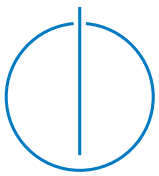
DEPARTMENT OF INFORMATICS

TECHNISCHE UNIVERSITÄT MÜNCHEN

Master's Thesis in Informatics

Mental Models of Cyber Security Attacks and their Influence on the Design of Cyber Security Dashboards

Janosch Maier



DEPARTMENT OF INFORMATICS

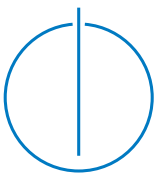
TECHNISCHE UNIVERSITÄT MÜNCHEN

Master's Thesis in Informatics

Mental Models of Cyber Security Attacks and their Influence on the Design of Cyber Security Dashboards

Mentale Modelle zu Cyber-Sicherheitsangriffen und deren Einfluss auf das Design von Cyber Security Dashboards

Author:	Janosch Maier
Supervisor:	Prof. Dr.-Ing. Jörg Ott
Advisor:	Dr. Wolfgang Wörndl
Submission Date:	May 15, 2016



I confirm that this master's thesis in informatics is my own work and I have documented all sources and material used.

Acknowledgments

Several people have contributed to the success of this thesis. I am grateful for all the support I got to realize this work.

Arne Padmos inspired me with a talk on the 31st Chaos Communication Congress about how software should be designed based on its users' minds. He asked me to work with him at Hogeschool Rotterdam. The talks with him showed me different views on my topics and pushed me to pursue this work. The environment at Creating 010 was always pleasant to work. I thank everybody who welcomed me warmly to Rotterdam.

Mortaza Shoae Bargh invited me to do the research at the Wetenschappelijk Onderzoek- en Documentatiecentrum [Research and Documentation Centre] (WODC). He showed me all important aspects of working in the Ministerie van Veiligheid en Justitie [Ministry of Security and Justice] (MinVenJ). With his support, I managed to bypass the challenges that come from the work in a security sensitive environment. Mortaza introduced me to several people involved in the Cyber Security Dashboard (CSD) project. These contacts helped me at all stages of my thesis. I thank him and all coworkers at WODC appreciating my work and supporting me.

I thank Kas Clark from Nationaal Cyber Security Centrum [National Cyber Security Center] (NCSC) to provide me with the data that allowed me to create a CSD. Without any data to show, a dashboard cannot fulfill its purpose. He helped me to transform the raw data to a meaningful form that hopefully will be helpful for the further development of public cyber security.

Prof. Jörg Ott supervised my thesis without exactly knowing what to expect. Wolfgang Wörndl established the contact and advised me during this thesis while I was hundreds of kilometers away. His comments helped me a lot to create a method that allowed to get insight into people's mental models and evaluate the finished dashboard. It is never granted to write such a thesis abroad and nevertheless keep in contact with the home university. I thank them for enabling my academic journey.

During talks with Prof. Thomas Eckert, he provided ideas on the connection of mental models and their visualization. I thank him for this input and the linkage to my second degree in pedagogy.

Seven experts and 20 students participated in my study. I asked them for their time and their help. All agreed without hesitation to support this work by taking part in my interviews or drawing a picture for me in their class. Those people generated the basis

of what this work produces. Without their input, this thesis could not carry any weight.

The Deutscher Akademischer Austauschdienst [German Academic Exchange Service] (DAAD) supported my stay with a scholarship of the FITweltweit program. This scholarship enabled me to focus on my research without worrying about the financial risks of such a journey.

Several people reviewed the many drafts of this work. In addition to everybody mentioned already, I thank Daniel Adam, René Milzarek, Tom Schosser and Leonie Tanczer for their constructive comments.

I am very grateful for my girlfriend, Anna Braukmann, who let me leave and welcomed me back the several times we parted during my time in Rotterdam. I thank her for the support in all the different stages of producing this thesis.

Abstract

A Cyber Security Dashboard (CSD) can be a tool for governmental operators, analysts and managers to monitor the cyber security state of a country. Such a dashboard should be designed to show all the important information at a glance to assess the cyber security state in a meaningful non-distracting way. Dutch governmental organizations want to use a CSD to simplify the work of their cyber security professionals. This research shall analyze how such a dashboard can be designed based on the mental models and data needs of its potential users.

Preliminary to the design, we interviewed seven Dutch governmental employees on their mental model of cyber security and their data needs for a CSD. To gather a point of comparison, we used a part of the interviews – a drawing exercise – with twenty students of a Dutch university of applied sciences. The data suggests that operators and analysts have a deeper understanding of cyber attacks, are more fluent with the domain language and can better describe such attacks. Due to the method, we cannot describe the students' mental model. Nevertheless, they show a difference in the attacks they mentioned. The experts described more social engineering attacks than the novices.

Based on the operators' and analysts' understanding of cyber security, we designed a CSD using an iterative approach. Members of the original interviews reviewed the prototypes during the design process. A final implementation of the dashboard shows data of incidents reported to the Dutch Nationaal Cyber Security Centrum [National Cyber Security Center] (NCSC) and their produced security advisories. The developed program imports data from two Comma Separated Values (CSV) files and uses a custom JavaScript module to draw the dashboard. The JavaScript module was created for the purpose of creating this CSD.

The evaluation of the dashboard shows that it is easy for the users to understand the dashboard content. They see it as a good possibility to review cyber security information. It may help to prioritize governmental efforts fighting cyber attacks. Although built based on the mental model of operators and analysts, managers were similarly able to work with the dashboards. Further dashboard development should include more analytic functionality to increase its usefulness. In the examined use case, a Cyber Security Analytic Tool (CSAT) might replace the CSD in the future.

Contents

1. Introduction

The threat of cyber security attacks has grown substantially over the last years. Current development of digitization enables attacks that have not been foreseeable several years ago. Information Technology (IT) solutions – especially internet based – have increased the speed and reliability of production and services. Devices such as watches, refrigerators, flat irons, heart rate monitors, diabetes measure pens [71] and others are continuously getting internet connections. This leads to the so called Internet of Things (IoT). For example heart rate monitors in the IoT can send health information to doctors or hospitals and help in keeping their patients safe. However an attacker remotely controlling one person's heart rate monitor may act maliciously and harm the device's owner. Current research shows that taking over medical equipment is not just a horror scenario [23].

There are many criminal actors in cyber space with various motivations [52] . Criminal organizations increased their professional knowledge on cyber security and offer "cyber crime-as-a-service" [69]. Digital espionage by foreign governments threatens intellectual property of economic drivers and national defense information. Governments, private organizations as well as citizens are targets of attackers. 67% of Chief Information Security Officers (CISOs) see a rise in threats on their applications [28, p. 7]. They approach the challenge to invest in countermeasures with limited budgets and competing priorities.

1.1. Motivation for a Cyber Security Dashboard

Governments are not autotelic organizations, they are supposed to give structure and security by laws and law enforcement to their citizens. Offering security from criminals while not restricting personal freedom seems to be even more important these days. Governments cooperate with private organizations to collect data and make sure that a country's infrastructure is prepared for cyber attacks [52, 68]. To make reasonable decisions, policy makers rely on cyber security related data that guides them. Laws that are based on wrong assumptions may have unforeseeable effects. Due to broad range and origin of cyber attacks and attackers (Figure 1.1), identification of meaningful data is difficult. However it is important to present correct and meaningful data on cyber security to the policy makers in an understandable way. A dashboard is a single screen

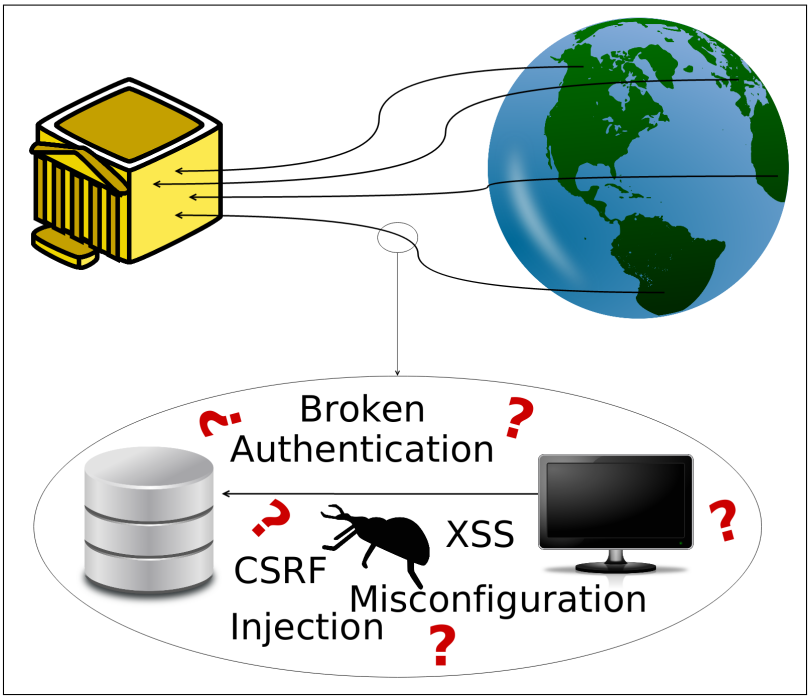


Figure 1.1.: Cyber attacks threaten governments, private organizations and citizens

that provides all important data to a user so that he can make decisions or act based on that data. It may be a suitable way of visualization in the cyber security domain. If policy makers have the most important information visible on one screen, they can easily monitor the cyber security status and make informed decisions. Visualizations enable a dashboard user to easily notice what is most important. For example a traffic light coded system can show whether a certain part of the monitored system needs a special focus. Therefore, a Cyber Security Dashboard (CSD) might be a suitable tool showing the most important information with the aim to guide policy makers responsible for cyber security. Other user groups might be operators or analysts in the field of cyber security. The data visualization in a dashboard can help them to have the important data for their analyses better available. In a first step, the CSD was developed in cooperation with two Dutch governmental organizations and a Dutch university to provide benefits to the cyber security of the public sector. Later similar dashboards may be implemented for policy makers of private organizations working with critical infrastructure.

1.2. Mental Models on Cyber Security

Mental models are internal representations how people perceive systems. Therefore they are the basis of how users interact with systems. The match between the design of a system and the users' mental model influences how the users can work with a software system [15]. In a domain, where there cannot be relied on the assumption that people share a similar mental model, it is important to take this into account while designing a system. In the relatively new area of cyber security – at least within politics – one can assume, that policy makers have different views of cyber attacks than computer scientists or researchers. Therefore a CSD has to be tailored to the target group. This work shall try to provide an overview of possible CSD users and their mental models on cyber security. These models with the identified information are the base for the development of a CSD.

1.3. Research Questions

This thesis shall explore the mental models in cyber security in order to see how they influence the design of a CSD. In particular this work shall answer the following questions.

As there are different user groups in the governmental setting that might work with a CSD, a closer look at those might give valuable insight. Before starting the design of a dashboard, we take a step back and try to understand how those people understand cyber security. This look at the mental model shall then provide the baseline for the later design. A close look means understanding what the users think about cyber security and how their knowledge is structured. This means asking:

Research Question 1 *What are the typical cyber security mental models of potential CSD users in a governmental institution?*

[27, p. 30] suggests that different user groups need different dashboards. However, it does not scientifically proof this at that point. Mental models on the underlying topic of the dashboard might be a reason why different dashboards are needed. Answering research question 2 shall show whether we can support this hypothesis based on mental models:

Research Question 2 *Do different user groups in a governmental institution need different CSDs based on their mental models?*

After identifying mental models, we try to create a dashboard that takes the users' needs into account. This means presenting the data that is important for them while

having their mental model in mind. People that perceive cyber security completely different than we can present it will not have a good chance to benefit from the dashboard. We ask ourselves:

Research Question 3 *How can a CSD built upon the mental model of users in a governmental institution look like?*

When we design a dashboard based on mental models, we need a critical evaluation afterwards. The evaluation has to show whether the design based on mental models proves to be useful. Even if it is useful, the complexity of a mental model study is high. The evaluation shall show if the dashboard outcome justifies the effort of this study. Therefore the last research question is:

Research Question 4 *How useful is the identification of a mental model on cyber security for the design of a CSD?*

Figure 1.2 shows how the research questions relate to each other. As they are based on each other, we try to answer each of the questions in the appropriate part of this work:

- Research Question 1 & 2 – 6 Identification of Users' Mental Models
- Research Question 3 – 7 Design of a Cyber Security Dashboard
- Research Question 4 – 9 Dashboard Evaluation

1.4. Methodology

For the study, we asked seven experts from Dutch governmental institutions to participate. We interviewed the experts in a semi-structured [19] way about their understanding of cyber security and the data of relevance to do their work. A drawing exercise in the interviews tried to give insights in the mental model the experts have on cyber security. We grouped the experts into operators, analysts and managers to see how their mental models differ. We used the drawing exercise also in a classroom setting to get a point of comparison of novices mental models. Based on the findings, we conceptualized the experts mental models and created a CSD targeting the operators and managers. The dashboard design followed an iterative process with the experts giving feedback on the prototypes. After the implementation of the dashboard, we evaluated the final design using an online questionnaire. The questionnaire contained the User Experience Questionnaire (UEQ) [41] as well as open questions to figure out whether the users understand the dashboard.

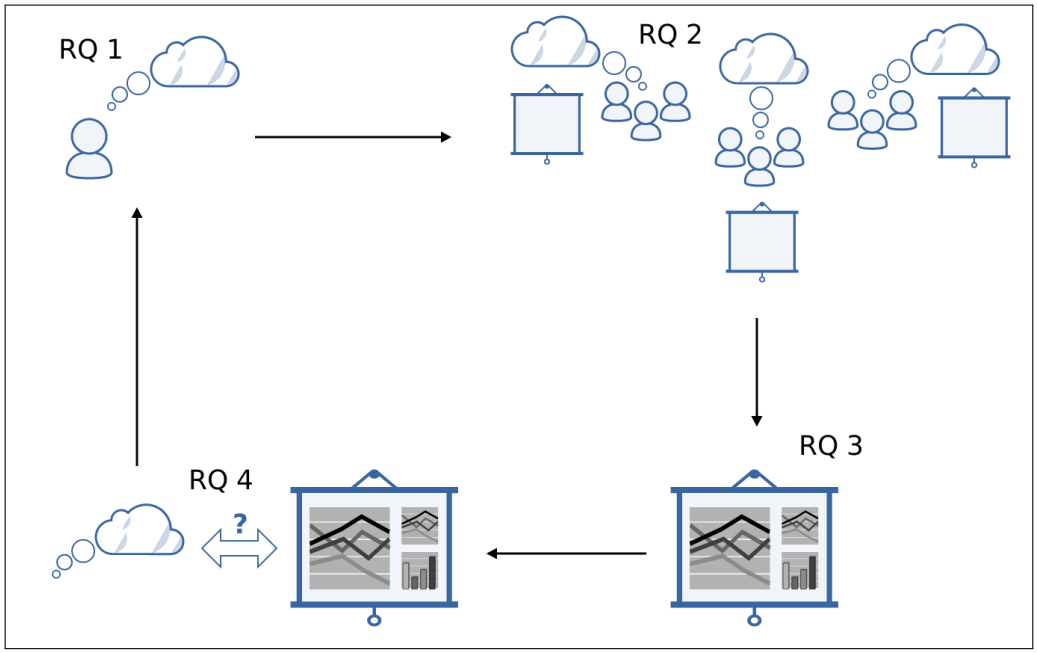


Figure 1.2.: Relationship of the different research questions

1.5. Outline

This work is structured the following: We present related work in the field of CSDs (Chapter 2). Then, we describe the problem setting at the partner organizations (Chapter 3). We try to present our understanding of mental models and depict conceptual models on cyber attacks based on risk assessment methods (Chapter 4). Furthermore we provide information on how dashboards should be designed to be pleasant and meaningful to their users (Chapter 5). We elaborate our interviews with experts and the shortened drawing exercise with undergraduate students. We draw conclusions on the mental models of our potential users from that (Chapter 6). Based on those mental models and the data the interviewees described as most valuable for them, we illustrate the design (Chapter 7) and implementation (Chapter 8) of our CSD. We present how we evaluated the CSD with the original experts and describe their impressions (Chapter 9). Lastly, we conclude this work, summarize its findings and address future CSD design and research (Chapter 10).

2. Related Work

Some work in the field of cyber security dashboards and a lot of mental model work already exists. The following sections shall first give an overview of these areas on their own. Later, we present work that relates to both fields.

2.1. Cyber Security Dashboards

Recently, several companies have started creating interactive cyber attack maps that visualize cyber attacks in realtime [14]. The media is also trying to visualize such attacks [73]. These maps mainly show attacks on honeypots. All traffic going there is treated as an attack, as they do not host any real services. Some of these visualization pages use a community approach to distribute the data collection [17]. One example of such a map is called "Sicherheitstacho"[1] (Figure 2.1) which translates to security-tachometer [16]. The tachometer is the most prominent example of a data representation on a traditional dashboard in a car or plane cockpit. Therefore the security tachometer tries to show important cyber security data. Some of these maps show e.g. the number of attacks originating from certain countries, which might be a useful indicator in a CSD. However these cyber attack maps do not aggregate the data sufficiently to monitor a system properly. It may be nice to watch attacks in realtime on such a map but it is difficult to base any cyber security related decisions solely on that. Therefore they cannot be seen as full CSDs.

The ECIR Data Dashboard[2] tries to gather and visualize national cyber security data gathered by several Computer Emergency Response Teams (CERTs) [43]. They add demographic data of the countries as well as data about the countries' IT infrastructure. Therefore the user can not only see absolute data of incidents but see them in relation to the countries' population or number of IT organizations. One of the challenges of this dashboard is the comparability of the CERT data used. Various organizations release their data differently. If a CERT asks for incidents, an organization that was attacked by 1000 different viruses will report 1000 incidents. If they ask whether they have experienced the problem "virus attack" they will only respond with one yes. Some

[1] http://www.sicherheitstacho.eu/

[2] http://coin.mit.edu:8080/Dashboard/

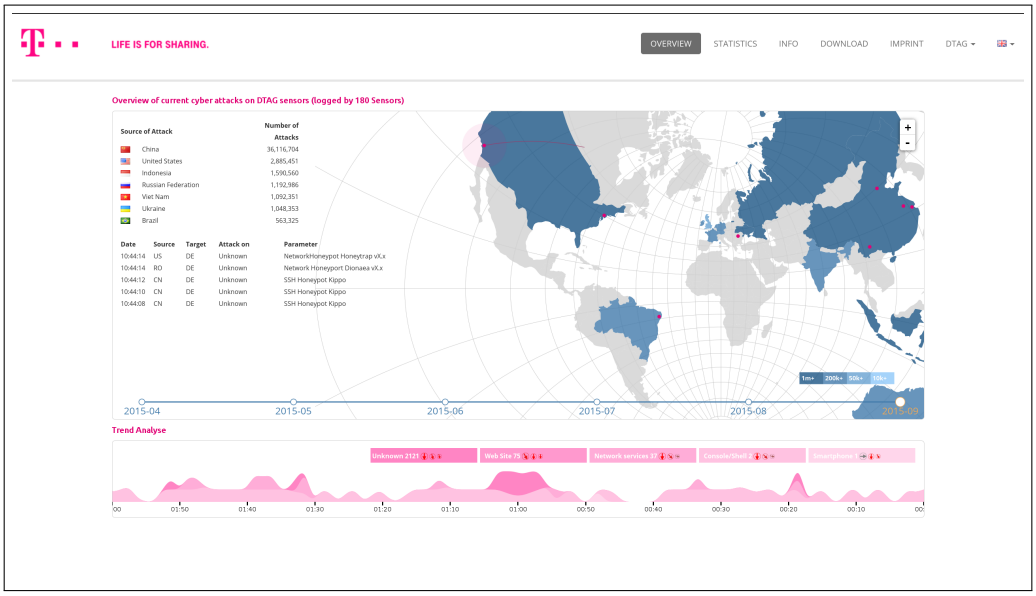


Figure 2.1.: Sicherheitstacho.eu showing honeypot collected data

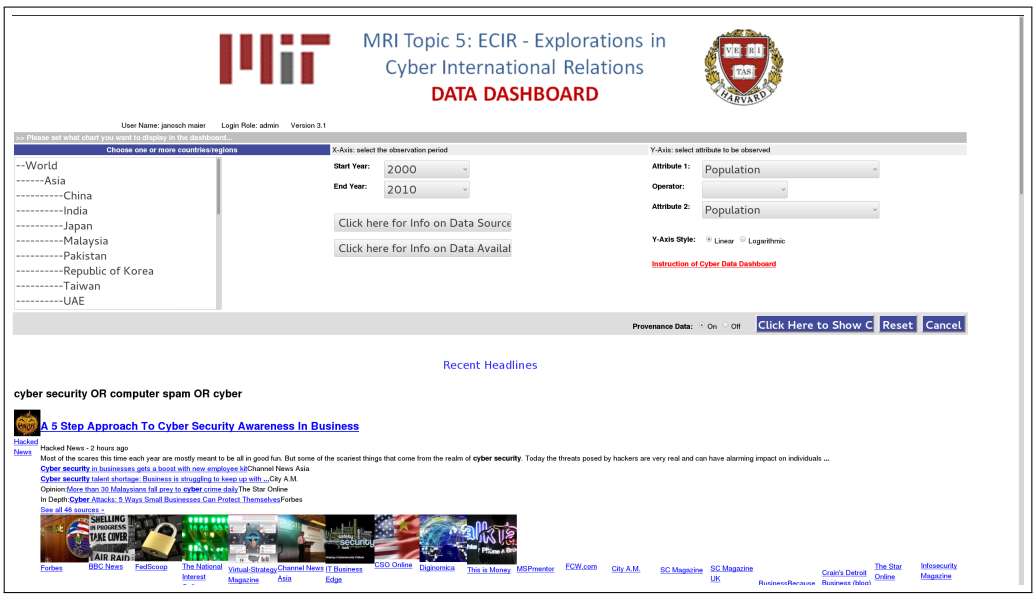


Figure 2.2.: Startpage of the Explorations in Cyber International Relations (ECIR) Data Dashboard

CERTs release their data in absolute (e.g. absolute number of attacks per sector), some in relative form (e.g. number of attacks per sector divided by total number of attacks). Despite these issues, the created dashboard provides benefit, as it not only shows absolute numbers, but sets them in relation to other important attributes. However this work cannot be seen as a dashboard by the definition used in this work (see chapter 5). It does not show its data to the user at a glance. The user sees several menu items to create one specific data visualization for himself, as shown in Figure 2.2. Therefore this is more a cyber security analysis tool than a CSD.

The Cyber Green initiative[3] (see Figure 2.3) of the Japanese CERT shows a cyber health index (Cyber Green Score) for countries. Their creators postulate that – analogous to humans health – not symptoms are important, but underlying causes. Therefore they do not visualize incident data, but try to assess vulnerabilities or misconfigured hosts to create a country-wide index. The Cyber Green Score takes the number of compromised nodes, unwanted traffic and vulnerable nodes into account. The measures are compared against the countries own past data. The goal of this measure is to show at a glance how secure the internet is [36]. Whether this can be really achieved with such a measure seems questionable. Looking at the measure itself, for an observer it is not clear how this score is calculated. Even though vulnerable and compromised hosts are good indicators for cyber security and a comparison with past data seems reasonable, the calculated number contains a lot of simplification. How this measure shall give any of their target groups – analysts, policy makers or Computer Security Incident Response Teams (CSIRTs) – meaningful insight on how to act or decide is not answered.

Comparing the existing solutions that try to visualize cyber security data with the dashboard definition from chapter 5, we notice that none of them qualifies as a proper dashboard. They present the data in a way that makes it complicated to work with it. One just visualizes the data in realtime without meaningful aggregations, another one aggregates all its raw data down to one number. Even if this number properly describes the cyber security state of a country, we cannot see how one can draw any meaningful conclusion from that. The third software provides meaningful information and a lot of analytic functionality such as filtering. However it does not display its information in a dashboard way that provides the important information to the user at a glance.

2.2. Mental Models Research

Mental model research is originally a psychological field. The referred work here mainly contributes to the construction of IT systems by exploring the mental models in the engineering context or for design reasons. In addition to these Human Computer

[3] `http://stats.cybergreen.net/`

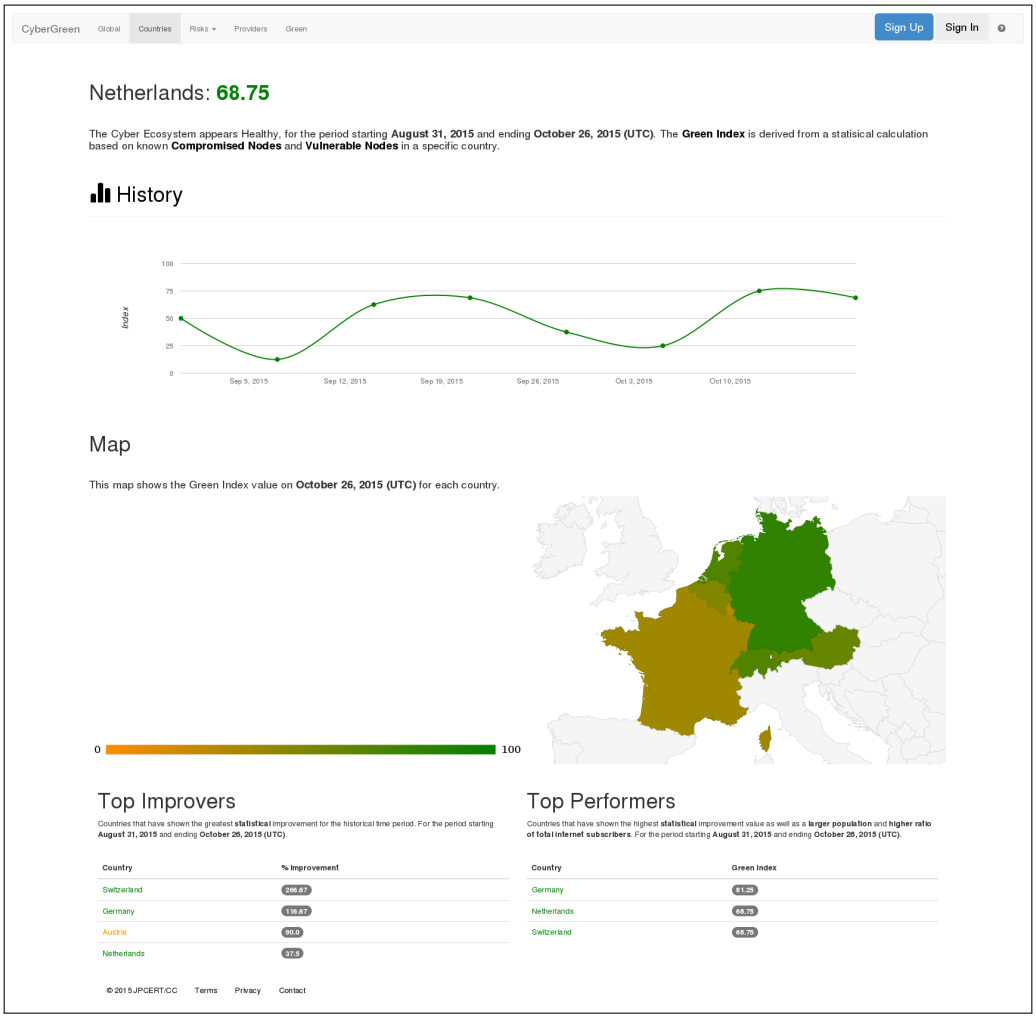


Figure 2.3.: Cyber Green Dashboard

Interaction (HCI) related works, some of the original mental model research is described in the theoretical chapter 4 on mental models.

