which will happen is not known, yet might be similar to other preceding attacks. But it can also be about some extrapolated version of preceding attacks. A crude example would be one based on the outlook discussion by Symantec in their security reports. Their latest security report discussed attacks being more targeted, increasingly using websites and a trickling down of sophisticated attacks (Symantec, 2013b, p. 54). It is rather crude, yet still identifies a *single* type of attack, it discusses the abstract target, the method of entry and the sophistication of the actual malware. The point of such a discussion is to discuss what actual weakness the attack exploits, what the attack is about and how to protect against this. In this crude example, the threat is about a new type of attack, the watering hole. It exploits the fact users are searching for information on the Internet and encounter a website which matches what they are looking (or at least peaks interest). And if they access the website or downloads a file, a more insidious type of malware gets activated.
The crude example, specifically the underlying outlook by Symantec, is based on preceding (one or, more likely, multiple) attacks. It can be about the sheer volume, but also because of the risk of such an attack. The watering hole is not that common, yet it is very powerful in that it catches more users more off guard. And the trickling down of malware is considered to be a matter of time given recent developments. Such a discussion is sparked by input from security experts and/or hands-on people. But the actual weight carrying decision will probably take place in the higher echelons. Decisions are not (or at least not supposed to be) about a constant reinforcing of barriers, but a more comprehensive focus on barriers, paths and hazards. The result could well be to train personnel to be more cautious about attacks using websites. But one could also think of ways to focus more on making the social engineering a bit harder. And like barriers, one could hereby focus on specific parties. For example, of witnessed cases by Symantec, R&D personnel got targeted most in 2012. The increase from 2011 was also the highest of all targeted attacks by role. (Symantec, 2013b, p. 17) That can be a reason to focus internally more on R&D. But also to focus on roles which might be the focus of next year, as adversaries will in all likelihood change their targets. Such decisions, on allocations of resources across departments[10] will thus at least (if not exclusively) require involvement of the higher levels of hierarchy in organisations.
With the outlined example the amount of roles involved is possibly also relatively large. This especially likely if the topic is brought on a more comprehensive level. Such a level refers to not solely discuss how to thwart the attack, but also about taking away the possibility, one way or the other. The AbuseHUB, as discussed in 4.6.1, is an example of thwarting attacks. In that collaboration the aim is to help Internet Service Providers (ISPs) improve their awareness of infections of machines of the subscribers which are in their respective networks. The improved awareness is supposed to allow ISPs to (indirectly, via customers) stop the negative consequences of such infections (such as sending spam). But aside from this rather focused, small scale collaborations larger collaborations can be envisioned to take away more opportunities of attacks by being able to take down a botnet

[10] Despite the fights, IT is still *considered* a department (or worse: a cost centre) in many organisations. However, in this report the allocation of budget to departments is not by any means a suggestion of IT being a department. The point is simply about whether to invest in IT or more of the people. And whether that means an organisation discusses about investing in the IT department, a department which than handles IT or whatever other way is of no importance in this research.

sooner. This could be by also inviting authorities, investigating organisations and providers to detect and stop the servers which send the infected machines instructions. The discussion on how to stop botnets in general (not a specific one), to better allow actually stopping one once it is detected, can and will thus involve more parties discussing the bigger picture. The result of all this is that the discussions can be rather abstract. No specific data, but primarily (more abstract) information if not knowledge will be exchanged. For example, authorities can inform other parties what type of information (evidence) they require to be able to be allowed to take down a server. Based on this, invited individual parties themselves can draw conclusions from this to understand what type of data they will have to acquire for that. In part the more abstract forms of data are exchanged as there not being an actual attack to discuss. But more importantly it is also necessary in terms of who is involved. With involvement of large quantities of parties and people, with highly diverse background by role and personal development, the discussion will have to be abstract. People will have to be able to follow discussions, not get lost in details and handing confidential data with large groups is troublesome. This mechanism was witnessed by NSS Labs in their analyses of Information Sharing and Analysis Centres. They tend to focus on the strategic level, a necessity from the amount and diversity of parties involved. The most mature and successful one in the United States is the one on Financial Services (FS-ISAC). It has over 4.400 member organisations, representing amongst others commercial banks, insurance companies and trade associations. Tactical threat related information thus has to be found elsewhere. (Rashid, 2013)

5.6.2 DURING THE ATTACK

Collaborations intended to discuss current attacks on selected targets on the other hand will have to discuss more detailed information. This does not necessarily equate to requiring the sharing of raw data (containing the most intimate information), but at least it requires pragmatic information. The level of sharing should be such that the attacks can be thwarted as productively as possible. In the end the goal is to thwart an actual, current attack. Possibly the collaboration has instated a risk level to determine whether all attacks are considered, but for those that are, the goal is to stop them.

In order to be able to share more confidential, pragmatic information on specific attacks for the purpose of finding solutions to stop the attacks the amount of parties involved will be lower to be able to share the information. The result will be a clearer set of participants and the trust of the parties will be more competence based.

All this will also be reflected in the type of response and the timeliness of the response. The collaboration will be about actively introducing countermeasures in a short time period. To actually keep up the pace this will thus require a balancing of the type information (data or information) shared and the volumes. There will not be a single perfect balance, as this one in all likelihood also depends on the involved roles and structure. Some collaboration centres might have analyses centres. An example is the Abuse Information Exchange initiative. They could receive large amounts of data, as the delivery can be fairly (easy) standardized, from known parties to known parties. However, in case such standardization is not possible and the ones with the right, compatible awareness are different each time the balancing act becomes harder. In contrast, predating the attack will be more about setting a course of action. Some careful evaluation of what might actually happen, which might result in a collaboration which can be more intermitted and sluggish at times.

5.6.3 AFTER THE ATTACK

Finally one or more attacks can give reason to some investigation on the handling of the attack. However, contrary to ‘predating the attack’, after the attack is, in this research, oriented towards the attacks themselves. Did the collaboration work as intended, was the information any good and were the actual solutions to satisfaction? With such questions the level of information required will vary, as will the level and types of participants. In case of a prosecution, for whatever reason, it is even possible that a set of objective outsiders will be involved. All in all, ‘after the attack’ will take place less often, with changing collaborations on different levels of involvement and topics discussed.

6 SCENARIO 3: DYNAMIC COLLABORATIONS

Thus far the only time dimension discussed was that of the positioning of *the* collaboration relative to the threat. As follows from the preceding chapters it is not necessarily the case of there being just a single, very large collaboration. If there will in fact be many smaller collaborations the question is how these relate to each other. After all, the collaboration itself is also affected by time. Furthermore, a collaboration itself was thus far in the research considered a static entirety of equal participants. Any dynamics in terms of development by means of changes of composition, development of participants and development or maturity of the collaboration was ignored.

With that said, the third and final roadmap, like the entire research, focuses on the level of a single collaboration. However, the third scenario, being at the basis of the third roadmap considers that participants are in fact different and that there different perceptions (or definitions of reality) are possible. Participants, that enter the collaboration, have different socially defined individual definitions of reality and different qualities to act upon those definitions. Similarly, these participants might be part of other collaborations, which allow them to affect the development of the collaboration. These notions provide *opportunities* for changes of collaborations. These opportunities are discussed, without discussing the impact of those opportunities, being a discussion of the actual, potential, developments of collaborations.

6.1 METHOD

As with in the preceding chapters in this chapter a theory is added to complement some of the underexposed aspects of a collaboration by Kowtha et al. In this chapter it is about the dynamic nature of the configuration of parties in a collaboration and its reciprocal relationship with *their* view of the cyber security problem. With that, the theory provides a frame of mind on who will have to share their awareness. It is more about the collaboration and the parties than the topic of discussion or the shared type of data. The theory of choice is called the configuration theory. Its use(fullness) is discussed in this chapter, the more comprehensive whole in A3.2.3.1 from page 101 onwards.

6.1.1 ASSUMPTIONS

The collaboration now does not presuppose equal participants anymore.

- For each instance of a roadmap there can be only one option selected per roadmap step
- Minimizing the amount of steps in the end state is preferred

6.1.2 ASSUMED RULES

Without the assumed orthogonality or roadmap steps the order of the roadmap steps is an important point of attention. In this research it is assumed both the order and the criticality of some steps is considered to be important:

- Some roadmap steps require a decision of all involved parties or even a sufficient level of maturity prior to discussing the next roadmap step
- With that the order of deciding upon roadmap steps and their division into phases becomes important and therefore the order and the phases have to be predetermined.

Based on the assumed importance of the order yet the variety of possibilities therein, the order and divisions in phases will not be discussed in this research.

6.1.3 THE CONSTRUCTED REALITY OF 'THE' CYBER SECURITY PROBLEM

The challenge of cyber security is the result of the increasing complexity and dynamics, resulting in definitions of reality being an accurate representation in time for a decreasing period of time. Motivated by the necessity to collaborate on the basis of consequences parties lack full situation awareness. They need each other to understand the situation at hand, to define reality. And given these (changing) interdependencies there is not a single party who can unilaterally define and decide on the course of action. A steering theory that respects such interdependence and that has the premise of reality being a socially defined construct is the configuration theory.

6.1.3.1 Configuration theory in a tiny nutshell

The goal of the configuration theory is to make 'sense' of the complex and dynamic whole of socially defined reality. This is supported with one of the important concepts of the configuration theory: the presumed double helix of the cognitive and social dimension. (A3.2.3.1) Parties which are part of a configuration represent the social dimension. Each of these parties brings in its own socially defined perception of reality. And together these parties define the, what they believe to be, common relevant reality they are confronted with. To say, this reality is the rather specific definition of the situation at hand. The resulting definition of reality represents the cognitive dimension. These two dimensions are connected in such a way that a change of one affects the other and a definition of one can be traced back to the other. The total of definitions and parties is a configuration, which is a merely a snapshot of a moment in time. It implies at an undefined later moment in time a configuration can be different. The possibility for such redefinition is even required as reality changes and so should its definitions. The theory favours instability with stability being allowed, but for the purpose of collaboration. In general definitions should be open to change and negotiable, unless there is a good reason to *temporarily* stabilize the definitions. But such stabilization should never develop into fixation of definition of reality and/or interaction, being that such definitions are fixed. (0)

6.1.3.2 Configuration theory as a frame of mind to information sharing collaborations

The approach of the configuration theory primarily and intentionally brings a frame of mind that is not only applicable but also relevant (A3.2.3.1.2) and of (limited) prior, practical use (A3.2.3.1.2) to the topic of cyber security. The theory (and its entire approach) 'forbids' interfering with definitions of reality. It is considered impossible and unacceptable (Twist and Termeer, 1991, p. 25). And with the prohibition of presenting predefined realities it is also prohibited to discuss who should be part of the configurations. With the introduction of parties definitions of reality are also entered, which means there is some interference with the content. The configuration theory only allows definitions regarding the process and even those are defined as forbidden process related matters *('un-values'*).
Despite the unacceptability of predefinition the theory still offers and allows for insights on how to approach information sharing collaborations. The configuration theory has an implicit suggestion to have some undefined limit on the size as a result of the combination of social constructed definitions of reality and multiple inclusion. The size of collaboration is actually a proxy to the scope of configurations. Configurations which are very intense in terms of interaction of some parties, with a clear common, socially defined reality can build on the basis of this clear reality. But at the same time, in all likelihood their reality is less contested, increasing the likelihood of groupthink. This is in particular a concern with cyber security because what was safe at some moment might not be a little later. With increasingly targeted attacks (as witnessed, not necessarily being the overall tendency) there are less signals worldwide indicating of such developments. From this follows the motivation to setup collaborations with a larger scope. But the expense is that the level of discussion moves to the abstract, to the strategic level. Nevertheless, with the configuration theory there is no such implicit suggestion given that many organisations will be multiple included. They do not fully agree with a single definition of reality and will be present in multiple configurations. This multiple inclusion allows for context variation. Definitions of reality of another configurations might be introduced posing a incongruence with the status quo of the configuration. Or an organisation part of another configuration might be introduced and with it enters a new definition of reality. It is this context variation that is at the core of development of configurations to redefine reality and it is this ability which has to be protected at all times. In terms of cyber security it comes down to the suggestion of considering various types intersections of cyber security aside from sectorial, such as size, assets, type of consequences. The various intersections result in multiple possible configurations which might appeal to a single party. With that, multiple inclusion provides an option to spread incongruent, socially constructed definitions of reality. Incongruence can disrupt configurations in a functional manner, as long as configurations consider definitions to be negotiable.

Inclusion provides another angle to categorize parties. Parties have some undefined level of inclusion in a configuration and might be part of one or multiple configurations. On the basis thereof three roles were distinguished by Termeer: *fixators, initiators* and *brokers* (Termeer, 1993, pp. 269–270). *'Fixators'* are the ones that are highly included in a configuration. They are the driving force of stability if not fixation of definition. At times of incongruence (by introduction of a party or definition) or a fixation (a development of the cognitive and/or social dimension becoming fixated, as in not open to change or redefinition) these are the parties to watch. These parties prefer to maintain the status quo. Specific targets of attacks can be a fixator. The aforementioned smaller organisations who consider themselves to be safe from attacks are an example. The reality could be that they are actually not safe from the attacks. Fixators might simply not be aware of the different reality and might 'successfully' ignore incongruent definitions of realities. The latter might be the result of that the 'wrong' organisation presents such realities (such as a security firm, larger firms or firms with highly specific assets). The cognitive distance might simply be too large. A second type of role of parties in configurations is that of an *'initiator'*. These parties present, possibly by themselves, new definitions on the basis of their formal role. But it could also be the result of their personal contacts. A security firm such as Symantec comes to mind as an initiator, being able to provide proof of smaller organisations also increasingly being targeted (Symantec, 2013b, p. 16). But it could also be a department or some parties of an organisation. The configuration theory does not focus on entities. Configurations can be a combination of all kinds of intersections of organisations, with a high degree of organisation and intensive contact. Configurations are platforms of information sharing and deliberation, not necessarily limited to formal boundaries (0 on page 103). An initiator has an undefined level of inclusion, yet will in all likelihood not be the highest included. The third and final distinguished role is the *'broker'*. Such a party is typically rather low included, meaning that the party is not really that concerned with the definition or does not really share the same view on reality as many of the others in the configuration. But brokers are multiple included. That has the advantage that they can easily introduce definitions and participants from one configuration to the other. Such a party could be thought of as the portions of security centres which are commonly presented as a hub in frameworks.

6.1.3.3 (Re)Definitions of reality and collaborations given a reality

The backcast (depicted with the 'development of the collaboration model') from the end state to the intended current state started off from the intention to have good *performance* on security on attacks in some to-be defined domain enclosed by an adversary and a target. And all this takes place prior to, at times of, or after those attacks, the decision regarding the time dimension. The plausible causal relation leading up to that performance reveals correct decisions on the basis of 'correct' situation awareness. And correct situation awareness implies perceiving the relevant elements in the current situation, comprehension of the current situation and understanding what it means for the near future. For that it is considered important to make sure situation awareness is 'correct'. Information exchanges are used to update or increase the level of situation awareness[11], supposedly allowing for better decisions and performance. This entire mechanism presupposes parties to know what the other needs to know, the Meta SA as identified by DSA[12]. But all this still leaves the situation, the reality, undefined. With the configuration theory the focus is on instability of definitions of reality. Knowing that more elements might be become part of the situation. For example, consider Flame which was able to spread using the trusted Microsoft Windows update

[11] The alternative of simply providing information on what to do and not actually improve the situation awareness is omitted. It would unnecessarily raise complexity and would not change the line of reasoning.

[12] SSA presupposes common perspectives or shared mental models (Salas et al., 1995, p. 129), which is akin to, yet not identical to what the configuration theory is about. The shared mental models are, according to Wellens, about definition of a shared understanding of the problem the group faces and agreement of the group on operation of the group. The latter referred to concept by Orasanu of mental models to allow for development of shared strategies (Salas et al., 1995, p. 130). In contrast, the configuration theory is about definitions of reality, the shared understanding of the problem of SSA. But what happens after this shared definition of reality of the configuration theory is a springboard to something else. Which could be SSA but also DSA, which does not presuppose shared goals and strategies.

(Zetter, 2012). The shared definition of reality provides on a definition of what the situation is, which elements have to be considered. In turn, situation awareness is about making sure participants have proper awareness about those elements.

With the introduction of the configuration some confusing or seemingly conflicting terminology is introduced. Specifically that of a configuration and a collaboration, and the roles in a collaboration and the roles in a configuration.
Strictly speaking a configuration refers to a group of parties which together define their definition of reality. And in that definition parties can play the role of an initiator, fixator and/or broker. The collaboration refers to a group of participants which work together to improve the situation awareness of (some) participants or some group (an analysis centre). And in such a collaborations every participant can have a different role, such as analysing log files or contacting some organisation in the environment about findings by the collaboration. This latter type of role is more the actual duty, or the formal role, the participant performs in the collaboration. Although configurations and collaborations could well be staffed by the very same people, it does not necessarily have to be that way. Possibly some superiors are part of a configuration defining the reality (of what is supposed to be collaborated about or achieved) and others have to actually do the collaboration.
The two types of roles, referring to the configuration or the collaboration, are to be distinguished from the function the participant performs on a day-to-day basis, outside the collaboration (the 'day job'). Although ideally participants of configurations and collaborations will be invited on the basis of their function, this is not necessarily the case. Depending on the ambitions of the collaborations and the staffing of the configuration or collaboration the participant might be asked to fill a void. Such as communicating to the environment, although outside of the collaboration that same party might perform analyses of data.

6.2 Definition of the end state

Thus far the end state was about collaborations to avoid, thwart or learn (time dimension) from attacks of an enclosed portion of attacks involving some target (target) and adversary (adversary), of some risk level (risk-value). The thing missing was *'influence'*. The focus can be on some party, but the collaboration has to truly reach such a party. Should the collaboration be about all attacks matching the aforementioned dimensions, period, or is the influence geographically limited. Such an influence should be backed up by some steps which are discussed in 6.3.

6.3 Development of the collaboration model

Whether a collaboration truly has some influence depends on its *external interactions, scale and maturity*. *External interaction* has to do with the ability of being able to influence and be influenced by other configurations. A collaboration might perceive reality as being such that more stringent regulation is in order to force organisations to make cyber security become part of their risk analyses. *Scale* has relations with the external interactions. Highly productive collaborations in terms of external interaction might be able to compensate for being rather small in scale, yet aiming for high influence. But scale can also be about defining maxima, for the sake of being able to maintain clear and detailed definitions of reality. *Maturity* is both a representation of the current status of the collaboration, a result, but it is also a result of decisions regarding requirements on participants of the collaboration. Thus far the assumption was that all participants are equal. But actually participants can be vastly different in many ways. Examples are their potential differences in their definition of reality, awareness of that reality and their ability to act upon that reality. And with that their weight in the credibility of the centre can vary.

6.4 SUMMARY

The third and final scenario underlying to the third roadmap is about the acknowledgement of the diversity of organisations. This notion puts some of the readily identified steps of the first roadmap in a slightly different perceptive. The participants of the collaboration have different capabilities of fulfilling roles, can provide different levels of information on the basis of different levels of situation awareness.

Current state ↓	Development of the collaboration model ↓							Desired end state ↓		
formulation goal of col- laboration	roles	collabo- ration structure	topic of shared information	level of information sharing	method of sharing information	type of response	timeliness of response	target	risk-value	adversary
	environ- ment							time dimension		
	external interaction	scale	maturity					influence		

Table 4: A visualization of the third roadmap. The steps that are identified specifically for the third roadmap are depicted on the third row. The readily identified steps of the first and second roadmap are depicted on the first row and second row.

With the third roadmap also four new steps are introduced, related to the type of participants taking part in the collaboration. The *influence*, as added to the desired end state, is about defining who has to be influenced (in some way) by the collaboration. For example solely organisation in the country in which the collaboration takes place. Such a decision has to be backed up by having the required *external interaction*, being able to actually influence the parties or getting being recognized by the external parties. Additionally, the *scale* and *maturity* are about the impact the collaboration can have. In all this, as mentioned, the capabilities of the participants in the collaboration are important. An ambition to influence a vast number of organisations without organisations being recognized as being capable has a challenge in getting recognized.

6.5 EXAMPLE OF

There are two forms of reality organisations are confronted with, an objective or actual reality and a subjective or perceived reality. These realities are in depicted in Figure 14 with circles one and two. Starting with the actual reality, not all organisations are confronted with the same adversaries. Some adversaries are opportunistic, attacking organisations pseudo-randomly. Other adversaries are actually targeting specific organisations. And of those, some are presumably actually targeting the smaller organisations. The defensibility of smaller organisations is lower, allowing for a more simple prey. The third type of reality is what the organisation thinks the reality is (circle number three in the Figure 14). This reality is affected by what is witnessed and what is believed to be the reality. What is witnessed could well be the result of the limited defences of the organisation. With limited resources invested in avoiding measures (such as installing patches), protective measures (such as firewalls and training of personnel) and detection measures (such as feedback by personnel but also proactive examination of activity on the systems) there is a chance those organisations have impaired awareness of the actual situation.

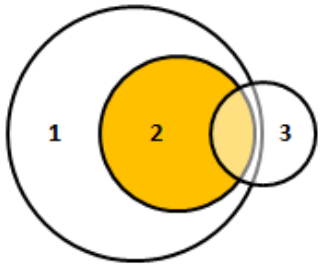


Figure 14: Forms of realities, with:
1. Reality
2. Reality of an organisation
3. Subjective reality of an organisation

In a collaboration, other organisations might affect the subjective, imagined reality by contrasting that reality with its own view of reality. To simplify the matter of realities from here on two types of reality are distinguished:

- objective reality (the reality an organisation is confronted with) and
- subjective reality (the reality the organisation thinks is in order).

Furthermore, the *objective is to improve the overlap of the objective and the subjective of a single organisation, with the help of other organisations*. Crucially, the objective reality is unknown to the organisations. It is subjective interpretation what that reality actually is. With the help of others, in a collaboration, the intention is to find out what the objective reality is. And hereby it is important to remember that organisations are confronted with their own objective reality.

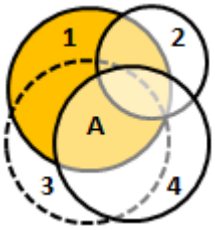


Figure 15: Objective and subjective realities of Bob (resp. 1 and 2) and Alice (resp. 3 and 4).

Because organisations in the collaboration are represented by persons names in this example, from hereon, persons are discussed. Bob and Alice are both in the collaboration. Of the two Alice has a better perception of the reality, meaning the subjective reality better overlaps the objective reality. The mechanism is illustrated in Figure 15, with the objective realities displayed using circles 1 (Bob) and 3 (Alice) and subjective realities using circles 2 (Bob) and 4 (Alice). The objective is to improve the subjective reality of Bob (circle 2) to better match the objective reality (circle 1) of Bob. For that, Alice tries to help out Bob by sharing her view of the subjective reality (circle 4). As illustrated part of her subjective reality (portion A) is actually unknown to Bob (not part of circle 2), yet it is part of the objective reality of Bob (circle 1).

In contrast, would the collaboration be different, such as with Bob and Carol as displayed in Figure 16, there is no way for an improvement of the awareness of Bob. Circle 6 (subjective reality of Carol) has nothing in common with circle 1 (objective reality of Bob). Even though Bob and Carol actually are confronted with, in part, the same objective reality B. With that it can be explained why the natural temptation is to include more participants. Participants might be of relevance to each other. Knowing the situation of Bob and Carol is as presented, in future the subjective realities might overlap and with that, Bob might uncover an unknown objective part of reality. A threat that Carol was aware of (but possibly not even relevant to Carol), could be of relevance to Bob. With that possible future development there is:

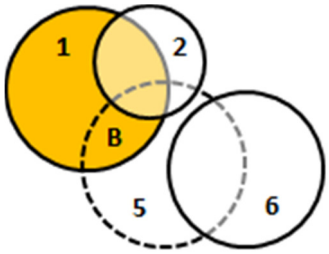


Figure 16: Objective and subjective realities of Bob (resp.1 and 2) and Carol (resp. 5 and 6).

- a motivation to collaborate (overlap of subjective reality) and
- there is an actual advantage (the two work together on an actual reality).