Staggers & Norcio investigated the mental models of nurses using the Software Package for Statistics and Simulation Extended (SPSSX) [65]. They interviewed five doctoral nursing students who had taken a two hours introductory course and one computer programmer. In comparison to novice users, experts had more developed mental models. They were able to define certain parts of the program and their differences. Experts were able to solve harder problems, committed less errors and worked faster as novices. Some of the users did not try to develop sophisticated mental

models of the system even though longer usage of the program. This research shows, how the existence of a proper mental model concerning an application is beneficial for using the software.

Using a think aloud method, Jonassen and Henning studied mental models of refrigeration technician novices [37]. They asked six participants to fix a broken refrigerator while telling the experimenter how they do their assessment. The participant who found the problem fastest had a sophisticated and highly structured mental model of the refrigerator. He tested the broken refrigerator systematically and found the fault within 45 seconds. The slowest participants tested the system randomly and took nearly five minutes to find the error. His mental model was more linear and less linked. The authors state that the differences troubleshooting ability depend on the mental model of the engineers. The engineers with better model were able to carry out their task better.

A different method to assess mental models has been used by [61]. They let their participants create instructions for other people how to make a bonfire or boil a cup of tea. The participants chose whether to draw a picture, create a flowchart or write instructions. Out of the instructions, the experimenters extracted process diagrams. These diagrams made comparison of the different models easier. To get an exact mental model of the participants was not needed. The diagrams showed differences in the detail of the instructions or the focus (such as safety for the bonfire). These characteristics show the differences in the understanding of those two tasks. Using this or a similar method for the design of a system can help understanding the aspects which are most important for the users.

Mental models have also been discussed in the usability context. The design of a software system should match the users' mental model to help him use the system [15]. The user interface design can support the users mental model using the common design methods: "Simplicity, familiarity, availability, flexibility, feedback, safety, and affordances" [15, p. 4]. Especially familiarity clearly shows the connection to the mental model. A technique that the user knows will be easier for him to replicate even in different environment. The importance of this aspect can be found in chapter 4.

This research shows how meaningful mental models are for the understanding of software and the fluency to talk about tasks. We therefore see the importance of mental models for the design of our CSD.

2.3. Mental Models on Cyber Security

Some research looks at mental models in the area of cyber security.

[2] compares the mental models of computer security risks between novices and experts. The authors use two card sorting experiments in which the 71 respectively 38

participants were asked to choose the category that a certain word belonged to. The categories were Medical Infection, Physical Safety, Criminal, Economical, Warfare or Can't Decide. These categories represent domains where analogies for computer science incidents are taken from. Such an analogy is the one of a computer virus. The words to be ordered were words of the single domains (e.g. Fever, Fence, Theft) as well as IT security related words (e.g. Phishing, Trojan, Exploit). Their experiments showed that novices and experts chose different domains for some of the words. For example experts were the only ones who attached any of the computer security words to the category warfare. The authors argue that talking about computer security risks, one should align its statements or recommendations at the mental models of the novice users. Using metaphors from the areas criminal and physical safety are most promising to be understood by large parts of computer users.

Wash and Rader studied mental models of computer owners in order to identify how and why they secure their computer in a certain way [70]. Depending on what mental model of hackers the users had, they were more or less likely to secure their computer. People who perceived hackers as teenagers trying to show off, were more likely to install firewalls than people who perceived hackers as criminals trying to make money. The authors argue that "[e]ven if the mental models are wrong, they can still lead to good security behaviors and more secure computers" [70, p. 58]. Therefore security specialists should not try to enforce correct mental models, but try to support mental models – even if they are wrong – as long as they lead to good security decisions.

Summers researched cyber security mental models by interviewing hackers. He interviewed 17 male and one female hacker between 20 and 50 that all had more than two years of experience in hacking. All participants were validated to have enough experience by their colleagues. Summers identified the themes Cognitive Patterns, Learning Patterns, Comprehension Patterns, Engaged Patterns and Predictive Patterns being prominent in the interviews. His interview partners described hacking as work in an uncertain domain. Most of the hackers said that they were comfortable solving problems that are not well defined and have a factor of uncertainty. All participants showed personal reflection to build up their mental model and maintain it. The hackers try to shape their mental model and build their understanding of the systems they are working with. They externalized their models with the help of diagrams to easier explore their situation. Hackers are strategists with a high tolerance for ambiguity that use personal reflection to build their mental models [66].

Other work tries to implement agents who simulate a user having a specific mental model with regard to IT security [7] or looking at implications of a certain mental model for privacy in the internet [38]. These papers give an impression of the different possibilities to look at and use mental models in the security domain. However this research is not clearly linked to the design of a CSD.

This work suggests some interesting questions and further research. What system design enforces good user decisions based on their mental models? We try to answer this question in the dashboard domain with our CSD design. We do not address other questions that might arise and are also interesting to think about. Can we foster the hacker attitude in people in the cyber security domain to support reflecting on one's own mental model? This might improve the work in this domain where the possibility to respond to new threats is important. This question and similar ones go beyond the scope of this work.

This chapter provided an overview on mental models and CSDs research. It gives an orientation on what questions are currently in research focus and what is missing in current research and CSD implementation. The following chapters shall introduce the concrete problem setting and provide the theoretical foundation for the later described study.

3. Problem Setting

As chapter 2 showed, the idea of measuring cyber security data and visualizing them is not completely new. Policy makers need an information basis to make decisions. Statistical information can be this information basis [12]. The project partners for this thesis have identified the need of such a dashboard in a preliminary study [11]. This chapter tries to describe the settings at the partners.

3.1. Project Partners

The Wetenschappelijk Onderzoek- en Documentatiecentrum [Research and Documentation Centre] (WODC) is the documentation and research center of the Dutch Ministerie van Veiligheid en Justitie [Ministry of Security and Justice] (MinVenJ). The center is analyzing data in different areas of social security and justice. It is trying to produce meaningful insights that can give advice for policy makers in these areas [50]. For the design of the CSD, the WODC is responsible in carrying out the research.

The Nationaal Cyber Security Centrum [National Cyber Security Center] (NCSC) is the Dutch national competence center for cyber security. It is part of the National Coordinator for Security and Counterterrorism and the Dutch governmental CERT [51]. As coordinator for cyber security, it is giving advice for critical infrastructure operators and also informing governmental institutions and the general public [52, 54]. The NCSC owns data that shall be visualized in the CSD.

Hogeschool Rotterdam is the university of applied sciences in Rotterdam. The university has over 30.000 enrolled students and about 3.000 staff members [33]. Its research center Creating 010 tries to encourage interdisciplinary research [32]. Hogeschool Rotterdam is a research partner in the field of mental model and cyber security research for this project.

3.2. Need of a Cyber Security Dashboard

A feasibility study by Capgemini Consulting, NCSC and WODC showed the need of a CSD. The CSD shall show quantitative measures on cyber security data such as trends of threats. Based on data presented by such a dashboard, reasonable decisions can be

made [11]. Operators might benefit by a real time presentation of current vulnerabilities, such as aggregated Common Vulnerabilities and Exposures (CVE) reports. Analysts could use current trends as starting point for further analyses. Policy makers and managers could see on one screen the current state of national cyber security and its development. The WODC is responsible for research on security data and has the possibility to work on such dashboard creation project. The NCSC has access to data gathered by certain governmental institutions as well as data concerning takedown notices or data from responsible disclosures. This data can be the basis for the CSD. The WODC is responsible for data processing and visualization. In the end, the raw data provided by the NCSC should be made available to them in a meaningful way.

In comparison to IT security dashboard, the term CSD highlights the focus on attacks that are carried out via the internet. For non-technical people it may attract more attention due to the buzzword "cyber" compared to IT.

3.3. User Groups

In this institutional environment, it is important to have several user groups on board. The operational people have access to current vulnerability data and resources to evaluate those. The analysts have the tools to generate a dashboard and the possibility to compare actual with older data. The management needs to see the benefit in the dashboard as an institutional project for granting people and money allocation as well as justifying the dashboard in front of higher management. The initial feasibility study tried to highlight the benefits of the CSD for all identified user groups [11]. Preliminary talks within the WODC and NCSC also suggested that potential users of the CSD belong to the following groups:

- Operational
- Analytical
- Management

Operational personnel is monitoring security data for the day to day business. For example they review new CVEs to ensure a secure own network or to give security advice for others on short term notice. Some NCSC operators produce security advisories for this reason.

Analysts analyze the development of data to look for connections or reasons of certain incidents. They try to make predictions for the future based on previous data and try to produce further information for meaningful decisions.

Strategical people such as policy makers or managers need information to back up decisions such as strategic business orientations, regulations or laws. We named this group managers, as they are prototypical for the strategic part in an organization. The proximity to the raw cyber security data increases from operators to managers. Therefore policy makers might only be an extreme form of the managers that hardly have any contact with the data. Chapter 6.1.5 shows how people of the different user groups describe their job.

[27, p. 30] suggests that different user groups need different dashboards. Our management group compares to what the author calls strategical in his classification. In the security dashboard research this has been picked up by [36], who name CSIRTs (operational), analysts and policy makers (management) as potential users. Whether this can be backed by different mental models is one of the questions (see chapter 1.3) of this work and a key contribution to the research on dashboard design.

3.4. Data Access and Limitations

As the NCSC is a national coordination center, it owns some countrywide data, but nevertheless its own data sources are limited. Currently available data concerns cyber security threats that are e.g. reported to them in responsible disclosure attempts. If somebody notices a vulnerability the infrastructure of a Dutch organization, he may use the NCSC as an intermediary for reporting his finding. This may protect him from the rage of the organization that somebody penetrated their system and at the same time attach more value to the report as it is delivered by the NCSC. Additionally there is data about security advisories (comparable to CVEs) that are created by the NCSC.

The raw dataset we worked with were anonymized in the way that the incident information did not contain any information where the incident happened. However it contained incident data in a form that included the date of the incident, the type of attack, the sector where it happened. As this data might still contain sensitive information that is not intended for public usage, the raw data could not leave the building of the MinVenJ. How this affected the creation of the CSD is discussed in chapter 10.2. Chapter 7.1 explains the data used for the dashboard design.

This work is embedded in practical research of Dutch governmental institutions who have a need of a CSD. We will present some theoretical background in the next chapters to understand on what basis we try to approach this research.

4. On Mental and Conceptual Models

How people think about cyber security depends on several factors [70]. The psychological work on mental models may provide insight on how people perceive such a field. For mental models, there exist several definitions and descriptions from different research fields. This chapter shall provide a basic understanding of mental models and a definition that is then used to describe mental models on cyber security. Several externalized conceptual models exist which may dominate people's thinking in this field. We will present such models at the end of this chapter.

4.1. Definition of Mental Models

The term mental model was first used by Craig in his book "The Nature of Explanation" [13]. There he discusses the following process: Humans translate external processes into internal representations. Then they reason based on this representation. The result of the reasoning can be retranslated by applying them to the external world. The internal representation is the person's mental model. This term was picked up later and is now a widely used term in cognitive psychology. However different researchers used similar but different terms to describe the same, or used the same term but having differences in their meanings.

Norman describes cognitive or mental models as the understanding a user has of a system [58]. Physical variables such as buttons or sliders are mapped to psychological variables. These psychological variables are the understanding of how interaction with the physical variable influences the system. If this understanding of the system does not match the actual behavior, interaction becomes difficult. The complexity of a system increases with the number of physical variables it has. Therefore specialized systems such as single purpose software might be easier to understand as general purpose software such as operating systems. He calls the mental model a user has the user's model, the mental model of the designer the design model and the model that is featured by the system the system image (Figure 4.1). Norman refers to mental models as conceptual models. In his later work he describes several system designs which were not designed with the users' mental models in mind [59]. Systems that highly differ from what the user knows from older systems with the same purpose might lead to users not seeing the full functionality. Labels implicating a certain function of a system

lead the user to build a mental model of the system. If this model does not match the actual system, proper use becomes difficult.

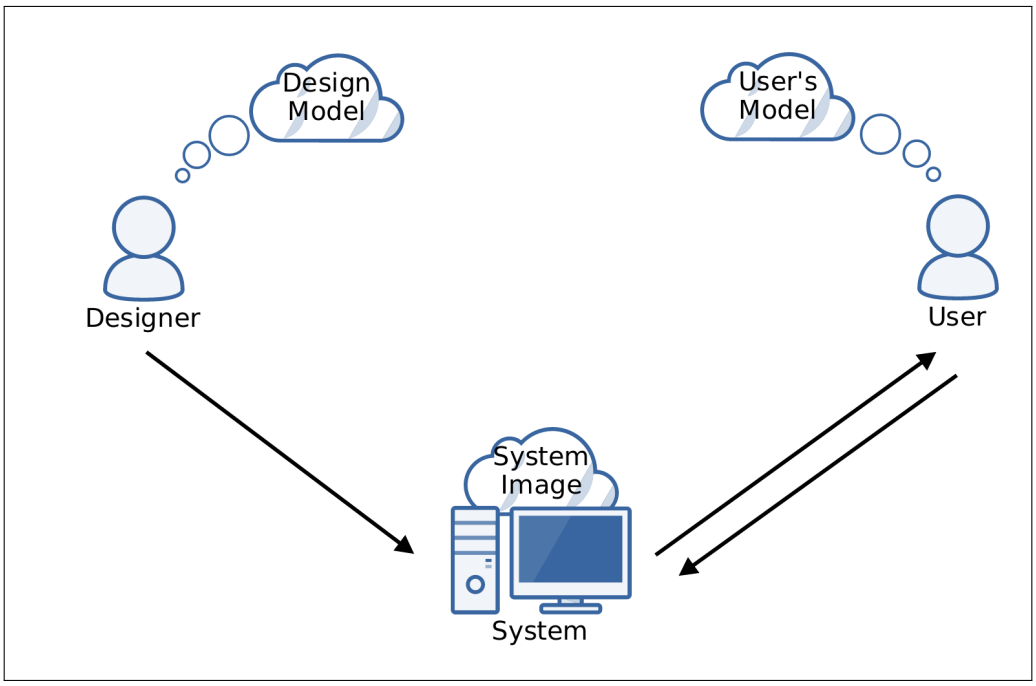


Figure 4.1.: Relation between different models according to [59]

A prominent example of such a system describes a fridge with a fresh food compartment and a freezer [59, p. 14ff.]. There are two control knobs to set the temperature of both compartments. If you look at Figure 4.2, you might think that both knobs control one cooling unit each and steer the both compartments independently. However the system only contains one cooling unit controlled by one of the knobs. The other knob controls a valve distributing the airflow to both compartments. Table 4.1 shows the fridge's manual. This system is even complicated to use whilst carefully studying the manual. One reason for that is that the mental model differs so much from what is actually happening in the fridge.

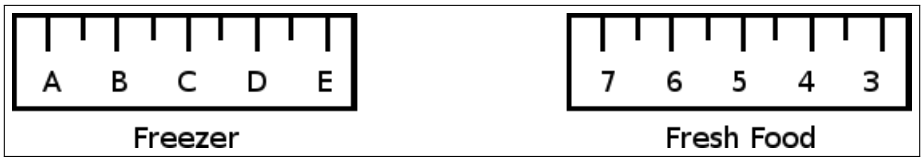


Figure 4.2.: Control knobs for the fridge described in [59]

Normal Settings	C	and	5
Colder Fresh Food	C	and	6-7
Coldest Fresh Food	B	and	8-9
Colder Freezer	D	and	7-8
Warmer Fresh Food	C	and	4-1
Off (Fresh FD & FRZ)			0

Table 4.1.: Fridge controls manual as described in [59]

Internalised	–	Externalised		
Structural	–	Distributed		
Generic	–	Instanciated		
General	–	Specific		
Descriptive	–	Analytic	–	Simulation
Static	–	Dynamic		

Table 4.2.: Model taxonomy by Nielsen

Nielsen [55] formalizes models and proposes to also take into account what model the designer has of the model a user might have of a system. This meta-model adds another level of complexity when talking about mental models. It influences how a designer will try to design a system. Nielsen describes a taxonomy of models with the dimensions, we present in Table 4.2. For this work such a formalization is not crucial. However the internalized feature of certain models is one of the key features of mental models used within this work.

Staggers & Norcio summed up existing research on mental models in the early 90s [65]. According to their summary, mental models in psychology and pedagogy refer to the cognitive structures people build while learning. This can be a network of subjects with relationships in between or analogies to knowledge areas the learners knew before. Exact definitions of the term differ and related terms like cognitive models may name the same thing differently. The definitions have in common that "Mental models are internal representations of systems in particular knowledge domain[s]" [65, p. 601] that are formed through learning or experience. A mental model is something personal. Therefore, it is unlikely to find exact matches of mental models of two people. The models relate to learning and system design. If a system is designed in a way that matches a user's expectation, it is easier to work with the system [65].

A prominent analogy in the computer world is the design of a workspace as a desktop with folders, files and bookmarks. This refers to mental models in the physical world that can be applied to this other domain. The design of interaction on mobile devices

also tries to imitate well known behavior e.g. by using swipe gestures to go to the next page in e-book reading applications. The used gesture is similar to the movement when physically turning a page [72]. Relevant aspects for such analogies are tightly linked to the system design [15]:

- Simplicity: The analogies must be simple to understand
- Familiarity: The analogies must compare a feature to something the user is familiar with
- Availability: The analogies must be something that the user does not need to think about but that is available to his mind effortlessly

For this work, the mental model refers to the cognitive model a person has in mind on a certain domain. We will use the following definition:

> A mental model of a dynamic system is a relatively enduring and accessible but limited internal conceptual representation of an external system whose structure maintains the perceived structure of that system [20, p. 17ff.].

Even though this definition describes a mental model as relatively enduring, this does not mean that there are no changes possible. McNeil shows how the mental models of industrial design students change whilst doing a collaborative project [47]. A learning experience might also be the use of software in a certain domain. Based on the constructivists view, learning leads to building a mental model [39]. This is also the case in learning computer science related topics [4].

The importance of mental models for the design process becomes again clear when looking at how software should be designed. Its features have to be simple and familiar so that the user can understand them easily. Important functions shall be available, therefore visible at a glance and need to give proper feedback to the user [15]. With such a design, the software can probably help the user to evolve his mental model on the domain that reciprocally fosters the usage of the software.

Indi Young proposed mental models as a grouped visualization of user behavior [74]. In the design process of a product, her mental models guides the designer to see the user's understanding. This method relies on ethnographic interview methods [74]. By focusing on the behavior and therefore the needs of a user, the design can provide specific aid to satisfy these needs. This method is used in understanding users in the design process of software [40, 42]. Using the term mental model in this way differs largely from the definition cognitive psychologists have. Her mental models describe a visualization method and not the mind representation of a system. A better name for this method might have been a task analysis [15]: Trying to visualize what goals a user

has for a specific task and trying to match existing tools against this. For this research the mental model design theory has no relevance. We only mention it shortly to prevent possible confusion for readers familiar with this method.

4.2. Conceptual Models on Cyber Security

For this thesis, we take the term conceptual model as a verbalized or illustrated model explaining a domain. Therefore, it is an externalized form of a mental model comparable to what [55] already names an externalized model. Such a conceptual model may or may not match the mental model different people have. The term conceptual shall show that such a model is a concept. It is a model somebody has explicitly thought about.

In research and practice, there exist several methods to assess cyber. Applying these methods in a particular scenario leads to the construction of a very specific security model. Results of the methods when evaluating a cyber security problem are therefore specific conceptual models on cyber security. But also these methods are (more generalized) conceptual models. They show how security risks can be assessed or how attacks relate to each other. Therefore they are a formalized way of thinking about security problems which matches the definition of conceptual models. Those models might provide a starting point when talking about mental models on cyber security attacks, as certain parts of mental models may refer to parts that are formalized in one or more of these conceptual models. How the identified mental models relate to these conceptual models is described in chapter 6.1.5.

4.2.1. ISM Risk Management

The Information Security Management (ISM) system guides through a risk management process that tries to identify and assess risks so that there can be proper risk mitigation. The ISM process is formalized in the ISO/IEC 27000 series [35]. The Information Systems Audit and Control Association (ISACA) defined risk management as

> the process of identifying vulnerabilities and threats to the information resources used by an organization in achieving business objectives and deciding what countermeasures, if any, to take in reducing risk to an acceptable level, based on the value of the information resource to the organization [34, p. 78].

IT security risks describe the risk of threats which use vulnerabilities to get access to assets. Accordingly, assets, vulnerabilities or threats are only parts of risks. Where there is a vulnerability without a threat or the other way round, there is no risk. Assessing a

Measure	Example
Avoidance	Avoid data getting stolen from laptops by not giving out laptops to employees
Mitigation	Mitigate risks related to computer viruses by installing anti virus programs
Transfer	Transfer the risk of a fire destroying assets by insuring against fire
Acceptance	Accept the risk of an earthquake destroying assets
Elimination	Eliminate the risk of a vulcano destroying assets by relocating to an area without vulcanos

Table 4.3.: Risk management measures

risk means evaluating its likelihood and potential damage. These evaluated risks are then documented e.g. in a risk register. For each assessed risk, management means implementing one of the measures from Table 4.3. The whole process is visualized in Figure 4.3.

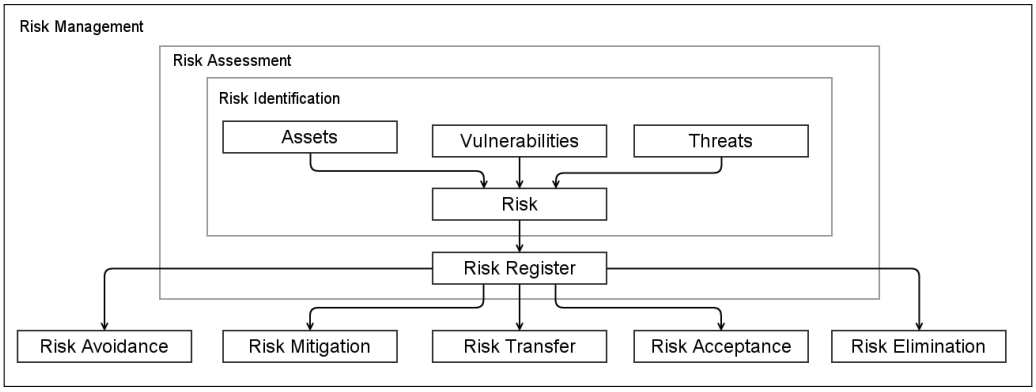


Figure 4.3.: Schematic version of the ISM risk management process

4.2.2. Attack Trees

Schneier proposed attack trees as an easy understandable method to evaluate attacks [64]. An attack tree consists of a root node representing the complete attack. Child nodes depict sub-attacks. An attack is successful if one of the children attacks is successful (OR-nodes). In contrast, attacks of AND-nodes are successful when all sub-attacks are successful. Each node can be labeled with values containing information about the attack such as the need for special tools or the cost to carry out the attack. Figure 4.4 shows a simple attack tree for getting root access of a webserver. To gain root access

on a webserver, an attacker can either try to get admin access via Secure Shell (SSH) or exploit a vulnerability. To exploit a vulnerability, the vulnerable software needs to be installed on the webserver and the attacker needs an exploit for it. If an attack tree is constructed correctly, it helps security officials to assess a potential attack. If no sub attack of an attack tree is possible, the parent attack is also not possible. Attack trees have been further formalized in [67, 44]. The risk that comes from a certain attack can be evaluated by looking at the steps needed to perform the attack.

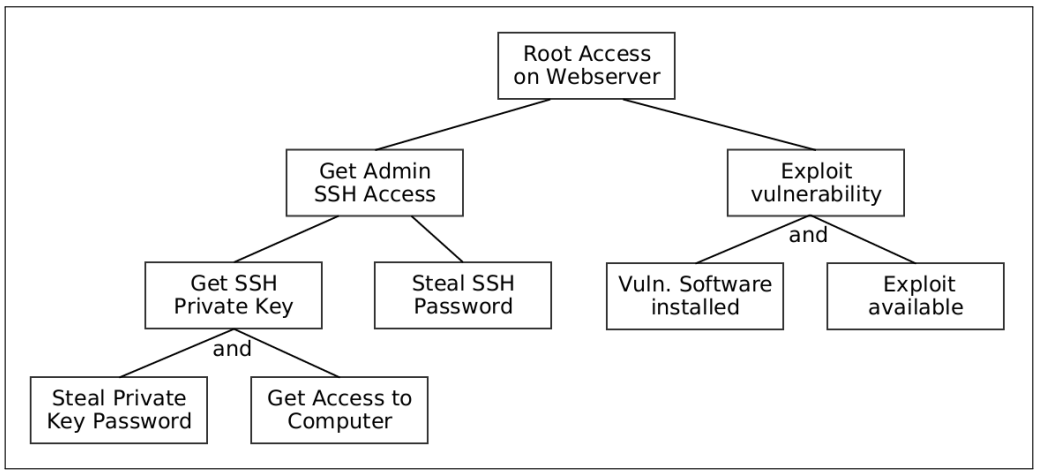


Figure 4.4.: Attack tree for getting root access on webserver

4.2.3. CSAN Core Assessment

The Cyber Security Assessment Netherlands (CSAN) describes in its core assessment how manifestations threaten interests when threats outperform resilience factors [54]. Figure 4.5 shows how manifestations exist in the triangle of interests, threats and resilience. This cyber security assessment takes only risks of cyber attacks into account and is therefore more narrow than the ISM risk management. However the interests are comparable to the assets that shall be protected. Threats such as cyber criminals using malware are similar to certain threats of the ISM risk management. Resilience is a combination of the presence of vulnerabilities and measures that mitigate those. Therefore this is a factor that relates to different phases of the ISM risk management process: The identification (vulnerabilities) and the real management of risks (measures). Manifestations describe real incidents. An interest is damaged, because of a carried out attack from a threat where the resilience is insufficient.

Based on the theory of mental and conceptual models, we created our study that we describe in chapter 6. For the design of the dashboard based on the study results

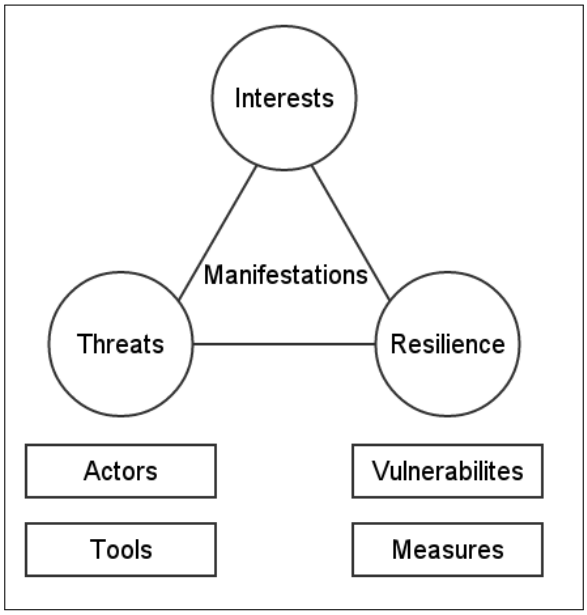


Figure 4.5.: CSAN core assessment on cyber security

we need some more theoretical background that is introduced in the next chapter on dashboard design.

5. Dashboard Design Theory

Besides the theoretical background on mental models, dashboard design theory is utterly important for the design of a CSD. This chapter introduces the concept dashboards and explains how a dashboard should be designed.

5.1. Definition

Originally, a dashboard is a piece of wood on a carriage or other horse pulled vehicle that should protect the drivers feet from mud thrown up by the horses feet [30]. Later, within cars, they developed from design elements to plain and functional parts containing the instruments for measuring the state of the car. This includes showing data of speed, fuel level or motor rotation. With this information, one can operate a car easily. He can e.g. make sure that he is not overspeeding and fill the tank before running out of fuel. Dashboards in the IT try to mimic these characteristics. For example the dashboard of the blog software Wordpress[1] shows important information about the state of the webpage to the administrator. As you can see in Figure 5.1, there are red highlights for pending updates. With respect to the security of the software those alerts are very important [22]. Those highlights are even visible on this scaled down picture. The dashboard also shows recently updated articles or the most recent comments.

Few [27, p. 26] defined a dashboard the following:

> A dashboard is a visual display of the most important information needed to achieve one or more objectives; consolidated and arranged on a single screen so the information can be monitored at a glance.

We will use this definition as the basis for dashboards for this work. As you can see in the example, this dashboard helps you achieving your goal of a properly functioning webpage. You instantly see, if you have any actions to take such as updating plugins or reviewing comments. Therefore the Wordpress dashboard enables the administrator to operate his webpage easily in a similar way that the car dashboard supports the car driver.

[1] `http://www.wordpress.org`

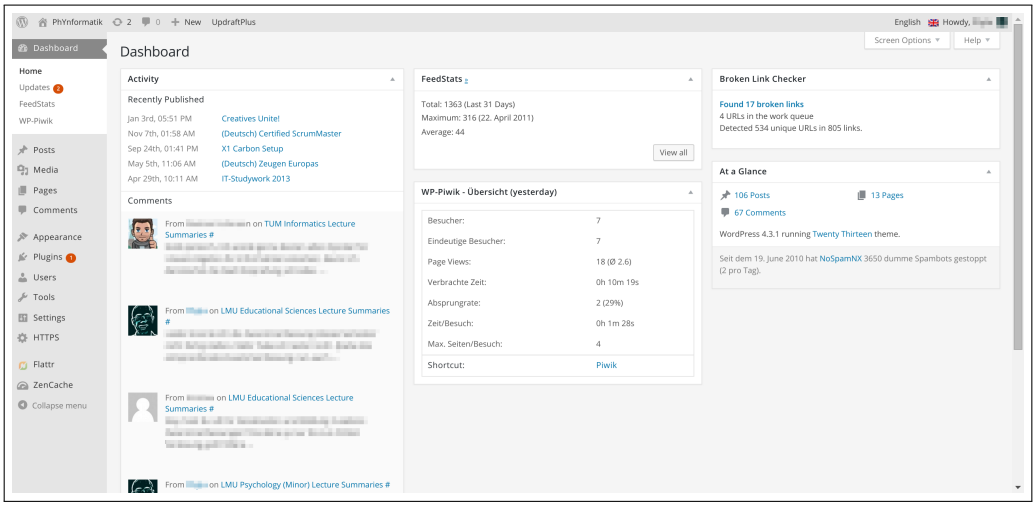


Figure 5.1.: Wordpress administrator dashboard

As one can see, a dashboard is not the same as an analytic tool. The main reason for a dashboard is monitoring. Deep analyses that rely on the comparison of many different kinds of data or the possibility to have specific data queries displayed are things that a dashboard cannot provide. For those tasks, one needs a fully fledged analytical tool. [26] gives some hints on how such tools which he names faceted analytical displays provide insight using proper visualization techniques.

5.2. Visualization of Cyber Attacks

For a CSD, the dashboard definition implies that the most important information on cyber attacks needs to be visualized for its users.

People base decisions on data which provides information. In the field of cyber security attacks, the amount of data is so big that humans have problems processing the raw data manually. It is therefore important to generate meaningful measures such as Key Performance Indicators (KPIs) out of collected raw data such as honeypot, netflow or incident report data. The data may focus on the attackers' side such as the number of attacks carried out or the usage of exploits. It might however also relate to the success of defense mechanisms such as blocked attacks. These measures should be helpful to make decisions based upon them. For this reason all such information like attack rates, origins, used exploits, implemented security features, detected or stopped attacks are treated as possibly useful measures.

The CSD will combine visualizations of these measures and present them to the user.

In the following design guidelines, we describe possible forms of visualization and how to make use of them.

5.3. Design Guidelines

While developing a dashboard, certain design guidelines should lead the design process. These guidelines are based on psychological principles that influence the human perception. Similar information is not always perceived similarly. Compare the color of the squares A and B in Figure 5.2. Square A seems to be darker. However both squares have exactly the same color (Figure 5.3). The mind is tricked to perceiving square A darker. When the brain detects the same color on two objects, but one of them lies in the shadow, it assumes the material of the illuminated object to be darker. Studying the mechanisms of perception and information processing resides in several psychological fields. A short description in a computer science context can be found in [49].

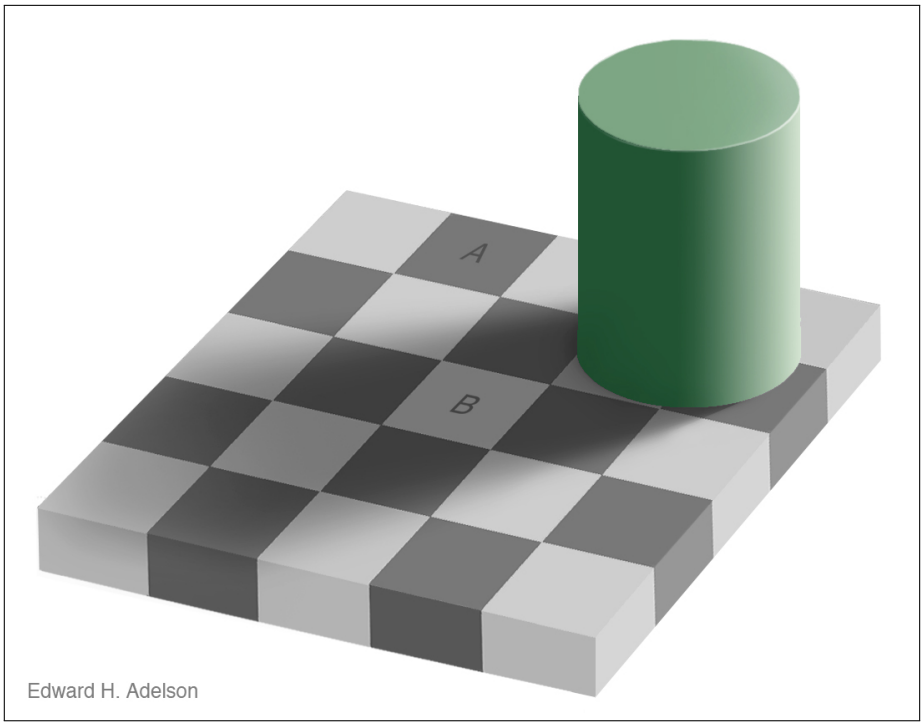


Figure 5.2.: Checkershadow Illusion [1]

For this work we will not take a deep look at the underlying psychological principles, but focus on what design principles are important and only some of the reasons for

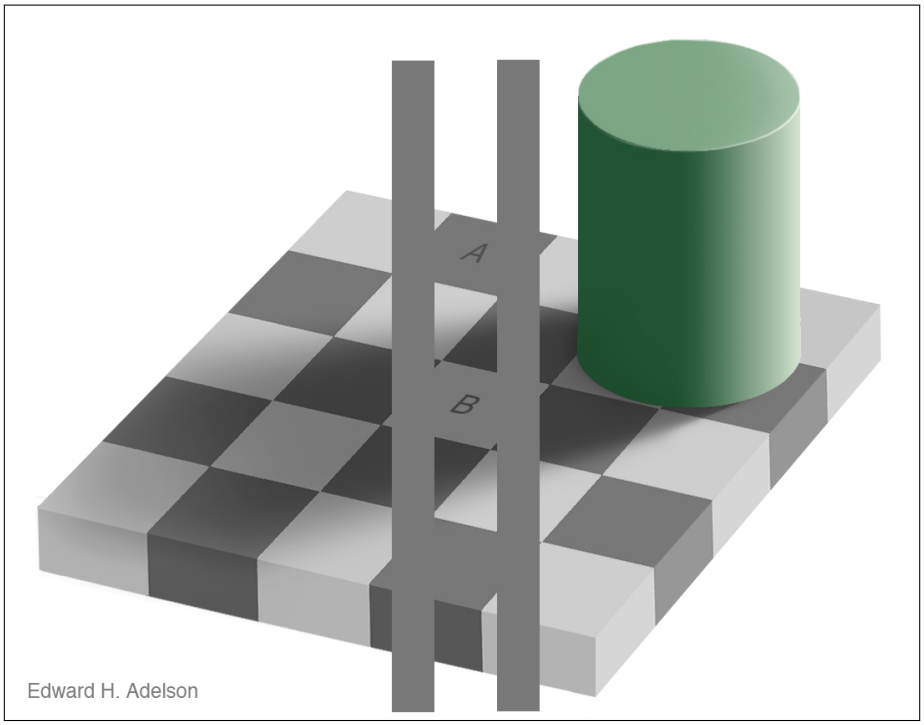


Figure 5.3.: Checkershadow Illusion Proof [1]

those. For a more thorough description of these principles and methods, we recommend reading [27].

Combining the different visualization types with highlighting methods, a dashboard should be able to focus the viewers' attention on the important information and enable them to make decisions based on this information easily. For the developed CSD, chapter 7 shows how we used these principles to create an easy to use dashboard providing meaningful information.

5.3.1. Simplicity

Simplicity and a clutter free design are important to identify important information [8, 27]. Everything visible in the dashboard will draw some attention of its viewer. Therefore it is important to keep the part that does not carry information as small as possible. [27, p. 84] calls these parts non-data pixels. Take a blank background and categorize all pixels that have a different color than the background either as data pixels or non-data pixels. Data pixels encode information. For example, they can be parts of a graph, textual information or indicator icons. The non-data pixels are e.g. frames,

grids or pictures that do not transport any information. When the non-data pixels are minimized, it is easier for the viewer to focus on the data pixels, that actually transport information. He is not distracted by unnecessary objects.

5.3.2. Information Visualization

When the dashboard is built in a way that minimizes non-data pixels, the next question is, how the data pixels should visualize data. Several possibilities exist that are suitable in different situations [27].

Graphs

Graphs have a visual explainability that enables viewers to compare different measures easily. Prominent examples are a bar graph where different numbers can be compared easily or line charts that may show a development of a single measure over time. A less known graph is the bullet graph introduced by Few [27]. It shows the current value of one measure with the possibility to give comparative measures such as a mean value and qualitative areas (Figure 5.4).

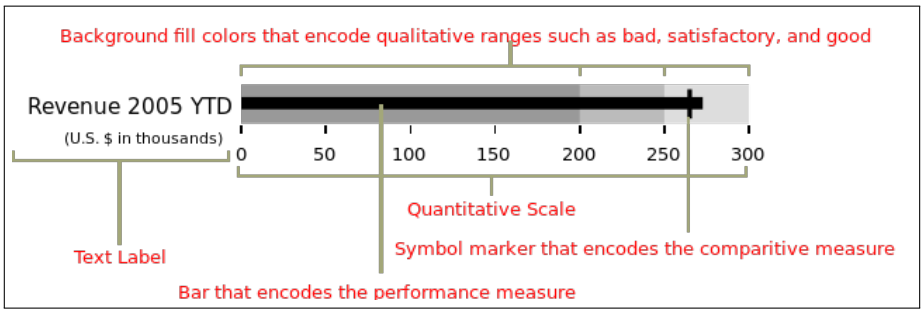


Figure 5.4.: Bullet graph with explanations [25]

Icons

Icons allow a symbolic representation of information. For example tiny arrows can show whether a measure has been rising or falling in the past where this is not visible from the graph itself.

Text

For categorical descriptions or labeling, text elements are needed. In Figure 5.4, the text on the left side denotes the meaning of the measure. Without this text, the whole

graph would be meaningless. When there are important numbers that stand out on its own, it is often easier to write them down instead of looking for a different meaningful way of representing. Artificially creating a graph for only one number contradicts its utility. Depending on the information, a group of numbers within a table may be a good visualization as it allows the viewer to see absolute numbers easily.

Images

Images should be used scarcely and only when they provide important information. A useful image might be a network plan on which devices are shown. In the case of an incident that happens at a particular device or part of the network, the corresponding part is colored. This can help an operator to easily identify which devices need a closer look when working on the problem. Pay special attention to non-data pixels when using images in a dashboard.

Drawing Objects

Arrows, curly braces or other drawing objects may show relationships between different parts of the dashboard easily. They may identify orders or show which individual numbers are part of a more summarized visualization.

Organizers

Organizers can be lines separating different parts of the dashboard or tables that guide the viewer. They may show which parts of a dashboard are categorical values (e.g. table headers) and which are actual data points (e.g. within a table).

5.3.3. Information Highlighting

To guide the viewer of a dashboard to certain points on the display, a designer can use several techniques.

Colors

Using colors is the most obvious technique. Coding a critical event with red, a minor alert yellow and normal status green will be easily understood even without prior training. One factor is the knowledge about those colors that people know from their normal life such as encounters with traffic lights. More important, colors draw attention when the dashboard is otherwise designed unobtrusive. One problem with color coding

information is that people with color blindness might have problems to extract all the information transported by the color or even recognizing that something is colored.

Saturation

In some situations color cannot be used for differentiation. Printing the dashboard in gray scale might be such an example. Saturation can be the solution as it is a more color-blind friendly version of coloring. Using the same color with different saturation can encode similar information as coloring and is also visible in gray scale prints and to people who have problems differentiating colors. When in doubt if saturation makes a good enough distinctive feature, look back at the checkers board in Figure 5.2.

Shapes and Orientation

Different shapes usually mean different things. Shapes do not draw attention as colors but can be easier differentiated as long as the shapes are sufficiently different. Therefore a combination of color/saturation and shapes is beneficial in many cases. When using the same shape – such as an arrow icon – the orientation of the shape might create the meaning. An arrow pointing up makes a different impression than an arrow pointing down.

Position

Position is another possibility of differentiation. Again a combination with color or saturation for highlighting and differentiation is suitable. A real life example is – again – the traffic light. Normal people get their attention drawn to it by the changing color. For color-blind people the position encoding is more important. Even if they are not able to differentiate the colors they can clearly see whether they can drive or not. This possibility enabling them to make a decision based on a short look is also important for a dashboard.

Motion

The last resort for a designer if he wants to draw attention is motion. Due to the human development, our brain is trained to focus on motion really fast. In earlier times, milliseconds to detect a snake on the pathway in the jungle could make the difference between life or death [62]. Therefore an animation will draw the viewers attention. This should be used scarcely so that the viewer is not flooded with too much attention requests. Graphs that draw themselves on the screen are nice to look at, but do not

provide any insight and only draw the attention of the viewer because of aesthetic reasons.

After the last two chapters, we now have the theoretical background to begin with the study which tries to identify mental models of our potential dashboard users. This is a vital part of this research as it leads to the design of the CSD afterwards.

6. Identification of Users' Mental Models

To get a knowledge about potential users' mental models, we conducted expert interviews. For a broader view, we reused a part of the interviews – a drawing exercise – and asked students to do this exercise during a lecture.

6.1. Expert Interviews

In chapter 3.3, we described the different user groups for the dashboard. We follow the idea from [27, p. 30] that different users need different dashboards. Reasons for that could be the need for different data but also a different mental model (see chapter 4). We interviewed potential users, trying to discover their mental models on cyber security attacks as well as relevant data for them. These findings are later on used to design a meaningful dashboard.

6.1.1. Sampling

Seven people (six male, one female) from two Dutch governmental organizations took part in the expert interviews. The interviewees were $M = 42$ ($SD = 6.3$) years old. Two participants belonged to each of the groups "operational", "analytical" and "management". One person stated, his job belonged to both the analytical and management group. Their jobs all included working in the cyber security domain. In order to gurantee the anonymity of our participants and for better readability, we refer to all our participants with the male pronouns.

6.1.2. Design

The interview consisted of questions grouped into three blocks and some demographic data asked in a questionnaire. The interview followed a semi-structured way [19]. Each block contained several questions that were asked after each other. If the participant did not answer the questions properly on their own, the interviewer tried to find follow up questions that would lead to a satisfactory answer. Where needed, the interviewer diverted from the guide to explore topics that seemed highly relevant to the interview partner.

Job Related Questions

The interview started with some questions concerning the participant's job. One of the questions was: "Can you please explain your job to me?" These questions mainly tried to check whether the classification in one of the three groups ("operational", "analytical", "management") fits. Another question in this block was: "What cyber security data do you regularly deal with?" This question tried to gain insight of what data is already available for the CSD and whether that data is sufficient to match the data needed for the CSD. We asked these questions of the beginning of the interview to help the participant to recall his work. This should enable them to answer the following questions according to their personal experience. We did not control the priming effects of these questions as this was a desired outcome.

Cyber Security Attacks

The second question block focused on mental models. As described in chapter 2, think aloud and drawing exercises can help to understand a user's mental models. Our method was a mixture that asked the participants to draw the message flow of two cyber attacks into a drawing template (Figure 6.1) while explaining their thoughts. The setting was as follows:

> Alice (A) works for a bank. Her regular work relies on accessing data from an application system (AS) on a bank application server. She can access this server via the internet. Look at this example (Figure 6.2): Alice makes a request to the application system. There her request is processed and the answer is sent back to her. The arrows describe where messages are sent. In this case this is the request and the response. Mallory (M) is a hacker that does not like the bank. He possesses a malicious system (MS), that he can use for attacks.

Both attacks described by the interviewees should enable Mallory to steal data from the application system. In a second step the participants drew counter mechanisms for both attacks into their pictures. This question block ended by looking at measures that could show whether the counter mechanisms worked correctly. One of the questions asked in this section was: "Please review your attacks. What can be done to deny the attack? Please draw additional messages or nodes/tools/hardware in your drawing. Please explain in detail how this prevents the attack."

Figure 6.1.: Interview drawing template

Cyber Security Goals and Data

The last interview block tried to figure out what data should be visualized in a CSD. Loosely oriented at the Goal Question Metric (GQM) method [3], first each person described his or her main cyber security goal (business goal). Then, we asked for three short term goals that supported their main goal. For each of those goals the participants then described questions to ask and in the end metrics that could answer these questions. One of the questions in this block was: "For each of the goals, that you have mentioned. If you would ask a colleague, if the organization has reached the goals. What questions would you ask?" Afterwards, we repeated the questions the interviewee had provided and followed these up by asking: "What measurements would you use to answer these questions? Please explain why you would use exactly this measure."

Demographic Data

The interviewees answered some questions on demographic information such as age or gender using a questionnaire. The questionnaire also asked them to assign themselves to one of the three groups "operational", "analytical" or "management". Appendix A shows the used questionnaire.

Figure 6.2.: Interview drawing example flow

6.1.3. Process

For each interview a meeting was set up with the participant in a Dutch governmental office. We reserved a time frame of one hour for each interview. Before starting the interviews, the interviewer introduced himself, the purpose of the study and the structure of the interview. The participants introduced themselves and agreed to taking part in the study and especially that the interview would be recorded. After the recording was started, we discussed all the interview questions. After the questions, the recording was switched off. The experimenter asked the participants if there was additional information they would like to add now that the recording was switched off. We offered this possibility, so they could tell about important confidential information that was needed to understand their answers. Then they filled the demographic questionnaire. To end the session, the experimenter asked how they liked the interview or whether they felt uncomfortable at any time.

The interviews took place between November 3rd and November 23rd 2015 and lasted between 24 minutes and 1:01 hours. All interviews were conducted by the same interviewer.

6.1.4. Data Analysis

For the analysis of the interviews, we used the qualitative content analysis according to Mayring [45, 46]. We structured the data into different analysis units. These units

Analysis Unit	Codes
job title	research, management, cert, development
typical day	read, write, meeting, partners, analysis, development, coordination
typical data	incidents, netflow, honeypots, vulnerabilities, malware, topography, actors, exploits, infections, dns, whois
recommendation	no, colleagues, partners, public, policy makers, government

Table 6.1.: Codes for the questions on the participants' work

usually corresponded to one question. Few units spanned more than one question. Some questions were covered by more than one analysis unit. For each analysis unit, we looked at the first interview and specified categories that the answer fitted in. Then, we tried to place the answers from the other interviews into those categories. Where the categories were not yet sufficient, we added a new category. For new categories, we went back to previous interviews to see whether any answer also fitted into this category. For the most cases the categories were non-exclusive. So an answer can belong to one or more categories. If a person did not answer the question, there may be no category attached to the analysis unit for this person.

The interviews were all coded by the author of this thesis. There was no second independent coding to compare the reliability of the coding scheme. Table 6.1 shows exemplary the possible codes for the first questions. The differences in the interviews were sufficiently large to have highly distinct categories. We present several extracts from the interviews in the next section 6.1.5.

6.1.5. Results

Due to the sensitivity of the study it is not possible for us to publish the full interviews. We hope that the excerpts presented in this chapter suffice to give an impression how the interviews were conducted and how the categorization originated from these interviews.

Group specification

To confirm whether the different user groups might need different dashboards, we first verified the user groups by their job description. This is a prerequisite to talk about user groups and not only about single users. See table 6.2 for the detailed results of the job descriptions.

One operational person (person 6) is handling incidents, the other one (person 5) said that his job position is a researcher position, but he is mainly doing development.

#	group	job_title	typical_day
1	analytical	research	read, write, meetings
2	analytical	research	meetings, analysis
3	analytical/management	management	Meetings, coordination
4	management	management	read, write, meetings
5	operational	research, development	develop, write
6	operational	cert	read, develop, coordination, partners
7	management	research, management	coordination, partners

Table 6.2.: Job description of the interviewees

The two participants who identified themselves as analytical people described their job position as a researcher. All participants that ticked the management field on the questionnaire also described their job position as a management position during the interview. One of those people (person 7) also mentioned his position to be partly a research position. Person 3 ticked both the analytical and the management field on the questionnaire, but only described his position as a management position.

The typical days of the operational people includes software development. No participant of the other groups mentioned this. The analytical people were mainly researchers. All people who classified themselves as managers on the questionnaire also told that management was their job title or at least part of it.

This suggests, that the classification into the three groups is appropriate. The members of a group described their job position similarly. Several tasks are mainly or solely used in one of the groups. Therefore a detailed comparison of the cyber security mental model and their data need seems reasonable.

Description of the Cyber Security Mental Model

The first cyber attack described by the operational and analytical people were all phishing attacks. The person who described himself as analytical-management and another management person described social engineering attack that were no phishing. One management person described a Distributed Denial of Service (DDOS) attack. Except the DDOS attack, all attacks were suitable to steal data from the application system. The operational and analytical people were able to describe the attacks more detailed and more fluently. Most participants used some type of technical terms and also used them correctly.

A description of analytical person 2 of a phishing attack is as follows (P denotes the participant, I the interviewer):

P: So, what I think is a very simple attack or what is shockingly one of the most common attacks, Mallory will send an e-mail to Alice. So that's one. Mallory sends an e-mail to Alice. This e-mail looks identical to the one from the bank. It has the bank's logo. Has the banks everything, house color, etc. Everything looks just like the bank. And it says: Mallory [sic!], your account has been attacked, but, you know, we have taken measures to secure it, but you need to login and make sure, everything is secure. So click on this link and you login. When she clicks on the link, she doesn't go to the bank, she actually goes to his, Mallory's malicious system, which looks identical to the bank. Maybe, there is one letter difference. I mean, it's also shocking, how easy that is to spoof. So, instead of Rabo Bank it will be Robo Bank. She doesn't notice that, because, again, it looks, the website looks identical to the bank. The same colors, the same everything. Just the way, she recognizes it. And there is a place to fill in your password and username. At that point, as soon, as she fills in her password and username, two things happen. The password and username is send back to Mallory. And Alice is send to the real bank. Then she is on the real bank. Everything looks normal. Her password and username is send to the real bank. She logs in, she looks in, she checks, everything looks ok. Nothing has happened. Ok, she can go back to bed. But now, Mallory has the username and password, so he sends that to, he uses that to login to the actual bank and transfers all the money to his own account or does whatever.

This description is very detailed. Each of the steps corresponds a line in the person's drawing (see Figure 6.3). The participant guided through the process of how an attacker creates a phishing e-mail that will trick Alice into clicking a malicious link which in turn gives Mallory access to her login data. In contrast (see Figure 6.4) the following description of the analytical/management person 3 is not only less detailed but also lacks important information of the attack. It describes a social engineering attack via telephone but does not say why Alice would be inclined to give Mallory her password.

P: Oh. The easiest one is to call Alice and say: "Hi, I am the helpdesk of Microsoft".

I: Ok. Then just make a line and number it one.