But currently, with the situation there is no motivation (no overlap) to collaborate, although the two are actually confronted with in part the same objective problem. And with that, the opportunity is lost for Bob to get recognition of the part B. A more productive collaboration is one in which parties collaborate on the basis of overlap of subjective definitions of reality. It is more straightforward in terms of motivation to collaborate and as organisations improve their understanding of cyber security, their subjective realities better approach the objective reality. This could mean that participants depart from collaborations and join others. The all-encompassing scenario is presented in Figure 17. For simplicity subjective realities are discussed, with the exception being circle 1 presenting the objective reality of Bob. The figure actually depicts two (or more) collaborations. One collaboration would consist of Bob and Alice. The other of Alice (circle 4) and Dave (circle 7). In the first collaboration, Carol still brings in the knowledge of A and D. But in the second collaboration, Carol with Dave she is confronted with D, but possibly also C comes to table, which is also of relevance to Bob. As discussed in the preceding, the likelihood of Bob (2) and Dave (7) to collaborate is limited, given their subjective realities do not overlap. But with multiple inclusion, Carol gets confronted with a reality of Dave (7) which is in conflict with her own and with Bob. And she might bring this notion to table in the first collaboration with Bob.

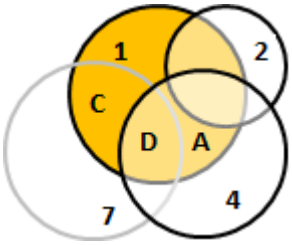


Figure 17: A collaboration on subjective definitions of reality by Bob and Alice (2 and 4) and one by Alice and Dave (4 and 7), with the objective reality of Bob being depicted by circle 1

Ultimately, what you're after is that the objective realities do not deviate too much, as it poses less relevant discussions. But the point is that such objective reality is not known, solely the subjective which has to improve to better approach the objective reality. And with that, it is to be expected that over time parties will depart from collaborations, as they have too little in common. Keeping these parties in anyway might result in a tendency to:

i. readjust the subjective definition given all those definitions, resulting in very large subjective definition of reality, possibly including the objective definition or simply heading off in the wrong direction not even including the objective definition,
ii. a more abstract definition of reality in which parties can find themselves (but one which is less practically usable), or
iii. unmotivated participants, ignoring other definitions, hereby possibly missing relevant definitions.

Practically, the result, in respective order, might in an exaggerated sense be that:

i. organisations with just three employees restoring artefacts think the reality is that highly advanced attacks are part of their objective reality and detecting unknown-unknowns is crucial. Even worse, in that collaboration, basic protective measures (such as patching and training) which might be relevant to them (such as updating the software behind their website) might not be discussed but implicitly presupposed. The organisation might be looking at the 'wrong' threats and hereby wasting resources on the wrong protective measures.
ii. It could also be that it is noticed that the realities differ far too much. In response, organisations might discuss more abstract aspects, such as the importance of training. On those aspects organisations might have more overlap. But it does not solve the real threat of organisations worried about the unknown-unknowns.
iii. Finally, organisations might simply ignore other definitions. But especially with large collaborations this might turn into default behaviour. The result is a collaboration which misses its point.

With all that, it goes to show that collaborations have to consider the scale and to truly consider whether an organisation has to be part of the collaboration. Instead, the collaboration could focus on liaisons, organisations which can form a bridge. Such decisions are particularly important as otherwise the collaboration might not be able to truly influence organisations. A collaboration which is very large, consisting of many organisations, can result in a collaboration which has abstract or too extensive definitions of reality. They badly represent the organisations they try to influence. It is the relation of scale, maturity and participants fulfilling some roles which affect whether a collaboration can actually influence organisations because the collaboration carries weight and represents the realities of the organisations the collaboration tries to influence. A large collaboration consisting of highly diverse organisations might cover in some way the objective reality of the organisation it tries to influence. But it does so inefficiently and possibly ineffectively.

7 DISCUSSION

7.1 INTERPLAY METHOD

Up to this point the focus was primarily on *separately* presenting components of collaborations, being the various roadmaps with their respective roadmaps steps and supporting theories. This chapter is more about the interplay of the components. Given the amount of options this has to be a coarse discussion. The main focus was on the interaction of the social and cognitive dimensions. The configuration theory is the most prominent embodiment of this, but situation awareness does the same. Situation awareness was used to discuss which cognitive aspects are of importance in the collaboration. What kind of topics would be have to be shared to improve awareness and *subsequently*, who is necessary for that. But adding parties also adds new required information. The configuration theory was added with the third roadmap, which pointed out that not all organisations are (to be considered) identical. Together they construct their view of reality, which dictates what they consider relevant to discuss. In line with the entire research, employing backcasting, the third roadmap and definitions of reality is the first step. The definition of reality defines what situations will be considered with some method of situation awareness (such as Distributed Situation Awareness). Additions of new definitions of reality, representing the cognitive dimension, are another way to affect the considered situations. With definitions of reality (re)considered (and with that considered situations), situation awareness is about identifying relevant elements for that new reality. After all, situation awareness is about of awareness at some level for some situation, a definition of some reality.

7.2 DISCUSSION OF THE MECHANISM BEHIND THE DEVELOPMENT OF *A* COLLABORATION

Based on the findings there is not much use in designing a collaboration which organisations can implement straightaway. But there is use in providing insights on the development process of such a design. Given the diversity of possible configurations a universal collaboration would be far too abstract and detached. Designing a productive collaboration for a specific reality might be much more convenient, yet time consuming and undesirable. A collaboration will have to evolve, it will not be perfect right from the start. It would take quite a study to also provide advice on how to get that step further from one of the possible initial collaborations. But more importantly, the reality is actually not a thing which can be designed for. Definitions of reality are (re)constructed by the participants themselves. Reasons are an improved understanding of the reality, but also as that reality changes. Such as because of a new threat or a new asset to protect. All this stimulates a 'designed collaboration' to be more abstract. What is possible is to discuss likely developments of designs in some extreme cases of configurations addressing some realities.

The ambition of information sharing collaborations was translated into improving situation awareness in general. It could be that the goal is to improve the situation awareness of a single organisation acting as a central hub or the awareness of all participating organisations. This improved situation awareness is considered to improve the possibilities for better decision making. And those better decisions are supposed to improve performance, being an increased cyber security of organisations. For that it is of importance to focus on the 'right' situation, by carefully defining the reality, and next sharing information about that reality to get a proper situation awareness of that reality.

For such definitions of reality on reality, determining what situation awareness is relevant for that reality and to be willing to share the information two constructs are important:

- *cumulative cognitive coverage*, which is sum of unique and compatible awareness; it is the extent to which the participants, together, are theoretically able to see the actual, relevant reality, and
- *cumulative cognitive distance*, which is the sum of the differences in awareness; it is the extent to which participants differ in such a way that it affects their ability to understand each other.

The ultimate goal is to minimize the cumulative cognitive distance and maximize the cumulative cognitive coverage. Several factors affect the values of the two constructs, such as the size of the collaboration, which will be discussed in the next sections.

The trade-off of larger versus smaller collaborations

A (very) large collaboration will in all likelihood have all required cognition to cover the relevant challenges. After all, the more participants are part of the collaboration, the more awareness is available in the collaboration, it has information on more aspects, there is a larger cumulative cognitive coverage. Alternatively, the focus could be on smaller collaborations, which, via multiple included participants, are supposed to complement each other. Despite this, the cumulative cognitive coverage in a configuration will be relatively limited (but larger than without multiple inclusion) compared to a single, large collaboration, but so will the cumulative cognitive distance.

Importantly, the cumulative cognitive distance is a factor that has to be reasonable. The configuration theory in particular demonstrates this. It stresses the interrelation with the social dimension and the importance of non-fixation of the two dimensions. Having limited cumulative cognitive distance is the result of participants which are (too) limited. And with that there are less vastly different (yet acceptable) definitions of reality, which can be the source of conflict. And conflicts are attributed as being a source of dynamics potentially leading up to redefinition (if the conflict is functional). If we were to assume as the configuration theory to at least bound the minimum of cognitive distance, situation awareness poses a maximum. Distributed Situation Awareness is the more liberal in terms of allowing different parties, focusing on compatibility. But if the cognitive distances become too large, situation awareness transactions become troublesome. Parties will still have to be able to understand each other. Increasing the cognitive distance will also affect the possibilities for organisations to acquire the required meta SA by means of empathy. Instead of empathizing, the organisation has to acquire the knowledge what the other needs to know. To some extent this might be acceptable, but it does limit the participants to anticipate on unexpected situations. The kind of situations which are common with cyber security. Similarly, cognitive distance affects the empathic abilities of parties to assess the trustworthiness of other parties. Again, the party will increasingly fall back upon more knowledge based forms of trust, as opposed to cognition based trust. And knowledge based trust relies more on assurance, knowing the trustee has limited opportunities to behave opportunistically.

With all that, larger collaborations tend to have larger cumulative cognitive coverage and larger cumulative cognitive distance. But given the larger coverage, keeping variables consistent, there will also be more information to be shared. Larger collaborations will thus technically allow for more effective cyber security, as more relevant information is available. But the information will also tend to be more abstract for all sorts of reasons. Although more information will be available, the information is part (to say hidden) in a larger network. The key with larger collaborations is therefore to get a Distributed Situational Awareness network with highly efficient networks, avoiding overloading of participants. And should the links not be of high quality there is a risk that parties can solely acquire the more abstract information and not find the right party who can discuss the abstract information in more detail. Another effect of the increase in cognitive distance, magnified if the collaboration becomes rather large, is an increase in knowledge based trust. Assurance, in terms of contracts, will become more important. In contrast, with limited cumulative cognitive distance, especially with smaller scale collaborations, there are more opportunities for cognition based trust.

Parties can empathize with the others to assess trustworthiness and rely less on assurance. With that, there are less disincentives to share more confidential data. All this comes at the expense of cumulative cognitive coverage though, which has to made up with multiple inclusions. In small configurations, more information on more limited topics will be exchanged. To still get the required coverage, organisations will have to acquire information from other configurations as well (either by asking for the information in another configuration, or receiving it from a multiple included party).

Lifecycle of collaborations

Another concern is the combination of the development and duration of the collaboration, which is primarily relevant to trust. The duration of the collaboration is not really a factor to the definitions of reality. Throughout the lifecycle the *un-values* are supposed to be governed, meaning that configurations should be open to entries of parties and new definitions (0). Situation awareness is only affected by time as setting up a model will take some time and quite possibly one or more iterations. Redefinitions of reality will necessitate changes of the collaboration model to improve situation awareness. But in the end the configuration theory and situation awareness are merely about identifying information to be shared and a method of sharing. *Whether actual, meaningful and valuable information is shared depends on the level of trust of the participants.*

Collaborations with a limited and highly paced lifecycle tend towards control based sources of reliance (Nooteboom, 2002, p. 131).This is especially the case if additionally there is no pre-existing trust to expand and cognitive distance is considerable. The expansion is a reference to the four place predicate of trust, as was discussed in 5.2. In case the lifecycle is rather long or has indefinite duration investments in trust based collaborations might be worth the effort. After all, trust has intrinsic and extrinsic values worth pursuing (A3.2.2.1.2). In case the lead time of the collaboration is long too, investing in trust might even be preferable. If the hypothesis of the circular causation of initial mode of interaction is employed, trust will be reinforced over time. Conversely, in that case opting for control based sources of reliance at the start of the collaboration would stimulate continuous use thereof. This is particularly important given the extrinsic value of trust. Of the two, control (such as contracts) falls behind compared to trust in terms of their use at times of uncertainty. Uncertainty might arise from unknown events (from the environment), new definitions of reality or new entries of participants. Trust handles these situations better as a trusting party trusts another within boundaries, as the trustworthiness of parties has limits too. The boundaries refer to the type of trust (the trust in some respect of the four place predicate) and the circumstances the trusted party is in. As long as the type of required trust and circumstances are within the coarse boundaries there is trust. Even though the exact activities of the trusted party within those boundaries are unknown. It is expected the party will not behave opportunistically regardless of incentives and opportunities to do so.

The presumed circular causation of initial mode of interaction might also explain for the potential growth of micro to macro forms of reliance. In an environment with some parties cooperating on the basis of trust the mode of interaction might increase to trust via third parties[13]. And eventually the trust might slowly develop into macro forms of trust, by means of norms, values and habits. Such type of trust might even scale towards system level, the type of trust in which participants trust that the system is such that it 'produces' trustworthy behaviour of its participants.[14] In contrast to all this, sources of control at the micro level are dependence, hostages and reputation. Development to the macro level leaves contracts and supervision.

[13] The mechanism is similar to that of the employment of a certification authority. If Alice trusts Carol (in a relevant respect) and Bob trusts Carol (in a relevant respect), there might be ground for Alice to trust Bob.

[14] All this is a rather black and white representation of what might happen. Typically even a relationship based on trust is still complemented with a level of control. This might be solely for the purpose of an aid of memory.

As a result of all this, in this research, it is assumed that for short term periods, contracts (especially the rather simple ones) are cheaper than investing in trust relationships. As collaborations last longer and the cognitive and social dimension are subject to changes (causing uncertainties), possibly due to incidents or accidents, trust based collaboration will actually be cheaper and allow for more intense information sharing. Assurance based collaborations will be more in a mode of distrust. And to solve distrust those collaborations for forms assurance like contracting and supervision. Such forms do not really become cheaper with uncertainty, which results distrust. There will always be an situation in which a party causes such distrust. Whether it is because of limited information sharing or requests for other types of information. Trust is subject to the same uncertainty caused by entries of participants, new requests for information and decreases in shared volumes of information. However, trust is better equipped for such uncertainties. It is just that the initial costs of participating in a collaboration are higher. This in particular the case if there is no trusted collaboration (yet) and a lack of prior (expandable) trust between organisations that are about to collaborate.

Notions on (changing) types of collaborations

Thus far the discussion focused on collaboration in general. But with different collaborations come different requirements regarding shared information and the level of sharing (being data, information or knowledge). Those are the actual assets which concern organisations and which are the reason of opting for some (combination of) sources to assess the trustworthiness of parties with those assets and not have them use it opportunistically.

Furthermore, depending on the type of collaborations there will be slightly different presuppositions of types of trust. There are actually different forms of trust in actors (A3.2.2.1.3). In collaborations which are focused on thwarting of attacks competence and material trust will be one of the more important forms of trust. If those collaborations increase scope by requiring more information intentional trust will become more important. In contrast, in collaborations focusing on the aftermath of attacks honesty trust is more than likely one of the dominant types of required trust.

All this touches upon the four place predicate of trust (Nooteboom, 2002, p. 38).. It poses that trust is about (1) a party trusting (2) another party (3) in some respects (4) depending on the conditions. The relevance is in that trust (as are contracts) is a uniform ‘thing’, which has consequences. A vast increase of scope of collaborations challenges trust, just like forms of control would have to change to be applicable to the new scope. For example, Bob might be trusted by Alice with raw data to help thwart an attack. The competences of Alice are trusted. But the very same Alice might not be trustworthy with some information in a role in a post-accident collaboration. The honesty trust might not be to the liking of Bob. (Such as Alice being too honest, too negative regarding the efforts to avoid accidents or simply biased in judgment.)

With these notions it is again stressed that a coarse descriptions of possible development can be defined, yet a predefined standard collaboration model in terms of content and participants is infeasible.

7.3 SUGGESTED DEFAULT FOR PRAGMATIC INFORMATION SHARING COLLABORATIONS

As a brief recap, this research is primarily oriented on information sharing collaborations on unknown-unknowns. (For convenience, the classification of security matters is depicted in Table 5.) The intention is to speed up the transition from unknown-unknowns onto known-unknowns. What matters is to minimize the time this transition takes.

Security matter	Manifestation of security matter	Cause
Unknowable-unknown	No harm (yet)	A vulnerability that is present, but not known
Unknown-unknown	Undetected incidents or even accidents	A vulnerability that is exploited by an adversary
Known-unknown	Avoidable accidents	A vulnerability that is known by defenders
Known	Avoidable incidents	A vulnerability for which a solution is available

Table 5: A simplified version of the classification that was depicted in 2.3. The cause was originally more extensive.

With an unknown-unknown there is a security breach, but this is not known yet. The result of the breach is a security incident, possibly a security accident. Two reasons for the security breach are distinguished, first and second order breaches. The first order breach is involves the initial or first attack, an exploitation of an unknowable vulnerability. The consequence of this first order attack is the acquisition of credentials by the adversary. Think of the attack on the servers of Adobe, in which adversaries were able to acquire a database with user credentials (Goodin, 2013d). A second order case of an unknown-unknown is one in which the acquired credentials are actually used. This is what Diapers and Facebook tried to avoid by revoking user credentials that were acquired in the attack on Adobe (Goodin, 2013d). Both, first and second order cases, are examples of unknown-unknown because it is unknown that negative consequences occur. In the second order breach there are no real indicators thereof. Somebody would log in using valid credentials. What matters with first and second order breaches is to quickly act to make sure it soon becomes a known-unknown. In the case of a first order attack the defender must become aware that there are negative consequences, indicating that there must be some vulnerability. With the second order attack the defender has to know that the credentials are acquired. The reason of it all does not have to be known, just that something is wrong. This step is all what information sharing on unknown-unknowns is about. It is initially about detecting and informing each other in case of incidents or accidents to avoid further damage and to investigate. The next steps would be to (work together to be able to) understand why it was possible, what the vulnerability was, how it could be exploited, and ultimately finding the solution.

7.3.1 *Small and focused collaborations of similar organisations as a default for improved situation awareness*

To set up a pragmatic information sharing collaboration of cases of unknown-unknowns, the default model is to use *multiple, focused collaborations of similar organisations which are part of multiple collaborations*. This statement consists of the following aspects:

- the collaboration has to focus on either avoiding, thwarting or learning from activities in specific parts of the Bow Tie, such as focusing on qualities of specific barriers, reconsideration of whether assets can or should be protected by barriers, or being able to detect suspicious activities earlier on.
- the participating organisations in the collaborations have to be similar, similarity can be found in for example the sector, used assets or the size of the organisations, to minimize the cognitive distance; with that it is likely that collaborations have to be small in terms of the amount of participating organisations, with small being limited to dozens of organisations and preferably even less, for the purpose of a minimization of the cognitive distances, and
- preferably the organisations are participating in multiple collaborations, allowing for limitation of the scope of individual collaborations on specific concerns (such as stopping a botnet with some participants) and yet still allowing organisations to acquire ‘all’ the relevant information which would otherwise be present in large collaborations.

All this is supposed to result in as small as possible collaborations of organisations which are affected by, if not targets of, the same adversaries. Expanding the scopes and sizes of collaborations might theoretically result in more relevant information being available in in the collaboration. But in practice the potential will be limited as there will also be more information shared that is not relevant. With more targeted attacks, less organisations get confronted with the same attacks. Furthermore, with larger collaborations organisations will refrain from offering intimate information. If more organisations take part in a collaboration it is harder for an individual organisation to keep an

overview of the collaboration. This overview is necessary to be able to assess trustworthiness of the participants and to detect (signals of) untrustworthy behaviour. With suspicion of freeriding behaviour can also increase by others. Such freeriding behaviour entails that organisations primarily consume information and do not provide useful information. And even if they do not, this could be the result of an impaired awareness that there is a need for information. Organisations might miss such requests if more, to the organisation in question, irrelevant information is requested and shared.
In all this, the size of the collaboration is merely to be considered a starting point. If deemed necessary, the collaboration might increase scale and scope. What is suggested is to 'start small' for the sake of development of trust and productivity of information sharing. Along this train of thought, the suggestion is to initially limit the level of information sharing to information and knowledge, not data. At least, this should be the default at first, until organisations are sufficiently prepared for the collaboration, by structuring the data.

The envisioned result is that organisations take part in multiple collaborations which can act quickly. Think of collaborations on:

- informing the origin organisation on negative consequences taking place (such as machines sending spam),
- informing others about assets were compromised (such as harvested user credentials),
- informing others of suspicious behaviour of services (cause by some vulnerability), but also
- discussing new defence mechanisms, or
- investigating and taking down botnets in joint effort.

With that, the 'time dimension' is yet again covered in that collaborations can focus on prevention, detection and response. Crucially, each of such collaborations requires preparation by the involved organisation to be able to quickly act upon required activity. For example, avoiding exploitation of zero day vulnerabilities or a machine sending spam is inevitable. But what matters is to minimize the time this is possible and that requires preparation and continuous improvement of the preparation.

7.3.2 LIMITED USEFULNESS OF DEFAULT COLLABORATION

The underlying focus of the default collaboration is on the unknown-unknown in that parties have to help each other improve their situation awareness. But the witnessed cases of cyber accidents are far from always the result of (un)known-unknowns. In many cases they are the result of known, avoidable, weaknesses. Examples are unpatched defences[15], lacking defences[16] or careless behaviour/mistakes by people. Having organisations struggling with more basic security measures take part in collaborations that intent to address unknown-unknown threats seems counterintuitive. The organisations have different primary concerns and different resources. For those having to deal with known issues there are patches available, best practices and trainings to deal with the what could be described as 'basic' security measures. And herein the information sources are rather straightforward. Think of a patch which is available from Microsoft, Adobe or Oracle. With that, there is no need for an advanced collaboration scheme, definitions of information requirements and so on.
However, there might be some value in having larger collaborations after all, albeit for a different purpose than the intended pragmatic information sharing. Larger collaborations tend to focus on

[15] Exploitation of vulnerabilities after a patch became available is 'typical occurrence' (Baccas, 2012, p. 2). Reasons for not installing patches range from unawareness of availability of a patch, laziness/busyness to install patches or the use of non-licensed software (Baccas, 2012, p. 10). Busyness is a simplification of what might actually be the result of wrong allocation of time ('make time) (Baccas, 2012, p. 10), not having sufficient resources to keep up with patch policy, or that the patch affects the functioning of the running system (Symantec, 2013b, p. 26)(Goodin, 2013e). In the end, whatever the reason, what matters is that the accident could have been avoided if the patch would have been installed.

[16] Organisations invest in new technologies (such as mobile, social and the cloud), posing additional risks for organisations (Symantec, 2013b, p. 32), yet a minority of organisations has a security strategy which anticipates the threats these technologies pose (PwC, 2012).

high level, strategic discussions. Such collaborations could be useful to discuss the underlying reasons of the vast amount of security accidents. Think of collaborations on topics such as:

- to reduce the reliance on passwords (FIDO Alliance)[17],
- to revisit the US credit and debit card system as it has inherent flaws and an unproductive security standard (Zetter, 2014),
- discussions on whether SCADA should be online given the current state of (in)security.

All three cases are in a sense about unknown-unknowns, but they fit in a larger discussion to not merely react to yet another unknown-unknown threat that materializes It is about solving the underlying causes. Such as inherent weaknesses of passwords and payment cards, or the lack of updating of technical systems. Such collaborations should take place in larger discussions at a strategic level. Such discussions do not need highly sensitive data, time is not that critical and more participants can actually result in the necessary support to make decisions with serious impact. Such as to force abandoning some payment cards, to ban passwords in favour of other technologies or to allocate the resources to update SCADA systems or take them offline instead.

[17] Passwords are often reused, allowing a hack of Adobe to affect users on other sites, necessitating sites as Facebook and Diapers to reset users credentials pre-emptively (Goodin, 2013d). Such an act by Facebook and Diapers is entirely what the collaboration is about, finding out credentials are compromised, and acting upon that information. However given its inherent underlying weakness, the passwords and their reuse, it is merely a treatment of symptoms.

8 VALIDATION OF THE FINDINGS

Lack of empirical/simulated validation of the findings, probable of findings being valid based on empirical & in extreme

Simplified to its bare essentials the main finding of this research is that a *pragmatic* information sharing collaboration requires a limited scale, a clearly defined goal and participants that are present in multiple collaborations. With that notion and given the amount of:

- different types of challenges (unknown-unknowns to the known risk related matters),
- different ambitions towards these challenges (preventing, thwarting or learning from those),
- different types of organisations involved (perceptions and capabilities),

there are quite a few potential collaborations to be thought of. And with the roadmap still at a conceptual level, implementation for the sake of validation is quite the leap. Falsification is easier . For that it would have to be demonstrated that large collaborations of different organisations can share pragmatic, valuable information and with that improve each other's situation awareness. Instead, to validate the findings it would require multiple implementations. Herein the actual implementation might be simulated. It is easier to change specific aspects of the collaboration and witness the result. But this presupposes that such a simulation model is built. Developing such a model might be troublesome, especially as a considered important factor in collaborations is trust. Modelling trust is no easy feat. Chan et al. developed a composite trust-based agent model (Chan et al., 2012). The purpose of the model is to have agents transmit information to improve their situation awareness, based on their assessment of the level of trust. Although very interesting, this model so far has limitations regarding collaboration structure, types of trust (willingness, competence and intent) and method of sharing, to name a few. With that, those researchers are not able to come up with the collaboration model, that is able to determine upon what should be shared with whom, just yet. And such a model would be necessary to validate the findings. This creates dependencies in that the simulation model has to be a valid representation of the collaboration mechanism in order for it to be useful in validating the findings of this model.