P: One. Two. This is her password. Now he is Alice. And he can do whatever Alice does.

I: Ok. Can you describe it a bit more in detail? When he says, he is the helpdesk from Microsoft.

P: He can social engineer into giving her credentials to him. And then he can just, as the server has no clue and he can – I suppose – remotely log on.

> And being, pretend to be her and just has her user rights and do whatever he wants.

The attack description of a DDOS attack is an interesting one to look at, as it is totally different from what all the other interview participants described. It is also the only attack that does not allow Mallory to steal data from the server. The description leads to the impression, that person 7 does not fully understand the effects of a DDOS attack. The drawing suggests that the person also tried to include features of a social engineering attack, as the application system should be used to trick Alice to enter sensitive data (see Figure 6.5).

> P: Well, there would be a line of course, Mallory to the malicious system.
> I: Number it with one, if it's the first message.
> P: I think, this would be number two. Well, probably, Alice would get that information and give some feedback to that, which would lead for instance.
> I: Can you elaborate that a bit more? So, what kind of attack is it? How do these messages influence what Alice sees, or...?
> P: Well, Mallory would give some input for the malicious system to start the attack. Then the system would try to hack or break into the application system. Of course, disguised. So Alice sees something, but does not realize, that it's malicious attack, or it's a malicious question or a malicious query. She then gives some input to the application system to send out information which would get back to the malicious system which would get back to Mallory.
> I: And can you tell. You said the malicious system would hack the application system, which is request two. What kind of attack could this be?
> P: For instance a DDOS attack. Or?

For the second attack, the field of attacks is a bit more widespread. The interview participants mentioned two injection, two Man-in-the-Middle (MITM), two social engineering and one malware attack. Again the descriptive quality of the management people was lower than for people of the other groups. Operational person 6 described that you first look for vulnerabilities on a system and then exploit them for example using a Structured Query Language (SQL) injection (Figure 6.6). He then elaborates other attacks that could be revealed by a vulnerability scan.

> P: Yah. You can of course scan and then search for SQL injection or something like that.
> I: And how does this attack work?

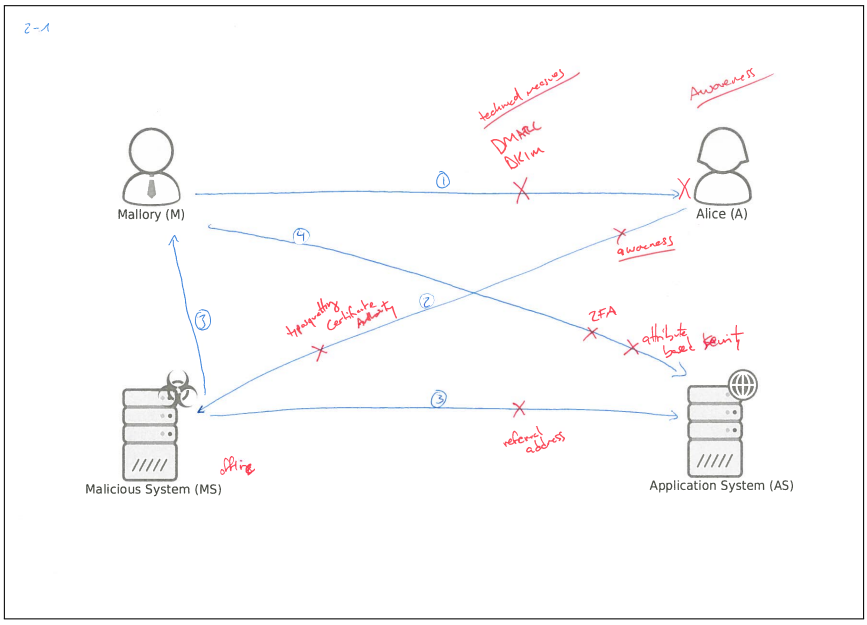


Figure 6.3.: Drawing of attack 1 for analyst 2 describing a phishing attack

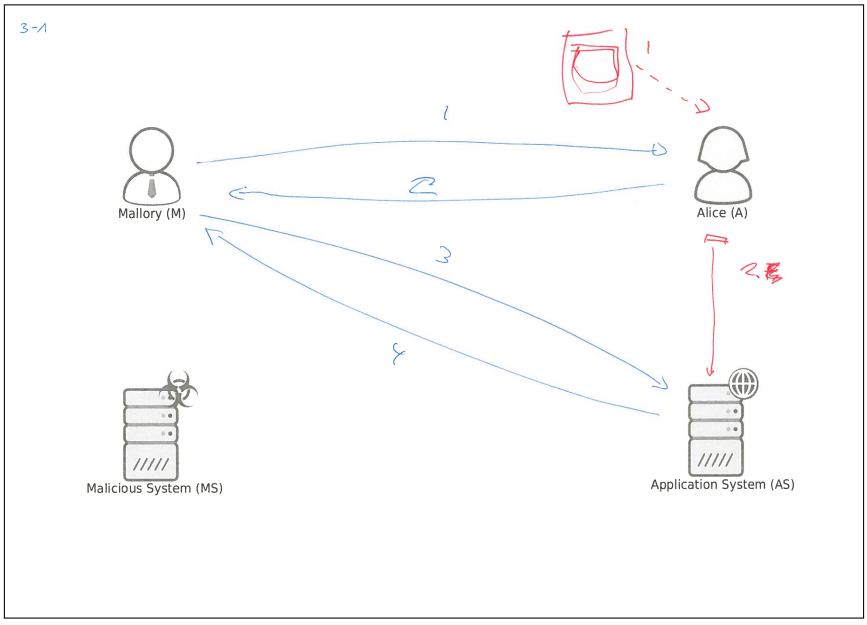


Figure 6.4.: Drawing of attack 1 for the analytical / management person (3) describing a social engineering attack

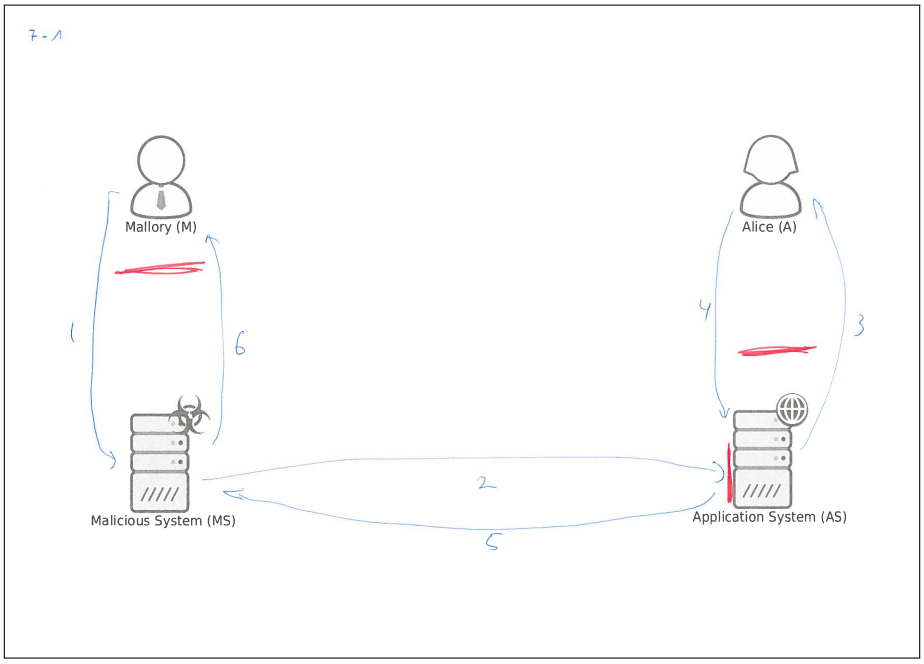


Figure 6.5.: Drawing of attack 1 for management person 7 describing a DDOS attack

P: I think, a bit shorter description, I think. Just there are more details [...]. Yah. Interaction with this machine.

I: Ok. Can you explain it a little bit?

P: Well, I mean first, you scan for vulnerabilities whatever the vulnerability is. It can be... Just scan for vulnerabilities. And then exploit the vulnerability. That can be SQL injection or maybe there is like a simple... Maybe FTP is open and you can just brute-force a simple login. Maybe there are default logins. And maybe, you have Wordpress running. A little bit an outdated plugin. Whatever, exploit a vulnerability and if it is a vulnerability that you can get the data out of the system with it, then you do so. And then you don't need Alice.

In contrast management person 4 described the attack asking for allowed queries. When he explains the attack more detailed, he talks about embedding malware into the system using a query. Although this is probably a similar attack to what the operational person described, the description is more difficult to understand. This comes especially apparent while comparing the two drawings (Figure 6.6 and 6.7).

P: A second attack could be, that Mallory go directly to the application

Attack	Amount	Percentage
Phishing	4	29%
Social Engineering (no Phishing)	4	29%
Injection	2	14%
MITM	2	14%
DDOS	1	7%
Malware	1	7%

Table 6.3.: Experts' attack drawings categorized by attack

system and just try to see, what she can ask the application system. To find out, what kind of queries are allowed. And if she can ask something to the system, well she can try to kind of SQL injection. So that could be one, another attack.

I: So that is also request number one?

P: Yea, and then, depending on what the systems give back, she has an idea, what she can do with the system. And, so, and then it depends, what answers Mallory gets. If Mallory gets, well say, I can ask all kind of questions and the system doesn't recognize anything. Well, she can ask all kind of questions and get a recognizance of the system. For what the system is actually doing. And then, if that is possible, what she could do is, to find out, how she could bring her malware into the system. Ok. That is the application system. So that's an, the other way to do that.

I: So, how... Mallory has found a vulnerability here?

P: Yes, exploit...

I: And then, how does the message go from his system to there?

P: Well, Mallory writes a query, where she embeds, where she may embed the malware into the system. For example an insured query or something like that.

I: Ok, and then the malware goes from here to here? Can you...

P: Via, via, ... Well it goes actually always via Mallory.

I: Ok, then this is like number two? And this is number three again? Can you just...

P: This is number one, this is number two. And then, bringing it in is number three. Two is brought in system.

Table 6.3 shows how often the interview partners described certain attack types.

The countermeasures that the participants proposed to deny the attack were mainly technical (e.g. two-factor authentication, firewalls or attribute based detection systems)

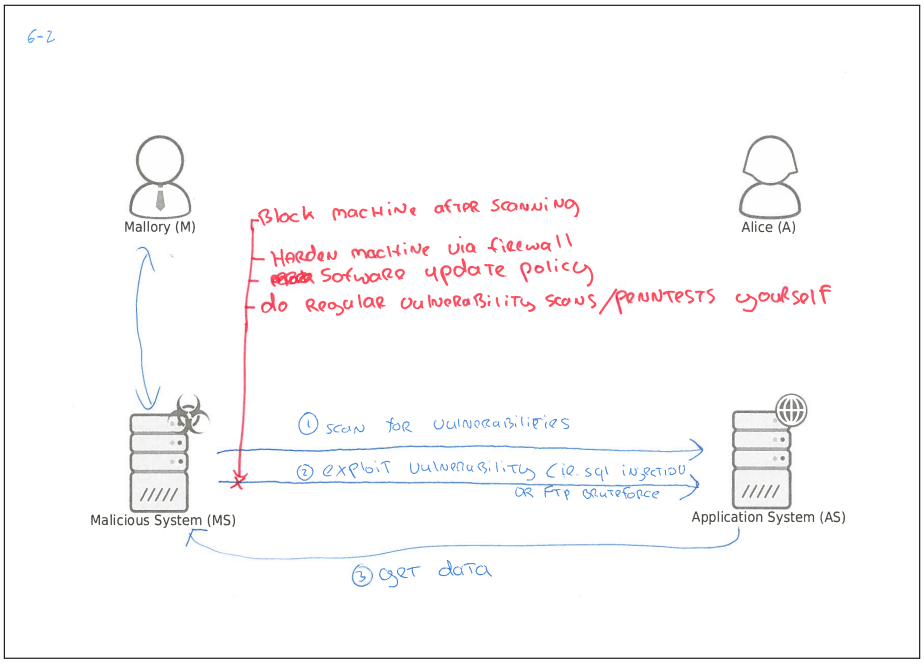


Figure 6.6.: Drawing of attack 2 for operational person 6 describing an injection attack

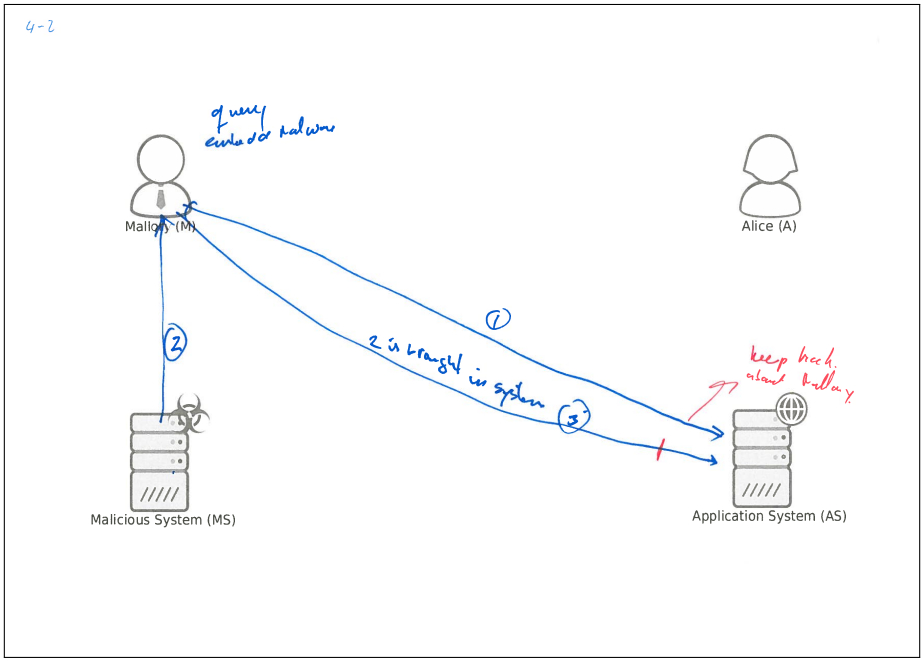


Figure 6.7.: Drawing of attack 2 for management person 4 describing an injection attack

and awareness related (e.g. awareness campaign). Few measures were also related to law enforcement (e.g. detention of Mallory or takedown of the system). The law enforcement related measures were mainly described by the two operational people as well as one analytical person. A comparison of the groups show that the management people described less countermeasures and the quality of their description was lower. An example for the quality of the descriptions are the countermeasures for the two injection attacks from Figure 6.6 and 6.7.

The Operator (person 6) said:

> P: If this [malicious system] is one machine. You can just block this machine after you see that machine is scanning for vulnerabilities, of course. You can harden the machine. Harden the machine via firewall or something like that. Software update policy. So that you don't have the outdated plugins for Wordpress. And do regular vulnerability scans, pentests yourself.

Whereas the management person 4 described:

> P: And the second one is a bit harder. Because, what you could do is, what you could do is, something, here. To say well, the system should recognize malware code. So what the systems could do is check, whether an, what the system could do is to check whether the things that are brought in, into the system, are they allowed. Is it allowed to store something in your system. So that's a solution direction for this one. And for this one, Mallory, what you could do, and what is done actually, is that we keep track about the queries. Keep track about what Mallory is asking. What you do is, you keep track, well maybe Mallory is asking a lot that she is not allowed to.

To sum up this comparison, the management people described cyber attacks and according countermeasures worse than the other groups. During their attack descriptions they mixed the order of steps, were less fluent with the language and missed important details that make out an attack. Their pictures were less detailed and it was more complicated to follow their explanations. Please refer to chapter 6.3.2 for an interpretation and discussion of these findings.

Comparison with the Conceptual models

Looking at the interviews, no participant outlined any of the conceptual models from chapter 4.2 during the interviews. Some mentioned or described terms that show up in some of the models.

The analytical/management person 3 talks about being "able to connect the raw data, so technical data to actors." The term actors appears in the CSAN core assessment from chapter 4.2.3 that is produced by the NCSC. One person even mentions the yearly CSAN report [52, 54] in which the core assessment is described. However he does not further link to the concepts from there.

Analytical person 2 talks a lot about the impact of an attack.

> P: Also here, it's really easy to count the number of unique malware the malware companies have found. That's really easy to count. So we can say, this year there were 13 million and last year there were only five million. That's really easy to count. What's missing from this thing is, trying to say, what does that mean? What is the impact? What is the damage? [...] If I am just counting the number of malware, saying, well this year it's 13 million and last year was like five million, that doesn't capture this impact. There is just one malware, Cryptolocker and there are several variations on this. But there is one malware that has an enormous impact. They can really change a company. And really change, it gets in the news, it makes the general public very scared. This is one malware.

Within the ISM risk management process (see chapter 4.2.1) part of assessing a risk is looking at the likelihood of it happening and the potential damage, the impact, of the risk. The combination of the two values then enables to take security measures that might prevent, mitigate or transfer the risk. So if one is not able to assess the impact of an attack, risk management becomes impossible.

Nobody referred to the attack trees from chapter 4.2.2.

Data Need

Looking at the creation of a dashboard, the data to be presented is important. Several questions aimed at the data, the participants of the interviews currently use and which data they additionally need. Except one management person, everybody identified several categories of data he works with. Incident data is the area that is most often used. People from all three groups use data from this area. Incident data mostly means measures how often attacks occur. One analytical person described "the data we're looking at is, counting the number of incidents". For example this could be incidents report that are responded by the NCSC. One operator mentioned six different categories of data whereas everybody else did not talk about more than three. The data usage mentioned decreases from operations to management. However there is no trend visible if any data is most important for one of the groups. Table 6.4 shows how often a certain area was mentioned.

Data	Amount
Incidents	4
Vulnerabilites	3
Honeypots	2
Netflow	2
Actors	1
Dns	1
Exploits	1
Infections	1
Malware	1
Topography	1
Whois	1

Table 6.4.: Data currently used by the experts

Table 6.5 shows what data the participants told they would like to have. Numbers about attacks is the data area that was mentioned most. This means how often attacks take place, who is targeted, etc. Of the four people who want to have numbers for certain attacks, two mentioned to already work with incident data. The management person who did not work with any data himself, also did not talk about data he wanted. Otherwise everybody specified some important data. The data was always different from the data they already possessed. Operator 5 said that "You need all the data to know, how to protect yourself". However, he did not further specify what "all the data" meant. Overall the data about incidents such as the number of successful attacks and their impact was most important to the participants. While looking at measures, several participants want to be able to compare or correlate data, look for patterns or unusual datapoints such as peaks. There was no visible difference between the three groups.

6.2. Student Drawings

For comparison, we asked students to draw attacks similar to the drawing exercise within the interviews. These drawings are supposed to be a point of comparison. The novices' mental models might produce further insight when looking at the experts' ones. We did not take any cohort effects into account that might occur due to different usage of technical equipment or tools.

Data	Amount
Number of attacks	4
Attack Difficulty	2
Impact	2
Patterns	2
Peaks	2
Actions	1
Actors	1
Attack Data	1
Code Overlap	1
Comparisons	2
Correlations	1
Countermeasures	1
Human vs. Automated Request Characteristics	1
Maturity Level of Organizations	1
Victimization	1

Table 6.5.: Data wanted by the experts

6.2.1. Sampling

For this part of the study, we asked students of the big data security lecture within the minor Big Data at Hogeschool Rotterdam to participate. 20 students handed in 23 drawings. The class consisted mainly of male students, however more demographic information is not available due to the embedding of the experiment in a classroom exercise. As students in a big data security course, they can be seen as novices. They do not have any formal training yet, so they are not on the same level of expertise as the government employees in the expert interviews. Nevertheless they are no absolute beginners as this course is part of a minor program that is typically part of the third year within the bachelor program.

6.2.2. Design

The design of the question was similar to the question in the expert interviews. The students got the same scenario of Alice as a bank employee and Mallory as an attacker. However they only got one question, how Mallory could steal data from the application system. They were also asked to only describe one attack but no security measures. This allows some insight into their cyber security mental model without taking too much time that was not available within the lecture that served as a frame for the exercise.

6.2.3. Process

This part of the study took place in a guest lecture within the big data security course. The topic was mental models and dashboard design. After the introduction, the lecturer told the students that he would like to do an introductory exercise and use the data as part of this work. After explaining the situation, he handed out the drawing templates and sheets for agreeing on participating in the study. Then he asked how Mallory could steal data from the application system. The students had five minutes to think alone and draw arrows on the sheets that describe one attack. Clarifications to all of the students included that each drawing should only include one attack and the arrows should be labeled to evaluate the drawings easily. After drawing, the experimenter asked the students to discuss the drawings with their neighbors. Finally some students shared their attacks with the whole class and described their drawings. During their description, the experimenter tried to show the attack flow on the projected drawing template to the whole class. However the descriptions have not been recorded and were provided independently from the drawings which the students handed in.

6.2.4. Data analysis

For the data analysis, the drawings were categorized into the attack categories that have already been used for the expert interviews. Where needed, additional categories were added. About half of the students did not only draw one attack but several. This complicated the categorization. For most of the drawings at least one of the attack schemes was identified to belong to one specific category. Eight attacks within the drawings could not be understood and are left out in the analysis. Some drawings included multiple attacks that we were not able to separate properly. These attacks are part of the drawings that are not analyzed.

6.2.5. Results

Within the 23 drawings we isolated 34 understandable attacks. Table 6.6 shows, how often each attack category appeared. The biggest part of the attacks were 14 social engineering attacks (including phishing). Eight direct attacks on the application system and six MITM attacks.

Figure 6.8 shows an example of a social engineering attack. The student even labeled it as such attack. Interpreting the drawing, Mallory tries to con Alice into gaining access to her computer. He might do this by flirting with her (see the hearts in the drawing, step 1 to 4). Then he installs a keylogger onto her system that sends her account data to his malicious system (step 5). He gets the data from there (step 6) and is able to access the application system (step 7) and steal her data.

Attack	Amount	Percentage
Direct	8	23.5%
Phishing	8	23.5%
Social Engineering (no Phishing)	6	17.6%
MITM	6	17.6%
Malware	2	5.9%
Botnet	1	2.9%
Cracking	1	2.9%
DDOS	1	2.9%
Phishing / MITM Mixture	1	2.9%

Table 6.6.: Students' attack drawings categorized by attack

6.3. Discussion of the Results

The results show different mental models between the expert groups. Comparison with the students also shows a difference. These results shall be the base for the CSD work.

6.3.1. Applicability and Limitations of the Method

The quality of the drawings in the expert method interview shows that it is possible to describe an attack with the help of this method. Instead of describing their abstract thoughts about cyber security, the single scenario made them focus on one particular setting. The drawing helped to visualize the thoughts. To create their drawings, the participants needed to structure the black box of their mind in a way that made sense for others. Some experts labeled their drawings in a way that makes them understandable even without their oral description. In the classroom setting it was not possible to get every student describing their own picture. Also many students drew more than one attack on the template. This had been prevented in the expert interviews. To understand how a person's mental model is, the drawing exercise helps, but only if the supervision is sufficient to guide and understand the drawing process. The think aloud part is therefore also important to externalize the mental model for the researcher.

Different researchers in HCI argue, few users such as five [57] or in some cases even only one [18] are enough for user studies. Whilst acknowledging that only seven participants in the expert interviews cannot give a very broad picture, we argue that for this work this is an acceptable number. With the qualitative data gathered, we do not want to get perfect numbers or distributions, but try to get an insight of how and why people think a certain way. Due to the governmental setting, the cost of acquiring participants is also important. There are only few potential users available for the

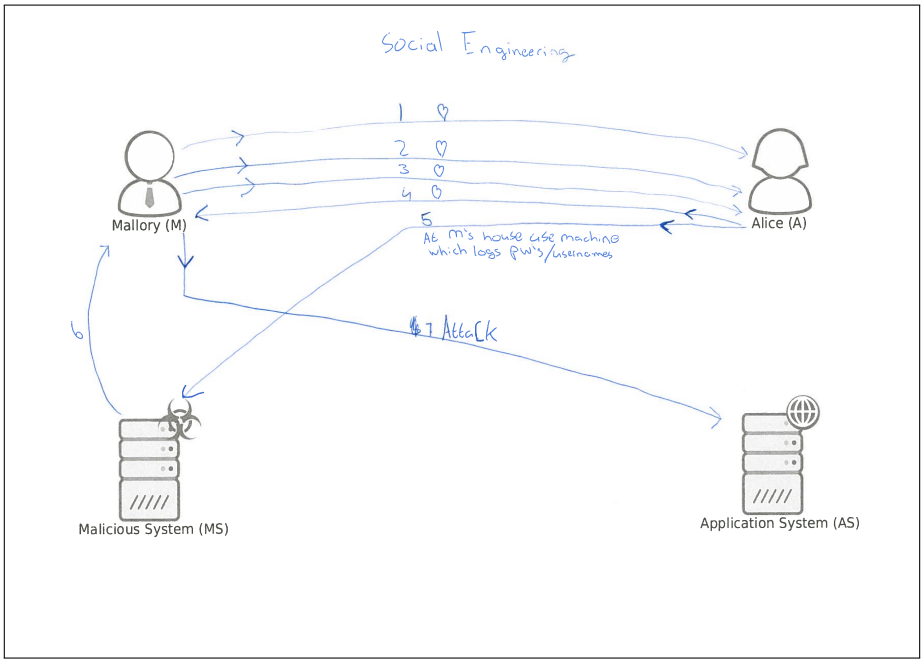


Figure 6.8.: Student's drawing of a social engineering attack

first version of the CSD. They should be governmental employees in the field of cyber security. For each interviewee, participating in the study means to use about one hour of their working time for a project not directly related to their work focus. They also had to share information that might probably be sensitive, as they did not know beforehand how the interviews would look like. Out of the seven participants, at least two of them are in each of the identified groups. This allows to not only differentiate between but also look for commonalities within the groups.

The students shall only provide a point of comparison for the more important expert data. Therefore the class chosen due to its availability to the researchers seems reasonable. The insufficient supervision due to the classroom setting during the drawing exercise, led to a high number of drawings that were not understandable. Trying to create a conceptual model based on this data might lead to a highly distorted view. Therefore, we do not provide a conceptual model of the students' view on cyber security. Nevertheless, we use the students' models for some comparison that seems applicable.

6.3.2. A Mental Model of Cyber Security

This section shall describe how we see the mental model of the potential users, thus trying to answer research question 1: "What are the typical cyber security mental

models of potential CSD users in a governmental institution?" We treat the operators and analysts as one group for this task, as we did not see differences between those groups during the interview.

Operators' & Analysts' Mental Model

According to our study, operators and analysts understand what cyber attacks are and how they work. They are aware of the different steps that are needed to perform a certain attack successfully. For each of the steps they understand in general how it works and how it contributes to the attack. If an attack contains several different sub-attacks, they understand the connection between them. Countermeasures target one specific step of the attack. Figure 6.9 shows how an operator or analyst might see the different steps in an attack building up on each other. They are able to elaborate different attacks. They understand what the effect of a successful attack is. These people know countermeasures from different areas such as technical measures, the role of awareness or legal actions.

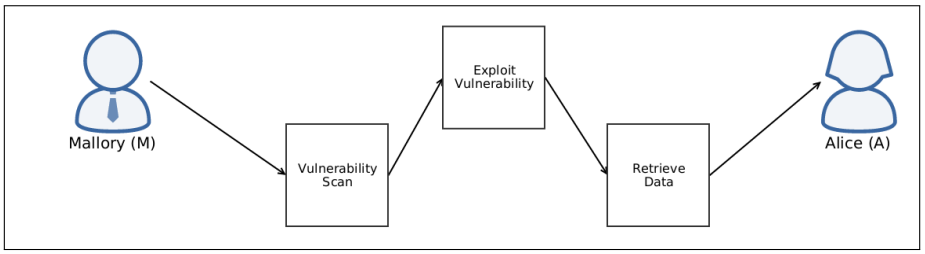


Figure 6.9.: Visualization how operators and analysts see cyber attacks

Managers' Mental Model

The managers from our study understand that there are cyber attacks which need different steps to be carried out. They do not completely understand which steps there are and how they contribute to the successful attack. When describing an attack, they do not elaborate on the specifics of the steps and how they relate to each other. The actual attack is more a black box with some single steps in it as shown in Figure 6.10. The managers know security measures to prevent attacks. Each countermeasure tries to prevent one of the steps that make out the attack despite the step being vague. The known measures are mainly technical. The managers are not aware of law enforcement measures such as detaining cyber criminals that are possible to prevent attacks. This might be due to the domain of the attacks. As the attacks are technical, the managers might focus on technical counter measures. They might think that technical solutions

are the only ones feasible to deny such attacks. Using our data, we cannot state why they did not think of any legal measure.

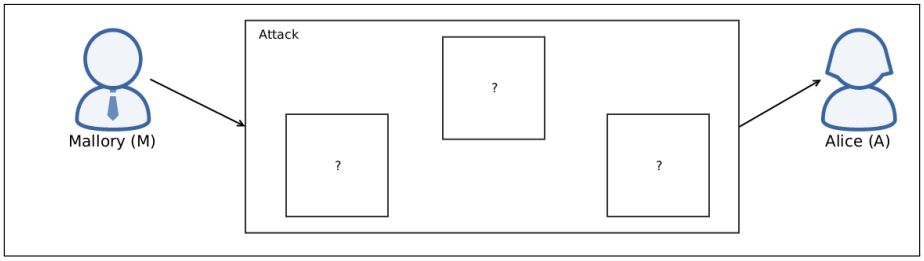


Figure 6.10.: Visualization how managers see cyber attacks

Comparison of the Mental Models

In line with research on mental models of novices and experts [10], we can see differences in the gathered data between the students and the professionals as well as the different groups of experts. Most of the experts described a social engineering attack as the first of their attacks. Several also described another social engineering attack as the second one. Of the students' drawings a bit less than half of the attacks were some kind of social engineering attack. So it seems that the knowledge about social engineering is more present in the experts' mental models. When comparing the operators, analysts and managers of our study, we have seen more superficial mental models for the managers. The mental model of the operators and analysts are more sophisticated than the ones of managers. We assume that this is also the case in comparison to the students. While not being able to support this properly with the students' drawings, we did not see any indicators that contradict this hypothesis.

6.3.3. Relation to Conceptual Models

Although there are some links between the experts' mental models to the described concepts from chapter 4.2, nobody linked to one of them thoroughly. The questions did not specifically ask for them, but a connection could have appeared much more prominent while talking about cyber attacks in general. This indicates that none of the described conceptual models controlled the participants' mental models while answering the questions. More general questions in further research are needed to draw conclusions here. Asking the people about how they deal with cyber security attacks in their organizations might show that those concepts are existing but only used in a different context.

6.3.4. Data Presentation for the Cyber Security Dashboard

For the design of the CSD, this means that we have two potential groups of users. On the one hand operators and analysts; on the other hand the managers. Their level of understanding cyber security has different depth. The comparison of the mental models is especially important, as the managers might not fully understand a CSD that is designed using technical language and data, operators and analysts are fluent with. More aggregated values such as KPIs might be more interesting for them. However it might not be too important that the mental model is correct, as long as the mental model (in combination with the data shown on a dashboard) produces good decisions [70]. For the dashboard design, this means finding ways to support this. Therefore, we excluded the managers and focused on the operators and analysts while designing the CSD. As the operators and analysts generally report to managers, a dashboard targeting those two user groups is believed to have the most impact. Even if it is not designed specifically for managers, the operators' and analysts' reporting should aggregate the data in a way that makes them useful for the managers. This answers research question 2: "Do different user groups in a governmental institution need different CSDs based on their mental models?"

The comparison of the mental models with the described conceptual models (Chapter 4.2) did not show strong links to any of those models. Using them for structuring the dashboard therefore does not seem essential. The focus on the needed data seems more appropriate. Using a bottom-up approach and grouping connected data to form the dashboard instead of taking an external format and trying to squeeze the data into it. Several interviewees mentioned attack data to be important. In connection with the deep understanding that most operational and analytical people have, showing number of recorded attacks in general as well as grouped by attack type seems important. In addition to generalizations, more detailed data will be interesting to look for connections in between different cyber security aspects. We also try to enable comparison as well as showing trends, spikes and anomalies.

Using the results of this study, we gave answers to research questions 1 and 2. We have identified mental models to base our dashboard upon. Still missing is the design of the dashboard and an evaluation of how useful the mental models are for the dashboard design.

7. Design of a Cyber Security Dashboard

We have showed that managers have a less detailed understanding of cyber attacks than operators or analysts. As the operators and analysts are the main people working with cyber security data, we start with the creation of a CSD for them. More important the managers' mental model might be less suitable for the first design of the dashboard. They might lack important vocabulary to understand the presented data. Although paraphrases and explanation of terms is possible, an accurate understanding simplifies the usage. For the interpretation of presented data, it is important to have a mental model that allows reasoning with the presented data. With a mental model that knows how different cyber attacks are composed, the operators and analysts may deduct how to take action against them using information from the dashboard. Without the details in the mental model, the CSD needs to do more interpretation than visualization. This is a highly desirable goal for further work but not addressed in this first design. The design tries to present the data, we identified in chapter 6.1.5 to be important for the interview participants.

7.1. Data Sources

For the creation of the CSD, the NCSC provided two datasets containing data about incident reports and produced security advisories. The data was provided in two excel files. File one contained data from March 2014 to March 2015. The first datasheet contained raw data on the incidents. The other sheets contained aggregated data such as the attacks per attack type or sector. One datasheet included how many security advisories were publicized in each of those months. The second file contained the incident data for the period from March 2013 to March 2014 only in aggregated form. This data was used for rapid prototyping. However it does not comply to a data format that can be produced automatically. Therefore the final implementation is based on the raw data which were also the basis for those excel files. As attack data was the most mentioned data which the interview partners were interested in, this data seems highly functional. We tried to include as much of this data as possible to provide meaningful information for the dashboard users. Chapter 8.3.1 provides a description of the raw data that was used for the final implementation.

7.2. Iterative Process

The design of the CSD followed an iterative process [56]. Each iteration includes several changes on the dashboard. After each iteration the current dashboard prototype was discussed with some of the original interview partners for additional tuning and further development.

7.2.1. Prototype 0

In addition to the computer scientist's way starting to count from zero, the numbering of this prototype is 0, as this was only a preliminary prototype to see whether the provided data was fitting for a dashboard and how this could be done with a business intelligence tool. Figure 7.1 shows this prototype created with SAS Business Analytics (see chapter 8.1.1 for more information).

The development of this prototype showed what functionality was important to create a meaningful dashboard. These findings have been crucial for defining the tool requirements in chapter 8.2. To reduce the non-data pixel part of a dashboard, customization of its content, especially the graphs, is important. It should be possible to easily remove grid lines and unwanted space in between the graphs. For comparison, several graphs should be placed nearby and use the same scales. Defining the grid independently from the data content is therefore necessary for the design.

Figure 7.1 shows some of the shortcomings when looking at the graphs on the right side. The axes do not match each other so that it is not possible to compare the graphs by just looking at them. The labels on the left do not seem to be connected to the graphs on the right.

7.2.2. Prototype 1

This prototype was created with the JavaScript framework Chartist (see chapter 8.1.5), that was chosen for the prototyping process and the final dashboard. The design is based on the need for attack data which the experts described in their interviews given the incident data for the last year's period.

Ideas

As all non managers named the number of attacks as one of their data needs, we chose this part of the data to be shown on this first prototype. Two people mentioned that they were interested in comparison over time and one asked "are we safer today in the cyber world than yesterday? Are we safer this year than we were last year?" (Participant 2). Therefore we chose a line graph to show the development of reported incidents over

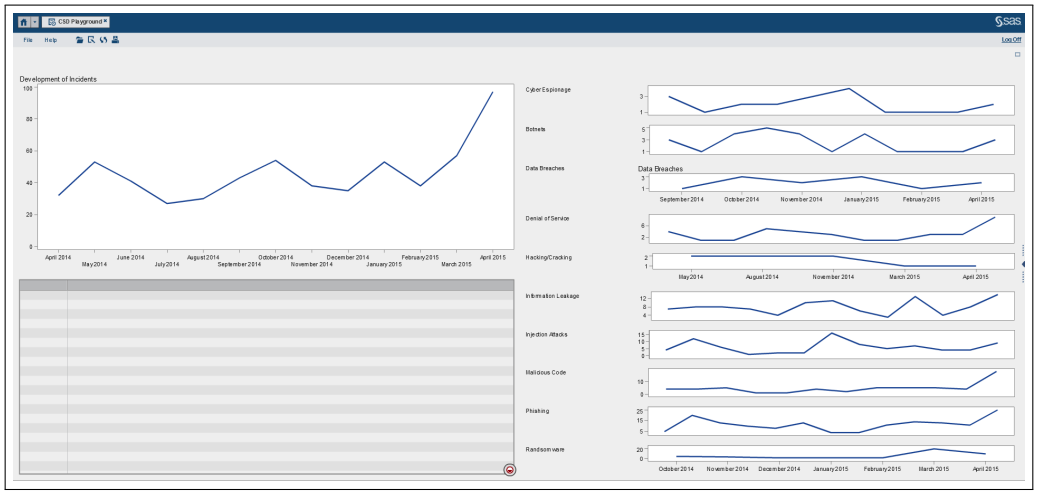


Figure 7.1.: Prototype 0 created with SAS

time. To show a detailed view of the attack development, we used the differentiation by attack type.

Looking at the right graphs in figure 7.2, one can compare the development of the different attacks by comparing the graphs. The graphs are aligned on the same months and the Y-axis has a maximum label of twenty even when there are no or few attacks detected. The labels on the X-axis are only shown on the last graph on the page and extend to all the graphs above. This reduces the amount of non-data pixels on the dashboard. The order of the attacks is the order in which they appeared in the raw data.

Looking at the line graphs it is also possible to identify peaks and search for patterns as requested by two interview partners each. They provide a comparative overview over the different attacks.

The big graph on the left adds up all the different attacks and shows the development of all reported incidents for the same time period. For this graph, we chose a simple summation of all the attacks to show the total number, as no weighting measures were available. For a graph that also depicts meaning in terms of impact or severity as described in chapter 6.1.5, further measures are needed. With such additional data it might be suitable to weight the different attacks to show a more meaningful graph or add some impact measure in addition to the single number of incidents.

To highlight data points, graphs are red whereas organizers and labels are black. Therefore the attention of a viewer is directly attracted to the points that carry the information on how many attacks happened [9, 48].

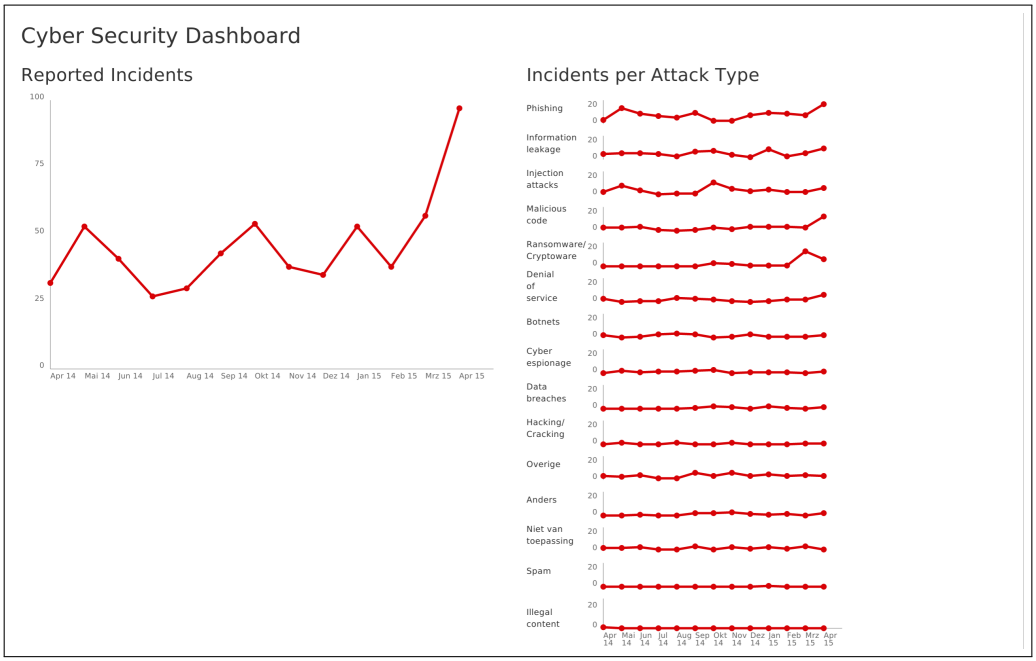


Figure 7.2.: Prototype 1 using Chartist

Review

One operator and one analyst reviewed prototype 1. They appreciated the good overview that these few visualizations already give. They commented several aspects that they would like to see in addition. The free space should be filled with additional information such as the comparison of incidents by sector they appeared in. They proposed a different visualization for additional data to make the dashboard also visually appealing. The graphs should provide possibility to extract more information such as magnifying the small graphs on the right or drill down functionality. For a quick view on the trend of a data series, they proposed a trend indicator next to each of the graphs. They also mentioned that it was not easily understandable that the reported incidents graph sums up all the incidents from the attack type graphs.

7.2.3. Prototype 2

The second prototype tries to implement the comments given during the review of the previous prototype as well as include data from an additional dataset.

Ideas

To clarify the composition of the reported incidents graph, we switched its order with the attack type graph series. So the summarized graph appears right to the graphs it is composed of. This is in line with the European reading direction from left to right.

We added the incident data for the last year to the reported incidents graph. A new bar graph shows how many incidents happened in each of the sectors contained in the dataset. The colors used are the standard colors of the charting library Chartist. The color difference also contains a saturation difference. Therefore the colors are also appropriate for color blind people and gray scale printing.

Trend indicators on the right of the attack type graphs and in the heading of the reported incidents graph show how the number of incidents developed. The trend indicator shows how the number from the last month (x) is compared to the mean of the previous months ($\overline{x}$) in the dataset. Table 7.1 shows which trend indicator is chosen when the corresponding formula is satisfied. If the last month deviates more than a specific factor of the standard deviation (σ) from the previous months' mean, the trend indicator shows the direction of that. The saturation of the trend indicators support its direction. Rising trends for attacks (which are bad) produce a full black arrow. Stable or falling trends are depicted with lighter arrows. We used the ColorBrewer [9] to create the gray scale values that do not distract and are also useful for colorblind people.

The original data included the categories "Overige", "Anders" and "Niet van toepassing". The differentiation between these categories is difficult to understand if one is not fully involved in the data gathering process. Therefore we reduced those categories into one which is called "Other".

To address the need for a closer look, clicking on a graph now shows that graph in full size on an overlay. With this overlay, users can look at the detailed numbers that are difficult to spot within the attack graphs. Figure 7.3 depicts the overlay. When an element on the dashboard is clickable, the mouse pointer changes from an arrow (🡅) to a hand (☝).

One study Participant said after the interview that he would like to have the possibility of an evolving dashboard. Therefore he wants to have the possibility to give suggestions on the further development of the dashboard and the data that is shown on it. To enable the users to contact the developers, we added two buttons that open an overlay with instructions on how to contact the developer. Figure 7.4 shows a screenshot of the prototype.

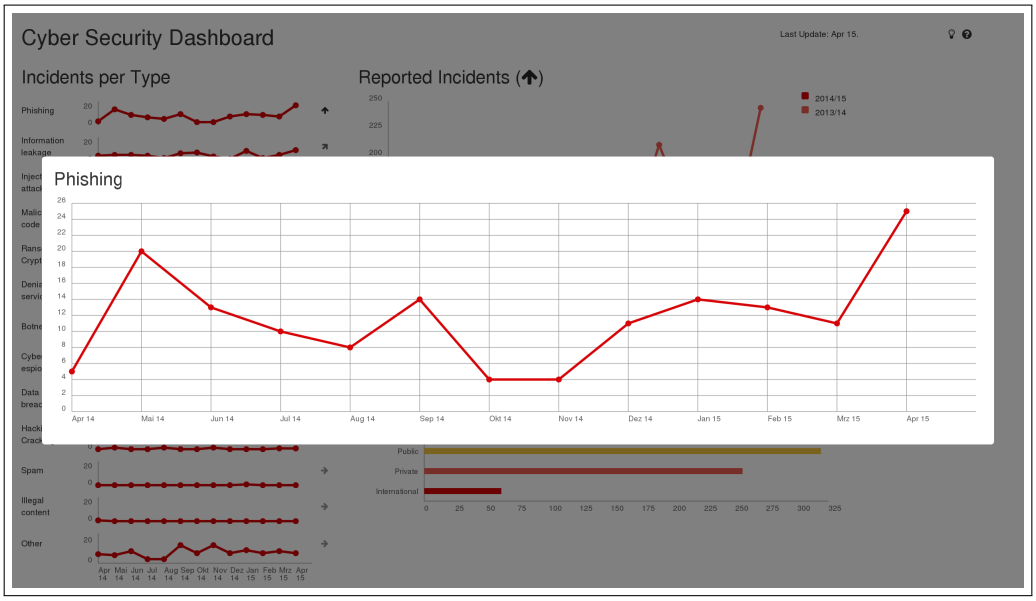


Figure 7.3.: Graph overlay in the dashboard

Indicator	Formula				
$\uparrow$	$\overline{x}+3\sigma$	$\leq$	x		
$\nearrow$	$\overline{x}+\sigma$	$\leq$	x	$<$	$\overline{x}+3\sigma$
$\rightarrow$	$\overline{x}-\sigma$	$<$	x	$<$	$\overline{x}+\sigma$
$\searrow$	$\overline{x}-\sigma$	$<$	x	$\leq$	$\overline{x}-\sigma$
$\downarrow$			x	$\leq$	$\overline{x}-3\sigma$

Table 7.1.: Calculation of the trend indicators

Review

An analyst different from the one who commented on prototype 1 reviewed the second prototype. Overall, he liked the insight the dashboard gave and the comparability of different attack types. He also acknowledged the cleanliness of the design that highlighted the important numbers and did not distract the user. He wanted area charts instead of line charts for the reported incidents by attack type. This should make a differentiation easier, whether some attacks are meaningful or not in terms of absolute numbers. The incident data is composed of automated and manually treated incidents. A distinction between those might be useful when showing the total number of attacks. He also proposed further division of the sector data. He wanted to see which part of the sector had how many attacks and which attacks were most prominent in a sector.

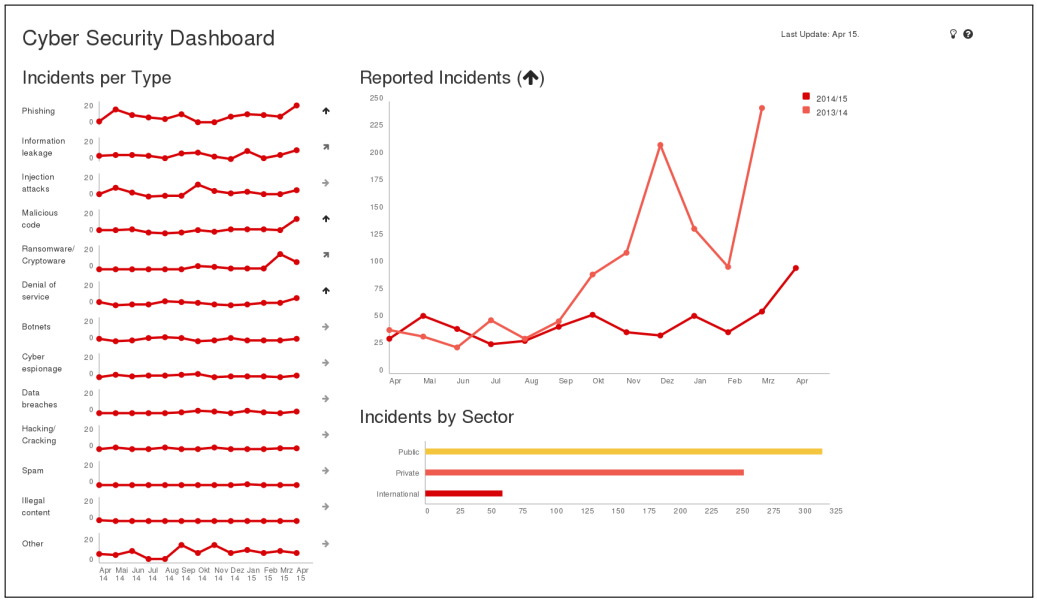


Figure 7.4.: Prototype 2 using Chartist

Additional data to show might be the security advisories the NCSC produces. The number of advisories that have a high rating in impact and likelihood can be an indicator for the workload of the NCSC.

Further discussion of the dashboard focused on the process of updating the dashboard. Due to security regulations a direct connection to the database is not possible. To have a CSD that shows recent data, the preferred way seems to have regular database dumps as a Comma Separated Values (CSV) file that the dashboard loads automatically. This process shall lead to a dashboard that can always show actual numbers instead of having a hardcoded time-period for the attack graphs. A description how we solved this issue is depicted in section 8.3.

7.3. Final Design

The last prototype adds additional graphs to utilize the space that has not been used so far. In addition, we tried to implement features mentioned in the last review as far as possible in the time frame of this research. Further data distinction was not taken into account. Figure 7.5 shows the final design which is our answer to research question 3: “How can a CSD built upon the mental model of users in a governmental institution look like?”

Ideas

To fill the remaining dashboard space, we added further information about the sectors. The table on the bottom of the page shows which attack types have been used most often to attack the sector within the last twelve months. Hovering the attack in the table, a tooltip shows how many attacks of this type have been reported from this sector.

On the right, we added a measure about the most critical security advisories. The number of security advisories created with a high impact and a high likelihood count for this visualization. The more advisories there are for a given month, the bigger the circle becomes. This form of visualization fills the space properly whilst giving indications of how many serious vulnerabilities in this month appeared. For a first glance the precise number is not important. The size of the circles allows comparing without knowing the exact number. When hovering the circle, a tooltip shows the exact number if it is needed.

According to the requests after the review of prototype two, we changed the charts showing incidents by type from line charts to area charts.

To address the wishes from the first prototype review, we added more interactivity to the dashboard. Hovering data points in the graphs now enables a tooltip that shows the exact number that the data point represents. In the graph overlay (that shows one graph in big) for the Reported Incidents graph, it is possible to disable one of the lines by clicking on the corresponding label in the legend on the top right of the graph. The titles of the different graphs now have an explanatory text, that appears when someone hovers the title with the cursor.

To visualize the evolution of the dashboard, Figure 7.6 shows all the prototypes next to each other.

Review

This final prototype has not been reviewed by individual members of the original interviewees. Instead, we invited the original seven study participants to evaluate the dashboard in a more formal way. Chapter 9 describes the complete evaluation process and its findings. There, we also discuss how useful the mental model research is for the dashboard design. Before focusing on the evaluation, we first describe the implementation of our software which is built in a way to be useful in the design of CSDs.