However, apart from demonstrating the validity using (simulated) implementations, validity of the findings could be made probably the other way around. And actually, they were. Quite a few examples were presented in the report. The examples were used to explain the notions. But crucially, the findings were not derived from the examples, they also demonstrated the point. A summary thereof:

- the impact of scale was demonstrated with the findings by NSS Labs which noticed that Information Sharing and Analysis Centres (ISACs) in the United States tend to focus on higher-level strategic discussions. Little pragmatic information is shared as that is harder with large memberships. (Rashid, 2013),
- the impact of scale and diversity was linked to trust, specifically that assuring and assessing trustworthiness becomes a challenge if the scale and particularly diversity increases. And all that was used to explain the way of working in the chemical and banking sector (p.43 & p.87) and between Internet Service Providers and other trusted complainers (p. 37). And to add to this, ENISA discussed trust in their recent report[18] on information sharing by CERTs[19] (ENISA, 2013, p. 3). Specifically they briefly discussed:
    - the detrimental effect of uneven flows of information on trust,
    - the impact of trust on the timeliness and sensitivity of shared information, and
    - the lacking scalability of trust (ENISA, 2013, p. 9), and

[18] The report is based on surveys, interviews and a dedicated workshop focusing on the information sharing by CERTs (ENISA, 2013, p. 3)

[19] Computer Emergency Response Team(s)

- the impact of a lack of a focus was also demonstrated by NSS Labs, ultimately leading to a generic that the private sector needs different information from what governmental organisation offers. (Rashid, 2013)

The findings are also probable based on a little though experiment. If we were to assume that organisations are able to develop a technical solution to deal with different realities of organisations. It is somewhat like what the combination of STIX, CybOX and TAXII is about: the automated sharing of information on analyses threats and defences (US-CERT, 2013). If such a solution of automated information sharing would be able to share relevant information from one organisation to the other where necessary, it could solve the need to define the focus in collaborations. The solution would encounter suspicious behaviour, know who (or what) to ask and with that detect the threat.

However, this still leaves the concern of trust. Just at a different level. With the solution the organisations would have to trust what the solution does. They would have to trust that it shares the relevant information and nothing more with the organisations that need the information. But the organisations would also have to trust the solution itself. And in such a way that it approaches that the organisation would have to have confidence in the solution. They would have trust that the solution is and will stay secure, from adversaries being able to successfully attack the solution. If they could, the adversary might be able to understand how to evade detection. And with organisations increasingly relying on its functioning the question is when they detect a decrease of effectiveness of the solution in stopping adversaries. But a successful attack on the solution could also mean that the information that the machine sends is replicated to the adversary. That way the adversary could receive information on what the environment of the target is like. And with that information they could more convincingly contact people of the organisation. By providing more intimate details in such moments of contact (such as in emails) those contacted persons would be less aware that they are actually dealing with an adversary. And with that they could be seduced to provide documents or pay money for ‘overdue’ bills.

Such concerns will still limit the willingness of organisations to collaborate with anybody and provide information to anybody. Possibly not for the sake of a lack of trust in organisations, but because that way the solution would have access to too much information. The organisation would have to trust that the solution does that what ‘it says that it does’. And that it will continue to do just that in the future regardless of what adversaries do. Such guarantees are hard if not impossible. And for that reason organisation will want to limit the functionality of the solution. Such a concern was already voiced by the Principal Technologist and Senior Policy Analyst of the American Civil Liberties Union, Chris Soghoian, after the presentation of ZeroPoint (Brandom, 2013). ZeroPoint is a technical solution that was recently implemented as a pilot on the network of an Internet Service Provider (ISP). At that level, it supposedly will automatically detect and stop malware that would otherwise affect customers of the ISP (Brandom, 2013). As expressed by Soghoian, the real concern might be that no one is trusted enough to solve the problem of malware entirely, as that would require access to the network which brings great possibilities to that ‘trusted’ organisation (Brandom, 2013).

9 CONCLUSIONS

9.1 CONCLUSION

In order for organisations to improve their cyber security it is considered that information sharing between parties in collaborations is essential. This information sharing is supposed to improve the situation awareness of organisations. With improved awareness organisations would be able to:

- better defend themselves against adversaries because they would be more aware of current modes of attack by adversaries, newly discovered vulnerabilities in assets (software and hardware), and newly developed methods to defends against attacks,
- take precautionary measures against consequences of an attack at another organisation, such as by revoking the user credentials, and
- discover unknown incidents and accident with the help of other organisations by detecting patterns in activity on systems.

The goal of the research was to determine what an information sharing collaboration should look like in order for organisations to actually improve their awareness. Herein the focus was on a collaboration that focuses on detecting unknown-unknowns. These unknown-unknowns are cases of incidents or accidents that took place, but the organisation was not aware of. Let alone know why they occurred.

The goal of the research was achieved by determining the critical decision factors in defining an information sharing collaboration. Collaborations would have to define the goal of the collaboration, the desired end state. And in order to meet that goal they would have to decide upon various aspects of the collaboration. Examples are decisions on who would be part of the collaboration, what information would be shared and how the information would have to be shared. The entire set of steps indicating aspects to decide upon is depicted in Table 6.

Current state ↓	**Development of the collaboration model ↓**						**Desired end state ↓**		
formulation goal of collaboration	roles	collabo-ration structure	topic of shared information	level of information sharing	method of sharing information	type of response	adversary	target	risk-value
	timeliness of response	environ-ment	external interaction	scale	maturity		time dimension	scope of influence	

Table 6: The roadmap consisting of steps representing aspects which organisations have to decide upon in developing a collaboration. Organisations have to first define the goal of the collaboration, as depicted on the left part of the diagram, before actually developing a collaboration. The goal is defined using the five steps on the right. The supportive collaboration to reach that goal is defined using the eleven steps depicted in the middle.

These aspects of the collaboration that organisations have to agree upon are not independent. Decisions on some aspects affect the other. For example, a desire to exchange more sensitive information requires participants to trust each other. Based on such dependences a default collaboration design for pragmatic information sharing is identified. A default with which the goal of this research is met. The default collaboration design entails that participants of collaborations would have to:

- focus in collaboration on either avoiding, thwarting or learning from *specific* attacks,
- be similar in the way they look at the reality that the collaboration focuses on, meaning that they share a similar perception of the challenge at hand, have a similar perception of what has to be done and have similar resources at their disposal to actually act upon those perceptions together, and
- preferably be multiple included in order for the organisation to acquire all the information it needs to get a proper understanding of what it has to know about the situation the organisation is confronted with.

The mechanism behind this is that in collaborations the differences between organisations should be as limited as possible. Maximization of the view on the reality comes second. If the differences between organisations are too great, the organisations:

- will not share the same view on cyber security, meaning they will look at reality different in terms of what happens, what has to be done to improve the security and how the security can be improved,
- cannot assess the trustworthiness of the other organisation, which make them seek forms of assurance such as the use of extensive contracts to limit opportunistic behaviour, or limit their willingness to share information with each other, and
- have a harder time sharing the information that is relevant to the other because they do not understand what the other needs or that the other knows, which results in limited improvement of situation awareness.

By starting small, with organisations that can empathize with each other, on a clearly defined collaboration topic can develop trust. This trust allows for a decrease of transaction costs in the longer run, but also for the possibility of slowly increasing the sensitivity and scope of the shared information.

9.2 RESEARCH LIMITATIONS

The three most important of this research are its binary methodological approach which affects the discussion of shared information, the fact that a custom methodology was developed and the lack of (empirical) validation of the findings.

Binary approach

The Bow Tie model was used to illustrate what participants could discuss. And as discussed in 3.1.2 the main reasons for using the Bow Tie for that are that the Bow Tie model is:

- relevant to illustrate the challenges of cyber security with threats and barriers,
- rather intuitive to use in systematically analysing risks, and
- that the findings on threats and barriers are easy to communicate in collaborations.

At the heart of this is the binary 'cause-effect' line of reasoning. For example, a barrier is secure or insecure. The binary approach is in that the barrier can protect against an event or not. That greatly simplifies the security challenge at hand in two ways. First of all, barriers, components and persons coexist and their specific combination might affect its performance. It is the tight coupling and the complex interaction which might present unique combinations of threats. This includes the type of couplings and interactions which were unknown to be possible. Think of the use of Windows update to spread malware. Which actually happened with the malware 'Flame' (Stevens, 2012). The second simplification is that the insecurity is not actually about insecurity of the environment per se. It could be that the conditions in the organisation are the actual reason of the insecurity. Perhaps circumstances, such as deadlines, result in a decrease of attention to security. Or it could be more systematically. This type of discussing was implicitly touched upon with unpatched system, but actually is more encompassing. The policies dictating requirements of testing of patches. But it is also about the adoption of techniques without proper support. Think of discussions of 'bring your own device', supposedly providing benefits to organisations. But with those options come challenges which security did not anticipate. Multi-layer defences oriented at the Internet connection are useless if the infection can already be brought in with machines which were outside the facility.

These type of challenges are cyber security challenges, they might be worth collaboration too, but were briefly and rather implicitly touched upon. With *such* challenges, the Bow Tie might not be the best model. As discussed in A2.1, systemic models are better equipped for studying the challenges resulting from couplings and interactions. But the systemic models are harder to use, visualize and as a result share in a collaboration. For that reason the Bow Tie was used, which certainly does have its use, it just is not capable of covering the entire scope of challenges.

Lack of a validated methodology

The lack of a standard, validated methodology made the selection of method and theories more critical and makes it harder to determine completeness of the findings. Methods and theories have their weaknesses and omissions and these could not be circumvented entirely.

This limitation is best explained by means of the way the model by Kowtha et al. was used. The model was developed on the basis of existing, American collaboration centres, presenting three weaknesses from the point of view of this research. First of all, the centres were in development, yet already functional, to say the characterisation is based on an end state. This research is about how to setup a collaboration from the start. The distinction is crucial as some starting conditions might not be noticeable after a while or be overlooked. Think of (what was omitted entirely) the development of trust or contracts to a macro level. Were contracts used or did the collaboration start small to minimize the stakes to allow trust to grow. Second, Americans have, compared to Europeans, different ways of working and are subject to different regulations, which might have consequences to the generalizability of the findings. Finally, the centres as studied by Kowtha et al. were more generic collaborations, whereas this research is about information sharing collaborations. For example, 'asset protection' is a by Kowtha et al. identified possible role. In this research asset protection is the goal, by means of sharing information on how to protect assets. As a result, not everything applies from the model by Kowtha et al. to this research. And possibly some relevant elements to this research might have been left out by Kowtha et al.. With these limitations of *the method* there was a need to assess the completeness of the model. The Bow Tie model was used, which is a rather coarse method, demonstrated by the fact it got extended on the basis of findings based on the model by Kowtha et al..

The lack of an overarching mapping necessitated to complement the model by Kowtha with additional theories and methods based on reasoned personal judgment. But this also necessitated to simplify methods and approaches in some way for the sake of time. For example in the development of the three roadmaps there was no guidance. The roadmaps are now merely three constructs, but are by no means the three possible constructs. At the very least there is a mirror option to the second one, which still considers collaborations in a secluded location, yet does not presuppose equal participants. For these reasons complete coverage, as far possibly anyway, is not considered to be the case.

Lack of empirical validation of the findings

Finally the findings were not validated. Actually validating the findings would require to use the findings and actually develop new collaborations using the findings. But the findings actually stipulate that collaborations would have to be focused. With that, quite a few collaborations would have to be developed, which would take time. Too much time for this research project. However, the findings were made probable by the extensive use of examples.

9.3 FUTURE RESEARCH

The first suggestion for future research would be to validate the findings of this research. Specifically that there is indeed not a single collaboration possible regardless of the challenge at hand. With that, the additional suggestions would be to deepen the understanding of successful information sharing collaborations. Specifically on:

- how to develop the best collaboration design to share information, and
- how to increase the scope of the shared information to improve the security of the environment (consisting of computer systems, but also people that use those systems), instead of just discussing aspects of attacks that organisations are confronted with.

These two suggestions are discussed in respective order in this section.

Knowing there are already some information sharing collaborations the main recommendation for future research is to compare these existing collaborations. And for that, parties need a way to look at such collaborations. The research by Kowtha et la. was a start in all this. This research presented a refinement thereof that is focused on information sharing. The intentions should be to compare the success of the different collaborations. As discussed in this research for that first a clear definition of success or the goal should be made. Collaborations that have the same goal should be compared to identify the constitutes of (a lack of) success. Herein the challenge will be to identify comparable collaborations. Collaborations of organisations that are confronted with a similar reality, consisting of similar threats.
What would be interesting to study in such a comparison is on whether pre-existing collaborations that added cyber security to the agenda are successful. Those collaborations would be able to benefit from pre-existing trust, yet have to expand the circumstances in which those organisations have to trust each other. The alternative would be to setup new collaborations. Such a collaboration is the implied route in this research. But it would take the organisations to build trust and slowly increase the sensitivity of the shared information.
In all this, whether it is on comparing or in developing collaboration models it is recommended to use theories, such as presented in this research, as a foundation. For example, as discussed and demonstrated using some examples quite a few collaborations seem to neglect concerns of trust.
Ultimately what would be of value would be to improve the theories on collaborations, such as those discussed in this research and to validate or falsify these using collaborations. That way collaborations could be custom, but still have a solid foundation, a clear line of reasoning on why the design of the collaboration is as it is.

Another suggestion for future research is to study how organisations can productively share pragmatic information about less 'binary' events. Some accidents are actually the result of the fact that systems are interconnected. Or the circumstances in which the those systems are positioned. Now there are tools such as guidelines, best practices and standards that are supposed to help organisations in keeping their environment secure. Not only by keeping the system secure by installing patches, installing defensive measures and separating systems where possible. But also by training personnel on how to work in a way that contributes to a secure environment. However, the tools are rather generic. Organisations have to use the tools and hereby translate the notions into actions that are applicable to their environment. The question is whether information sharing collaborations can help organisations in improving the security of the actual environment. With such an environment organisation might be able to detect incidents, before they become accidents.

10 REFERENCES

Abuse Information Exchange, 2013. AbuseHUB van start: botnets aangepakt [WWW Document]. URL http://www.abuseinformationexchange.nl/mm_uploads/AbuseHUB_van_start_botnets_aangepakt-1.pdf

Anderson, N., 2012. Confirmed: US and Israel created Stuxnet, lost control of it [WWW Document]. Ars Technica. URL http://arstechnica.com/tech-policy/2012/06/confirmed-us-israel-created-stuxnet-lost-control-of-it/ (accessed 6.5.13).

Anderson, R., Barton, C., Boehme, R., Clayton, R., Eeten, M.J.G. van, Moore, T., Savage, S., 2012. Measuring the cost of cybercrime.

Baccas, P., 2012. A time-based analysis of Rich Text Format manipulation: a deeper analysis of the RTF exploit CVE-2010-3333.

Brandom, R., 2013. ZeroPoint is the malware cure that could be worse than the disease [WWW Document]. The Verge. URL http://www.theverge.com/2013/12/9/5191858/secdev-zeropoint-could-destroy-malware-for-good-rohozinski (accessed 2.8.14).

Bruijn, M., de, Wal, F., van der, Swuste, P., 2011. Learning from HSE-MS based incident investigation.

Cabinet Office, 2013. Government launches information sharing partnership on cyber security [WWW Document]. URL https://www.gov.uk/government/news/government-launches-information-sharing-partnership-on-cyber-security (accessed 9.10.13).

CGE, 2013. The History of Bowtie [WWW Document]. URL http://www.cgerisk.com/knowledge-base/risk-assessment/the-bowtie-methodology (accessed 10.23.13).

Chan, K., Cho, J.-H., Adali, S., 2012. Composite Trust Model for an Information Sharing Scenario, in: 2012 9th International Conference on Ubiquitous Intelligence Computing and 9th International Conference on Autonomic Trusted Computing (UIC/ATC). Presented at the 2012 9th International Conference on Ubiquitous Intelligence Computing and 9th International Conference on Autonomic Trusted Computing (UIC/ATC), pp. 439–446.

Committee on National Security Systems, 2010. National Information Assurance (IA) Glossary.

Daley, R., Millar, T., Osorno, M., 2011. Operationalizing the coordinated incident handling model, in: 2011 IEEE International Conference on Technologies for Homeland Security (HST). Presented at the 2011 IEEE International Conference on Technologies for Homeland Security (HST), pp. 287–294.

Dongen, H.J. van, 1991. Some Notions on Social Integration and Steering, in: Veld, R.J. in 't, Schaap, L., Termeer, C.J.A.M., Twist, M.J.W. van (Eds.), Autopoiesis and Configuration Theory: New Approaches to Societal Steering. Springer Netherlands, pp. 47–54.

Dongen, H.J. van, Laat, W.A.M. de, Maas, A.J.J.A., 1996. Een kwestie van verschil: conflicthantering en onderhandeling in een configuratieve integratietheorie. Eburon, Delft.

Emisoft, 2012. Malware and viruses – What's the difference? [WWW Document]. URL http://blog.emsisoft.com/2012/03/08/tec120308/ (accessed 2.5.14).

Endsley, M.R., 1995. Toward a Theory of Situation Awareness in Dynamic Systems. Human Factors: The Journal of the Human Factors and Ergonomics Society 37, 32–64.

Endsley, M.R., 1995. Direct measurement of situation awareness in dynamic systems : Situation awareness. Human factors 37, 65–84.

Endsley, M.R., Jones, W.M., 1997. Situation Awareness Information Dominance & Information Warfare (No. 97-01).

ENISA, 2013. Detect - SHARE - Protect [WWW Document]. URL https://www.enisa.europa.eu/activities/cert/support/data-sharing/detect-share-protect-solutions-for-improving-threat-data-exchange-among-certs/at_download/fullReport

FAA, 2008. Air Traffic Organization - Safety Management System Manual - Version 2.1 [WWW Document]. URL http://www.faa.gov/air_traffic/publications/media/atosmsmanualversion2-1_05-27-08_final.pdf

Finkle, J., Menn, J., 2012. Keith Alexander, NSA Chief, Asks For Hackers' Help In Making Internet More Secure [WWW Document]. Huffington Post. URL

http://www.huffingtonpost.com/2012/07/28/keith-alexander-nsa_n_1712185.html (accessed 2.5.14).
Florêncio, D., Herley, C., 2011. Sex, Lies and Cyber-crime Surveys [WWW Document]. Microsoft Research. URL http://research.microsoft.com/pubs/149886/SexLiesandCybercrimeSurveys.pdf
Fricke, M., 2008. The knowledge pyramid: a critique of the DIKW hierarchy. Journal of Information Science 35, 131–142.
F-Secure, 2013a. Threat Report H2 2012 [WWW Document]. URL http://www.f-secure.com/static/doc/labs_global/Research/Threat_Report_H2_2012.pdf
F-Secure, 2013b. Threat Report H1 2013 [WWW Document]. URL http://www.f-secure.com/static/doc/labs_global/Research/Threat_Report_H1_2013.pdf
Goodin, D., 2012a. Discovery of new "zero-day" exploit links developers of Stuxnet, Flame [WWW Document]. URL http://arstechnica.com/security/2012/06/zero-day-exploit-links-stuxnet-flame/ (accessed 6.5.13).
Goodin, D., 2012b. Crypto breakthrough shows Flame was designed by world-class scientists [WWW Document]. Ars Technica. URL http://arstechnica.com/security/2012/06/flame-crypto-breakthrough/ (accessed 6.5.13).
Goodin, D., 2012c. Mystery malware wreaks havoc on energy sector computers [WWW Document]. Ars Technica. URL http://arstechnica.com/security/2012/08/shamoon-malware-attack/ (accessed 6.5.13).
Goodin, D., 2012d. Hack attack on energy giant highlights threat to critical infrastructure [WWW Document]. Ars Technica. URL http://arstechnica.com/security/2012/09/hack-attack-on-energy-giant-highlights-threat-to-critical-infrastructure/ (accessed 6.5.13).
Goodin, D., 2013a. Revealed: Stuxnet "beta's" devious alternate attack on Iran nuke program [WWW Document]. Ars Technica. URL http://arstechnica.com/security/2013/02/new-version-of-stuxnet-sheds-light-on-iran-targeting-cyberweapon/ (accessed 6.5.13).
Goodin, D., 2013b. Why Red October malware is the Swiss Army knife of espionage [WWW Document]. Ars Technica. URL http://arstechnica.com/security/2013/01/why-red-october-malware-is-the-swiss-army-knife-of-espionage/ (accessed 6.5.13).
Goodin, D., 2013c. Massive espionage malware targeting governments undetected for 5 years [WWW Document]. Ars Technica. URL http://arstechnica.com/security/2013/01/red-october-computer-espionage-network-may-have-stolen-terabytes-of-data/ (accessed 6.5.13).
Goodin, D., 2013d. How Adobe's messy password breach can spill to sites like Diapers.com [WWW Document]. Ars Technica. URL http://arstechnica.com/security/2013/11/how-adobes-messy-password-breach-can-spill-to-sites-like-diapers-com/ (accessed 1.27.14).
Goodin, D., 2013e. How hackers made minced meat of Department of Energy networks [WWW Document]. Ars Technica. URL http://arstechnica.com/security/2013/12/how-hackers-made-minced-meat-of-department-of-energy-networks/ (accessed 1.27.14).
Hathaway, O., Crootof, R., 2012. The Law of Cyber-Attack. Faculty Scholarship Series.
Hevner, A.R., March, S.T., Park, J., Ram, S., 2004. Design science in information systems research. MIS Q. 28, 75–105.
Hollnagel, E., 2004. Barriers and accident prevention. Ashgate, Aldershot, Hampshire, England ; Burlington, VT.
Hollnagel, E., 2008. Risk + barriers = safety? Safety Science 46, 221–229.
Hutchinson, J., 2013. McAfee regrets "flawed" trillion dollar cybercrime claims [WWW Document]. Financial Review. URL http://www.afr.com/p/technology/mcafee_regrets_flawed_trillion_dollar_msQ2WFkVLEZKx7Yv7ZCMQI (accessed 2.5.14).
Hyppönen, M., 2012. Cyber War [WWW Document]. URL https://www.youtube.com/watch?v=iwiqgm4-HR4&feature=youtube_gdata_player

ISP Today, 2012. Nederland loopt voorop met Abuse IX [WWW Document]. URL http://www.isptoday.nl/nieuws/nederland-loopt-voorop-met-abuse-ix/ (accessed 11.3.13).

Jacob Appelbaum: NSA's FoxAcid/Quantum Programs "Like the Military Occupation of Entire Internet," 2013.

Jeffries, A., 2013. Edward Snowden is now a gimmick to sell security software [WWW Document]. The Verge. URL http://www.theverge.com/2013/8/1/4577376/edward-snowden-is-now-a-gimmick-to-sell-security-software (accessed 9.10.13).

Kaspersky, 2013. Malware, spam, and phishing: the threats most commonly encountered by companies [WWW Document]. URL http://www.kaspersky.com/about/news/virus/2013/Malware_spam_and_phishing_the_threats_most_commonly_encountered_by_companies (accessed 2.5.14).

Kirk, J., 2011. "Night Dragon" attacks from China strike energy companies [WWW Document]. Network World. URL http://www.networkworld.com/news/2011/021011-night-dragon-attacks-from-china.html (accessed 6.5.13).

Klijn, E.H., Teisman, G.R., 1991. Effective Policy Making in a Multi-Actor Setting: Networks and Steering, in: Veld, R.J. in 't, Schaap, L., Termeer, C.J.A.M., Twist, M.J.W. van (Eds.), Autopoiesis and Configuration Theory: New Approaches to Societal Steering. Springer Netherlands, pp. 99–111.

Kowtha, S., Nolan, L.A., Daley, R.A., 2012. An Analytical Model For Characterizing Operations Centers.

Kroes, N., 2012. Cyber-security – a shared responsibility [WWW Document]. URL http://europa.eu/rapid/press-release_SPEECH-12-774_en.htm (accessed 3.5.13).

Laat, W. de, Maas, A., 2003. Syllabus configuratieve integratietheorie, 3rd ed. Eburon, Delft, Netherlands.

Lange, R. de, 2013. Onvrede over aanpak cybercrime. Het Financieele Dagblad.

Léger, M.-A., 2008. Bow Tie [WWW Document]. URL http://www.leger.ca/GRIS/BowTie.html (accessed 11.10.13).

Mandiant, 2013. M-Trends 2013: Attack the Security Gap [WWW Document]. URL https://dl.mandiant.com/EE/library/M-Trends_2013.pdf

Maude, F., 2013. Cyber Security Information Sharing Partnership [WWW Document]. URL https://www.gov.uk/government/speeches/cyber-security-information-sharing-programme (accessed 8.16.13).

Milan, M., 2008. Backcasting 101 [WWW Document]. URL http://www.slideshare.net/mmilan/backcasting-101-final-public (accessed 9.10.13).

National Cyber Security Alliance, Symantec, 2012. New Survey Shows U.S. Small Business Owners Not Concerned About Cybersecurity; Majority Have No Policies or Contingency Plans [WWW Document]. URL http://www.symantec.com/about/news/release/article.jsp?prid=20121015_01 (accessed 12.24.13).

National Infrastructure Advisory Council, 2012. Intelligence information sharing [WWW Document].

NCSC, 2013. Cybersecuritybeeld Nederland [WWW Document]. URL https://www.ncsc.nl/binaries/nl/actueel/nieuwsberichten/cybersecuritybeeld-nederland-kwetsbaarheid-van-ict-onverminderd-hoog/1/NCSC%2BCSBN%2B3%2B3%2Bjuli%2B2013.pdf

NIST, 2012. SP 800-30 Rev. 1 - Guide for Conducting Risk Assessments.

Nooteboom, B., 2002. Trust: forms, foundations, functions, failures, and figures. E. Elgar Pub, Cheltenham, UK ; Northampton, MA.

Nordgård, D.E., 2008. Quantitative risk assessment in distribution system maintenance management using bow-tie modeling. Presented at the 16th Power Systems Computation Conference, Glasgow, Scotland.

Ponemon Institute, 2013. 2013 Cost of Data Breach Study: Global Analysis.

Proposal for a Directive of the European Parliament and of the Council concerning measures to ensure a high common level of network and information security accross the Union, 2013.

PwC, 2012. Cybersecurity: The new business priority [WWW Document]. PwC. URL http://www.pwc.com/us/en/view/issue-15/cybersecurity-business-priority.jhtml (accessed 1.27.14).

Rashid, F.Y., 2013. Report Shows "Uneven Progress" in Cybersecurity Information Sharing [WWW Document]. SecurityWeek. URL http://www.securityweek.com/report-shows-uneven-progress-cybersecurity-information-sharing (accessed 10.20.13).

Rodgers, J.L., Nicewander, W.A., Toothaker, L., 1984. Linearly Independent, Orthogonal, and Uncorrelated Variables. The American Statistician 38, 133.

RPS, 2012. BowTie Methode [WWW Document]. URL http://www.rps.nl/content/downloads/RPS%20BowTie%20Methode.pdf

Salas, E., Prince, C., Baker, D.P., Shrestha, L., 1995. Situation Awareness in Team Performance: Implications for Measurement and Training. Human Factors: The Journal of the Human Factors and Ergonomics Society 37, 123–136.

Salmon, P.M., Stanton, N.A., Walker, G.H., Jenkins, D.P., 2009. Distributed situation awareness theory, measurement and application to teamwork, Human factors in defence. Ashgate, Farnham, England ; Burlington, VT.

Schellevis, J., 2013. "Ict-beveiligers lopen achter op hackers" [WWW Document]. URL http://tweakers.net/nieuws/87585/ict-beveiligers-lopen-achter-op-hackers.html (accessed 3.5.13).

Schneier, B., 2013. Understanding the threats in cyberspace [WWW Document]. Europe's World. URL http://europesworld.org/commentaries/understanding-the-threats-in-cyberspace/#.UoikseIUY5g (accessed 11.17.13).

Security.nl, 2012a. Hoe providers weten dat je computer besmet is [WWW Document]. URL https://www.security.nl/posting/38285/ (accessed 11.3.13).

Security.nl, 2012b. XS4ALL zet 1000 klanten per maand in Walled Garden - Security.NL [WWW Document]. URL https://www.security.nl/posting/38306/XS4ALL+zet+1000+klanten+per+maand+in+Walled+Garden (accessed 11.24.13).

Seppänen, H., Mäkelä, J., Luokkala, P., Virrantaus, K., 2013. Developing shared situational awareness for emergency management. Safety Science 55, 1–9.

Sklet, S., 2006. Safety barriers: Definition, classification, and performance. Journal of Loss Prevention in the Process Industries 19, 494–506.

Skopik, F., Ma, Z., Smith, P., Bleier, T., 2012. Designing a Cyber Attack Information System for National Situational Awareness, in: Aschenbruck, N., Martini, P., Meier, M., Tölle, J. (Eds.), Future Security, Communications in Computer and Information Science. Springer Berlin Heidelberg, pp. 277–288.

Solutionary, 2013. Solutionary Research Reveals that 58 Percent of Vulnerabilities Targeted by Well-Known Exploit Kits Are Over Two Years Old; 70 Percent of Exploit Kits Originated in Russia [WWW Document]. URL http://www.solutionary.com/news-events/press-releases/2013/01/sert-2012-q4-intelligence-report/ (accessed 10.14.13).

Sorensen, L., Stanton, N., 2012. Is there a relationship between team organisational structure, distributed situational awareness and performance?, in: Anderson, M. (Ed.), Contemporary Ergonomics and Human Factors 2012. CRC Press, pp. 151–158.

Sorensen, L.J., Stanton, N.A., 2011. Is SA shared or distributed in team work? An exploratory study in an intelligence analysis task. International Journal of Industrial Ergonomics 41, 677–687.

Sorensen, L.J., Stanton, N.A., 2013. Y is best: How Distributed Situational Awareness is mediated by organisational structure and correlated with task success. Safety Science 56, 72–79.

Stanton, N.A., Salmon, P.M., Walker, G.H., Jenkins, D.P., 2010. Is situation awareness all in the mind? Theoretical Issues in Ergonomics Science 11, 29–40.

Stanton, N.A., Stewart, R., Harris, D., Houghton, R.J., Baber, C., McMaster, R., Salmon, P., Hoyle, G., Walker, G., Young, M.S., Linsell, M., Dymott, R., Green, D., 2006. Distributed situation

awareness in dynamic systems: theoretical development and application of an ergonomics methodology. Ergonomics 49, 1288–1311.

Stevens, M., 2012. Technical background of the Flame collision attack [WWW Document]. CWI. URL http://www.cwi.nl/nieuws/2012/cwi-cryptanalist-ontdekt-nieuwe-cryptografische-aanvalsvariant-in-flame-virus (accessed 12.18.13).

Stevenson, A., 2013. UK government's anti-hacker CISP initiative failing to support SMBs [WWW Document]. URL http://www.v3.co.uk/v3-uk/news/2304171/uk-governments-anti-hacker-cisp-initiative-failing-to-support-smbs (accessed 2.9.14).

Symantec, 2013a. Stuxnet 0.5: Disrupting Uranium Processing at Natanz [WWW Document]. URL http://www.symantec.com/connect/blogs/stuxnet-05-disrupting-uranium-processing-natanz (accessed 6.5.13).

Symantec, 2013b. Internet Security Threat Report 2013 [WWW Document]. URL http://www.symantec.com/content/en/us/enterprise/other_resources/b-istr_main_report_v18_2012_21291018.en-us.pdf

Termeer, C.J.A.M., 1993. Dynamiek en inertie rondom mestbeleid: een studie naar veranderingsprocessen in het varkenshouderijnetwerk. VUGA.

Termeer, C.J.A.M., Kessener, B., 2007. Revitalizing Stagnated Policy Processes Using the Configuration Approach for Research and Interventions. Journal of Applied Behavioral Science 43, 256–272.

Twist, M.J.W. van, Termeer, C.J. a. M., 1991. Introduction to Configuration Approach: A Process Theory for Societal Steering, in: Veld, R.J. in 't, Schaap, L., Termeer, C.J.A.M., Twist, M.J.W. van (Eds.), Autopoiesis and Configuration Theory: New Approaches to Societal Steering. Springer Netherlands, pp. 19–29.

US-CERT, 2013. Information Sharing Specifications for Cybersecurity [WWW Document]. URL https://www.us-cert.gov/Information-Sharing-Specifications-Cybersecurity (accessed 2.8.14).

Verizon, 2013. 2013 Data Breach Investigations Report [WWW Document]. URL http://www.verizonenterprise.com/resources/reports/rp_data-breach-investigations-report-2013_en_xg.pdf

Von Solms, R., van Niekerk, J., 2013. From information security to cyber security. Computers & Security 38, 97–102.

White, G.B., 2011. The community cyber security maturity model, in: 2011 IEEE International Conference on Technologies for Homeland Security (HST). Presented at the 2011 IEEE International Conference on Technologies for Homeland Security (HST), pp. 173–178.

Zetter, K., 2012. Flame Hijacks Microsoft Update to Spread Malware Disguised As Legit Code [WWW Document]. Threat Level. URL http://www.wired.com/threatlevel/2012/06/flame-microsoft-certificate/ (accessed 2.9.14).

Zetter, K., 2014. Target Got Hacked Hard in 2005. Here's Why They Let It Happen Again | Threat Level | Wired.com [WWW Document]. Threat Level. URL http://www.wired.com/threatlevel/2014/01/target-hack/ (accessed 1.27.14).

Zhao, W., White, G., 2012. A collaborative information sharing framework for Community Cyber Security, in: Homeland Security (HST), 2012 IEEE Conference on Technologies for. Presented at the Homeland Security (HST), 2012 IEEE Conference on Technologies for, pp. 457–462.

A1 TERMINOLOGY

The definitions that were used in the research might not be identical to those most commonly used. In the end the main purpose of the definitions is to clearly and consistently convey a principle, explanation or the scope of term.

Accident

An occurrence (such as an attack) that results in a negative consequence (such as a loss of data).

Asset

The thing (such as devices, data or other valuables) one tries to protect

Attack

A threat launched by some adversary on the assets of some target, resulting in an incident or even an accident

An attack is always coming from some party, the adversary, and affects the assets of some party, the target. The simplified definition considers what is relevant from the point of view of this research. It is a simplification from the definition of a cyber-attack, as presented in the introduction. In the introduction a cyber-attack is defined as:

> "A hostile act using computer or related networks or systems, and hereby affecting and/or disrupting and/or destroying an organisations' cyber systems, assets, or functions. The intended effects of cyber-attack are not necessarily limited to the targeted computer systems or data themselves. The activation or effect of a cyber-attack may be widely separated temporally and geographically from the delivery."

This definition is based on the original definition by Joint Chiefs of Staff, defining cyber-attacks as:

> "A hostile act using computer or related networks or systems, and intended to disrupt and/or destroy an adversary's critical cyber systems, assets, or functions. The intended effects of cyber attack are not necessarily limited to the targeted computer systems or data themselves—for instance, attacks on computer systems which are intended to degrade or destroy infrastructure or C2 capability. A cyber attack may use intermediate delivery vehicles including peripheral devices, electronic transmitters, embedded code, or human operators. The activation or effect of a cyber attack may be widely separated temporally and geographically from the delivery." (Hathaway and Crootof, 2012, p. 824)

This definition is more extensive, but focusses more on military and nation state related elements. In this research cyber-attacks are considered from a more encompassing whole. Organisations can be affected by cyber-attacks financially.

Cumulative cognitive coverage

The sum of unique and compatible awareness; it is the extent to which the participants, together, are theoretically able to see the actual, relevant reality.

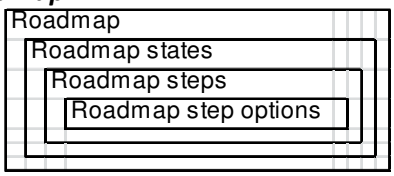
Cumulative cognitive distance

The sum of the differences in awareness; it is the extent to which participants differ in such a way that it affects their ability to understand each other.

Incident

An occurrence (such as an attack) that does not have negative consequences to the organisation (such as a scanning attack for weaknesses by the adversary).

Roadmap



There are three, numbered roadmaps: first-, second- and third- roadmap

Each roadmap consists of a current state, development of the collaboration model state and an end state

Each state consists of one or more steps, a step represents a decision moment, depicted on a roadmap, at which organisations, together, have to decide and agree upon a specific aspect of an information sharing collaboration and that will help further define the design of the information sharing collaboration.

For each step one or more step options are identified

The different steps and step options are presented in A3.2. They are separated by the roadmap and roadmap states they belong to.

Technique (in the context of being part of threats)

An undefined method such as a specific type of malware which harvests credentials or to log on to a system using the harvested credentials which potentially could cause harm to the organisation.

Threat

The technique(s) adversaries use which affects the target in some way.

Threats are defined as the technique(s) adversaries use which affect the target in some way. This definition is inferred from the following extensive definition by the Committee on National Security Systems (CNSS) (Committee on National Security Systems, 2010, p. 75) is used:

> "Any circumstance or event with the potential to adversely impact organizational operations (including mission, functions, image, or reputation), organizational assets, individuals, other organizations, or the Nation through an information system via unauthorized access, destruction, disclosure, modification of information, and/or denial of service.".

This definition got reiterated by the National Institute of Standards and Technology (NIST) later on (NIST, 2012, p. B–13). Furthermore, of the two by NIST offered options (NIST, 2012, p. 8), threats refer in this report to single events, actions, or circumstances. The second option, sequences of relation actions, activities and/or circumstances are referred to as a or the 'sequence of threats'.

Threat sources

An adversary

In general, types of threat sources include:

(i) hostile cyber or physical attacks;
(ii) human errors of omission or commission;
(iii) structural failures of organization-controlled resources (e.g., hardware, software, environmental controls); and
(iv) natural and man-made disasters, accidents, and failures beyond the control of the organization (NIST, 2012, p. 8).

In this research, following from the scope focusing on intentional actions, solely the first type of threat source is considered. That first type is referred to as the adversarial type by NIST. This type could be represented by an individual, a group, an organisation or a nation-state. The primary characteristics of the adversarial type are capability, intent and target(ing). (NIST, 2012,

p. D–2) Given the scope, this research solely mentions adversaries, despite there actually being different types adversaries. Examples of such types are: criminals, hacktivists and governments.

Unknowns

A security state of for example vulnerability, consequence or defence.

Security state	Cause	Initially resulting in
Unknowable-unknown	A vulnerability which is present	No harm (yet)
Unknown-unknown	A vulnerability which is exploited	Undetected incidents or even accidents
Known-unknown	A vulnerability which is detected	Avoidable accidents
Known	A vulnerability for which a solution is available	Avoidable incidents

Vulnerability

A weakness or gap in the protection efforts which the technique is able to exploit.

A2 METHOD

A2.1 ACCIDENT MODELS

Hollnagel (Hollnagel, 2004, p. 66) identified three types of accident models: sequential models, epidemiological model and the systemic model. Of the three types of models the latter two are most appealing from the point of view of cyber security with adversaries actively trying to exploit unknown vulnerabilities. A high level characterisation of the three models is, as presented by Hollnagel (2004, p. 66), is displayed in Table 7.

	Model type		
	Sequential models	Epidemiological models	Systemic models
Search principle	Specific causes and well-defined links	Carriers, barriers, and latent conditions	Tight couplings and complex interactions
Analysis goals	Eliminate or contain causes	Make defences and barriers stronger	Monitor and control performance variability
Examples	Chain or sequence of events (domino), tree models, network models	Latent conditions, carrier-barriers, pathological systems	Control theoretic models, chaos models, stochastic resonance

Table 7: Types of accident models as distinguished by Hollnagel (2004, p. 66).

The first model is an oversimplifies the challenge of cyber security because sequential models are about elimination or containment the causes of accidents. This would mean the adversary, the threat or the vulnerability would have to be removed or eliminated. But with cyber security the challenge is that continuously new possible causes appear. And some of those, threats, cannot really effectively be avoided. To still remove or eliminate the cause of the accident would be to not use assets that are, in any way, approachable from the outside. An option which will often not even be possible or acceptable.

The second type of model is used in this research, in the form of the Bow Tie model (which is discussed in more detail in 0). The Bow Tie model is both applicable, intuitive and easily usable in a collaborative environment. Like the first type of accident model it is about cause and effect thinking. But the second type puts great emphasis is put on introducing better defences and barriers (such as firewalls and malware scanners or information and training to detect specific threats), whereas the first is about disarm hazardous causes. The second type of accident model, in this research represented by the Bow Tie model,

The third type of accident also deserves some attention. Organizations' environments can be considered to be so complex with all kinds of applications, devices and users it becomes too hard to secure (by means of patching) the systems. To resolve these situations, the third model represents a focus on determining whether the system behaves normally. Herein normally is a dynamic state, accounting for possible desired changes. This possibility distinguishes the systemic model from epidemiological models, which aim to maintain the status quo. The focus of systemic models is on the conditions in which problems can emerge. Hereby any suggestion to an explanation using consecutive series of events (let alone an order, sequence of events) is avoided. Given this nature, Hollnagel considers systemic models to be inherently 'difficult to represent graphically'. (Hollnagel, 2004, p. 65) Furthermore, in the cyber environment system models are more of an implicit concept or aim, still requiring solid implementations. As a result, epidemiological models are selected, yet systemic models are not be ignored entirely. They are discussed in chapter 0.

A2.2 BOW TIE MODEL

Two versions, base model and extensive model focused on collaborations.

A2.2.1 THE BASE MODEL

The Bow Tie model is a qualitative, event based, risk analysis method to structure and systematically analyse the risks and their mitigating measures in an organisation. (Nordgård, 2008)(RPS, 2012) It is a high level method, allowing for further (detailed) refinement, but by design not intended to use for risk calculations (CGE, 2013). Instead, it is used to illustrate the relationships between hazardous activities, causes leading up to undesired events and potentially resulting in undesired consequences (FAA, 2008, p. 33)(RPS, 2012). To help stop these hazardous events from ultimately resulting in those consequences barriers can be employed.

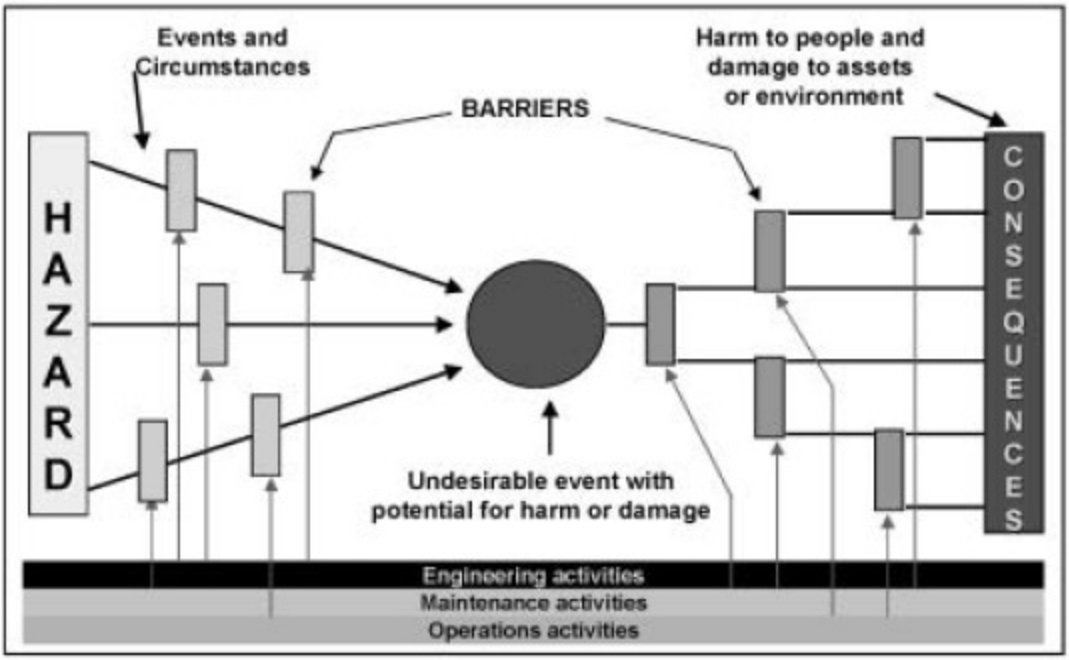


Figure 18: One of the visualizations of the Bow Tie model by Shell International Exploration & Production (Léger, 2008)

Allegedly it came from a chemical industry company (ICI Australia), and got real popular in in the oil and gas industry, before spreading to other industries as aviation, mining and the maritime industry (CGE, 2013). Possibly as a result, there exists multiple slight deviations surrounding the two same concepts: analysing chain of events and an identification of (potential) control measures (more commonly referred to as barriers).

Control measures (or barriers)

To prevent events related to hazardous activities from resulting undesired outcomes barriers can be employed. The concept of barriers and limitations thereof is commonly explained by using the Swiss Cheese model by James Reason. The main principle is to install multiple layers of barriers, as none will be perfect. Each will have its weaknesses (the holes in their defensive measures). (CGE, 2013)
These barriers can be installed on either side of the knot of the bowtie to either protect against losing control over an asset (reflected by the top event, being the knot in the model) or by minimizing the consequences in case control is lost.

Chain of events

The concept of chain of events in the Bow Tie model was allegedly inspired by three methods: fault trees, event threes and causal factors charting (CGE, 2013).
In the Bow Tie model the fault trees are represented by (and commonly referred to as (Léger, 2008)(Bruijn et al., 2011)(CGE, 2013) the left hand side of the model, including the so called 'top

event', the knot of the Bow Tie. The fault tree of the Bow Tie model is actually a simplified form (CGE, 2013). It describes how, related to a hazardous activity, (one or more) events and circumstances can result in an undesirable event with the potential to harm people, damage assets or the environment (Léger, 2008). But contrary to the actual fault tree which is characterised by its completeness and throughout quantification for the determination of risk values the Bow Tie favours readability (CGE, 2013). In this article events exclusively remaining on the left hand side of the Bow Tie will be referred to as incidents, event which do not result in (negative) consequences.
The event tree is covers the remainder of the Bow Tie model on the right, the top event (the knot) excluded. Events which reach to this side of the model will have some sort of consequences. The severity thereof is depending on protective measures to stop progression of damaging consequences (Léger, 2008).
Causal factors charting is the likely origin of the escalation factors (CGE, 2013) and comes back as causality mappings in the Bow Tie model. Whereas the former is mainly used for incident analyses the latter is more about proactive risk analyses, considering all possible causal event paths. Escalation factors have causal paths to specific barriers and pinpoint its weaknesses in those barriers. (CGE, 2013) With that, escalation factors represent the second way in which a barrier can fail to stop an hazardous event, the other being an inherent weakness of the barrier.

Bow Tie model elements in detail
The discussion of the Bow Tie is the result of quite a few publications. But in the end it is my interpretation as there are quite a few conflicting statements or ambiguous statements posed. Particularly the escalation factor is somewhat troublesome. CGE (2013) discusses the fact that barriers are not perfect and that they can fail. Both result in holes in the defences. But a real or inherent weakness is a hole and thus there is nothing to escalate from. For that reason in this reason I distinguish weaknesses in, (inherent) weaknesses and failures.

Hazard
A hazard is something or some activity, in or around the organisation which can result in one or more undesired (subsequent) events (FAA, 2008, p. A–2) (CGE, 2013). An example of a hazard is the storage of sensitive data.

Threats
Threats and events are often used interchangeably in explanations of the Bow Tie model.
An example of a threat is a website open to an SQL-injection which possibly allows an attacker to access unintended portions of the database.

Top event
The top event demarcates the moment of loss of control over the hazard, illustrated by the knot in the Bow Tie model. The actual definition of when control is lost is subjective. Crucially though, with the top event taking place there is no damage yet, it is an incident, not yet an accident. (CGE, 2013)
An example would be the moment somebody who was not supposed to be able to (like an attacker) did query for the user database with account details, but has *not* retrieved the results *yet*.

Consequences
In case control is lost an accident (as in damage to something) is imminent, being defined as some of the possible negative consequences actually taking place. Unless the consequence is actually the case, the event which takes or took place is considered an incident.
As conveyed by CGE (2013) consequences are specific cases of damage. As such, reputation damage and asset damage are just broader categories. An example of a potential consequence related to the reputation category is a change in people's behaviour or willingness in the long run. For example a decrease in customer 'intimacy' by not sharing all details or bogus data in surveys is a possible consequence if the organisation turns out to have lost control over its customer data. Short term

possible consequences are financial penalties for a 'loss' of confidential data to the organisation itself and confidential ending up on spamlists affecting the customers of the organisation.

Barriers

To control the scenarios barriers can be introduced on either side of the top event. Its purpose is to prevent events related to hazardous activities from resulting in top events and ultimately undesired consequences (Bruijn et al., 2011). A barrier is short for a barrier system, which is designed and implemented for one or more barrier functions (Sklet, 2006)(Hollnagel, 2008, p. 225). Which functions are to be distinguished differs per author, as identified by (Sklet, 2006) and displayed in Table 8. This function can be delivered using different types of barrier systems, of which also different classifications exist (Sklet, 2006, p. 500). Each of these systems have their own qualities and inherent weaknesses (irrespective of the actual quality of implementation) (Hollnagel, 2008, p. 228).