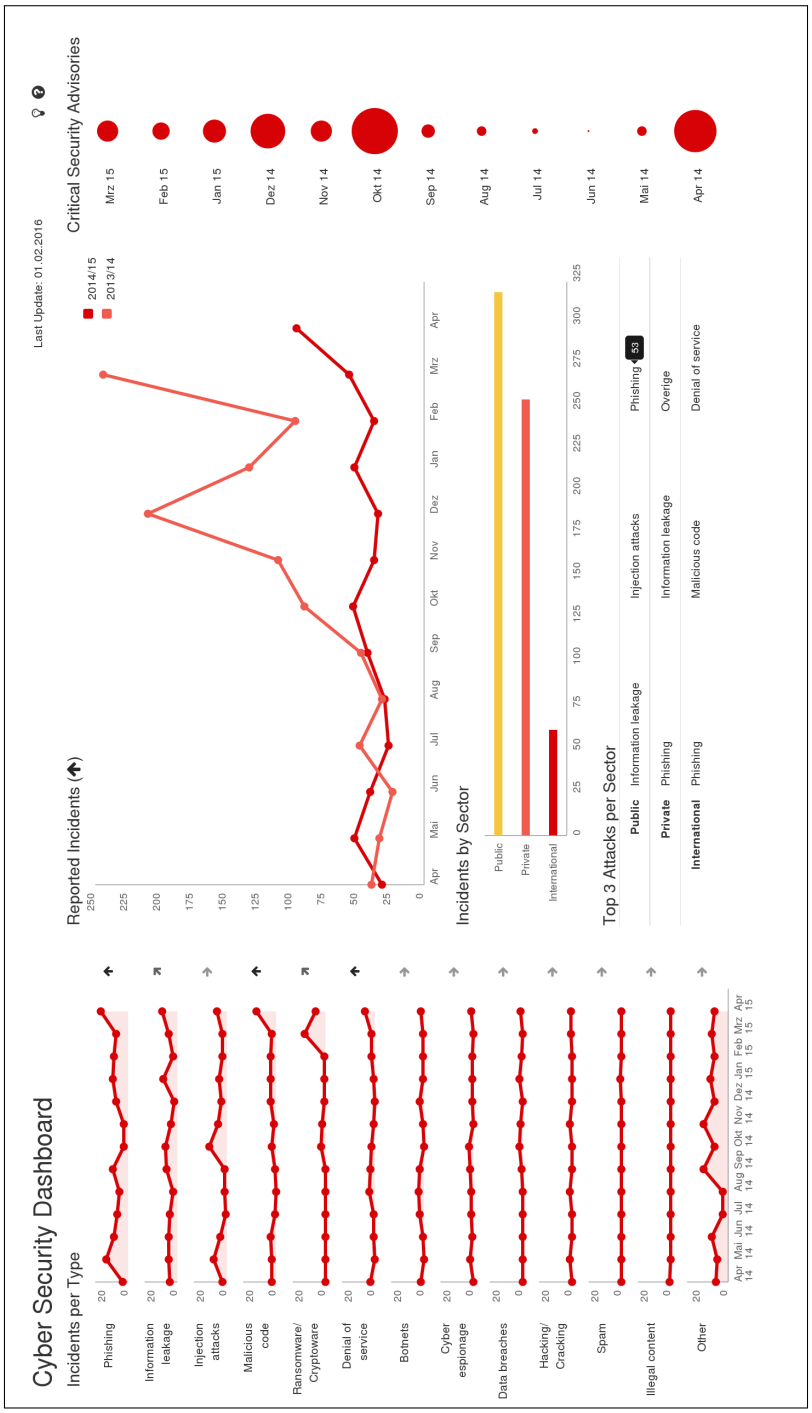


Figure 7.5.: Final prototype

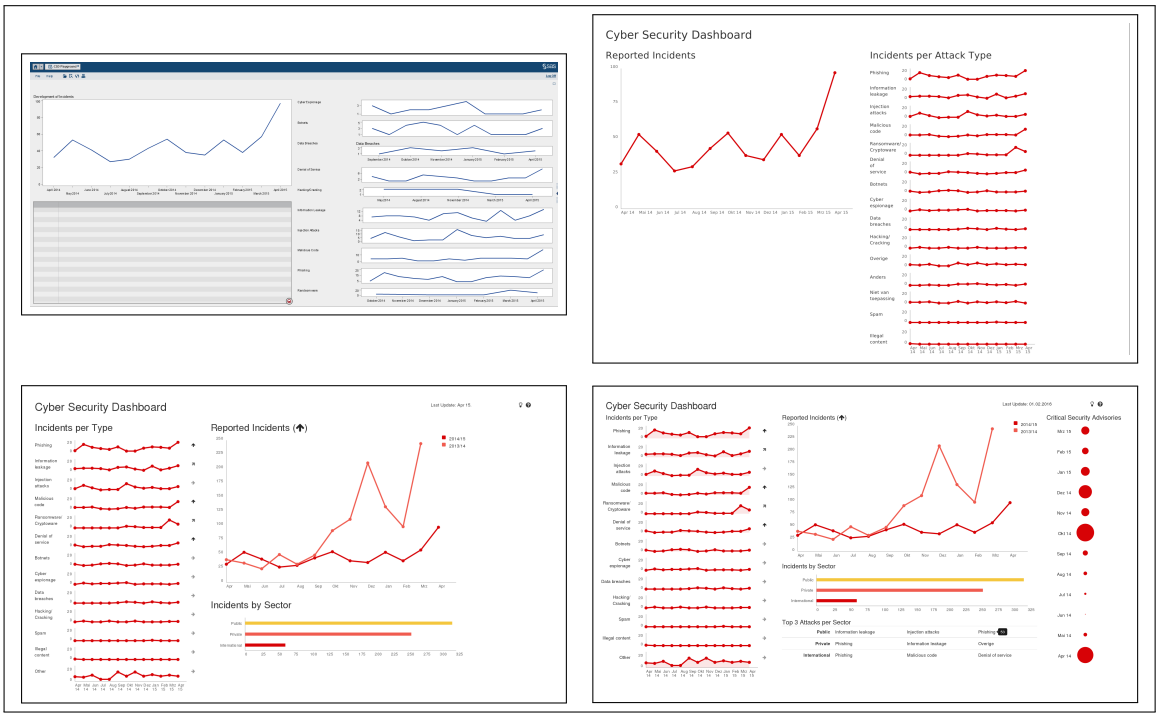


Figure 7.6.: Evolvement of the dashboard prototypes

8. Implementation

Before the CSD implementation, we compared different software and frameworks. This chapter describes the evaluation of those tools and depicts the final implementation of the dashboard.

8.1. Software Toolkits

We examined several charting toolkits and one business intelligence suite regarding their appropriateness to create our dashboard with them. This section describes all those software.

8.1.1. SAS Visual Analytics

SAS Visual Analytics is a business intelligence software with database access. The software is installed on a customer's server. It allows to create dashboards in a drag & drop web interface. Fine tuning of the created visualizations is not easily possible. The license type is commercial.

SAS Visual Analytics[1] was available at the cooperation partner and used by their analysts. Therefore, we evaluated this software first. A created dashboard would integrate in their existing infrastructure. Access to the data was easy, as the used database already contained most of the needed data. The very first prototype described in chapter 7.2.1 was created with this tool. Due to policy changes, this tool became unavailable during the development process.

8.1.2. RazorFlow

RazorFlow is a PHP and JavaScript Framework for the creation of dashboards. A free evaluation license of the library can be downloaded. Commercial licenses are also available. The source code is available[2] using the MIT license [60].

As a JavaScript framework the data can be loaded from arbitrary sources. This is also true for all of the following frameworks. Using functions, one can draw box elements

[1] `https://www.sas.com/en_us/software/business-intelligence/visual-analytics.html`
[2] `https://github.com/RazorFlow/framework`

containing graphs. The data is inputted as two arrays, one containing the X-axis labels and one containing the corresponding Y-axis values. Using the Quick Start template, one can set up a dashboard with simple JavaScript commands easily. Graphs are ordered within a grid system. Customization of the charts is only possible to a certain point without the need to modify the framework's source code.

8.1.3. Chart.js

Chart.js is a pure JavaScript framework for chart creation. It is also open-source[3] under the MIT license [60].

The framework functions allow to easily draw six different types of charts with many possibilities for customization. However it is not easily possible to define the maximum value on the Y-axis, which is crucial for an easy comparison of graphs. Graphs are created using simple JavaScript commands similar to RazorFlow. For prototyping, the framework functions need to be embedded in a webpage.

8.1.4. Canvas.js

Canvas.js[4] is another charting JavaScript framework. It is a commercial product with a free license for non-commercial use.

This library contains 24 different types of charts. It uses a different data format than the other evaluated software. Data needs to be supplied as an array of data points instead of two arrays, each for one axis. Charts drawn by the free version of the framework contain an origin text in the bottom right corner. Due to those non-data pixels as described in chapter 5.3.1, this reduces the functionality rating.

8.1.5. Chartist

Chartist is the last of the reviewed tools. It is an open-source framework[5] using the very permissive WTFPL license [31].

It's usage is similar to the one of Chart.js. It includes only the most important chart types (line chart, bar chart, pie chart) but more importantly, those are highly customizable using the options the framework provides. Customization of the appearance is done mainly via Cascading Style Sheets (CSS). To create a dashboard this framework also needs to be embedded into a manually created webpage.

[3] `https://github.com/nnnick/Chart.js`
[4] `http://canvasjs.com/`
[5] `https://github.com/gionkunz/chartist-js`

8.2. Tool Evaluation

Due to the development of the dashboard as a research project in cooperation with a governmental institution, the used tools needed to satisfy certain requirements. These requirements should ensure, that the dashboard could be built in a secure, privacy preserving way by the ministerial analysts. Most of the evaluated tools were JavaScript libraries, as we only reviewed tools that we had access to during the project's development phase. Access in this case includes the possibility to use the tool on our Linux development machine. The requirements are ordered by importance starting with the most important one. Comparable to [29], each category is rated with one of four ratings from $--$ to $++$. For the creation of the CSD, the rating means the following:

$--$ The software is not usable

$-$ The software is usable with major flaws

$+$ The software satisfies most needs

$++$ The software satisfies all needs

Table 8.1 shows the comparison of all the reviewed tools.

Availability The software must be available to governmental operators and analysts.

Data Security It must be possible to store the data in the governmental network.

Functionality The tool must allow to create tailor-made solutions with regard to the data visualization. This includes creating at least line graphs and bar graphs with the possibility to fully customize their appearance.

Usability & Prototyping The creation of the dashboard should be easy for the developer in order to allow fast prototyping.

License For further usage and to share the results, open source tools were preferred.

8.3. Prototypical Implementation

Starting from prototype 1 (see chapter 7.2.2), we chose a JavaScript implementation using the Chartist framework for development. For the website that embeds the charts, we used Bootstrap[6] to create the base template. Figure 8.1 shows how the software is

[6] `https://getbootstrap.com/`

Category	SAS	RazorFlow	Chart.js	Canvas.js	Chartist
Availability	−−	++	++	+	++
Data Security	++	++	++	++	++
Functionality	+	+	+	+	++
Usability & Prototyping	+	++	+	+	+
License	−−	++	++	−	++

Table 8.1.: Requirements assessment for the reviewed tools

structured. The main components are the index.html, csd.js and dashboard.js. The full source code is available on GitHub[7].

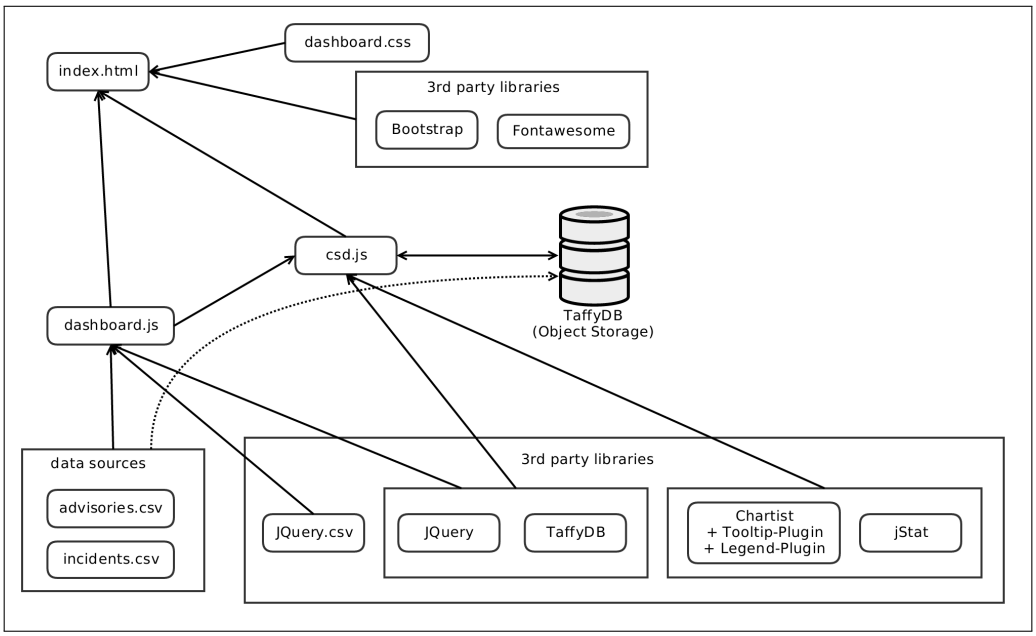


Figure 8.1.: The architecture of the dashboard

8.3.1. Input Data

To automate the process of updating the dashboard information in the secure environment of the NCSC, the data input should take raw data from the NCSC databases. A direct database connection was not possible. Instead the program reads the data from

[7] `https://github.com/Phylu/csd`

two input files. Both files are generated by exporting the content of one database table as a CSV file.

advisories.csv

Listing 8.1 shows the structure of the file advisories.csv. In the current dashboard implementation only the datetime field as well as likelihood and impact are used. With this information the software can group the advisories by month and filter how many advisories with high likelihood and high impact were produced.

```
datetime;id;version;description;likelihood;impact
1-3-2016 13:37; ABCD-2016-0001 ; 1.00 ; Big vulnerability in SSL/TLS ; L ; H
2-3-2016 13:37; ABCD-2016-0002 ; 1.00 ; Small problem with Apache ; M ; M
```

Listing 8.1: Input file advisories.csv

incidents.csv

The structure of the incidents.csv file is shown in listing 8.2. Important information items for the dashboard are the date field and the custom fields Hulpmiddel and Sector. We use those columns for aggregation and filtering by attack type and attacked sector.

```
id,date,CustomField.{Classificatie},CustomField.{Hulpmiddel},CustomField.{Impact},
 ↪  CustomField.{Sector}
000001,01.03.2016,RD-melding,Injection attacks,Toegang tot gevoelige gegevens,Secundaire
 ↪  doelgroep
000002,02.03.2016,NDS-Melding,Malicious code: worms/trojans,Onbekend,Rijksoverheid
```

Listing 8.2: Input file incidents.csv

8.3.2. index.html

The index file contains the Document Object Model (DOM) structure to which the JavaScript files attach the dashboard elements. The DOM elements use the Bootstrap grid layout to distribute the dashboard nicely on one page. Ultimately the file loads the JavaScript files responsible for drawing the dashboard.

8.3.3. csd.js

The csd.js is the main JavaScript module that draws the dashboard. It uses jQuery[8] to manipulate DOM elements, TaffyDB[9] for data storage and manipulation, Chartist to draw the graphs and jStat[10] to make statistical calculations.

```
/**
 * Add a new database to store data
 * @param name
 * @param database
 */
csd.addDatabase = function (name, database) {
    databases[name] = database;

    /**
     * Run queryies on database using DataQuery Object
     * @param filterObject
     * @returns {*}
     */
    csd.DataQuery.prototype[name] = function (filterObject) {
        this.filterObject = filterObject;
        this.database = name;
        return this;
    };

    /**
     * Run Query on database internally
     * @returns {Query}
     */
    Query.prototype[name] = function () {
        this.database = databases[name];
        return this;
   };

};
```

Listing 8.3: Add database to module

[8] https://jquery.com/
[9] http://www.taffydb.com/
[10] https://jstat.github.io/

Initialization

Several methods provide functionality to load data into the database. The data is loaded into TaffyDB and the columns that define attack type and sector are registered with the JavaScript module. These columns can be later used for easy data access and manipulation. Listings 8.3 and 8.4 show the corresponding code. Adding the database `'incidents'` and then running `CSD.addFilterable('sector');` creates a new function `csd.DataQuery.incidents()` that can use a `filterObject` like `{sector: 'public'}` for querying data. Section 8.3.3 shows how to use these functions to retrieve data.

```
/**
 * Add a column that can be used for filtering
 * @param column
 */
csd.addFilterable = function (column) {

    /**
     * Filter querys by column
     * @param value
     * @returns {*}
     */
    Query.prototype[column] = function (value) {
        var compObj = {};
        compObj[column] = {'likenocase': value};

        this.filter.push(compObj);
        return this;
    };

};
```

Listing 8.4: Add filtering column to Query object

Getters

The getters return data like months or years that are used to describe labels within the dashboard.

Database Manipulation

The DataQuery object provides functionality to query the database and retrieve the information needed to draw dashboard elements. The command in listing 8.5 queries

the incident database, filters all attacks that appeared in the public sector and aggregates the last year's values. Therefore it returns the number of incidents happened in the public sector in the last year.

```
new CSD.DataQuery().incidents({sector: 'public'}).yearly()
```

Listing 8.5: DataQuery to the incidents database

Internally the DataQuery uses a Query object that applies the filter onto the correct database. Listing 8.6 shows exemplary how the `yearly()` function works. The `incidents()` function has set the correct database and a `filterObject`. The `yearly()` function creates a new Query object and applies all the filters. Then it runs the query on the correct database and counts the entries for the last year.

```
/**
 * Get data for the last year
 */
csd.DataQuery.prototype.yearly = function () {

    // Calculate date values
    [...]

    var query = new Query();
    for (var key in this.filterObject) {
        query[key](this.filterObject[key]);
    }

    var queryResult = query[this.database]().after(1, startMonth, startYear -
↪   1).before(31, latestMonth, latestYear).count();
    return queryResult;
};
```

Listing 8.6: Query object filtering and counting

Drawers

To create the dashboard elements, several functions create chart and attach these to the DOM structure. These methods make intensive use of Chartist. Take listing 8.7 as an example. It takes the DOM element identified by `selector` and attaches a heading including the tooltip describing the chart. It creates a new chart using Chartist with the given `labels` and the `series` as datapoints. It creates a new DOM element that is

uniquely identified by the id `chartSelector`. It uses the `createOverlayChart` function to create the same chart within an overlay and attaches a click event to `chartSelector`. Therefore the overlay is shown when somebody clicks on the chart in the dashboard. Finally, it removes the grid to get rid of some non-data pixels.

```
/**
 *
 * @param selector
 * @param name
 * @param labels
 * @param series
 */
csd.bar = function (selector, name, description, labels, series) {

    var chartSelector = createUniqueSelector();

    // Add Heading
    var heading = $('<h2>').html(name).addTooltip(description, 'bottom');
    $(selector).append(heading);
    $(selector).append($('<div>').attr('id',
↪    chartSelector).addClass('ct-chart autoscaleaxis clickable'));

    var chart = new Chartist.Bar('#' + chartSelector, {
        labels: labels,
        series: series,
    }, configBar);

    createOverlayChart(chartSelector, name, 'bar', labels, [series], configBarOverlay);

    chart.on('draw', removeGrid);
};
```

Listing 8.7: Creation of a bar chart

8.3.4. dashboard.js

The dashboard.js file puts everything together. First it loads both data sources for incidents and advisories using an Asynchronous JavaScript and XML (Ajax) requests. Each of the CSV files from chapter 8.3.1 is read using the jQuery-csv[11] library. The JavaScript code creates a database and forwards the database to the CSD module. It defines the columns for filtering on the database and groups some data such as combining several incident types into the category 'others'.

[11] `https://code.google.com/p/jquery-csv/`

It then uses the CSD module to draw the dashboard elements. In listing 8.8, you can see the creation of the bar chart showing the incidents of the different sectors. For each sector the yearly data is retrieved. Then the drawer function is used to attach a bar chart to the element with the id `sector-incidents` in the index.html file.

```
var sectors = ['Public', 'Private', 'International'];
var attackNumbersSector = [];
for (var sector of sectors) {
    attackNumbersSector.push(new CSD.DataQuery().incidents({sector: sector}).yearly());
}
CSD.bar('#sector-incidents',
        'Incidents by Sector',
        'Incidents by Sector shows how many attacks were reported in each sector in the
 ↪  last year.',
        sectors,
        attackNumbersSector);
```

Listing 8.8: Creation of the sectors bar chart

8.4. Deployment

One can get the CSD source code and deploy the dashboard from GitHub[12]. A prerequisite is a working Node.js[13] environment as the CSD uses the Node Package Manager (npm) to get its dependencies.

To install the dashboard run the commands from listing 8.9. This will grab the source code from GitHub and install all dependencies using npm. Then it uses browserify to compile all dependencies needed by the dashboard.js file into bundle.js. The last command starts the npm http-server. One can then access the dashboard by accessing `http://127.0.0.1:8080/` in the browser.

```
git clone https://github.com/Phylu/csd.git
cd csd
npm install
node_modules/browserify/bin/cmd.js -t browserify-css js/dashboard.js > js/bundle.js
node_modules/http-server/bin/http-server
```

Listing 8.9: Install the CSD

[12] `https://github.com/Phylu/csd`
[13] `https://nodejs.org`

Changes on the dashboard source code require to re-run the browserify command to include any changes at the JavaScript or CSS code into the files that are used by the index.html file.

The data displayed by the dashboard after this installation is randomly generated data, as the raw data used for this research cannot be distributed. Refer to chapter 10.2 for more information on this issue. We created a Python script that creates data samples to be shown in the dashboard. To generate new random data, use `python sampler/sampler.py`. This will recreate the files advisories.csv and incidents.csv with different entries.

To deploy the CSD in a production environment, one needs to replace this input files with ones that include real gathered data instead.

Still missing is the evaluation of our dashboard software. The next chapter takes the software described here and presents how the original experts assessed the software.

9. Dashboard Evaluation

For the evaluation of the CSD, we used two different methods. At first, we showed the dashboard to the seven experts from the interviews and asked them to evaluate the dashboard. This led to a user experience evaluation as well as an assessment of the functionality. Furthermore, we compared the single graphs from the CSD with corresponding figures from the CSAN [53, 54] which are based on the same dataset.

9.1. Expert Evaluation

All seven experts from the interviews also took part in the evaluation of the dashboard. Information on the participants is presented in chapter 6.1.1.

9.1.1. Design

The evaluation task consisted of three parts. We used the constructed software to present the dashboard to the participants. Working with the real dashboard for some minutes enabled them to answer the evaluation questions. The dashboard displayed randomly generated data. Therefore the users familiar with the real data had no advantage on answering some of the questions.

For an user experience assessment, we used the UEQ [41]. This questionnaire is a standardized measure to gain insight in the user experience of software. It asks the participants to rate the software with the help of word pairs. The participant states on a scale from one to seven which of the words describes the software better. A one means that the first word totally describes the software. A seven that the second word fits completely. A four shows indifference and the numbers between show further graduations. Example word pairs are 'easy to learn' vs. 'difficult to learn' or 'valuable' vs. 'inferior'. The answers are grouped to provide measures for the scales attractiveness, perspicuity, efficiency, dependability, stimulation and novelty [41]. We used a slider in the online questionnaire tool LimeSurvey[1] to simplify the rating of the words.

Open questions tried to see how the participant understands the dashboard. To answer these questions properly, the participants needed to appropriately interpret the dashboard. One of these questions was:

[1] `http://www.limesurvey.org`

> Recently, hackers managed to encrypt the data of several German hospitals and demanded Millions of Euros ransom in exchange for the data. A Dutch member of parliament is scared that something similar happens in the Netherlands. He asks you if this could be a problem for the Dutch healthcare system as well. If you look at the data of the dashboard. What can you tell him?

Additional open questions looked at the usefulness and the future perspective of the dashboard.

We used a structured questionnaire [19, p. 398] to allow a faster processing of the answers compared to a real interview situation. Nevertheless we invited all participants to a personal session to ensure that the questionnaire is filled out completely and the participants are able to use the dashboard in an environment that allows questions. You can find the full questionnaire in appendix B.

9.1.2. Process

We invited all participants to attend to a personal demonstration and evaluation session. The sessions took place between 23rd February and 3rd March 2016. About half of the participants had seen the dashboard design at that moment already. Either while commenting on one of the prototypes during the prototyping phase or during an internal presentation of this study. None of the participants had used the dashboard before.

In the evaluation setting, we welcomed the participant and thanked him for taking part in the evaluation. Then we asked him to take a look at the dashboard on a laptop and play around with it. The participant should try to understand everything the dashboard showed him. We answered questions that occurred whilst trying out the dashboard.

Afterwards we opened an online questionnaire and asked the participant to fill out the questionnaire. After entering his participant code – to allow a connection to the original interviews – the participants rated the dashboard on the UEQ scale. After finishing this questionnaire part, we handed a printout of the dashboard to the participants to have a point of reference when answering the open questions.

9.1.3. Results

User Experience Questionnaire

Due to the small number of participants, we do not use these measures for any statistical test. When looking at Cronbach's alpha for the scales (table 9.1), one can already see, that the users did not rate very consistently especially on the efficiency and the dependability

Scale	α	Mean	SD
Attractiveness	.57	1.333	0.500
Perspicuity	.68	2.179	0.535
Efficiency	.25	1.571	0.607
Dependability	.31	0.869	0.721
Stimulation	.46	1.179	0.572
Novelty	.83	0.357	0.911

Table 9.1.: Descriptive values of the UEQ evaluation

scale. Nevertheless we think that the descriptive values might give some insight how our experts see the CSD. A generalization to a larger population is not possible even for the alpha values [75]. We did not try to compare the values for the different user groups as this would shrink each sample size to two or 3 users only.

Table 9.1 also shows the mean and standard deviation of the scales for all participants. The data is transformed from the initially described scale from 1 to 7 to a scale from -3 to 3. A three denotes the best possible value on this scale. Due to the avoidance of extreme values, extremely high or low values are unlikely. Social desirability makes very low values even more unlikely [5]. For a reasonable interpretation of these numbers, one needs a point of comparison. Luckily, the creators of the UEQ provide a benchmark set to compare our dashboard data to ratings of 4818 people from 163 different studies such as business software or web shops. Even if the single software in the benchmark dataset may not be directly comparable to our dashboard, we argue that a comparison with the mean values of more than 150 different software products is reasonable. Figure 9.1 shows how the CSD compares to the data from the benchmark dataset. Table 9.2 shows what the ratings in comparison to the benchmark mean.

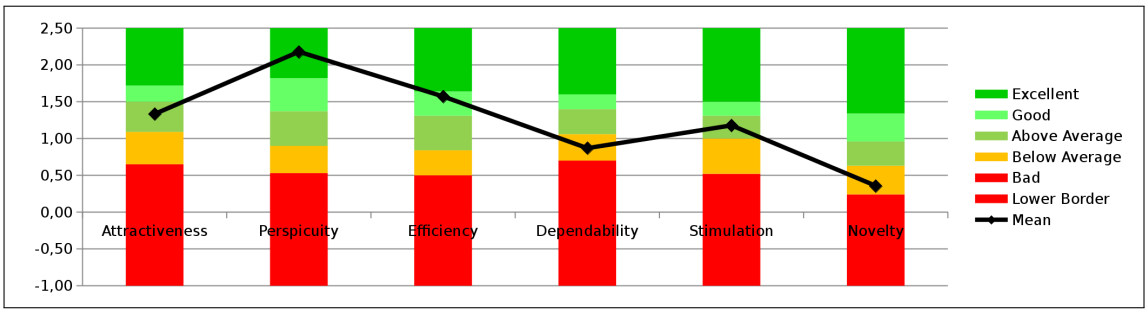


Figure 9.1.: Comparison of dashboard rating with benchmark data by [41]

The comparison with the benchmark shows that our experts see the CSD as out-

Excellent	In the range of the 10% best results
Good	10% of results better, 75% of results worse
Above Average	25% of results better, 50% of results worse
Below Average	50% of results better, 25% of results worse

Table 9.2.: Meaning of ratings in comparison to benchmark dataset

standing clear respectively understandable and very efficient. It is quite attractive and stimulative. However, they seem as if they cannot depend too much on the dashboard and do not see the design as very innovative or new.

Open Questions

The open questions focused on three different points: (How) Did the participants understand the dashboard? How useful is the dashboard to them? What future work is beneficiary to the dashboard?