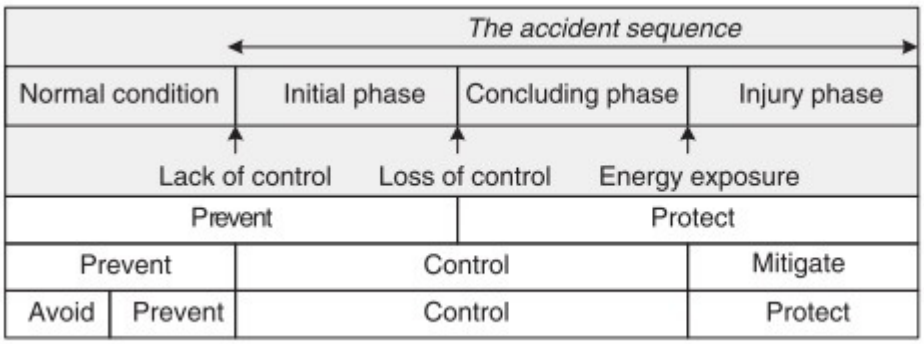


Table 8: Barrier functions (Sklet, 2006, p. 498).

Hollnagel makes a case of not solely introducing barriers in response to known risks of events which have happened (and possibly also had to result in an accident). This reactive behaviour makes safety become a constant battle of catching up with new threats. According to Hollnagel risks should be taken by proactively installing barriers, at possible expense of installing barriers against threats which turn out to never happen. (Hollnagel, 2008, p. 229)

Westrum (Hollnagel, 2008, p. 229) decomposes the threats further in three types of threats: regular threats, irregular threats and unexampled threats (defined as unexampled events). Regular threats occur relative often allowing for a standard response from the system. These threats are imaginable and frequent. Irregular threats occur not often enough to make it practically possible to develop a standard response. These threats are imaginable, yet infrequent. The third type of threats are virtually unimaginable and infrequent. According to Hollnagel (5) the latter two types of threats, both being infrequent, 'cannot be treated in the conventional way' using barriers(Hollnagel, 2008, p. 229). The distinguishing feature appears to their emergence from a condition. Hollnagel suggests to address such conditions, by means of performance variability management and resiliency engineering. These suggestions refer to the systemic type of accidents models. (Hollnagel, 2008, p. 229)

An example of a preventative barrier function is to separate the customer database from the website database, or alternatively not allowing the user account of the website access the customer database. A (ex-ante) mitigating barrier would be to have parts of the customer data stored as a one-way hash (such as passwords). Importantly, at that moment, from the point of view of this report, loss of control is the case and an accident has happened. But the barrier limits the damage to some degree, not all information is available in readable (plain) text to the offender.

Escalation factors
Besides inherent weaknesses of barriers, escalation factors pose an alternative way in which a barriers can fail to stop a threat is by an actual failure of the barrier (CGE, 2013). The contrast with a weakness can be tricky as the end result of an event passing through a weakness (a hole in the barrier) or a failure of the barrier is the same.
An example of an escalation factor is a collision attack on a hashing algorithm. The algorithm is supposed to deliver a unique hash for unique pieces of data. With a collision attack this quality is compromised by being able to alter the data, yet still get the same hash. That way the data (be it a document or a password) could be changed without the owner noticing this on the basis of the hash file. By the time this type of failure becomes *common and under control* this would have to characterised as a weakness of the hashing algorithm.

Incidents versus accidents visualised on the Bow Tie model.

- incident: breaches of the left side bowtie, including the top-event (fault tree)
- accident: breaches of the right side of the bowtie, the top-event excluded (presupposes incidents) (event tree)

A2.2.2 BOW TIE MODELS IN THIS RESEARCH

In this research the base Bow Tie model is considered from a collaborative stance. It is about Bow Tie models representing generic environments of targets, which are the focus in collaborations. And similarly, the infrastructure would in this research be translated to the 'collaboration' side of things. With this alternate view on Bow Tie models suddenly 'the flanks' become of importance. The flanks are the adversary and the target. Normally the target is covered implicitly, it is the organisation which developed the Bow Tie. The adversary (or in general the threat-source) is not identified either, just the corresponding threats to some hazard.

Whereas the Bow Tie model provides a solid and intuitive structure regarding risk management (as in stimulating an identification and omissions in risk related elements), it is rather weak on the helping out on identifying the (required) supporting infrastructure. Versions of the Bow Tie model at least identify such an infrastructure as being a (supposed to be) supportive engineering, maintenance and operation activities. (As displayed on Figure 18 on page 77.) And the collaborative version of the Bow Tie model would also be weak in the (supposed to be supportive) collaborative activities. But in the end it is a simplification of merely a supportive task. For that reason, the supportive activities are merely considered to be of importance, but the how-to will be filled in using the roadmaps and the corresponding theories.

A3 IDENTIFICATION OF ROADMAP-STEPS

The main contributor to the roadmap model is the model by Kowtha et al., which its contents being renamed, restructured, complemented using the Bow Tie model, the roadmaps and personal judgement. More about the method in 3.1, the result in 3.2.

A3.1 METHODOLOGY

In their analyses of collaboration centres Kowtha et al. used a nested approach to consistently characterize collaboration centres. The researchers distinguished a total of seven dimensions, each having multiple factors which help characterize the operation centres. Each of those factors in turn consist of multiple attributes by which a centre can be evaluated. Finally, these attributes are represented by values to quantify or qualitatively attached to describe the attributes.
In this research a similar nested approach was used to structure the roadmap. It resulted in roadmaps, roadmap phases, roadmap steps and finally roadmap step options. Herein the roadmap steps bears closest resemblance to the by Kowtha et al. identified factors. The roadmap step options often come down to the attributes by Kowtha et al.

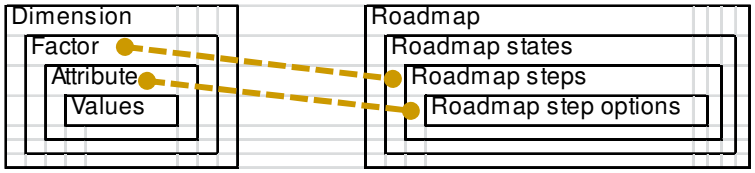


Table 9: The nesting by Kowtha et al. and Spijkervet and closest resemblance of the two (dashed lines)

The in Table 9 illustrated resemblance is a resemblance of meanings of the constructs. There is no direct translation of the two. Kowtha et al. is used to directly or indirectly provide steps and options. But whether, how and when the (typically) factors and attributes of Kowtha returned depended upon the Bow Tie model, the roadmap models and personal evaluation.
Regarding the shared information the Bow Tie model on a collaborative level is used. Although the Bow Tie model is considered to be leading, it is not considered to be perfect. Based on Kowtha et al. an extension is made to the Bow Tie model. For the collaboration related factors and attributes in Kowtha et al. the roadmaps with the corresponding theories were used. Finally, just personal evaluation was used as a last resort for the more troublesome cases.
In all cases the Kowtha et al. model is following. The Bow Tie model is considered to lead as the model by Kowtha et al. is a descriptive model based on how collaboration centres are set-up. The Bow Tie model, abstract as it is, is prescriptive in the sense of providing a way of working and thinking. It remained rather untouched and has a rich history in a vast amount of industries, which suggests it being a simple, yet useful, method. For that reason, in this reason the order suggested by the Bow Tie takes precedence over the Kowtha model. The roadmaps are also leading to further structure the model from a collaborative point of view. With each roadmap (and the corresponding theories) comes another factor of complexity, be it for example the influence of the environment, maturity of the collaboration or maturity participants. Such complexities are all put together on one pile in the model by Kowtha et al. Finally, some factors and elements were not represented in this research after the two preceding activities. Those required more personal judgement than the judgment required to map factors and attributes to the Bow Tie model or the custom nested roadmaps. Some factors and attributes turned out to be out of scope of this research, add a totally different level of analysis or some other reason. Those will cases will be discussed more extensively.

Kowtha			Spijkervet		
Dimension	**Factor**	**Attribute**	**Roadmap**	**State**	**As**
Scope	Impact focus	minimal - catastropic	1	end	risk-value
(OPL	Sector	commercial, federal/civilian, ..	1	end	target
reach)	Influence	local, regional, national, ..	3	gap	influence
	Scale	n.a., small, medium, large, ..	3	gap	scale
Scope	Roles	governance, analysis, collab	1	gap	roles
(function)	Functional abstraction	policy, SA clearinghouse, …	1	gap	topic of shared info
	Type of response	alert information, analysis, …	1	gap	type of response
	Timeliness of response	seconds, minutes, hours, …	1	gap	timeliness
Activities	Protection	preparedness, design, …	2	end	time dimension
	Incident management	detection, response & rec, …			
	Analysis	information extraction, event/..			
Organi	Growth operation center	negligible, minimal, low, …	3	gap	maturity
-zational	Organizational longevity	negligible, minimal, low, …			
dynamics	Organizational change	negligible, minimal, low, …			
	Mission transformation	negligible, minimal, low, …			
	Funding source	none, discontinued, partial, ..			
Facilities	Space size	very small, small, medium, ..			
	Number of desks	small, medium, large, very ..			
	Surge capability	minimal, low, medium, high			
	Center hours	24x7, business hours, event..			
	Layout type	boardroom, mission control, ..			
	COOP scope	none, minimal, some, full			
	COOP readiness	none, days, hours, in real-time			
	Coordination methods	periodic notification, ticket ba..			
Process	Training and certification	initial, managed, defined, …			
-mgmt	Active use of SOPs	initial, managed, defined, …			
	Production	initial, managed, defined, …			
	Analytics	initial, managed, defined, …			
External interactions	Emergency services	customer, supplier, peer, …	3	gap	external interaction
	Government	customer, supplier, peer, …			
	Law enforcement	customer, supplier, peer, …			
	International	customer, supplier, peer, …			
	Commercial	customer, supplier, peer, …			
	Intelligence	customer, supplier, peer, …			
Environment	Visibility	mission, networks, servers, ..	2	gap	environment
	Reach	mission, networks, servers, ..			
	Data handling	mission, networks, servers, ..			
	Capability	mission, networks, servers, ..			
	External stability	mission, networks, servers, ..			
	Community coordination	mission, networks, servers, ..			
			1	end	adversary
			1	gap	collab. structure
Scope	Roles	governance, analysis, collab	1	gap	level of sharing
Scope	Functional abstraction	policy, SA clearinghouse, …	1	gap	method of sharing

Table 10: Identified steps from Kowtha (and redefined using the Bow Tie model where necessary). The gap is short for the 'development of the collaboration model'-state. The name is a references to, at least at first, the gap between the current state and the desired end state, the goal. To close the gap several decisions have to be made.

The solid black rectangles indicate larger deviations than just restructuring and a possible renaming from the original publication by Kowtha et al.:

- of dimension scope, attribute roles, some of the attributes were split of onto step 'level of sharing' and others remained 'roles',
- of dimension scope, attribute functional abstraction, some of the attributes were split of onto step 'method of sharing' and other attributes are represented as 'topic of shared information',
- of dimension activities all factors were put into a new special step activity of the end state, but it has an impact on all other steps, just like the other steps defined in the end state, activities is a dimension in the terminology of this research too.

The motivation for these changes are covered in the discussions of the respective steps.

A3.2 NET RESULT

In the report a brief overview was provided of the steps per roadmap. A more extensive discussion of the different steps is provided in this chapter. It first presents an overview of all identified steps per roadmap(phase), followed by a discussion of the individual states and the respective roadmap steps. Per roadmap steps, where applicable, some possible options are presented. Some of these are (renamed or restructured) versions of the attributes as identified by Kowtha et al. The purpose of the list of options per step is to demonstrate the scope of the step, not to provide an exhaustive list of all possibilities. In some cases more distinct options can be identified, in other cases combinations can be made to create new distinct options.

A3.2.1 FIRST ROADMAP

In the table below the first roadmap is depicted.

Current state ↓	Development of the collaboration model ↓							Desired end state ↓		
formulation goal of col- laboration	roles	collabo- ration structure	topic of shared information	level of information sharing	method of sharing information	type of response	timeliness of response	target	risk-value	adversary

Table 11: The first roadmap

A3.2.1.1 END

The end state is shaped by the target, risk-value and the adversary. Those three are considered to be the minimum required steps which have to be defined to enclose specific Bow Tie models. The adversary and the target are selected to discuss threats with some intention on some specific targets. With the risk-value specific cases can be selected, to not have to focus on all enclosed threats. Organisations face all kinds of threats, but some are more likely than others, others are likely but result in no negative consequences. Such threats could be ignored to not overload the collaboration with these less important issues.

With the selected end state steps plenty of options remain. In the discussion one is not limited to discussions of *the* threat of a certain risk. The threat represents a path in the Bow Tie model. The discussion could also be about a part of that threat, any aspect which is part of the Bow Tie model. This part has a risk, a risk which contributes to the overall risk of the threat. For example, the discussion could be about the likelihood of some line of defence mechanism failing to protect against some threat.

An alternative approach in defining the end state would be to solely pursue collaboration on large risks (in general defined as likelihood times impact) regardless of adversary or target. However, for that purpose the collaboration would have to be rather large, covering many target sectors, scenarios and types of threats. The other way around, by focusing on specific aspects of the risk value the discussion would be more focused. In a sense this comes down to defining collaboration on specific portions of the actual Bow-Tie. Although this would focus the discussion (such as by solely discussing attacks using exploits of Oracle's 'Java'), the list of potential participants will be rather opaque and the entire collaboration purely reactive. Java is still used by many, many organisations, so who should be part of collaborations. Furthermore, collaboration would be on detecting new

exploitations of unknown vulnerabilities in Java. A reactive mode. An alternative would be to start a large discussion discussing on how to be less dependent on Java. But such a collaboration is not the focus of this research. Such a discussion is a discussion on a 'known', the fact that Java is in need for frequent patches.

A3.2.1.1.1 TARGET

Targets are the recipients of an attack, with target in this case referring to some type of organisation. Following from the problem description attacks are increasingly targeted. With that, there are less organisations confronted with the same attack. And thus, for the most advanced attacks to be detected, all of the similar organisations would be considered *the target*.

There is plethora of target options to focus on. The first high level distinction are options like commercial parties, governmental organisations or research laboratories. And from these more specific, yet open specified, targets could be defined. The target might be defined as (public and/or private) organisations from the critical infrastructures, and where applicable just the critical infrastructure part of that organisation. Some telecom provider might be identified as being a critical provider. However, not all of its activities are about the critical infrastructure.

Following from the assumptions one specific target has to be defined. This can be a rather high level target, as governmental organisation at nation state level, or commercial banks. What matters is to select a single target and to consider all organisations falling under the distinction.

In practice

In current collaborations there are some mixed, high level decisions made regarding the targets. In the United States and the Netherlands the targets of a collaboration are organised per sector. And per sector the level of involvement varies. In the United States, the financial sectors involve a vast amount of parties, whereas the chemical sector has more involvement of the larger organisations. In contrast to all this, in the United Kingdom there is no clear target defined, neither regarding sector, nor regarding size. A pilot project, Auburn, ultimately was a collaboration of 160 companies across five sectors. And its successor, CISP (Cyber security Information Sharing Partnership) which started in March of 2013 followed suit, opening up to larger organisations at first.

Kowtha referred to the target as the sector and described it as "the specialization or primary focus of the operations centre's mission". Possible attribute values are commercial, federal/civilian, defence/intelligence, state & local and other. In this research the step is more openly specified. The sector is just one of the possibilities to focus on. Alternatives are to focus on specific critical infrastructures, organisations of a certain size or with a specific mission which can attract more advanced attacks.

A3.2.1.1.2 RISK-VALUE

The final decision variable to *potentially* decide upon is the portion of the attacks (or components thereof) participants will consider. An all-encompassing method would be to use the risk value of an attack. The risk value is commonly defined as the probability of an event taking place times its consequence. Alternatives are to select on the basis of the probability of attacks or to focus on the impact of the attacks. Participants of the collaboration would have to agree on the selection criterion (risk, probability or impact) and define a bandwidth of the considered attacks. Thwarting of attacks which meet a minimum risk level will be discussed in the collaboration. But possibly the participants also determined a maximum risk level. The motivation for that could be that certain attacks require involvement of different parties, a situation of escalation.

The definition of a risk value in a collaboration is rather difficult because the probability is difficult to assess with certainty and the impact even more so. This applies to risk management in general, but in collaborations in particular. The troubling factors can be distinguished in objective and subjective factors. Of the objective type there is the issue of differences of possibility and possible impact between organisations. Organisations do not have identical environments containing identical valued resources and loss is also has a different value. Even if a software organisation and a grocery store are both confronted with a data leak, spilling their entire, yet equal sized customer database, the

impact will in all likelihood be higher to the software organisation. The result is a difference in the *actual* risk values per organisation for an *identical* attack. On top of the objective part, comes the subjective. This makes a consistent risk analysis difficult. The participants might experience the resulting risk value differently. This payoff could be by not introducing the better protection method, which saves money, and accepting some cases of (unresolved) fraud[20]. And in the end, systems have vulnerabilities, what matters is whether they provide a positive return. Although online banking introduces new types of fraud, it reduces costs in other respects (postage and processing). With all these difficulties coarse estimates of levels risks are probably the most attainable meaningful construct of selecting types of attacks to consider. An example is the use of a five point scales ranging from minimal to catastrophic levels, as suggested by Kowtha.
However, with impact being particularly difficult to assess and compare, possibly probability is the highest attainable selection criterion, at least on a collaborative level.

In practice
The use of the ***impact*** of the risk value is used by many. Zhao and White, for example, defined four 'threat' alert levels (guarded, elevated, substantial and severe). Their 'threat' alert levels are the result of the incident impact, scope and severity to a community. The authors consider their threat alert level to be the counterpart of the National Cyber Risk Alert Level for the National Cyber Incident Response Plan and Cyber Alert Level for the Multi-State Information Sharing & Analysis Centre. (Zhao and White, 2012, p. 460).
In their security report of 2013, NCSC focuses on the impact, in a reactive manner, defined as 'relevance'. Relevance is attaches to the relationship of threats, adversaries and targets. (NCSC, 2013, p. 113) The relevance is a three point ordinal (low-high) scale which is the result of detections of threats, presence or absence of relevant barriers and the manifestation of incidents (NCSC, 2013, p. 9).
Kowtha distinguished the factor impact focus, and characterized it as "the types of incidents that an organisation's mission focus". The corresponding possible attribute values for the impact focus are minimal, moderate, significant, critical and catastrophic.

A3.2.1.1.3 ADVERSARY
The adversary is not identified by Kowtha et al. in some respect in the characterisation of (the focus of) collaboration centres. The adversary is added based on insights from the Bow Tie on a collaborative level, as discussed in A2.2.2.
Many distinctions of adversaries are possible, such as the coarse distinction in criminals, hacktivists and governments (Hyppönen, 2012, 3m40s-4m20s)(Schneier, 2013). More extensive is the distinction by NCSC in nation states, terrorists, criminals, cyber vandals & scriptkiddies, hacktivists, internal actors, cyber researchers, private organisations (NCSC, 2013, p. 9). The point of defining an adversary is to consider the motivation and intentions of the type of adversary. It is not about actual attribution as that is, at the moment of an attack, of subordinate importance. Whereas criminals try to make money in some way, hacktivists try to protest against or embarrass their target, and finally governments try to acquire information (NCSC, 2013, p. 9) or sabotage activities. Hereto information can be a mean to a launch of another attack later on. An of this is the assumed relationship of Flame as the information acquiring piece of malware for the purpose of the launch of the highly targeted disrupting piece of malware called Stuxnet. (Hyppönen, 2012, 3m40s-4m20s).
The practical implementation of identification of the (type of) adversary behind an attack proved troublesome. Not only is it hard to pinpoint a perpetrator, it is quite possible it is not the actual perpetrator. Symantec already discussed this in their security report stating the techniques used in

[20] An example of such a trade-off is the case of the public transport chipcard in the Netherlands, at least after its introduction. The card uses a Mifare Classic chip, which can easily be tempered with to adjust the balance. Switching to a different, more secure card would cost quite a bit of money. Instead Trans Link Systems sticks with the card and focussed for the time being on (automatic) detection of cases of fraud followed by a block of the card.

the attack might obfuscate the source. Highly advanced techniques could be developed by state funded parties, yet used in some form by organised crime. (Symantec, 2013b, pp. 19–20)
In the end, the determination of an adversary is a simplified proxy to the thwarting of the accomplishment of the intentions of the adversary. Whereas the adversary is more encompassing and straightforward to discuss, being the counterpart of the target (/defender), the practical implementation comes down to determination of motivations and intentions. The determination of the adversary is thus more about determining which attacks coming from some motivated party having some intention, will be considered in the collaboration.

In practice
As mentioned, the identifications and distinctions of the adversaries are plentiful. Yet, the topic appears to be more about the intentions and (the difficulty of) attribution (NCSC (NCSC, 2013, pp. 21–26), Symantec (Symantec, 2013b, p. 19)) than actually (suggesting) on acting upon the insights. NCSC discussed limitation of vulnerabilities and defensibility of the target, yet there is no direct link of intentions mentioned (NCSC, 2013, pp. 31–42). Similarly, Symantec stresses overall protection by improving defensibility (Symantec, 2013b, p. 22). Similarly, representing actual collaborations, Kowtha et al. distinguished the impact focus, yet there is no mentioning of focus on the adversary (Kowtha et al., 2012).

A3.2.1.2 DEVELOPMENT OF THE COLLABORATION MODEL
The 'development of the collaboration model' is all about compositing the mean to the end of thwarting the attacks as enclosed in the end state. The main construct for that is improvement of situation awareness of participants in a collaboration by means of information sharing arrangements between the participants. The individual steps to improve situation awareness, the source for the steps and the motivation for inclusion are displayed in Figure 19.

step	source	rationale
roles	Kowtha, DSA	compatibility and complementarity of participants to improve SA
structure	DSA	supportive structure to productively exchange information
level of	Personal	level of information exchanges
type of	Kowtha, Bow Tie	the topic of considered SA (and of exchanges)
response	kowtha	information exchanges need some impact
timeliness	kowtha	SA has to reach level three in time, prior to the future

Figure 19: The identified steps for the base roadmap with its source and the motivation for the usefulness of the step.

Following from the theory on situation awareness in a collaborative environment the importance of supportive roles and structures is identified. The actual shared information is another consideration, being the topic of shared information and the level of information. The topic is covered by Kowtha and the Bow Tie model. The level of information sharing is a personal evaluation to clarify the ambiguity of the term information. Finally, the type of response (the intended impact) and the timeliness are provided by Kowtha. Those two steps are more like the conditions of situation awareness. Not having timely exchanges or without sufficient impact the recipient will not be helped enough for performance to flourish on the basis of situation awareness.

A3.2.1.2.1 ROLES
The roles are about selection of participants which fulfil specific roles in a collaboration in such a way to end up with compatible and complementary participants. Compatibility refers to the compatibility of situation awareness of participants. They have to be able to work with each other. Given the inherent differences of participants (even those with identical backgrounds due to personal history) they will have slightly different perspectives on topics. That way those participants might be able to improve each other's situation awareness. Complementarity is added, on the basis of Kowtha, to stress the importance of differences of perspective for the sake of situation awareness. Having all

banking firms send representatives to a collaboration will result in great compatibility, yet possibly still limited complementarity. Those representatives would also have to perform the analyses role to not just present issues, without chance of solutions or even explanations.
In light of the assumption of a single option being allowed to be selected, in this case it refers to a single non-conflicting option per participant. Depending on the topic some specific roles might be required, such as analysis, providing the material for analyses or the one acting upon the analyses. The purpose of the assumption was to satisfy a clear responsibilities and to avoid possible conflicts of uniting counterparts in one role.
The importance of such an assumption at the extreme was recently brought forward by Jacob Appelbaum (one of the developers of the Tor Project) to the European Parliament at the Privacy Platform on the 15th of October. Appelbaum questioned the role of defensive organisations with an offensive mission. Those organisations would receive or discover information on new (potentially zero day) vulnerabilities and they could use the information for their offensive mission. The net result would be an incentive to never report such vulnerabilities. (Jacob Appelbaum, 2013, pt. 7:09–7:45)

In practice
Kowtha identified some roles organisations can fulfil to use and defend an environment. These are governance, analysis, collaboration & information sharing, incident management and protection. Given the scope of this research collaboration & information sharing is not really optional. Incident management and protection are not a complete representation. The extended Bow Tie model on a collaborative provides these and additional roles in an implicit manner. Incident management are loosely related to all roles dealing with active attacks which deal with the left side of the Bow Tie model. Protection is rather ambiguous and could cover all roles dealing with making sure attacks will at a maximum result in incidents, not accidents. It could also be about the roles which are about the barriers. For these reasons, the Bow Tie model is used as a frame of inspiration on roles which could be present in a collaboration. Which roles to choose depends on the goal of the collaboration, which can just be about avoiding accidents, as opposed to avoiding incidents.

A3.2.1.2.2 COLLABORATION STRUCTURE
The collaboration structure is about how participants can inform each other. There are actually multiple ways in which organisations can be arranged and the structure turns out to have an impact on the performance.

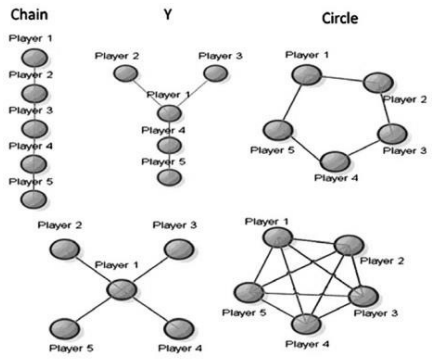


Figure 20: Some of the options for structures of collaboration (Sorensen and Stanton, 2013)

Sorensen and Stanton tested the presumed impact of organisational structure on performance by means of a simulation of a realistic command training. For that they compared performance of teams which were organised in five different organisational structures: Chain, Y, Circle, Wheel and All-connected (illustrated in Figure 20). It turned out there is indeed a discernible difference in the performance of teams, in terms of speed and accuracy of the teams, depending on the organisational structures. (Sorensen and Stanton, 2013, p. 77) They found the 'Y-structure' to perform better than the other organisational structures in their simulation (Sorensen and Stanton, 2013).
However, although the Y-structure performed best in *that* simulation the authors researchers said literature suggests there is no single structure optimal for all conditions. And the authors themselves even suggest that "teams may benefit from working in more than one organisational structure" depending on the "classes of tasks" at hand (Sorensen and Stanton, 2013, p. 78).

In practice

In cyber security collaborations there appear to be different collaborations, such as the wheel and the all-connected structure. The wheel approach is used with the Dutch initiative AbuseHUB (as implied with second part of the name). AbuseHUB is an initiative in which information on botnet infections is gathered and analyses centrally. The result of the analyses is sent to the connected Internet Service Providers, where relevant. Similarly, the Dutch Cyber Security Centre (NCSC) informs (or at least informed) all relevant parties themselves if they know those organisations have been affected by some attack. A recent example was the notification of organisations by the NCSC on computers which were part of the Pobelka-botnet. For that NCSC compared a list of IP addresses from which computers were connected to the Pobelka-botnet with the list of IP addresses which in possession of organisations the NCSC protects. In the United States there is a contrast to be noted, with banking and finance most heavily relying on the Financial Services Information Sharing and Analysis Center (FS-ISAC) (National Infrastructure Advisory Council, 2012, p. 37). The FS-ISAC has a *central* role in intra-sectorial information sharing (National Infrastructure Advisory Council, 2012, p. B–16), suggesting yet another wheel structure. In contrast, the chemical sector relies more on personal relationships, in small circles, and a variety of networks. The security managers the National Infrastructure Advisory Council interviewed "rarely rely on a single mechanism for receiving information on threats, intelligence, and security trends" (National Infrastructure Advisory Council, 2012, pp. C–6). A motivation for not widely sharing vulnerabilities is the due to the risk of exploitations of recently discovered zero day vulnerabilities. The type of vulnerabilities for which no solution is readily available. (National Infrastructure Advisory Council, 2012, pp. C–6) The consequence of all the scattering of information which has to be acquired using (personal) networks is the difficulty of smaller organisations to acquire this information (National Infrastructure Advisory Council, 2012, pp. C–6). Based on the preceding it appears the chemical sector primarily relies on multiple, smaller all-connected networks.

The organisational structure is not covered by Kowtha. The factors which are most akin to the organisational structure are the coordination method or the layout type. However, the possible attributes for the coordination method are more related to technology (ticket based, video based, or web based coordination). The layout type is about the physical configuration of furniture, equipment and staff in the operations centre. An organizational chart, let alone a depiction of the actual structure, is missing.

A3.2.1.2.3 LEVEL OF SHARING

Although many collaboration programs are about the sharing of information, it is often undefined and possibly an ambiguous term. Undefined refers to the topic and is covered by the step 'topic of information sharing' in A3.2.1.2.4. The level of sharing deals with the troublesome ambiguity, covered in this section.

Information can indeed refer to information, but also to data, knowledge or some other level of information. Quite often these three (or variations thereof) are put in some relation to each other. In this research the specific relation (being hierarchical or not) is not really of interest, neither is the value *typically assumed* of each type of data. (Fricke, 2008) In this research it is solely and simply about the distinction in types of data, in terms of rawness or richness of the data. Raw data is about unprocessed, to say untouched, log files. This data can redacted to remove logins or such, that would make it data. Data can also be interpreted, processed and redacted, with which it is commonly referred to as information.

The distinction is of importance because different types of data have different characteristics in terms of costs, volume, usefulness, applicability and required trust, just to name a few partly overlapping characteristics. Actual data can refer to log files (anonymized or not), which potentially contains everything ranging from not applicable to applicable to some situation. In case an organisation notices 'odd' behaviour at their systems a log file might be helpful. Information is about a subset of the data which appears to be relevant to explain the behaviour. And knowledge is the transcending option, knowing why it behaved odd (and actually, that feeling of 'odd' behaviour is

based on knowledge). Although using log files might be the most effective way of solving issues, it can potentially also take far more time to analyse. This is especially the case if the analysis thereof is not standardised in some way. Additionally, it might also present a challenge to the organisation providing the log file. The log file might contain confidential data. Removing those will cost resources, leaving it in requires trust in the recipient of the log file.

In practice

To increase the efficiency of the exchanges of *technology related* information different techniques are in development, such as STIX and CybOX (ENISA, 2013, pp. 18–21). STIX is about structuring threat related information in a standardized manner, one which is readable by computers and machines. And CybOX is about standardizing the representation of events and properties in organisations' environments. But this still leaves the less to non-technological related exchanges of data, information and knowledge wide open.

A3.2.1.2.4 TOPIC OF SHARED INFORMATION

Figure 21: The orange line splits itself somewhere, indicating the first part of two different attacks is identical, yet the remainder differs (some different barriers) with ultimately a total different asset being hit (of a different bow-tie). The green part is monitored part for informing-others purposes. As soon as some successive barriers are passed by the attack (inside of the green

Aside from the richness of the data which will be shared the actual topic is another thing to decide upon. One could think of warning others about a (successful) attack which took place or inform about individual elements of the bow-tie. The start of this Bow Tie is the adversary and the end the target. Some high level, early-on type of sharing could be about sharing possible activity using, for instance, attack trees. An attack tree is all about identifying potential target goals and identifying the path to these targets. This will, hopefully, result in multiple steps to be taken (by the attacker) to reach this goal. Organizations could monitor for activity on parts of their system corresponding to each of the steps. In case some 'step' or 'steps' are triggered the organisation could inform the others of the attack. The level of detail can vary, such as some vague characterisation of the adversary, the used techniques and the possible intentions. Importantly, this still leaves a degree of freedom as to at what level the organisations will warn the others. A triggered barrier (leaving many target options), series of triggered barriers (more clear target of the attacker) or just in case the actual target is undoubtedly clear and the adversary is 'known'. In the latter case one could think of informing ISPs if the adversaries are in some way grouped.

In practice

What in this research was dubbed as the topic of shared information is referred to as functional abstraction by Kowtha. The functional abstraction is considered to "capture the purview of the center within its operational scope". And it consists of five attributes, as displayed in Table 12.

Kowtha	**Spijkervet**
Policy	Paths in Bow Tie (what-if…analyses)
Command & control	Active recovery to left state Bow Tie
Continuity of operations	Introduction of barriers & recovery measures
SA clearinghouse	Addition of newly *identified* escalation factors & vulnerabilities
Asset protection	Protection- and reporting- activities of asset protection

Table 12: Mapping of the identified 'steps' by Kowtha to the options for this report

The attributes are considered to come down to aspects of the (inside of) the Bow Tie model.

A3.2.1.2.5 METHOD OF SHARING

The method of sharing is about the interactions between organisations, specifically about the purpose of the sharing and the method thereto. It is, like the entire research, about what medium to use. The entire research is about improving the capabilities of organisations by means of information

sharing. The underlying concept is the presumed causal relationship of situation awareness, resulting in better decision making opportunities, which result in the potential for better performance (represented by fewer incidents and accidents). Importantly, the two presumed causal relationships appear to never have been satisfactory demonstrated. Just correlations were demonstrated and the entire relationship is considered plausible. This section is all about Situation Awareness and schools thereof, which is the main method considered of this research to improve performance.

Situation awareness of the individual

Situation Awareness (SA) is described by Endsley (M.R. Endsley, 1995, p. 36) as a state of knowledge about the situation in the considered environment. She separates it from the process leading up to SA, which is defined as situation assessment. The decision making processes on the basis of SA is not part of SA. The importance of Situation Awareness is stressed by some scholars as they link high levels of SA to high levels of organisational performance (Sorensen and Stanton, 2012).

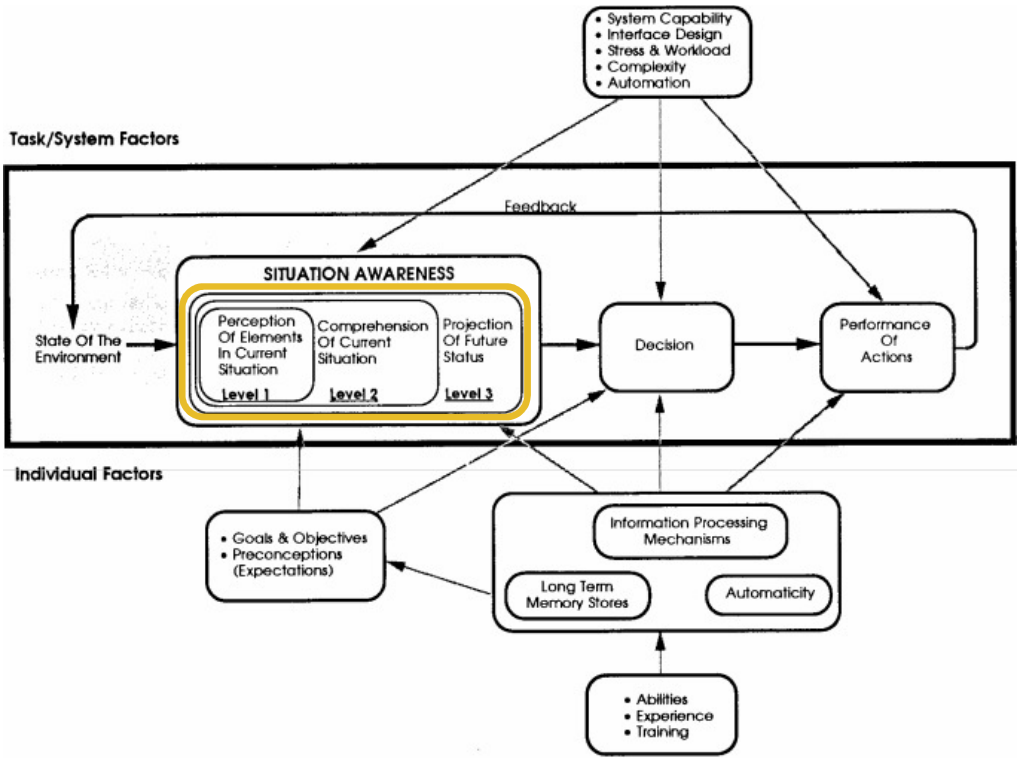


Figure 22: Model of situation awareness in dynamic decision making [26, p. 35] (orange overlay added by author). The orange overlay demarcates the three levels of situation awareness. Herein the higher level of situation awareness presupposes the lower level of situation awareness. Hence, for example, reaching level two awareness (understanding the current situation) is impossible without level one awareness (knowing the elements of the current situation).

According to Endsley (M. R. Endsley, 1995, p. 36), as illustrated in orange in Figure 22, SA can reach three levels: perception, comprehension and projection. Whereas perception is about awareness of the presence of elements (e.g. amount of traffic to a specific server on specific ports), comprehension is about attaching meaning to the values of those elements (e.g. unusual amount of traffic from specific locations). The final level, projection, is about being able to understand the future status (e.g. understanding it will saturate the amount of resources and render the service unavailable). SA in all this refers to "the perception of the elements in the environment within a volume of time and space, comprehension of their meaning, and the projection of their status in the near future" (Seppänen et al., 2013).

The level of situation awareness of the agent affects the type of information an agent can share. An agent at level three *on something* is theoretically able to share information or knowledge with somebody else. But still, even if an agent (possibly the same one) fails at some other aspect to reach the second level, sharing the things the agent knows based on its level one situation awareness

might be of value. This would, in a way, mimic how antivirus software, in part, works. It detects some behaviour which is similar to what somebody else witnessed and flags the behaviour as being suspicious.

Situation Awareness at the collaborative level

In an inter-organisational setting this concept of SA gets a new dimension as in all likelihood, at times, information has to be shared to improve the overall SA. Endsley named this overall SA 'Team Situation Awareness' (Team SA). According to Endsley it is the degree to which every team member possesses *their* SA required for *their own* responsibilities (M.R. Endsley, 1995, p. 39). Similarly, Salas et al. (Salas et al., 1995, p. 131) defines Team SA as 'The shared understanding of a situation among team members at one point in time'.

The fact that some partner organisation has the right piece of information available is not sufficient. The timely delivery of the right information to the people of organisations that need that information is critical to be able to reach the third level of SA: projection of the future status (Salmon et al., 2009, pp. 144–145).

Of SA there are, according to Stanton et al. three schools of thought (Stanton et al., 2010, p. 30)[21], with SA being defined as:

- a psychological phenomenon experienced in the mind of the individual person (represented by Fracker, Sarter & Woods, and the aforementioned Endsley),
- a phenomenon situated in the world, by means of displays and all (represented by Ackerman),
- an emergent property that arises from the interaction between people and their environment (represented by Stanton et al.).

The first school is referred to as the cognitive- or psychological- approach and the second school as the engineering approach. Finally the third is referred to as the system ergonomics approach (Stanton et al., 2010, p. 30). According to Walker et al. the latter is consistent with human factors and a socio-technical viewpoint (Stanton et al., 2010, p. 30).

The method of actually improving 'Team SA' varies per school. Endsley, a proponent of the cognitive approach introduced Shared Situation Awareness (SSA). She defines it as "the degree to which team members have the *same SA* on *shared SA requirements*" (Endsley and Jones, 1997, p. 54). SSA presupposes shared requirements and purposes between team members on some situation. Only in case team members have a similar understanding of a situation can they meaningfully share information. (Stanton et al., 2006).

The system ergonomics approach, represented by Distributed Situation Awareness (DSA), does not presume *shared* -requirements and –purposes of participants of a collaboration or team. On the contrary, they *might* be different yet compatible(Salmon et al., 2009, p. 191). Individual team members have their own *unique* SA. Together the SA of individual team members and *other systems* makes up the total of DSA. (Salmon et al., 2009, p. 59) Crucially, the system ergonomics approach actually leverages these different perspectives and unique SA of involved parties. (Stanton et al., 2006, p. 1291) Parties[22] do not share their SA, yet are involved in SA transactions. Parties, with and using their own SA, can give other parties information about a current status of something (Salmon et al., 2009, p. 193). Actually sharing SA, as suggested with SSA is considered impossible as the SA of an individual is the result of their unique position, in part as a result of their own personal experience. Despite being unique, the SA is still considered compatible as all SA is 'collectively required for the system to perform collaborative tasks successfully'. (Salmon et al., 2009, p. 190). With DSA the links between nodes (parties) are actually more important than the nodes themselves (Stanton et al., 2006, p. 1308). The level of DSA depends on the effectiveness of finding (or being

[21] All the identifications of persons representing the schools of thought are also presented by Stanton et al. (2010, p. 30),

[22] Stanton refers to the parties as 'agents', with an 'agent' being a human or an artefact.

found by) the right parties which can help interpret information in a meaningful way. Knowing what other parties have contained in the system is called meta SA (Salmon et al., 2009, p. 58).
In the end, what one would have to decide upon is what levels of situation awareness are required in the collaboration and how the agents in the collaboration will work together. Although agents with level one of situation awareness could be of value it could also overload the collaboration. But if agents are only able to share information if they have level three awareness might result in quite a bit of incidents, possibly accidents.
And the actual method of sharing information amongst agents, be it Shared Situation Awareness (SSA) or Distributed Situation Awareness (DSA), has its effect on the productivity. SSA will result in high quality assessments, whereas with DSA there will be more assessments.

Measuring (Distributed) Situation Awareness
Measuring the level of SA in is difficult, and team SA in particular. Although there are over thirty techniques (Salmon et al., 2009, p. 39) available to assess the level of SA:

- most measuring tools lack validation, with the exceptions of SAGAT and SART (Salmon et al., 2009, p. 56), and
- the most commonly used tool is SAGAT (Salmon et al., 2009, p. 56), however it 'may be less sensitive when applied to assess team SA' (Sorensen and Stanton, 2011, p. 685),
- of the by Salmon et al. identified thirty measuring techniques only four are capable of measuring Team SA (one of those can solely be used to measure team SA) (Salmon et al., 2009, pp. 52–54).

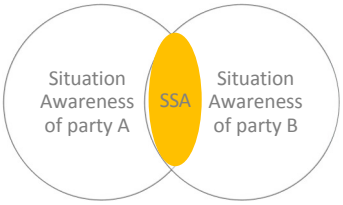


Figure 23: Team Situation Awareness, with the orange portion, Shared Situation Awareness (SSA), representing the portion that causes trouble with measuring Team SA. This part will, in this case, be counted twice. [23]

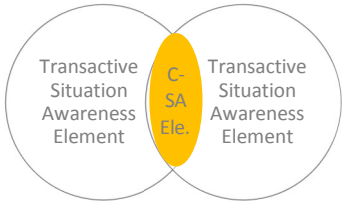


Figure 24: Situation Awareness according to Distributed Situation Awareness [24], the Compatible-Situation Awareness (C-SA) is still the 'overlap' as with SSA. However, with C-SA there are actually two versions. Both parties have the same information, but they look at it differently (but in a compatible way).

Measuring Team SA is even more difficult as it is not simply the sum of SA of individual parties. Some SA is useful and applicable to multiple parties. That portion, the Shared Situation Awareness, directly affects the level of Team SA. This is visualised in Figure 23. But not only does presence of SSA increase the value of Team SA upwards, it is also capriciousness. Depending on who enters or leaves, specifically how much SSA they *ultimately* introduce, the level of Team SA will vary. At the moment of entrance the SSA is not immediately increased, but over time Situation Awareness will be shared, resulting in more SSA. But the maximum amount of SSA depends on how much the newcomer differs from any of the incumbents. Importantly, up to this point it was not discussed to what extent the participants actually have the to them required Situation Awareness. Team SA is perfect if all

[23] This is actually a simplification. As extensively discussed in section 6.5 there are multiple ways of looking at SA. For starters with organisations need and what they think they need.
[24] Meta Situation Awareness ('knowing who knows what and knowing who has to know what') will typically not be a form of Situation Awareness that is of interest. However, depending on the situation it might actually be a form of situation of interest. Think of cyber security centres that have the intention of merely bringing collaborations in touch with each other. In such a case suddenly Meta SA is a form of SA of interest. Multiple security centres might be aware that some collaboration discusses some situation. Their Meta SA is thus compatible and of interest. But again, typically Meta SA is considered by the author to not be part of Transactive or Compatible SA, it plays out on a different level in such cases. Because regardless, Meta SA is still important with DSA.

participants have the SA they require to execute their task. And with that the required SA of a party could in that case be less than the SA the party actually has. Measuring is also troublesome with DSA. The main concept of DSA is the compatibility of SA. Individuals might even have the same information, but their resultant awareness is different. However, it is compatible in that it is required for the success of the collaboration. (Salmon et al., 2009, p. 190). With that definition, Team SA would come down to the summation of all Compatible Situation Awareness. But to actually measure the levels of compatible Situation Awareness is troublesome. Measuring the flow of information as a proxy will skew the results as some information is just an indication of Transactive SA. It is the piece SA of individual which forms the basis for a piece of information which is exchanged in an SA transaction with another individual. With that information the other is supposed to be able to improve its own CSA. (Salmon et al., 2009, pp. 192–193)
Reverting to the performance[25] of the individual or team straight away is undesirable due to the 'unstable relationship' between the level of SA of a person and the task performance of that person (Salmon et al., 2009, p. 49). According to Orsanu (Sorensen and Stanton, 2013, p. 72) a positive correlation was found between information exchange in teams and levels of SA. The author also discovered the level of SA has a positive correlation with performance of teams. And according to Sorensen and Stanton (Sorensen and Stanton, 2013, p. 73) these findings are in line findings by Cooke et al. and Endsley. But in the end, good SA can (solely) 'be viewed as a factor that will increase the probability of good performance but *cannot* necessarily *guarantee* it' (M.R. Endsley, 1995, p. 40). The level of SA, decision making and performance are distinct elements, influenced by different factors which are assumed to have some underlying, yet unknown causation. As a result, at for the time being, the three elements are considered to require separate attention. (M.R. Endsley, 1995, p. 36) And given the line of reasoning, the situation awareness has to be measured as that is considered to improve performance. The performance of an individual or team can also improve without a change in situation awareness (better decisions). And quite possibly the performance might improve due to a limited situation awareness, yet awareness of that fact. Reliance on less, but high quality information might be more valuable than more information, which turns out to be outdated or slightly off.

Improvement of the design to improve situation awareness
Developing a system to improve the situation awareness will take multiple iterations. It will always be a matter of some form of cyclic development on:

- what has to be known,
- designing a system for exchanges of information,
- measuring the results or the impact to the level of situation awareness, and
- improving the method in some way to further improve the situation awareness.

Based on the results it might be necessary to revise the requirements or to change the system that is used to exchange information. (Salmon et al., 2009, p. 217).

In practice
Situation Awareness is most elaborated upon in other sectors, such as aviation and the army. But despite the rich history, only the individual Situation Awareness is rather mature. There are some rather commonly, universally accepted definitions and a plethora of tools to measure the level of awareness in specific environments and conditions. At the level of collaborations measuring the level of Team Situation Awareness lacks solid methods, even in (and applicable to) sectors which relatively intensively worked with the concept. In the cyber security domain Situation Awareness is more of

[25] And in the domain of cyber security (or similar domains of protecting against something which might not happen or so) measuring performance is yet another issue.
For example: If an organisation informs another organisation on how to defend against some attack which does also take place in a later stage and got successfully defended against using the information could be called a success. But what if such an attack does not take place? And what if the actual attack is slightly different and as such does result in some damage or so. Is that bad performance?

loosely defined concept along the likes of knowing what goes on. Operationalization and practical implementations are limited. And typically the cognitive approach (specifically the implementation by Endsley) is considered, not the Distributed Situation Awareness approach, despite it appearing to be a better match in an environment with more heterogeneous organisations. This is particularly an issue in larger collaborations, in which organisations from different sectors have to collaborate.
Kowtha did not focus on information sharing collaborations, but more about cyber security collaborations in general. As a result, Situation Awareness is just an attribute of the functional abstraction. From the point of view of this research such a stance is far too limited, as SA touches upon the capabilities of participants. Furthermore, the corresponding values of SA are limited to the reception, processing and dissemination of the (cyber) situational awareness information. This is a rather abstract representation of what could be about anything.

A3.2.1.2.6 TYPE OF RESPONSE

In this research the purpose of information sharing is to enhance situation awareness of participants. But that is still a little ill definition of what can be expected of the collaboration. The purpose can just be about making sure participants have an improved situation awareness. But it could also be less ambitious by not attempting to improve the situation awareness of all participants. Instead, the participant could share its current position, based on non-complete situation awareness in a context, and get input from others on how to proceed. Strictly speaking the situation awareness is not improved. It is about knowing what to do by getting some type of input of somebody who has improved its situation awareness beforehand. A motivation and example for the latter would be an organisation which is confronted with some consequences, but is not able to determine why this was possible. A potential cause is an unknown (or unpatched) vulnerability in an application. By sharing data (such as log files) another organisation might be able to help. The initial piece of information is thus about improving the awareness of another organisation of some issue (possibly also affecting that organisation). That organisation might discover the source of the consequence, such a previously unknown or unpatched vulnerability. Possibly that organisation can even point at an available patch. With all that, the awareness of the first organisation is not necessarily improved (in the example there is no exchange of how the vulnerability was detected), just information on what to decide upon or even do. Although actually improving the situation awareness of the first organisation might be preferable, this will not always be an option. Especially with more sophisticated attacks such activities might not stretch much further than detecting the same type of attack. But the value is in making other organisations aware something happened, or something is off. In the end, the consequence of more targeted attacks is that security organisations will need help to know what happens.