Understanding The first questions related to specific parts of the dashboard. Four questions asked for a specific answer regarding data of the dashboard and an explanation of that answer. Five participants provided a reasonable explanation of their statement that is backed up by the data of the dashboard for all those questions. Two participants only provided a reasonable explanation for three of the four questions. One of these wrong explanations is similar to the other peoples' explanations. According to this answer, however the question before should have been answered differently. This might be a mistake in reading the dashboard labels. The other wrong answer uses background information on the frequency of attacks that is not displayed in the dashboard.

As an example, we present some answers for the question concerning ransomware in hospitals. Analyst 2 says:

> yes, this can be a problem for the Dutch healthcare system. according to the dashboard this trend is increasing and it is now the top attack in both public and private sectors.

Manager 7 has a similar reasoning:

> While the healthcare system as such is not specified in the dashboard, ransomeware [sic] makes out a high number of attacks in the public sector, so it could be a problem

These both answers show how the participants read the data from the dashboard and use it for their answer. Although the second answer is completely reasonable, it also shows that some of the nomenclature of the dashboard is not completely clear without background knowledge. The public sector data does not include hospitals. Those belong to the private sector.

The answers to the last question concerning the dashboard understanding showed similar problems with the graph labels. The experts noted the terms that they did not understand in the dashboard. Three participants mentioned, they did not understand information leakage especially in comparison to data breaches. Other mentions concerned mainly further attack types.

Usefulness Two questions asked the participants which new insight they got from the dashboard and how fitting the visualization of the data was for their purpose. For the better legibility, the following excerpts of the questionnaire answers are formatted using bullet points.

Person 2, an analyst, says:

- a quick overview of the current top threats/attacks. this allows quick decision making e.g. determining how to spend budgets, which new factsheets and whitepapers are needed, etc.
- a feeling for the developing trends. this provides background knowledge to advisors/consultants.
- the dashboard is very clean and tidy. this is appropriate to provide 'just the facts' to the analist.

The analytical/management person 3 states:

- Insight in trends, insight in past months, insight in relative occurence of attacks, trends in reporting.
- Helpfull for
  - prioritizing [t]ype of attacks to ad[d]ress
  - helpfull in informing public and parlement [sic] about current status
  - helpfull for focusing futu[r]e policy priorities based on assumed upward and downward trends
  - reporting production and activity to management
- easy to read. numerical pop/ups usefull for more detailed comparison
- timescale is usefull

- different visualisations of comparable data is insightfull for global indication of status quo.

One operator (person 6) says:

- It gives a general overview, from an operational point of view its usefulness is limited because of the high level view.
- It would be interesting to see how the criticial security advisories graph looks like over multiple years.
- Quite appropriate. It is w[h]at can be expected from a bigscreen overview.

These answers show how the dashboard provides a high-level overview on the incidents and security advisories of the last months. The information can be used for prioritizing and decision making. Person 3 (analyst/manager) mentioned orally, that the reported incidents are meaningful because of a new law that requires critical infrastructure companies to report any security incident [21]. This law might lead to significant raises of the reported incidents in the private sectors. Effects on the absolute numbers of reported incidents will be visible within the dashboard. Analysts might therefore use the dashboard to measure some effects of this law.

Perspective The last open questions asked the participants to give hints for improvement of the CSD and generally comment the dashboard and the performed study. The main points mentioned were:

1. The lack of possibility to zoom in to have a better view on smaller time periods
2. The wish for more details in the dashboard especially the description of attack types and sectors
3. The need for a way to correlate the advisories and incidents.

There are several wishes for further development in these answers which we further describe in chapter 10.3.1. The following examples give insight into how the experts state their ideas. Operator number 5 states:

> It is a nice global view. On the other hand there is no details. For example what kind of public sector, what sort of private sector, where are the attacks coming from. I also can't see any relation between [the] reported incidents [and the] critical security advisiories. The seem [sic!] with that I can see the relation between [the] sector [and the] reported incidents and between the sector [and the] incidents types.

The other operator (person 6) comments:

- The resulting graph is something [...] quite generic, understandable if it has to appeal to both management and operations
- There is more data that could be useful on our bigscreen, like which incidents are open the longest, which botnets are most popular etc. Depends a bit on the scope of the dashboard what could be added. Now it is limited to incident types and critical advisories.

Person 2 (analyst) says:

> dashboards must walk a fine line between on the one hand providing a quick overview for management and on the other hand providing a useful tool for analysts. i think [...] this dashboard provides management with a very clean, fast tool, but i think analysts could be helped with additional details, zoom-in pages, etc.

During the work with the CSD several participants also mentioned where they would like to have more detailed information. Analyst 2 mentioned that a click on the advisories graph could produce a list of the security advisories. Hovering the attack types should give a description of the attack types as categorized by the European Union Agency for Network and Information Security (ENISA) [24].

9.1.4. Interpretation of the results

We already argued that the descriptive scales of the UEQ should not be used for generalization and can only give hints on how our experts see the dashboard. The high rating on the perspicuity scale suggests that the visualization focuses on the relevant data. The low rating on the novelty scale might also result from this focus. There are no visual gadgets that are obstructive for the user. All the graphs are things that the users know and feel comfortable with. There is nothing surprisingly new in the design which could lead to a high novelty rating. The good rating on the efficiency scale suggests that the users can use the dashboard efficiently, e.g. see the numbers they like to see very fast, do interesting comparisons and get all needed information at a glance. The stimulation and attractiveness give hints that the users like the dashboard in terms of how it is designed. The below average rating on the dependability scale is difficult to interpret. On three of the four items measuring this scale at least one participant did not give any rating. One of the items in this scale (unpredictable vs. predictable) has no correlation with one item and a negative correlation with the two of the other items of this scale. The CSD is seen as unpredictable. This might be to the fact that the dashboard

shows unknown numbers which the users cannot predict. Another explanation might be the use of randomly generated data in the evaluation setting. This data did not match what some of the experts might expect from their prior knowledge.

Every user was able to interpret the dashboard reasonably. Only few mistakes in the interpretation showed minor mismatches that could be caused by unfitting mental models. One example is the wrong classification of hospitals into the public sector. Based on these results, further explanation of the used terms in the dashboard seems useful to foster a better understanding. Also a more detailed view on the data might improve the dashboard experience. Nevertheless, the participants describe the dashboard as a clean view that is easy to read. It can help them in making decisions and prioritizing their work. The cleanliness described while answering the open questions matches the good rating on the perspicuity scale of the UEQ.

We were not able to see any difference in how the different groups understood the dashboard. A member of each group had some comments. An analyst (person 2) said that additional zooming functionality would be good for a nice analytic dashboard. Operator 6 stated that the CSD's usefulness for operators is limited due to the high-level visualization which tries to "appeal to both management and operations". Therefore it seems that the development of a dashboard for analysts and operators based on their mental model has not been completely successful. We try to answer research question 4 "How useful is the identification of a mental model on cyber security for the design of a CSD?" as follows: We did not see any indication that the focus on mental models improved the dashboard design for the targeted user group. None of the findings suggests that the focus on the analysts' and operators' mental models led to a design that is more suitable for them than for the managers. Therefore this research technique can provide insight in mental models, but cannot be seen as a suitable method for dashboard design. For this research, the data to be shown in the dashboard was limited by the available data. Therefore we could not put significant focus on the data need even though some questions in the first expert interview tried to provide insight in this area.

Chapter 10.1 will put the findings of this evaluation in relation with the theory and the results of the expert interviews.

9.2. Comparison with CSAN 2015

The CSAN 2015 [53, 54] is based on the data used for the CSD. It presents the data in Appendix 1 with graphs and a textual description of those graphs. This comparison shall show how the dashboard presents the same data in a different way and what information the viewer can grasp easily with each of the presentation forms. The scope of the two visualizations is different: The CSD tries to present its data on a single screen

as compact as possible for experts. The CSAN includes further explanation of graphs to make the data available to the general public. Keeping this in mind, the comparison can show how the same data can be visualized differently and what that means for the people working with the data. We only used the CSAN graphs and not the explanatory texts for the comparison.

9.2.1. Security Advisories

For the security advisories, the CSAN has two figures that show the development of the advisories over time (Figure 9.2). Graph one shows how many security advisories are newly produced and updated each quarter since 2003. The second figure shows for the last twelve months how severe the advisories were rated. A stacked bar graph shows for each month the percentage of the advisories that has a certain impact and likelihood rating. This can show, how the total amount of advisories has developed within the last years. A splitting by severity is not possible. Looking at the last years data, one can clearly see how the distribution of the advisories changes. The normalization to total number of advisories is important for this view. However an absolute development is not possible.

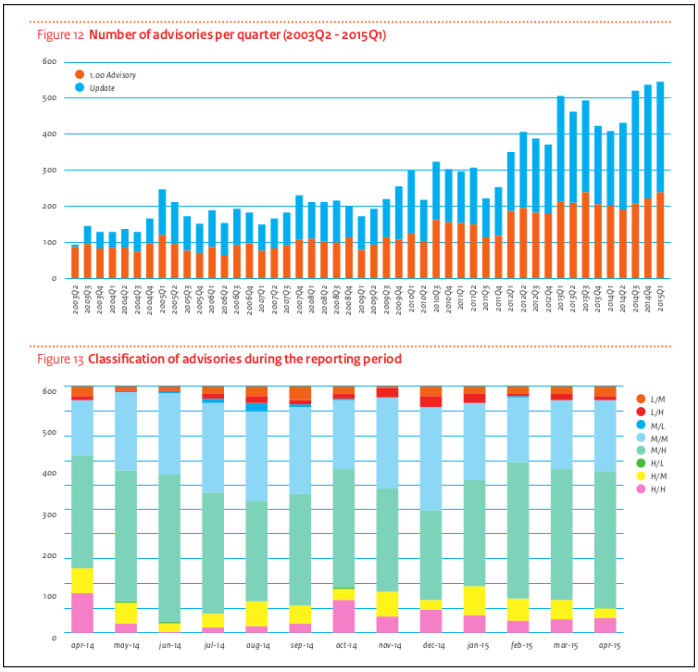


Figure 9.2.: CSAN graphs concerning security advisories [53]

The CSD figure concerning advisories can neither show the development for a period as long as the CSAN does not show the distribution of the advisories. In contrast it shows only the advisories that are rated as high both in impact and likelihood. This focus gives the user a measure how busy the corresponding month was for the NCSC personnel. These high/high advisories in general need much more focus which generally means more time spent working on the underlying vulnerability.

9.2.2. Incident Handled

The CSAN splits the handled incidents into two categories. Manually handled incidents and automated controls. Automated controls may be the sending of an email to the responsible tech contact after an automated scan identified a botnet run on Dutch servers. This separation allows a more differentiated examination of the incidents. The CSD does not differentiate the incidents in this way as this information is not available in the received raw data. The data only included the manually handled incidents.

One graph shows the development of (manually handled) incidents over time which the reported incidents graph does in the CSD. The graph shows the development of incidents for the last two years in a bar graph. Each month denotes a point on the x-axis. The information is not shown in different series as the two lines in the dashboard do.

One pie chart shows the incidents handled by tool. In the CSAN, tool means the same as type in the CSD. For the last year this graph allows the comparison which attack was reported most. The CSD does not explicitly show this information. One part of this information is shown by the incidents by attack type graph. It shows the development of all the attack types for the last year. The top 3 attacks per sector table shows the information of the CSAN pie chart grouped by sector in a different form of visualization.

One graph in the CSAN shows the incidents handled by type. Type here means how the incident was handled by the NCSC. This might for example be an incident response – helping the organization to resolve the incident – or creating a Notice and Take Down (NTD) request. This kind of information is not displayed on the CSD.

Another graph shows the incidents handled per month per sector (Figure 9.3). For the last twelve months, a stacked bar graph shows how many incident reports concerned the public, private and international sector. For each month the three corresponding numbers are stacked bars of different colors. The CSD shows this kind of information only aggregated for the whole last year in the incidents per sector graph.

Three pie charts (Figure 9.3) show the number of incidents per attack type for each one of the three sectors. This information is partly included in the table showing the top 3 incidents per sector in the CSD.

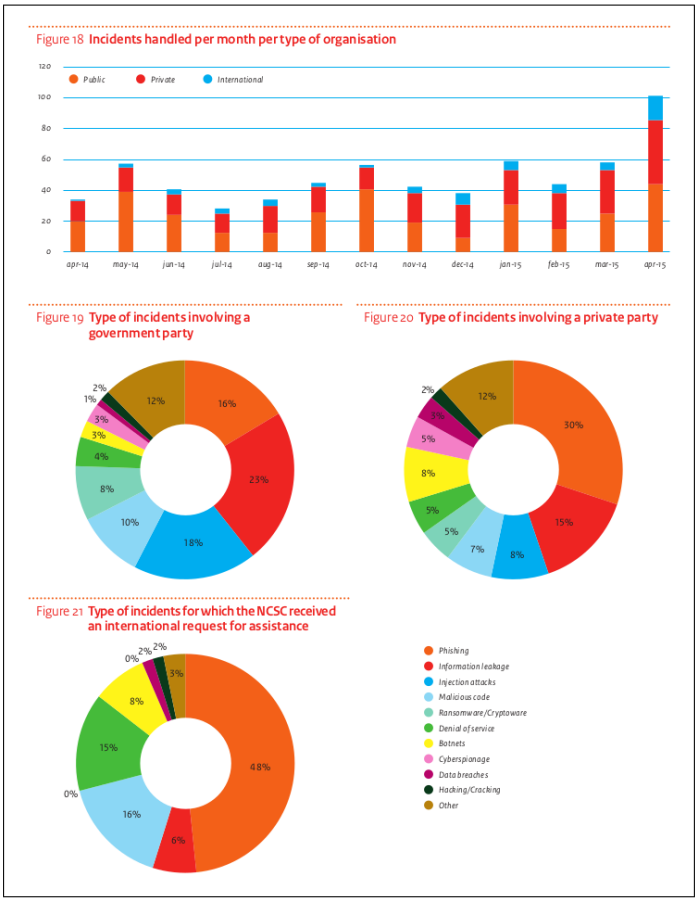


Figure 9.3.: CSAN graphs concerning incidents [53]

9.2.3. Discussion of the Comparison

The comparison showed that the same data is displayed in quite different ways in the CSAN and the CSD. The visualization in the CSAN is six pages long whereas the graphs span approximately three full pages. The CSD displays all its data on one screen. Both have several visualizations that show the development over time for the last one or two years. However the percentage of those is higher in the CSD where more than two third of the space is taken by time-wise graphs and only a bit more than half in the CSAN. The only data where the time-wise development is shown more detailed in the CSAN is the development of attacks per sector. The CSD is more compact and uses the space more efficiently without showing less information. Some parts of the dashboard – especially the development of the incidents per attack types – are much more detailed

than the information in the CSAN.

The focus in the CSD is stronger on the development over time as this was one of the most important data needs of its users. Due to the compactness and the arrangement of the graphs in the dashboard, it is easier to compare the numbers of different attack types or between the last two years for all reported incidents with each other. Nevertheless, the CSAN might be an inspiration for the further development of some graphs in the CSD. Stacking attacks per sector or differentiating between manual and automated incident handling were also mentioned by some of the experts. The graphs in the CSAN might be examples of how this information can be presented within the existing graphs of the dashboard efficiently.

This software evaluation has answered the last of our research questions. Though providing a possible basis for dashboard design, a mental model as single driver for dashboard design seems unwise. The following and last chapter summarizes all our findings and gives ideas for further work in this field.

10. Conclusion

This last chapter shall conclude the work and put the different parts of the study in relation to each other. After summarizing the work, we will present some obstacles of security relevant research in a governmental institution. We finalize the work by giving ideas on further developing the CSD and going on with research in this area.

10.1. Summary

This work described the design, implementation and evaluation of a CSD based on the mental models of potential users. The CSD provides beneficiary insights for governmental researchers. We carried out this research in cooperation with NCSC and WODC. This work answers the research questions whether there are typical cyber security mental models of governmental employees (RQ 1) and if they need different CSDs based on their mental models (RQ 2). Based on the answers to these questions, we described, how a CSD can be built upon such models (RQ 3) and critically discussed whether the identification of the mental models is useful for the design of a CSD (RQ 4).

The theory on dashboard design suggests to produce different types of dashboards for different people. Different mental models dominating in different user groups might be one reason for that. Mental models describe how people understand a certain domain and provide a base for the creation of a CSD. We interviewed seven experts from Dutch governmental institutions that work in the cyber security domain. Within the interview we used a drawing exercise to get insights in the participants' mental models on cyber security. Additionally we asked questions about their data need for a CSD. We compared their mental models with a group of 20 novices who did a drawing exercise in a classroom setting to get a better understanding of possible differences in the mental models. The expert interviews showed that there exists a difference in the perception of cyber attacks. Managers have a more superficial understanding of such attacks than operators or analysts. The most important reason might be the day to day work which influences the familiarity with cyber attacks and the data describing those. The experts were more focused on social engineering attacks than the novices. Answering research questions 1 and 2, we can say that different user groups in the examined governmental setting have different mental models. The operators and analysts have a similar model which is different from the one of the managers. The models differ essentially in the

depth of the understanding. Therefore managers might need a different CSD as the operators and analysts.

Based on those findings, we focused on the operators and analysts for the design of the dashboard. We excluded any needs that might arise from the managers mental models or to the unfamiliarity with the data. One example is the nomenclature in the dashboard which is highly technical.

For the design of the CSD, we created several prototypes that two analysts and one operator commented on during the development process. The final dashboard design is implemented as a web application. A JavaScript module using several open source libraries for data storage and charting creates the dashboard. It uses data about security advisories and incidents and visualizes them in different charts. We released the JavaScript module as an open source library. We presented the final CSD in Figure 7.5 which is the answer to research question 3 how a CSD could look like.

The evaluation of the CSD showed that it is highly usable and provides meaningful insight. It visualizes data in an comprehensive way and can be used for prioritization the governmental focus of cyber security. The UEQ evaluation showed that the dashboard is clearly arranged and easy to use. A comparison with the CSAN showed that the dashboard makes excellent usage of the space available while presenting meaningful figures. Despite the focus on operators and analysts, the experts pointed out limitations especially for operators and analysts. One operator mentioned that the data is too high level to be useful for the day to day operators work. The experts asked for further details and analytical functions such as zooming into smaller time frames. The evaluation showed that there was no difference between the groups concerning the understanding of the dashboard. The dashboard provides value to all the user groups and enables them to base decisions on the dashboard content. Therefore we need to revise our answer on research question 2 and state that users with different mental models are able to understand one CSD similarly well. This could mean that the focus on mental models does not provide additional insight, our operationalization of cyber security mental models did not catch the important differences which lead to a different cyber security understanding. A final answer to this research question 2 is difficult. If the dashboard design included all the features additionally requested by the operators and analysts, this might change the perception also for the managers. Unfortunately further investigation was not possible within the limited time of this thesis.

These findings suggest that mental models for the purpose of creating a CSD does not add extra value to the iterative design process. The questions on the data need during the interviews were the ones that were most helpful during the design of the dashboard. In combination with the high amount of resources needed to carry out and interpret the interviews, we do not see this method as suitable for developing CSDs. The answer to research question 4 is that mental models can not be more than a

starting point for dashboard development. Nonetheless, questioning different experts provided highly valuable insights for the development phase such as the identification of attack development as a highly important figure. Further research should investigate mental models of cyber security more detailed but more separated from the dashboard design aspects. Dashboard design research needs to focus more on the needed data. A combination of these subjects may seem more useful after further exploring them on their own.

The current CSD design shall be further developed during its usage within the project partners.

10.2. Reflection Security Research in a Governmental Institution

Security research in a governmental institution has several impediments. IT security research in governmental institutions is a relatively new field. The attack on the German Bundestag recently showed, how a government is unable to cope with cyber attacks properly [6]. The NCSC produces a yearly report, the CSAN to inform the Dutch parliament and population on cyber security. Their categorization of attacks changed over the last years from a self-made classification scheme to one from ENISA [24] which is used in the CSD. This will likely change to a taxonomy proposed by the Portuguese national CERT [63]. This requires flexibility in the design of the CSD and makes comparison of data over the years difficult.

Security policies shall guarantee that sensitive data is not leaked. Therefore the raw data for the CSD could not leave the governmental office building. With an internal IT structure that does not provide proper development tools a guest researcher, this complicated the development process of the dashboard. Due to this fact we first used a tool of the internal network which had easy access to the data. The tool was abandoned during this research. Therefore, we needed to shift to a self administered development environment without violating the confidentiality of the data.

The raw data is stored in a database within a secured environment. It is not possible to install additional software in this environment to prevent unauthorized access due to misconfiguration or vulnerabilities of the additional software. Therefore the workflow that provides the CSD with up-to-date data is complicated. A semi-automatic workflow shall enable the operators within the NCSC and WODC to have current data in the dashboard shown every month. Using some scripts, they can create a database dump and sanitize it properly. Then it needs to be moved from the secured environment into the less secure environment where a webserver provides the CSD.

Despite those obstacles, we claim this research to be important. The design of this

CSD shall not be the end here. The dashboard will be further improved and better integrated in the existing systems to enable the governmental researchers with a data visualization that helps them doing their work.

10.3. Future Work

Future work in this area includes the future development of the CSD to foster its use by governmental researchers. Possibly this development might lead to a useful dashboard also for critical infrastructure providers. Such utility should also be scientifically supported. Additionally future research in the cyber security mental models and CSD domains seems promising.

10.3.1. Dashboard Development

During the prototyping phase several experts mentioned ideas that could not go into the final implementation due to the limited scope of this work. Some of these ideas were also mentioned during the evaluation phase. In the evaluation the experts also requested other additional features. The following list shall give an idea of how the dashboard can be improved regarding to the need of its users.

- The terms of the dashboard should be meaningful by itself. Due to the use of categories from external sources [24], it might not be possible to choose the nomenclature. In those cases, it is important to describe the used wording appropriately. In the case of the current dashboard design, it might be useful to provide tooltips that explain all attack types and sectors.
- More details can enrich the dashboard experience. Several ideas can enhance the dashboard at several points
  - A selector for the sector might restrict the dashboard data to all incidents in this sector. The dashboard graphs then show the attack development for this sector only. This can give additional information on the IT security state in this sector.
  - The sector data in the attacks per sector graph could be stacked. This means showing in different colors which sub-sector is responsible for which part of the attacks. For example in the private sector, sub-sectors might be communication, healthcare, transportation, finance, ...
  - Clicking on data points in the graphs could show additional information. For example in the advisories graph, this could be a list of all the advisories that were produced in this month.

- Further analytic functions such as zooming in might provide the possibility to focus on a certain timespan. For example framing a time in the reported incidents graph might apply a filter on that timespan to all the graphs.
- Embedding further data sources may provide additional information to the user. For example the inclusion of CVE information might show whether the produced advisories correspond to the CVEs. The inclusion of a Twitter stream might provide additional information on certain vulnerabilities or reported incidents.
- A connection to the database that contains the raw data could provide always up-to-date data to the dashboard.

Some of these functions are easy to implement using the developed JavaScript framework. Integrating additional tooltips or using stacked graphs only needs minor changes in the display logic. Filtering the data accordingly can be done using the appropriate filters of the DataQuery method. Functions like interactivity by zooming can enhance the framework but need significant additions to the module. The JavaScript code needs to register the mouse events, calculate the range to which should be zoomed. According to this zoom, the dashboard needs to show different labels and a different granularity of the data. Possibilities to show these must be added to the framework.

Important to notice is how these feature requests relate to the definition of a dashboard from chapter 5. The analytic features are things that a dashboard does not try to provide. It is something for monitoring and not an analytic tool. The wishes suggest that some of the users would like to have such an analytic tool instead of a dashboard. Further development needs to take this aspect into account. Instead of rejecting those comments based on the definition of a dashboard, the focus of the project might be shifted. If this fosters public cyber security, we need to think of transforming the CSD to a Cyber Security Analytic Tool (CSAT). For dashboard development, terminological clarifications as part of the requirements engineering process may prevent disagreement over the functionality between the designer and the users.

With the open source release of the framework, we encourage the governmental users to adapt the dashboard to their need and push their changes back to implement those functionality based on the actual need.

10.3.2. Future Research

For future research, we see two important points following this study. With cyber security as a growing topic, people's understanding becomes more important. Mental models can provide insight into how somebody understands this area. Drawing exercises in interviews provide benefit, as they externalize somebody's model. A problem with

the drawing exercise in the interview setting is the scalability. Performing this exercise with one person takes about 45 minutes for the interview, 4.5 hours for the transcription plus time for the coding. This results in about 6 hours of work per participant. For larger studies, this method should be refined to minimize this. More abstract models that are not directly linked to one specific attack may give further insight. How do people perceive cyber security in general? What attackers do they fear [70]? How can experts talk about cyber security in an understandable way [2]? Combining such research in a way that enables cyber security researchers to talk to non technical people meaningfully can enable them to better secure their IT life. Governments can use this information when designing awareness campaigns and have a common ground to talk about this topic with institutions involved in critical infrastructure. This may help to make IT security education more efficient.

For the design of CSDs, we recommend to focus on the data need of the users. Unfortunately, it is not always clear to them what their needs really are. Adjusting interview techniques to explore this area may provide valuable information to designers. With more information being available, it is important to present the information that is needed to achieve a certain goal. We suggest to compare different techniques in this domain to see how people can tell what data is important for them. Common methods from social sciences [19] should be compared with different forms of user experience research such as Young's task analysis [74]. Mental model research may provide background information to the researchers on the users' general understanding of cyber security.

Appendix

A. Expert Interviews Questionnaire

This appendix contains the questionnaire used after the first interview to get some demographic data of the participants.

1 Interview Information

We are planning to do a follow up interview in several month. To connect the data from this with the future interview, we ask for a subject code. This code ensures, that we can connect the datasets, while guaranteeing your anonymity. Your participant code consists of the first letter of the name of your mother, the first letter of your fathers name, your age and the day. So if your mother is called Anne, your father called Peter and you are born on 16-May-1972, then your participant code is AP4316. The code will never be published.

1.1 Participant Code:

2 Demographic Data

2.1 What is your gender?

2.2 How old are you?

3 Job Description

3.1 Which of the following categories describes your job best?

- ☐ Operational
- ☐ Analytical
- ☐ Management

B. Software Evaluation Questionnaire

This appendix features the full questionnaire used for the software evaluation. The appearance of the questionnaire differs from the one presented using the online tool. For better readability, we show the printout version of the questionnaire here.

Cyber Security Dashboard Evaluation

Thank you for evaluating the Cyber Security Dashboard. The study tries to get insight on the design of a meaningful and easy to use Cyber Security Dashboard.