In practice

The type of response is one of the factors by which Kowtha et al characterise the scope of operations centres. They distinguished some possibilities of types of response, being: direction, requested action, tailored information product, analysis & recommendation, general alert & information. These options come down to what happens after information exchanges, after the situation awareness of organisations is improved. The result could be to notify all participants of some threats. But it could also be about requesting participants to perform some action, such as to replace some type of software. A more specific, dictating option is direction, defining courses of action and timelines for those actions.
An example of a research project which focuses on a star/hub-like network with a centre providing advisory on the basis of relatively sophisticated situation awareness is CAIS. CAIS is an Austrian project focused on protection of critical infrastructures against cyber-attacks (Skopik et al., 2012). Its architecture has a National Cyber Defence Centre as a hub which received input from all kinds of organisations in the nation (Skopik et al., 2012, p. 281). The intention is to achieve situational awareness at (central/)nation level of the state of cyber security of critical infrastructures (Skopik et al., 2012, pp. 281–282). The centre analyses the retrieved, detected anomalies, and captures it in

simulation models to detect potential cascading effects and rolling breakdowns. On the basis of the insights those studies provide national, strategic decision making is considered to be possible. (Skopik et al., 2012, p. 284) Additionally, the centre can offer feedback and advice on the basis of the provided information (Skopik et al., 2012, p. 283). With all that, CAIS is about improving situation awareness of the centre and allowing it to give organisations better advice. It is not the main intention of CAIS to also improve the situation awareness of organisations themselves.

A3.2.1.2.7 TIMELINESS OF RESPONSE

Finally the timeliness of the response has to be decided upon and has to backed by the collaboration, which has to be in line with the intended type of response. An ambitious response, such as a recommendation, will require more active involvement of relevant roles than purely informing everybody on threats. What the response will be depends on the collaboration in question and was covered in A3.2.1.2.6.

In practice

Kowtha et al. define the timeliness of the response as a the time interval it takes for the centre to perform its task. And the options range from seconds to months. With that Kowtha et al. use timeliness of response as an 'output', not an ambition. Instead, in this research the timeliness of the response is considered to be another decision option. In the end, the goal is to thwart some type of attack (if the collaboration focuses on current attacks). Having a timeliness which is too long results in more cases of attacks, meaning the intention of the centre fails. In the end, that is what is supposed to matter to the collaboration, whether by means of the collaboration attacks can be stopped. That requires an assessment of required typical response time, and that has to be backed up by the setting of the collaboration (roles, ambition and structure).

A3.2.2 SECOND ROADMAP

In the table below the second roadmap is depicted.

Current state ↓	Development of the collaboration model ↓							Desired end state ↓		
formulation goal of colllaboration	roles	collabo-ration structure	topic of shared information	level of information sharing	method of sharing information	type of response	timeliness of response	target	risk-value	adversary
	environ-ment							time dimension		

Table 13: The second roadmap

A3.2.2.1 TRUST

A3.2.2.1.1 DEFINING TRUST

There are many things parties can trust, such as people and organisations. Other things, such as institutions or higher powers, are also things some people have (no) 'trust' in. However, Luhmann defined such type of trust as confidence, given the inability of parties to influence those things. Either by actively influencing or walking away from those things. (Nooteboom, 2002, p. 55) In this research the focus is on a specific type of actual trust, behavioural trust. The object of behaviour trust is a person or organisation. (Nooteboom, 2002, pp. 49–50) Behavioural trust has two sides, the trusting party (the trustor) and the trusted party (the trustee) (Nooteboom, 2002, p. 8). A motivation for a trustor to trust the trustee in some respect and situation is if the trustor is trustworthy in those conditions.

In his discussion of trust Nooteboom distinguishes three forms of expectations: reliance, assurance and trust in the strong sense. Reliance is the most comprehensive form of expectations, covering all bases of expectations. (Nooteboom, 2002, p. 11). Reliance is considered by the author to cover trust in the wide sense; trust as an umbrella term to an expectation that 'things will not go wrong', regardless of the basis of that expectation. With that, reliance consists of assurance and trust in the strong sense (Nooteboom, 2002, p. 49). Assurance is about an expectation of the execution of desired behaviour by a trustee. The basis for that is contributed by some control measure (such as legal coercion or dependence) stimulating a trustee handling self-interest to not behave

opportunistically by taking the room for opportunism away. (Nooteboom, 2002, p. 49) Finally, trust in the strong sense is defined as the expectation that trustees can be trusted, even if there are perceived opportunities and incentives to the trustee to behave in self-interest. (Nooteboom, 2002, p. 48) The crucial difference with the other forms is that with trust in the strong sense the trustworthiness of the trustee goes beyond self-interest (Nooteboom, 2002, p. 49).
In the research the term trust referred to the status of the relationship, or trust in the strong sense. Assurance was also covered, yet referred to as contracting, as one of the better known method of assurance or control. In the end, the contracts could be theoretically be such that there is no reasonable option to the trustee to behave opportunistically. In such a situation trust is simply based on the knowledge of the trustee having to behave moderately.

A3.2.2.1.2 IMPACT OF TRUST

Part of the main goal of this research is that the collaboration is based on trust. A more realistic version would be that it is largely based on trust, as will be discussed later on. Trust has extrinsic and intrinsic values (Nooteboom, 2002, p. 2). The extrinsic value is in that it allows to reduce transaction costs (Nooteboom, 2002, p. 2) by reduction of relational risk (Nooteboom, 2002, p. 108) more productively than assurance (such as contracts) *in the long run*. Trust also has intrinsic value, in that many prefer to work on the basis of trust, as opposed to having to work on the basis of distrust. (Nooteboom, 2002, pp. 2–3) Even more, trust is often even seen as being more preferable than assurance even at the expense of profit (Nooteboom, 2002, p. 200). With that, trust is considered to be the preferable option to control, with contracting as the most prominent, extreme case.
Assurance is the type of expectation that focuses on the expectation of behaviour of self-interest by the trustee. And to protect against such behaviour, the trustee seeks some source of assuring desired behaviour. Examples are presented in Table 14. In cases of inherent or perceived uncertainty effective contracting might be relatively costly. Especially over longer period of time the contract would probably have to adjusted to cover the current status entirely and without exception. And with that the question is whether under such uncertainty contracts can even be effective. If the trustee truly acts in pure self-interest the contract, under full assurance, should not allow any options for opportunistic behaviour, at any time. Although the level of uncertainty might differ per situation it will be always be there to some extent. And with that, the productivity (the ratio between effectiveness and efficiency) of the contract will be also uncertain.
Trust (in the strong sense) is equally exposed to the uncertainty of long term collaborations. Often trust is an (implicit) rational evaluation of trustworthiness of the trustee, but it is not entirely calculative. Regarding the uncertain aspects of situations trust is based on assumptions, instead of rational evaluation. This blind trust is not unconditional and there are limits to this kind of trust. The fundamental principle of trust is that it is to certain extent a rational calculation, in terms of being an evaluation of the trustworthiness of the trustee. But it is never fully calculative by nature, given the uncertainties. Not all future options, preference and states can be known beforehand. Trust is about imposing limits in the trustworthiness of trustees in situations, it is about leeway of the trustee. Within those limits there is trust. And those limits and trust in general are conditional. It comes down to the four place predicate of trust. A (1) trustor trusts a (2) trustee in (3) some respect (4) depending on the conditions ((Nooteboom, 2002, p. 38). And that trust has limits, as has the trustworthiness of the trustee.

	Macro	Micro
Control (assurance)	Contracts, supervision	Partner's dependence on value, hostages, reputation
Trust (in strong sense)	Norms, values, habits	Habituation, empathy/identification, friendship

Table 14: Sources of reliance, adapted from (Nooteboom, 2002, p. 65)

However, trust does not operate by itself and is not without risk either. It requires some combination of modes of governance, including the option of using contracts. Other forms are mutual dependence, reputation mechanisms and network structures. (Nooteboom, 2002, p. 201) The difference is in the intention, for example contracts can be intentionally used *complementary*. Their

purpose is as an aid to memory to the participants. Its use is not based on distrust. In the end, there will therefore be a mix of trust and control (/assurance), with some balance of the two types. But as suggested in this chapter, what matters is to opt for a balance in the favour of trust in the strong sense.

But even with modes of governance, trust can be betrayed. The consequences of betrayal has to do with relational risk, distinguished by Nooteboom in the risk of dependence and the risk of loss of knowledge. According to Nooteboom the main source of dependence is switching costs. The sunk costs participants have to incur in starting relationships with partners and to be able to understand the partners (related to the cognitive distance). Risk of loss of knowledge can affect the competitive position. Herein the focus of Nooteboom shows, with innovation as a primary focus. In this research the 'loss' of knowledge can also refer to the undermining of the position of an organisation. Especially with exchanges of raw data confidential information might leak, which can harm the reputation of the organisation.

A3.2.2.1.3 FORMS OF TRUST

Thus far trust (in the strong sense) in parties was considered as an entirety. But behavioural trust is complicated as there are many forms, related to different types of causes of action by parties. With that trust is about necessary levels on forms of behavioural trust depending on the situation at hand. On the basis of the course of action by Aristotle several forms of behavioural trust were distinguished by Nooteboom. The forms are depicted in Table 15.

Form of trust	**Object of trust**
▪ Material trust	▪ Means, inputs
▪ Competence trust	▪ Ability, skills, knowledge, to use technology, methods, languages, etc.
▪ Intentional trust ▪ Dedication trust ▪ Benevolence (/goodwill) trust	▪ Aims, intentions ▪ Dedication/care ▪ Benevolence, goodwill, lack of opportunism
▪ Conditional trust	▪ Outside enablers, constraints
▪ Exemplar trust	▪ Role models
▪ Informational trust ▪ Honesty trust	▪ Information ▪ Trustfulness

Table 15: Forms of behavioural trust, adapted from (Nooteboom, 2002, p. 50)

In this research this distinction is important because the different required forms and levels thereof are considered to be depend on the situation at hand. For example, in setting up a collaboration focused on thwarting attacks the trustee would have to the satisfactory trustworthy in terms of intention. But at stressful times of serious attacks, the leeway, the limits of trust in the trustworthiness will larger than in normal attacks.

A3.2.2.1.4 SOURCES OF COOPERATION BY THE TRUSTEE

To stimulate the trustworthiness of the trustee some sources of cooperation can be aimed for. These are presented in Table 16. The more altruistic sources will be more of a challenge is absence of a system. But over time, as cyber security collaborations become more common, these forms can become feasible. This explains for the expectation of trust being to reduce transaction costs to reduce over time. First with longer term collaborations participants will start to behave normal to expect good behaviour in return. But over time, if proven successful, the collaboration might develop into a system, the way of working.

	Macro		Micro
Egotistic	Coercion, fear legal obedience		Quid pro quo Economic reciprocity
	Social hierarchy Trust guardians	Reputation, Social reciprocity	
	Social conformance		Disinterestedness
Altruistic	Social obligation		Spontaneous sociality

Table 16: Intermediate sources of cooperation (Nooteboom, 2002, p. 74), which is an extended version of the sources of cooperation, as identified by Wiliams (1988).[26]

The connection of the sources of reliance, representing the side of the trustor, and the sources of cooperation which affects the trustworthiness of the trustee is an imperfect one. Although there is a connection, it is still an assessment. Potentially the trustee is indeed, in a situation, actually trustworthy, yet the trustor fails to recognize this. Although this is not covered any further in this research, this is important to remember. For example, a trustor noticing (or being confronted with) untrustworthy behaviour will grow more suspicious in general. Or a lack of continued meaningful information being fed back might grow suspicion, despite there being 'legit' reasons (nothing to provide).

A3.2.2.1.5 Motivation of focus on (type of) trust in this research

With the extensive division of trust in all kinds of aspects and characteristics a more extensive definition of the focus and ambition of this research regarding trust is possible. The research initially focuses on (1) trust, specifically (2) rational reliance in (3) parties, with the intention to have this develop in (4) trust in the strong sense on a systemic level.

(1) Confidence is ignored as potential participants are considered to have the opportunity to not join a collaboration. (At least this research does not suggest making collaboration obliged.) And thus, the collaboration is not a thing pushed to the participants in which they are supposed to have confidence.

(2) It is not supposed collaborations start off with purely trust in the strong sense, more of reliance, as some participants will join out of self-interest. But the main focus is on not having to use control extensively, specifically not by means of contracts. From thereon, slowly, trust in the strong sense would have to develop, based on development of more altruistic sources of cooperation (Table 16).

(3) At the start there is no system, no real renowned collaboration on cyber security. As a result, there is no system to trust, it has to come down to parties (representatives and respective organisations).

(4) Ultimately the goal is to get a situation in which participants trust the system. A system which stimulates a strong altruistic style of cooperation by participants. That in turn would make it easier for participants to trust each other. It would also for a less rational form of trust, based on routinized behaviour, empathy and with that influence the limits of trustworthiness and tolerance levels of trust.

[26] Contrasting the three prime examples in this report based on publications of collaborations, the cooperation of financial parties in the United States appear to be more egotistic-macro nature (with "well-developed roles and responsibilities in the information sharing process") (National Infrastructure Advisory Council, 2012, p. B–17) to economic reciprocity ("Trusted partnerships are built over time, as the participants learn each other's needs, capabilities, and commitment to honouring confidentiality. Each side of an information sharing partnership must see a benefit from the partnership's success,") (National Infrastructure Advisory Council, 2012, p. B–18). The chemical parties more on reputation (as personal relationships and networks are important (National Infrastructure Advisory Council, 2012, pp. C–7) and the AbuseHUB in the Netherlands of the economic reciprocity (as receiving and providing information about infections is their own interest [4.6.1].

A3.2.2.2 END

A3.2.2.2.1 ACTIVITY

The time dimension does, in some way, resemble the activity dimension by Kowtha. But it's a bit of departure as Kowtha et al. distinguish centre's activities in:

- protection (preparedness, scanning & data capture and design),
- incident management (detection, response & recovery and assessment), and
- analysis (information extraction, event/incident correlation and cyber threat discovery)

The difference is in their addition of the analysis and the absence of the post incident activity.
The analysis activity is considered to be a rather common activity, at least in this research. Analysis is not really an optional, separate, activity in the characterization model of Kowtha. For example, the detection factor by Kowtha and its attribute intrusion detection / event correlation presupposes the attribute event/incident correlation of the analysis activity.
The post incident activity is added as at times there will be a need to discuss accidents or potential ways to improve the thwarting of an attack. Such activities are tied to (series of) specific attacks. In contrast, the activity prior to attacks is about what might happen. This could be based on types of past attacks, but also about trends and developments.
The net result as suggested in this research is:

- identifying a future attack (what might happen and what can we do about it?),
- thwarting an attack (what happens and how can be recover?), and/or
- learning from an attack (what has happened?).

The time dimension of a collaboration centre is an actual decision to be taken as it affects the filling of the options for some of the identified steps of the 'development of the collaboration model' phase. For instance, a collaboration setup prior to the actual attack it is focused on will be about what the attack might be like. What technique will be used to, what vulnerability might be exploited and/or who will be the target? Or even more proactive, it could be about the intentions of the adversary against some target. What elements are required in the range of possible attacks all resulting in satisfying the same intentions. This explains for the addition of the adversary as a roadmap step in the end state of the initial roadmap. Although discussing who the actual adversary is troublesome, an identification of the type of adversary with respective intentions and available resources of the adversary could well be worth the efforts. It could result in a discussion of for example the removal of some vulnerability by not having the hazard in the first place. An example would be to no longer have some system with weak or no protection connected to the Internet.
Aside from the impact of the choice of the positioning of the collaboration relative to the attack, the decision has an impact on the backcasting steps as well, making it a true dimension affecting all phases. For example the removal of a hazard by not having the hazard online will in all likelihood require involvement of higher level management. After all, there was some benefit attached to having the system online in the first place, for example a cost saving of having to travel to the system. On the other hand, collaboration setup for the attack phase will possibly have 'lower' levels of hierarchy involved to actually thwart the attack, instead of mindfully avoiding it.
Finally, a decision on the time dimension is not the only influence on a roadmap step, the roadmap steps will affect and be affected by the other roadmap steps too. Prior to an attack the required (type of) information and time sensitivity will be different from the situations an attack takes place. And *that* requires involvement of different parties.

Implicit criticism on the time dimension

The approach of making the time dimension of a collaboration an actual decision in the definition of a future end state will meet resistance. In the field there appear to be collaboration centres that perform all three activities (Kowtha et al., 2012, p. 35). However it is uncertain if the centre changes in some way depending on the activity at hand.

Zhao and White are more explicit about the sequential nature in their collaboration framework (Zhao and White, 2012, p. 460). Their *proposed* collaboration framework suggests a changing collaboration structure depending on whether there is an incident or not. Skopik et al. propose an 'extended incident response cycle' that consists of a preventive and reactive part in series. The preventive part is (amongst others) about identifying potential attacks and the reactive part about short-term tackling of the attacks. (Skopik et al., 2012) Although some kind of circular, feedback-loop-like, way of improving is common, the question is whether this is no oversimplification, if it is even possible in the first place. As argued in this research with a shift from preventing to responding to a threat it could be different levels of hierarchy of participants are involved, particularly with larger organisations. This is not necessarily impossible, but it requires an effort from the participating organisations *to translate* the discussion. After all, management level speaks a different language and has more (high level) means than more hands-on levels. Hands-on levels have less degrees of freedom and require different information. As a result, a response cycle or some changing collaboration structure is an oversimplification at best in this research.

A3.2.2.3 DEVELOPMENT OF THE COLLABORATION MODEL

A3.2.2.3.1 ENVIRONMENT

The environment is not entirely a given, it is about the relevant environment, which depends on the intentions of the collaboration. By adding a participant to the collaboration the corresponding organisation will have a different position in the environment. But also with a different ambition, such as preventing types of attacks instead of thwarting specific attacks, the environment will be different. And with that change of the environment, some of the steps will be affected. The question is what participants are willing and are able to share in a collaboration. This depends on the environment and the participants of the collaboration.

With all that participating organisation have to consider on how to 'work' the environment, especially their own organisations. The organisations might impose what information can be shared. But also the organisation can allocate better resources to the collaboration. And there are other, similar considerations. Such considerations are discussed below in the changed set of 'sub-steps' as identified by Kowtha et al.. These 'sub-steps' are refinements of the step 'environment', and the corresponding options are not undefined because of the amount of degrees of freedom.

One specific concern about the environment deserves more attentions: confidentiality of the shared information collaboration. In collaborations organisations might share information that is confidential or sensitive. But in the environment organisations might want to know what information is shared (for the sake of knowing whether privacy is at stake or insight in the processes). Collaborations have to be aware of whether such requests can and have to be fulfilled. For example, recently, in the Netherlands this was reason for the Rabobank to ask for classifying information a 'state secret'. Otherwise, if the company intends to share information like vulnerabilities the environment (such as the media) could ask for such information. (Lange, 2013) Especially in case of involvement of governmental organisations in collaborations this is a concern. In various countries there is a law for the 'Freedom Of Information' (FOI law). Such FOI law stipulate that the general public has a 'right-to-know' what data is held by national governments. Because of that law the public could ask from the government what is shared and with that might get to know of the vulnerabilities. Should it be impossible to avoid this, than participants might refuse to share information with the government or even not allow them in a collaboration.

In practice

Kowtha et al. identified quite a few, what in this research would be called 'sub steps' for the environment. In this research 'environment' is about those options that are affecting the collaboration from the outside. On those aspects the collaboration has to decide on how to handle these concerns and if possible and necessary try to influence to get the required support. With that description of the purpose of 'environment' in this research the 'options' called 'visibility' and 'reach' are removed. But still there are 'sub-options' to consider, such as:

- Data handing, indicating restrictions on handling and use of the data. What is acceptable to a collaboration in that regard? This should be defined early on in order to avoid problems in a collaboration. An example could be that suddenly one organisation is not willing to share some specific information, whereas other organisations did.
- Capability, Kowtha et al. presented this as organizational- and policy considerations, in this research it is tightly linked to data handling. Data handling is about what can be shared, capability is about what can be done, in what way and so on.
- External stability is somewhat troublesome as it is both a given, but also the result of choices. Some domains are simply less 'stable' in that the developments are more active. And similarly some organisations might be less 'stable'. Think of organisations that expanding rapidly or are highly attractive for whatever reason. The question is on how to handle such instability. It could be a reason to not include organisations because the collaboration wishes stability. Again, with this the dependence between the several options is demonstrated.
- Community coordination, as part of a step which does not consider collaborations to be impartial to its environment the collaboration has to communicate with that environment. How do activities reach the collaboration, how is this coordinated and so on.

A3.2.3 THIRD ROADMAP

In the table below the third roadmap is depicted.

Current state ↓	Development of the collaboration model ↓							Desired end state ↓		
formulation goal of col- laboration	roles	collabo- ration structure	topic of shared information	level of information sharing	method of sharing information	type of response	timeliness of response	target	risk-value	adversary
	environ- ment							time dimension		
	external interaction	scale	maturity					influence		

Table 17: The third roadmap

A3.2.3.1 CONFIGURATION THEORY

The configuration[27] theory is in part an approach to steering which respects autonomy and interdependence of parties. Herein the steering is not directed on the parties or content, but the interaction processes. (Twist and Termeer, 1991, p. 26) Characteristic to the configuration theory is the focus on the relation of the social environment and the interrelation with the definition of reality by that social environment (Termeer, 1993, p. 27). Reality is considered to be a social construct. This reality is constructed and reconstructed in ongoing interactions of parties (Termeer, 1993, p. 325), in configurations, for the purpose of making sense of *their* environment (Termeer and Kessener, 2007, p. 258).

[27] Mintzberg first popularized the concept of configurations in organizational science. But its meaning is different from the way the configuration theory uses the concept (Termeer and Kessener, 2007, p. 271). Mintzberg distinguished (initially) five configurations defining fundamental characteristics of organisations(Dongen et al., 1996, p. 84). The underlying combination of variables of those configurations are considered to be 'pre-coded' (Termeer and Kessener, 2007, p. 271). In contrast, the configuration theory implicitly opposes to there being pre-coded variables underlying the configurations. Furthermore, the configuration theory adds the social-descriptive element along with Mintzberg his cognitive element. (Dongen et al., 1996, p. 86)

A3.2.3.1.1 INTERDEPENDENCE OF PARTIES WITH LIMITED, SUBJECTIVE VIEW

Reality as a social construct

The configuration theory is positioned in scientific-philosophic debate by its representatives[28] (Termeer, 1993, p. 23). The debate boils down to a discussion on three themes[29] and on the basis of those three themes three positions can be distinguished (Laat and Maas, 2003, p. 92). The first position considers there to be one reality, one which is objective. The second considers reality to be a subjective perception, a construct which is the result of the beholder. The third position considers definitions of reality to be the result of social configurations, the result of interaction. (Laat and Maas, 2003, pp. 92–94),(Termeer, 1993, pp. 24–25) The configuration theory is based on the third position. Configurations consisting of different (types of) parties exchange definitions of reality with the purpose of agreeing upon one common definition of reality (Twist and Termeer, 1991, p. 20). A reality which is constructed, perceived and reconstructed by that specific configuration. (Termeer, 1993, pp. 24–25) With that, the main intention of the configuration theory is for parties to make sense of the world by discussing with others what happens, what it means and what is unknown (Termeer and Kessener, 2007, p. 258).

Options for steering

Collaborations are amongst different participants, fulfilling different roles and having interdependencies which might change over time. Therefore this research opts for 'steering' in network like fashion (pluricentric approach[30]), as opposed to a hierarchical environment (unicentric approach). A characterisation of these two extreme approaches on steering is presented in Table 18

	Unicentric approach	Pluricentric approach
Principle to define common interest	The politicians on the central level decide upon goals and means	Decisions on goals and means are taken in a network
Structure of the public sector	The public sector is hierarchically divided into functional parts	Government is a complex constellation of actors
Process in the public sector	Starts with problem on national agenda, an optimal solution is adopted there and implemented	Several actors in the network can take initiatives and others can support or oppose them
Most adequate metaphor	The system, in which every element has a functional task in the whole	A network, in which actors struggle to influence policies

Table 18: The "main axioms about the policy field from two perspectives",
as distinguished and defined by Klijn and Teisman (1991, p. 101)

[28] At the same time the theory is positioned in the organisational theory debate. However, structuring organizational theories turns out to be hard, as justly characterisation of the theories is no easy feat. The various 'futile' classifications of theories do not appear to have mutual cohesion (Laat and Maas, 2003, p. 94). Additionally, there is a severe risk in losing the important nuances of underlying positions of the different theories, resulting in caricatures (Termeer, 1993, p. 25). The representatives present a rather coarse classification of organisational theories in three clusters. These are the classic organisational theories, interpretative theories and the process- and organisation theories (Termeer, 1993, pp. 25–27). These will not be discussed any further as this classification is strongly akin to the scientific-philosophic positions. A rationale therefore is presented by Voogt (one of the representatives of the configuration theory). According to Voogt the position of choice of the scientific-philosophic debate affects the position in the organizational debate (Termeer, 1993, p. 25).