The study is carried out as part of a research thesis from Technische Universität München in cooperation with Hogeschool Rotterdam and Wetenschappelijk Onderzoek- en Documentatiecentrum. Responsible for the study is Janosch Maier, B.Sc.

As with the data from the former interviews, the following privacy regulations apply: The data is only stored and processed for the study. The anonymity of the data will be guranteed at all times.

There are 16 questions in this survey

Personal Information

To connect the data from this evaluation with the former interview, we ask for a subject code. This code ensures, that we can connect the datasets, while guaranteeing your anonymity. Your participant code consists of the first letter of the name of your mother, the first letter of your fathers name, your age and the day of your birth. So if your mother is called Anne, your father called Peter and you are born on 16-May-1972, then your participant code is AP4316. The code will never be published.

[]Participant Code *

Please write your answer here:

User Experience

For the assessment of the Cyber Security Dashboard, please fill out the following questionnaire.

The questionnaire consists of pairs of contrasting attributes that may apply to the dashboard. The slider between the attributes represents gradations between the opposites. You can express your agreement with the attributes by using the slider to show what most closely reflects your impression.

Please decide spontaneously. Don’t think too long about your decision to make sure that you convey your original impression. Sometimes you may not be completely sure about your agreement with a particular attribute or you may find that the attribute does not apply completely to the particular product. Nevertheless, please move the slider in every line.

If you want to rate the middle value, please move the slider to the side and then back there. A number between 1 (leftmost value) and 7 (rightmost value) will appear above the slider when your answer is registered.

It is your personal opinion that counts. Please remember: there is no wrong or right answer!

[]Please assess the cyber security dashboard now by moving the slider each line.

Please write your answer(s) here:

|annoying|enjoyable

|not understandable|understandable

|creative|dull

|easy to learn|difficult to learn

|valuable|inferior

|boring|exciting

|not interesting|interesting

|unpredictable|predictable

|fast|slow

|inventive|conventional

|obstructive|supportive

|good|bad

|complicated|easy

|unlikable|pleasing

|usual|leading edge

|unpleasant|pleasant

|secure|not secure

|motivating|demotivating

|meets expectations|does not meet expectations

|inefficient|efficient

|clear|confusing

|impractical|practical

|organized|cluttered

|attractive|unattractive

|friendly|unfriendly

|conservative|innovative

Cyber Security Attacks

In the next part you will have to answer some questions concerning the Cyber Secuirty Dashboard and cyber attacks. Please tell the experimenter to give you a screenshot of the Cyber Security Dashboard. You can use this screenshot if you need to remember a detail of the dashboard. You may also switch to the dashboard by pressing Ctrl+Tab on the keyboard. Press Ctrl+Tab again to switch back to the questionnaire.

It is important for the evaluation that you answer all questions in detail. Please take your time to fill out the answers carefully.

Assume that you are working for the Dutch National Cyber Security Center (NCSC) which is operating the Cyber Security Dashboard. Try to use the data from the dashboard to answer the following questions.

[]According to the data on the dashboard: In which month during the last year (March 2015 - Febuary 2016) were you and the other NCSC personal most busy?

Please write your answer here:

[]Why were you and the other NCSC personal most busy in this month?

Please write your answer here:

[]You are planning an awareness campaign on Cyber Security as part of your NCSC work. Your campaign focusses one attack type. Which type of attack do you choose for the campaign?

Please write your answer here:

[]Why did you chose this particular attack type for the awareness campaign?

Please write your answer here:

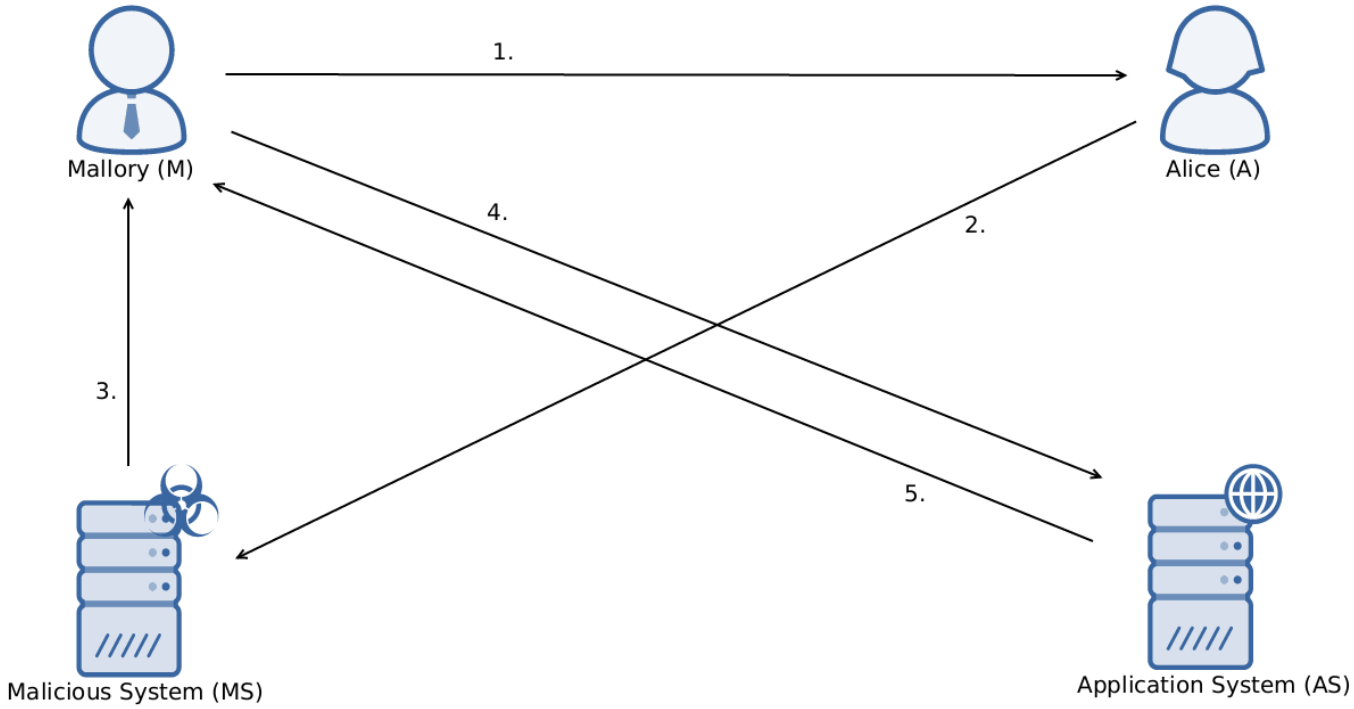
[]Recently, hackers managed to encrypt the data of several German hospitals and demanded Millions of Euros ransome in exchange for the data. A Dutch member of parliament is scared that something similar happens in the Netherlands. He asks you if this could be a problem for the Dutch healthcare system as well. If you look at the data of the dashboard. What can you tell him?

Please write your answer here:

[]

Please try to remember the interview setting for the next questions. We were talking about Alice – a bank employee – and Mallory – a hacker. Assume that Mallory is attacking this Dutch bank. The attack is detected and reported to you as an employee of the NCSC which operates the Cyber Security Dashboard.

Mallory is trying to attack the bank with the attack shown in the picture. What kind of attack could this be?



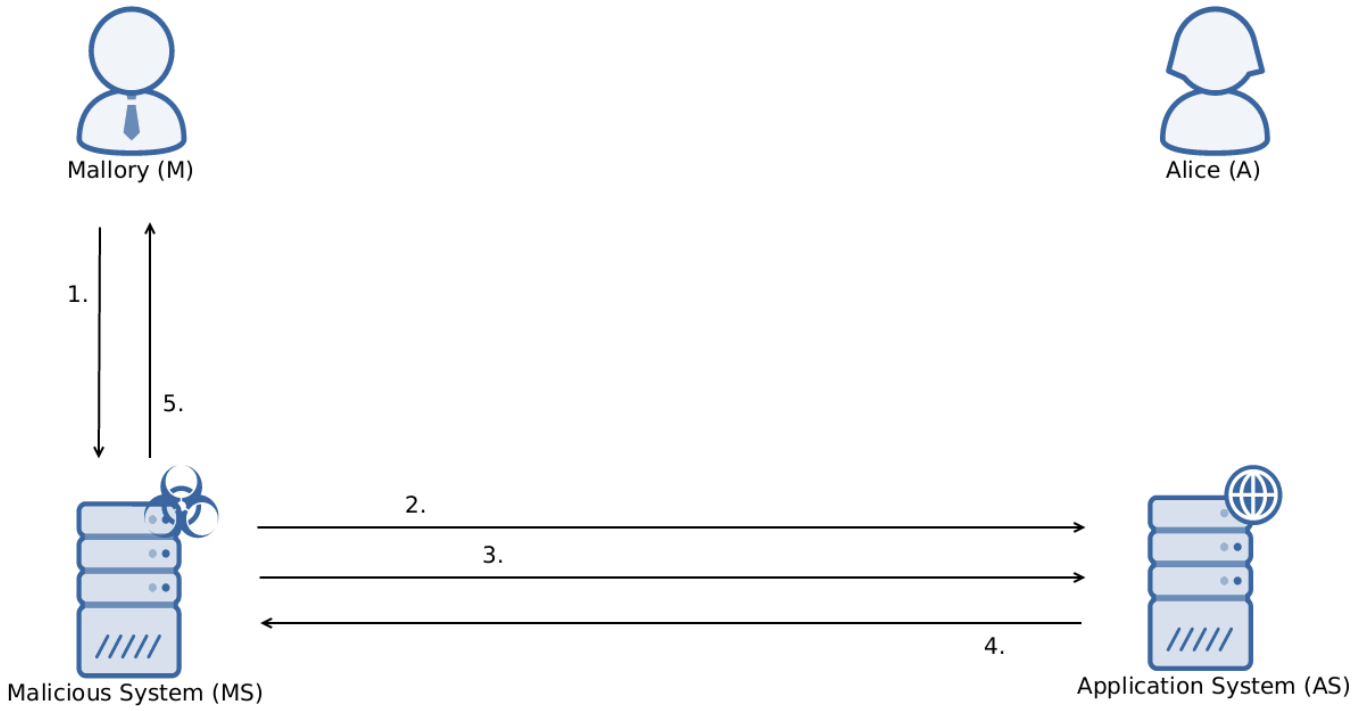
Please write your answer here:

[]The bank reports this attack to the NCSC. How would this attack influence the appearance of the Cyber Security Dashboard?

Please write your answer here:

[]

Mallory is trying to attack the bank with the attack shown in the picture. What kind of attack could this be?



Please write your answer here:

[]The bank reports this attack to the NCSC. How would this attack influence the appearance of the Cyber Security Dashboard?

Please write your answer here:

Personal Impression

Please describe your impressions concerning the Cyber Security Dashboard by answering the following questions.

It is important for the evaluation to have precise answers to all questions. Please take your time to think about the questions and answer them in detail.

If you like to look at the Cyber Security Dashboard before answering one question, you can always look at the screenshot or switch to the dashboard by pressing Ctrl+Tab on the keyboard.

[]Please name all terms from the Cyber Security Dashboard that you did not understand (if any).

Please write your answer here:

[]What insight can the Cyber Security Dashboard provide to you? Please think especially about the benefits it can give you for your further work or research.

Please write your answer here:

[]How appropriate is the visualization of the data in the Cyber Security Dashboard?

Please write your answer here:

[]What can be improved about the Cyber Security Dashboard?

Please write your answer here:

[]Do you have any comments about the Cyber Security Dashboard or the study?

Please write your answer here:

Thank you very much for taking your time evaluating the Cyber Security Dashboard.

Submit your survey.
Thank you for completing this survey.

Acronyms

Ajax Asynchronous JavaScript and XML.

CERT Computer Emergency Response Team.

CISO Chief Information Security Officer.

CSAN Cyber Security Assessment Netherlands.

CSAT Cyber Security Analytic Tool.

CSD Cyber Security Dashboard.

CSIRT Computer Security Incident Response Team.

CSS Cascading Style Sheets.

CSV Comma Separated Values.

CVE Common Vulnerabilities and Exposures.

DAAD Deutscher Akademischer Austauschdienst [German Academic Exchange Service].

DDOS Distributed Denial of Service.

DOM Document Object Model.

ECIR Explorations in Cyber International Relations.

ENISA European Union Agency for Network and Information Security.

GQM Goal Question Metric.

HCI Human Computer Interaction.

IoT Internet of Things.

ISACA Information Systems Audit and Control Association.

ISM Information Security Management.

IT Information Technology.

KPI Key Performance Indicator.

MinVenJ Ministerie van Veiligheid en Justitie [Ministry of Security and Justice].

MITM Man-in-the-Middle.

NCSC Nationaal Cyber Security Centrum [National Cyber Security Center].

npm Node Package Manager.

NTD Notice and Take Down.

SPSSX Software Package for Statistics and Simulation Extended.

SQL Structured Query Language.

SSH Secure Shell.

UEQ User Experience Questionnaire.

WODC Wetenschappelijk Onderzoek- en Documentatiecentrum [Research and Documentation Centre].

List of Figures

List of Listings

List of Tables

Bibliography

[1] E. H. Adelson. *Download Checkershadow Illusion*. 1995. URL: http://web.mit.edu/persci/people/adelson/checkershadow_downloads.html (visited on 01/05/2016).

[2] F. Asgharpour, D. Liu, and L. J. Camp. "Mental Models of Computer Security Risks." In: *Workshop on the Economics of Information Security*. Pittsburgh, 2007, pp. 1–9.

[3] V. R. Basili and H. D. Rombach. "Tame Project: Towards Improvement-Oriented Software Environments." In: *IEEE Transactions on Software Engineering* 14.6 (1988), pp. 758–773.

[4] M. Ben-Ari. "Constructivism in computer science education." In: *ACM SIGCSE Bulletin* 30.1 (1998), pp. 257–261.

[5] D. Bertram. *Likert scales are the meaning of life*. Tech. rep. 2013.

[6] A. Biselli. *Wir veröffentlichen Dokumente zum Bundestagshack: Wie man die Abgeordneten im Unklaren ließ*. 2016. URL: https://netzpolitik.org/2016/wir-veroeffentlichen-dokumente-zum-bundestagshack-wie-man-die-abgeordneten-im-unklaren-liess (visited on 03/11/2016).

[7] J. Blythe and L. J. Camp. "Implementing mental models." In: *IEEE CS Security and Privacy Workshops, SPW 2012*. 2012, pp. 86–90.

[8] M. Bodt. *The Many Ways of Creating Dashboards Using SAS®*. Tech. rep. 2014.

[9] C. A. Brewer. *ColorBrewer: Color Advice for Maps*. 2016. URL: http://colorbrewer2.org/ (visited on 01/20/2016).

[10] S. Carey. "Cognitive Science and Science Education." In: *American Psychologist* 41.10 (1986), pp. 1123–1130.

[11] R. Choenni, E. Leertouwer, M. Spaans, E. P. Hoorweg, P. de Graaf, C. F. M. Sadée, and R. M. Oudeman. *Eindrapport Cyber Security Dashboard – Haalbaarheidsonderzoek fase 1 SBIR*. Tech. rep. Capgemini Consulting, WODC, NCSC, 2013.

[12] R. Choenni and E. Leertouwer. "Public safety mashups to support policy makers." In: *Lecture Notes in Computer Science*. Vol. 6267. 2010, pp. 234–248.

[13] K. J. W. Craig. *The Nature of Explanation*. Cambridge: Cambridge University Press, 1943.

[14] CTF365 Blog. *Interactive Cyber Attack Map*. 2014. URL: https://blog.ctf365.com/interactive-cyber-attack-map/ (visited on 10/14/2015).

[15] M. Davidson, L. Dove, and J. Weltz. *Mental models and usability*. Tech. rep. Chicago: Depaul University, 1999.

[16] Deutsche Telekom AG. *Sicherheitstacho.eu*. URL: http://sicherheitstacho.eu/ (visited on 10/06/2015).

[17] Deutsche Telekom AG Honeypot Project. *T-Pot: A Multi-Honeypot Platform*. 2015. URL: https://dtag-dev-sec.github.io/mediator/feature/2015/03/17/concept.html (visited on 10/14/2015).

[18] A. Dix. "Human-computer interaction: A stable discipline, a nascent science, and the growth of the long tail." In: *Interacting with Computers* 22.1 (2010), pp. 13–27.

[19] N. Döring and J. Bortz. *Forschungsmethoden und Evaluation in den Sozial- und Humanwissenschaften*. 5th ed. Berlin, Heidelberg: Springer, 2015.

[20] J. K. Doyle and D. N. Ford. "Mental models concepts for system dynamics research." In: *System Dynamics Review* 14.1 (1998), pp. 3–29.

[21] Eerste Kamer der Staten-Generaal. *kst-33662-A*. 's Gravenhage, 2015.

[22] R. Eikenberg. *CTB-Locker: Krypto-Trojaner befällt hunderte Webserver*. 2016. URL: http://www.heise.de/security/meldung/Admins-aufgepasst-Krypto-Trojaner-befaellt-hunderte-Webserver-3116470.html (visited on 02/26/2016).

[23] S. Erven and S. Merdinger. *Just What the Doctor Ordered*. 2013. URL: http://campustechnology.com/Articles/2010/01/01/Just-What-the-Doctor-Ordered.aspx (visited on 10/06/2015).

[24] European Union Agency for Network and Information Security. *ENISA Threat Landscape 2014*. December. 2014.

[25] Extemporalist. *Labelled Bullet Graph Example*. 2015. URL: https://commons.wikimedia.org/wiki/File:Labelled_Bullet_Graph_Example.svg (visited on 01/05/2016).

[26] S. Few. "Dashboard Confusion Revisited." In: *Perceptual Edge* (2007), pp. 1–6.

[27] S. Few. *Information Dashboard Design*. 1st ed. North Sebastopol: O'Reilly Media, 2006.

[28] T. Gondrom, M. Morana, S. Tan, and C. Watson. *CISO Survey and Report 2013*. OWASP Foundation, 2014.

[29] J. S. Harris. "New product profile chart." In: *C&EN* (Apr. 1961).

[30] J. Hinckley, J. Hinckley, and J. G. Robinson. *The Big Book of Car Culture: The Armchair Guide to Automotive Americana*. St. Paul: Motorbooks, 2005.

[31] S. Hocevar. *About the WTFPL*. 2016. URL: `http://www.wtfpl.net/about/` (visited on 01/12/2016).

[32] Hogeschool Rotterdam. *Creating 010*. URL: `https://www.hogeschoolrotterdam.nl/onderzoek/kenniscentra/creating-010/` (visited on 02/25/2016).

[33] Hogeschool Rotterdam. *Rotterdam University of Applied Sciences*. URL: `https://www.rotterdamuas.com/about/rotterdam-uas/` (visited on 02/25/2016).

[34] Information Systems Audit and Control Association. *CISA Review Manual 2007*. Illinois, 2007.

[35] International Organization for Standardization and International Electrotechnical Commission. *ISO/IEC 27000*. 2014.

[36] Invincea Labs. *Cyber Green Research Paper*. Tech. rep. 2015.

[37] D. H. Jonassen and P. Henning. "Mental Models : Knowledge in the Head and Knowledge in the World." In: *Technology* 39.3 (1996), pp. 433–438.

[38] R. Kang, L. Dabbish, N. Fruchter, and S. Kiesler. ""My Data Just Goes Everywhere:" User Mental Models of the Internet and Implications for Privacy and Security." In: *Symposium on Usable Privacy and Security*. 2015, pp. 39–52.

[39] K. Knorr-Cetina. *The Manufacture of Knowledge: An Essay on the Constructivist and Contextual Nature of Science*. 1st ed. Oxford: Pergamont Press Ltd., 1981.

[40] N. Kolb. "3, 2, 1, meins! Analyse der Nutzerziele beim Onlineshopping mittels der "Mental Model Diagramm"-Methode." Master's Thesis. Technische Universität Darmstadt, 2015, p. 233.

[41] B. Laugwitz, T. Held, and M. Schrepp. "Construction and Evaluation of a User Experience Questionnaire." In: *4th Symposium of the Workgroup Human-Computer Interaction and Usability Engineering of the Austrian Computer Society*. Ed. by A. Holzinger. Graz: Springer Berlin Heidelberg, 2008, pp. 63–76.

[42] J. Liu, C. Zhao, and H. Zuo. "Engineering Psychology and Cognitive Ergonomics." In: *Engineering Psychology and Cognitive Ergonomics, 12th International Conference, EPCE 2015, Held as Part of HCI International 2015*. Vol. 9174. 2015, pp. 129–139.

[43] S. Madnick, X. Li, and N. Choucri. "Experiences and Challenges with using CERT Data to Analyze International Cyber Security." 2009.

[44] S. Mauw and M. Oostdijk. "Foundations of Attack Trees." In: *Lecture Notes in Computer Science* 3935 (2006), pp. 186–198.

[45] P. Mayring. "Qualitative Inhaltsanalyse." In: *Handbuch Qualitative Forschung in der Psychologie*. Ed. by G. Mey and K. Mruck. 1st ed. Wiesbaden: VS Verlag für Sozialwissenschaften, 2010, pp. 601–613.

[46] P. Mayring. *Qualitative Inhaltsanalyse Grundlagen und Techniken*. 12th ed. Weinheim und Basel: Beltz, 2015.

[47] S. McNeil. "Visualizing mental models: Understanding cognitive change to suppor teaching and learning of multimedia design and development." In: *Educational Technology Research and Development* 63.1 (2015), pp. 73–96.

[48] B. P. Meier, P. R. D'Agostino, A. J. Elliot, M. A. Maier, and B. M. Wilkowski. "Color in context: psychological context moderates the influence of red on approach- and avoidance-motivated behavior." In: *PloS one* 7.7 (Jan. 2012), e40333.

[49] R. Milzarek. "Analyse des Mehrwerts von innovativen Usability- und dynamischen Visualisierungskonzepten für die Darstellung von KPIs der Siemens AG und prototypische Implementierung einer nativen iOS-Applikation." Bachelor's Thesis. Technische Universität München, 2013, p. 104.

[50] Ministerie van Veiligheid en Justitie. *Organization | Research and Documentation Centre (WODC)*. en-GB. URL: https://english.wodc.nl/organisatie/ (visited on 10/26/2015).

[51] Ministerie van Veiligheid en Justitie. *What is the NCSC? | NCSC*. URL: https://www.ncsc.nl/english/organisation (visited on 10/26/2015).

[52] National Cyber Security Center. "Cyber Security Assessment Netherlands 2014." 2014.

[53] National Cyber Security Center. "Cyber Security Assessment Netherlands 2015." 2015.

[54] National Cyber Security Center. "Cybersecuritybeeld Nederland CSBN 2015." 2015.

[55] J. Nielsen. "A meta-model for interacting with computers." In: *Interacting with Computers* 2.2 (1990), pp. 147–160.

[56] J. Nielsen. "Iterative Design of User Interfaces." In: *IEEE Computer* 26.11 (1993), pp. 32–41.

[57] J. Nielsen. *Why You Only Need to Test with 5 Users*. 2000. URL: https://www.nngroup.com/articles/why-you-only-need-to-test-with-5-users/ (visited on 01/22/2016).

[58] D. A. Norman. "Cognitive engineering." In: *User centered system design*. Ed. by D. A. Norman and S. W. Drapper. 1st ed. Hillsdale: Lawrence Erlbaum Associates Inc., 1986, pp. 31–61.

[59] D. A. Norman. *Design of Everyday Things*. New York: Basic Books, 1990.

[60] Open Source Initiative. *The MIT License (MIT)*. URL: `https://opensource.org/licenses/MIT` (visited on 01/12/2016).

[61] R. Phillips, D. Lockton, S. Baurley, and S. Silve. "Making instructions for others: exploring mental models through a simple exercise." In: *Interactions* 20.5 (2013), pp. 74–79.

[62] J. Pratt, P. V. Radulescu, R. M. Guo, and R. A. Abrams. "It's Alive!: Animate Motion Captures Visual Attention." In: *Psychological Science* 21.11 (2010), pp. 1724–1730.

[63] Rede Nacional CSIRT. *Proposal for a Common Taxonomy for the National Network of CSIRTs*. Tech. rep. 2014, p. 14.

[64] B. Schneier. "Schneier on Security: Attack Trees." In: *Dr. Dobb's Journal* (1999).

[65] N. Staggers and a. F. Norcio. "Mental models: concepts for human-computer interaction research." In: *International Journal of Man-Machine Studies* 38.4 (1993), pp. 587–605.

[66] T. C. Summers. "How Hackers Think: A Study of Cybersecurity Experts and Their Mental Models." In: *Third Annual International Conference on Engaged Management Scholarship, Atlanta, Georgia* June (2013), pp. 1–25.

[67] T. Tidwell, R. Larson, K. Fitch, and J. Hale. "Modeling Internet Attacks." In: *2001 IEEE Workshop on Information Assurance and Security*. 2001.

[68] Trend Micro Incorporated. "Report on Cybersecurity and Critical Infrastructure in the Americas." 2015.

[69] D. S. Wall. "Dis-organised Crime : Towards a Distributed Model of the Organization of Cybercrime." In: *The European Review of Organised Crime* 2.2 (2015), pp. 71–90.

[70] R. Wash and E. Rader. "Influencing Mental Models of Security: A Research Agenda." In: *Proceedings of the 2011 workshop on New security paradigms* (2011), pp. 57–66.

[71] S. Weibelzahl, E. Herder, M. Rokicki, D. Heckmann, K. Müssig, and J. Schildt. "Personalized Advice and Feedback for Diabetes Patients." In: *Mensch und Computer 2015–Workshopband*. Ed. by A. Weisbecker, M. Burmester, and A. Schmidt. Berlin/Buston: Walter de Gruyter GmbH & Co KG, 2015.

[72] S. Weinschenk. *The Secret to Designing an Intuitive UX : Match the Mental Model to the Conceptual Model: Match the Mental Model to the Conceptual Model*. 2011. URL: https://uxmag.com/articles/the-secret-to-designing-an-intuitive-user-experience (visited on 10/13/2015).

[73] Wired UK. *Infoporn: Cyberattacks have created an invisible but vast war zone*. 2015. URL: http://www.wired.co.uk/magazine/archive/2015/10/start/infoporn-cyberattacks-state-sponsored-hacking (visited on 10/14/2015).

[74] I. Young. *Mental Models: Aligning Design Strategy with Human Behavior*. New York: Rosenfeld Media, 2008, p. 299.

[75] H. Yurdugul. "Minimum Sample Size for Cronbach's Coefficient Alpha: A Monte-Carlo Study." In: *Hacettepe University Journal of Education* 35 (2008), pp. 397–405.