[29] The three themes are ontology, epistemology and methodology. Ontology is about the question to what extent reality can be known or understood. Epistemology is about the relation of object and researcher, specifically the question on how can be known what one knows. Methodology is about applied research of the first two themes in different fields. (Termeer, 1993, pp. 23–24)(Laat and Maas, 2003, p. 92)

[30] By itself the pluricentric approaches to steering leaves quite a few types of management options. Termeer further divides the options in the focus of steering, being on (1) the relations between the different actors (who), (2) the rules of interaction of the actors (how) and (3) the definitions of reality (what). In configuration management these three options for steering are linked. (Termeer, 1993, pp. 283–284)

Unit of analyses: (Social-Cognitive) Configurations

A configuration is a snapshot, an empirically derived moment, in time of intensive interaction patterns of parties, with shared interaction rules and a shared definition of reality (Termeer and Kessener, 2007, p. 258)(Termeer, 1993, p. 326). Maas (Dongen et al., 1996, p. 95) characterised a configuration extensively as having:

- High degree of organisation (different actors, intensive contact),
- Close, intensive contact between actors (formulation of strategies),
- Configurations are nodes of information and deliberation,
- Boundary crossing formal consultation (discussion documents are present),
- Configurations are recognizable and open to others from outside the configuration,
- Configurations are platforms or links, not necessarily limited to the boundaries of a department, division or organisation,
- Configurations are strongly embedded in the direct social environment.

A configuration is considered to be an extreme on a scale of describing social situations. The counterpart is an aggregate, with particular characteristic the low degree of organisation and infrequency and informality of interaction. (Dongen et al., 1996, p. 95) But with interaction the internal homogeneity of the configuration can increase (Termeer and Kessener, 2007, p. 258).

Underlying mechanism of reality as a social construct: Double helix

Configuration management is about the cognitive dimension ('what') and the social dimension ('who' and 'how' of interaction (Termeer and Kessener, 2007, p. 271)(Termeer, 1993, p. 325). These two dimensions are interconnected (Dongen et al., 1996, p. 86)(Termeer and Kessener, 2007, p. 258) in a way as background and foreground (Termeer and Kessener, 2007, p. 271) or sides of a medal, with one side being visible at a time (Dongen et al., 1996, p. 87). The coherence is also compared to the double helix of a DNA-structure (Termeer, 1993, pp. 34–35). It is presumed (Dongen et al., 1996, p. 90) that on the basis of the social structure the cognitive can be traced and the other way around (Dongen et al., 1996, pp. 86–87). An example thereof is presented by Maas of the description of a banana crate. One person called it a representation of a little goldmine, a product delivering the majority of operational results. The other person called it a back breaker, suggesting this person represented a different configuration with a different definition of reality. (Dongen et al., 1996, p. 88)
The cognitive dimension is about the definition of reality as considered by a social configuration and the interpretation of the what this reality entails. For that interacting parties do exchange their vision and redefine their vision (where necessary) based on input of others. With that they define the agreed upon reality, which forms the basis of further action. (Termeer, 1993, pp. 30–32) Simply put the cognitive dimension is the 'what' dimension (Termeer, 1993, p. 325). The definition of such an agreed upon reality is active process of frequent (re)definition by the configuration (Termeer, 1993, p. 325).
The complementary strand of the cognitive dimension is the social dimension of interactions. It deals with the 'who' and 'how' (Termeer, 1993, p. 325). The social dimension entails the configuration of parties. Given the limited resources of parties participants will have to decide upon who they will interact with. The resulting configuration, with interaction of parties with other parties influences the cognitive dimension. The configuration will construct its definition of reality. As part of the social dimension the configuration also defines and redefines the rules of interaction between participants. As examples of what such rules can be about Termeer mentions (non-)allowed participants of a configurations, expected and accepted roles of those participants, but also time and place of interaction. (Termeer, 1993, pp. 32–34)

Functional conflict as the source of development

The interrelation with double helix of interaction is presumed to be the basis of development. A social structure is the result of a cognitive dimension yet also the basis for *potential* further development by a change in the cognitive dimension. (Termeer, 1993, p. 326) This can be in a reinforcing direction but also in a change of course. People with certain beliefs of reality interact with

people with similar beliefs. And this increase in contact is likely to reinforce the similarity of belief (Termeer and Kessener, 2007, p. 258). But an introduction of a different party with a different perception of reality (on the basis of prior social interactions) might change cognition. It introduces a (slight) context variation. This different perception might be the result of that party also being present in a different configuration (multiple inclusion), which has a different definition of reality. The incongruence can result in a conflict of definitions. And such conflicts, assuming it to turn out to be used in a functional manner(Dongen et al., 1996, p. 108), are considered to be an important source dynamics, of a change of course of definitions. (Termeer, 1993, pp. 246–247)

An important condition of the configuration theory is to 'always' allow for possibilities of changes of definitions. The possibility to reflect on and actually change definitions is supposed to be safeguarded by norms. Those norms purely refer to the process, not to the composition of the configuration or dimensions (such as required parties or definitions) (Termeer, 1993, pp. 38–39). This attention to not interfering with the contents of discussion is made explicit by defining the norms as un-values. Un-values or negative steering are put in the ethical space, the kind of steering supposed to protect the possibility of reflection. The un-values are supposed to define in-justice, not what justice is. These kinds of negative steering should clearly be distinguished from positive steering. Positive steering focuses on 'negotiable order'. (Dongen, 1991, p. 54) The kind of steering for which there is no room in the configuration theory. The principle of using un-values or negative steering is that it is easier for people to agree upon what is undesired than to define what is desired (Termeer, 1993, p. 39). For that reason, the required continuous possibility of reflection is translated in un-values as for example non-blocking, non-discrimination and non-excluding processes of ongoing interaction (Termeer, 1993, p. 325). These three examples of un-values affect the cognitive and social dimensions of interaction. A motivation to protect these dimensions, in undetermined form referred to as thirdness[31] (Dongen, 1991, p. 52), is that each provides angles for dynamics. Non-discrimination and non-excluding processes of ongoing interaction are about offering opportunities of context-variation. Non-blocking refers to actually leveraging from a confrontation as the opportunity of a confrontation by itself is not sufficient.
Termeer identified three moments of blockage (Termeer, 1993, p. 261). The confrontation has to be considered a confrontation of incongruence, this incongruence has to be considered problematic (or at least worth an investigation) and finally the configuration has to be willing to reflect on existing definitions based on the incongruence. According to Termeer the autopoiesis theory provides interesting insights if incongruence is not detected or not considered(Termeer, 1993, p. 261). Despite it being of importance, these two will not be considered in this research[32]. The third type of blocking

[31] Van Dongen distinguishes two types of thirds: a third and thirdness. A third is commonly used to refer to a third party, one which has its own socially constructed definition of reality. With the introduction of this third (party) to a configuration a possibly conflicting reality enters the configuration. Thirdness is the more encompassing yet undetermined option of introducing a social and/or cognitive structure to a configuration. Similarly Termeer mentions thirdness (as 'thirds' or a third) referring to a different definition of reality ('what'), a different rule of interaction ('how') or a different actor ('who') (Termeer, 1993, p. 37). In this research the 'how', as put forward by Termeer as another angle for dynamics (Termeer, 1993, p. 286) is not really considered to be another option of the undetermined thirdness. This research limits itself to the double helix, with one helix providing the 'what' and the other the 'who' yet without the 'how' which Termeer added to that helix(Termeer, 1993, p. 325). The motivation for that is that it breaks with the double helix metaphor, whereas the 'what' and 'who' can be traced back to each other, the 'how' is a bit of an anomaly. Furthermore, the separate discussion of the 'how' seems to distinguish Termeer from the other representatives, raising the complexity of grasping the concept of the configuration theory. A contributing factor in the complexity is of seemingly slightly different notions of configuration theory. What distinguishes the configuration theory from the configurative integration theory. The latter adds the social integration theory, which puts the norms of interaction parties use outside the configuration theory. Finally, if the 'how' is actually an integral part of the double helix, interfering with the 'how' would come down to interfering with the contents. The thing the configuration theory intends to avoid at all costs by solely focusing on process of interaction, of which the interaction rules are a likely part.

[32] The motivation is that a line has to be drawn, the main concern in this research is to find a way to end up with a collaboration with reasonable chance of being usable for the more challenging type of cyber security attacks. Although failing to detect or considering differences to be worth the discussion is a real threat, it is considered for that there first

of configurations not being willing to act upon incongruence is called 'fixation'. With fixation the social and/or cognitive dimension are fixed. With the cognitive dimension it means definitions of reality are declared unchangeable and non-negotiable. With the social dimension it means that rules of interaction and the list part participants is fixed and not up for reflection. (Termeer and Kessener, 2007, pp. 258–259) Social fixation might also be expressed by pseudo conflicts of definitions of reality. Instead of discussing issues at the social level, the terrain of discussion is the content. (Termeer, 1993, p. 262) Although a fixation can be limited to a single dimension, due to the interconnection, fixation of one can result in fixation of the other (Termeer, 1993, p. 262). With that, 'groupthink' or 'dialogues of the deaf' can be thought of a form of fixation. (Termeer and Kessener, 2007, p. 258). Conditions for groupthink (small fixed groups expressing ritual behaviour and being unsusceptible to outside influences) are similar to excluding entry of thirdness (Termeer, 1993, p. 261). More specific, the *combination* of social and cognitive fixation can thus be thought of as the description of groupthink. From now on in this research, fixation refers to fixation of one of the two dimensions, groupthink to actual (non-pseudo) fixation of both dimensions.

Balance of dynamics and stability, option for inertia without fixation

Earlier redefinition was considered to 'always' be possible. This possibility is more of a possibility in time, at least redefinition is not supposed to be a continuous activity. There should be a balance of dynamics and stability, with dynamics as in openness the preferred and distinguishing mode of the configuration theory (Termeer, 1993, pp. 38–39). Dynamics refer to a change of definitions of reality, interaction patterns and/or rules of interaction. Stability refers to a fixation of definitions of reality and interaction. Not only is such fixation at times allowed, it is even necessary condition for collaboration[33]. In moments of stability parties in a configuration can build on agreement of definitions of reality (cognitive dimension) and interaction (social dimension). (Termeer, 1993, pp. 38–39) To distinguish acceptable from unacceptable stability inertia and fixation are used. Inertia is het resistance to change, but resistance is not the same as impossibility. Fixation is an impossibility to change. The norm, operationalized using some un-values, of 'always leaving option to change' ((Termeer, 1993, pp. 38–39)) is therefore confusingly defined. The suggested norm is that definitions should always be subject to inertia or negatively formulated 'definitions should never be fixated'. The positive formulation is a more realistic representation, with the resistance still being an undefined value ranging from zero to some high, yet non-infinite, value to separate it from fixation.

The main source of dynamics is confrontation. And with that, the un-values are aimed at avoiding fixation of either dimension. Should fixation (as opposed to merely inertia) of either dimension be the case it is of importance to de-fixate the respective dimension. As an important consequence of the presumed dynamic coherence the path of least resistance is to affect the non (or less) fixated dimension. For example, if definitions of reality are considered to be non-negotiable, introducing a new party is the best bet. With that new party a new definition of reality will enter. (Termeer and Kessener, 2007, p. 259) Such 'context variation' (distracting from a fixated dimension to the other (Termeer and Kessener, 2007, p. 259)) has its limits, a too extreme incongruence might not work (Termeer, 1993, p. 291). Introducing a party which has no common elements with the configuration might not be accepted.

---

should be some reasonably mature collaborations to be out there. Until that moment in time, failing to detect incongruence if it is right in your face (what the first two types of blocking are about) should not yet be the real concern.

[33] Termeer considers stability to be essential for communication ((Termeer, 1993, p. 39)). In this case the communication refers to communication having defined the reality and rules of interaction. It is the type of communication which is deemed necessary on the basis of those definitions. In this research such communication would be to collaborate given some socially defined perception of reality. To avoid confusion such 'communication' is replaced by collaboration

Management of configurations

Thus far the two[34] non-mutually exclusive methods of management of configurations were implicitly presented: facilitation and intervention. Facilitation is the most important form, it is about facilitating the conditions for continuous interaction with the intention of avoiding fixations. This comes down to making sure third parties can enter, stimulating reflection of definitions and to make sure the resulting incongruence results in functional conflicts (Termeer, 1993, pp. 285–288). A functional conflict means definitions of reality do change (either in something similar by reinforcement or confirmation, or actually change in some different reality). Intervention is in order if fixation settles in. The basic principle of intervention is context variation: focusing on the dimension which is (least) fixated. It is important to focus on fixators, the participants which are highly invested in the configuration and are exposed to limited different views on reality. (Termeer, 1993, pp. 288–294)

Empirically derived role of parties in configurations

Termeer distinguishes three roles parties can fulfil in a configuration: brokers, fixators and initiators. These roles are not pre-determined but empirically derived, but those roles might be conditional on the formal position of the party and the role already fulfilled in another configuration. (Termeer, 1993, pp. 269–270) After all, some parties might be multiple included, as they are not in full agreement with a single definition of a single configuration. This applies to a broker in particular. These parties are by definition multiple included and, knowingly or unknowingly, bring in definitions from one configuration to the other. These brokers are will never be included very high as they are not highly invested in a configuration. The counterpart is the fixator, a party which is responsible for the inertia (or even fixation) of definitions. These parties prefer the status quo and are highly included in typically a single configuration. Given these characteristics interventions focus on this type of party as they are a likely stimulating source of fixation. The third type is the initiator. A party which is responsible for bringing in thirdness (definitions of reality or interaction and new parties). The thirdness could be the result of their formal position or their contacts (which might be in different configurations). The formal position justifies them bringing in thirdness. Their level of inclusion varies, but they will not be the highest included party. These three distinctive roles can in reality be more difficult to distinguish. It is possible for parties to fulfil multiple and different roles over time. (Termeer, 1993, pp. 269–270)

A3.2.3.1.2 USE OF THE CONFIGURATION APPROACH

The configuration approach in the configuration theory was already considered to be applicable, but it is also useful. It provides a way of looking at collaboration that is useful to the design of information sharing collaborations..

[34] Actually, on the basis of the used norms (or un-values) Termeer distinguished two forms and added one form of management: management (presented as facilitation), intervention and route management. The third, aimed at influencing substance in some *desired direction* instead of the process, is omitted in this research. It was added by Termeer to the configuration theory as it was missing compared to network management and steering. But despite the attractiveness some difficulties apply, rendering the form of management to a limited and intangible extent usable. (Termeer, 1993, pp. 294–295) More importantly it breaks with the consideration of reality being a social construct, as not a single party has *the* overview. And with that comes the question what the actual desired direction is, whether it even exists. It is in essence the embodiment of the dilemma of prescription as discussed by Termeer (Termeer, 1993, p. 298).

Usefulness of configuration approach

At the start of the discussion of the configuration theory in this appendix the position of the theory was presented as being *applicable* to the challenge of cyber security as discussed in this research. In brief the attractive concepts of the configuration theory are:

- its post-modern position in the scientific-philosophic debate, considering a definition of reality the result of a process of sense making by some configuration at some moment in time,
- its pluricentric approach to steering, considering interdependence of parties which might be also be part of other configurations, and
- its focus on continuous interaction to redefine definitions of reality unless stability is required.

These concepts are particularly interesting because the challenge of cyber security is closely related. Cyber security is defined as being a challenge:

- as no party has full situation awareness and at least the known attacks suggest a decreasing amount of parties will be aware of the attacks, given the increase of targeted and more insidious attacks,
- there is an independence with parties having to fulfil different roles at different moments in time in a collaboration, making it harder to use a unicentric approach to steering[35], and
- using the same definition of reality might not suffice over time as unknown, unimaginable attacks or different targets might necessitate a redefinition, for example the social definition of reality by configurations on smaller organisations not being a target might be in need for redefinition.

Aside from the configuration theory being oriented on some of the challenges cyber security suggesting the theory to be a *usable* approach, there are also an implicit notion on the size of a collaboration. The philosophical position of reality being a social construct is less of not that hard to imagine with cyber security collaboration. The data breach investigation report by Verizon opens with: ‘ “Some organizations will be target regardless of what they do, but most become a target because of what they do. If your organization is indeed a target of choice, understand as much as you can about what your opponent is likely to do and how far they are willing to go.” ‘ (Verizon, 2013, p. 2) It demonstrates organisations have to make sense of what they are confronted with and what the resulting reality is. Add the consideration of organisations having to collaborate as they act on the basis of consequences and socially constructed reality is suddenly not just a philosophical discussion. But such an organisation will not be confronted with one reality. There are different adversaries and some organisations will be confronted with more. Based on the aforementioned activity (the sector) and size of organisations, but possibly also based on the use of an application with a known vulnerability. With that different potential configurations could be thought of. Configurations with a shared reality at some moment in time, such as a simple reality of a piece of software being vulnerable or even untrustworthy given its track record. The result might be, possibly externally imposed, to face reality and condemn the use of some application.

The purpose of the example is not to suggest a development of configurations for each and every application, every sector, every size organization. It is about considering smaller configurations and benefiting from multiple inclusions of parties as opposed to opting for a large scale, single collaboration. A sectorial configuration provides more opportunities for a common configuration, but it is still a rather rudimentary configuration. And in such a configuration the definition of reality will (and turns out to be) limited to a strategic level. A low level/detailed definition of reality is easier to accomplish in a smaller configuration. And such a more detailed reality in a smaller configuration will also allow for more impact with context variations.

[35] As discussed it is possible, for example by introducing stringent regulation, forcing organisations to cooperate, but the assumption of this research is that if force is the main driver organisations will be less willing to cooperate. The organisations might cooperate to comply, yet might oppose the regulation or not share the real valuable type of data.

Limitations regarding practical recommendations

The respective sections on the configuration approach are not riddled with recommendations on recommended configurations or even its possible result, various definitions. The reason is twofold, it is undesired from the point of view of the configuration theory and difficult in general to provide these recommendations. The undesirability has to do with the prescriptive dilemma affecting both recommendations regarding content and process. Should a definition be provided from the outside that definition would be more akin to that of one party than to some other. (Termeer, 1993, pp. 297–299) With that, providing a *predefined* definition can disrupt the processes of ongoing interaction. It might even limit the chances of future intervention by a reduction of acceptability. (Twist and Termeer, 1991, p. 26) Aside from the possible detrimental effect of providing predefined definitions, based on the foundation of the theory, the question is to what extent such definition is any good anyway. Parties in a configuration will start off with a personal concept of reality and together they will have to define the socially agreed upon reality. That process of definition is considered necessary, not necessarily desired, given the presumed complexity making it impossible for a single party to understand (relevant) reality. Interfering or presenting with that process by presenting definitions would therefore be the height of arrogance from the perspective of the theory. The result is that predefined definitions factor some parties in a configuration, possibly freezing functional interaction, and the quality of the definition is questionable. Even the more acceptable level of interference, process related, is confronted with the prescriptive dilemma. What qualifies an outsider to diagnose fixation, definitions on that can vary per party in a configuration. Following up such a diagnoses with intervention (which is actually more content related than presented anyway) by bringing in a party or definition poses the prescriptive dilemma yet again.

Besides the undesirability presenting definitions is difficult given the complexity and dynamics (Termeer, 1993, p. 299). It is a bit like anticipating a chess match, considering all subsequent possible moves and countermoves, and advising up the basis of that knowledge. Accumulating the required knowledge would take a long time and the result would be a long set of conditional statements. But in all fairness, a comparison to chess is even an understatement. The complexity is limited because with chess the amount of options are large yet limited, contrary to the challenge of cyber security. With cyber security new chess pieces appear and chess pieces at unknown moments in time are (in)capable of different moves. It is the complexity and dynamics in time which is at the foundation of the considered necessity to collaborate. Termeer considers the combination of complexity and dynamics (in general) to be even harder than overcoming the prescription dilemma. The suggestion is to opt for the middle ground of abstract and concrete recommendations, with the intention to stimulate the configurations themselves to approach the problem in a similar way. (Termeer, 1993, p. 299)

A3.2.3.2 END STATE

A3.2.3.2.1 INFLUENCE

The step called influence is about a definition by the collaboration of who the organisations intend to reach. The collaboration could be about influencing organisations in a nation of some sector, but also a more localised initiative. With that the options are like with Kowtha et al. ranging from local to global. But more refined options are possible, such as a subsets of a sector (such as trying to influence organisations that are in control of the power grid).

In practice

In the model by Kowtha ‘influence’ is also distinguished. The options for the influence range from local to global.

A3.2.3.3 DEVELOPMENT OF THE COLLABORATION MODEL

A3.2.3.3.1 SCALE

The scale has relations with the external interactions. Highly productive collaborations in terms of external interaction might be able to compensate for being rather small in scale, yet aiming for high influence and high productivity. Scale is in this research primarily about defining maxima, for the sake of being able to maintain clear and detailed definitions of reality.

As extensively discussed the scale of the collaboration is a trade-off of cumulative cognitive distance and cumulative cognitive coverage. Having many organisations in a collaboration will result in increases of both the cognitive- distance between and - coverage of the collaboration. But also having large scale collaborations makes it harder to share information. Large collaborations tend to focus on strategic information sharing. And with that, the defined goal of stopping some type of attack could be in danger. Stopping an attack requires pragmatic information (and a group of organisations that can act fast).

In practice

The scale of the collaboration was discussed frequently in the report, including several examples of the consequences of the scale. An example was the impact of the amount of participants connected to an Information Sharing and Analysis Centre. Given their typical large scale in the United States they typically focus on rather strategic information sharing. (Rashid, 2013)

A3.2.3.3.2 MATURITY

Maturity is both a representation of the current status of the collaboration, a result, but it is also a result of decisions regarding requirements on participants of the collaboration. Participants can be vastly different in many ways. Examples are their potential differences in their definition of reality, awareness of that reality and their ability to act upon that reality. And with that their weight in the credibility of the centre can vary.

The question is whether in a collaboration participating organisations should have a certain level of maturity. Reasons for this could be efficiency and effectiveness. If the organisations are less ‘mature’, being that they are less capable in terms of cyber security they might need more assistance of other organisations in the collaboration, unless the organisation has resources at its disposal to overcome this maturity. Whether the organisation in question is able to quickly act based on the, or share the required the information affects the efficiency of the collaboration. The effectiveness is related to the goal. Depending on the goal, a lacking efficiency can decrease effectiveness. The most straightforward case is if the collaboration wishes to thwart attacks. But maturity and effectiveness are also related if the goal of the collaboration is to influence organisations. A collaboration which consists of organisations that are fairly young or ‘frequently’ affected by security incidents has a harder time influencing other organisation.

Whether maturity is really a relevant step depends on the goal of the collaboration. If the collaboration wishes to influence a group of organisations there will not be a real requirement of maturity. And actually, it could be valuable for the sake of representation. If the collaboration wishes to influence small organisation that think that cyber security is not a concern to them representation can be more important. Adding organisations that are similarly ‘immature’ about security, but represent the other organisations can be of more value than adding an organisation to the collaboration that is very mature but can be claimed to be a different type of organisation.

In practice

Maturity options can in part be operationalised using the levels of awareness, but this is rather abstract given that the awareness depends on the situation. A situation that first has to be defined With the Community Cyber Security Maturity Model a model is presented by White the maturity can be operationalised. The model depicts five levels of maturity on four areas, being ‘awareness’, ‘information sharing’, ‘processes and procedures’ and ‘integration’. (White, 2011) It goes too far to

discuss these different areas in detail in this research. But what matters is that this consideration of maturity of communities might be translated to an operationalization of the maturity of an organisation that wishes to join a collaboration. Only if they are sufficiently aware, have specific qualities to share information or some other relevant indicators that make them mature enough can those organisations join the collaboration.

A3.2.3.3.3 EXTERNAL INTERACTION

The external interaction has to do with the ability of being able to influence and be influenced by other configurations. A collaboration might perceive reality as being such that more stringent regulation is in order to force organisations to make cyber security become part of their risk analyses. This is yet another step for which numerous possible options can be imagined. The external interaction could indeed with other collaborations. But the collaboration could also be with intelligence agencies or law enforcement. All depends on the intentions of the collaboration. A collaboration that wishes to stop botnet attacks might be in need for collaboration with the authorities.

In practice

Kowtha et al. identified some options for external interactions with "emergency services, government, law enforcement, international, commercial, intelligence". But far more can be imagined and it is considered that collaborations would have to increasingly collaborate with each other. An example is the call by ENISA for Computer Emergency Response Teams (CERTs) to collaborate with each other. The CERTs will remain intact, but they would have to collaborate in the form of information exchanges to globally prevent cyber-attacks. (ENISA, 2013, p. iv)