Chapter 15
Best Practices and Recommendations for Cybersecurity Service Providers

Alexey Kirichenko, Markus Christen, Florian Grunow, and Dominik Herrmann

Abstract This chapter outlines some concrete best practices and recommendations for cybersecurity service providers, with a focus on data sharing, data protection and penetration testing. Based on a brief outline of dilemmas that cybersecurity service providers may experience in their daily operations, it discusses data handling policies and practices of cybersecurity vendors along the following five topics: customer data handling; information about breaches; threat intelligence; vulnerability-related information; and data involved when collaborating with peers, CERTs, cybersecurity research groups, etc. There is, furthermore, a discussion of specific issues of penetration testing such as customer recruitment and execution as well as the supervision and governance of penetration testing. The chapter closes with some general recommendations regarding improving the ethical decision-making procedures of private cybersecurity service providers.

Keywords Data handling · Data sharing · Penetration testing · Threat intelligence · Vulnerability disclosure

A. Kirichenko (✉)
F-Secure Corporation, Helsinki, Finland
e-mail: alexey.kirichenko@f-secure.com

M. Christen
Digital Society Initiative, University of Zurich, Zürich, Switzerland
e-mail: christen@ethik.uzh.ch

F. Grunow
ERNW. Enno Rey Netzwerke GmbH, Heidelberg, Germany
e-mail: fgrunow@ernw.de

D. Herrmann
Privacy and Security in Information Systems Group (PSI), University of Bamberg, Bamberg, Germany
e-mail: dominik.herrmann@uni-bamberg.de

© Springer Nature Switzerland AG 2019

M. Christen et al. (eds.), *The Ethics of Cybersecurity*, The International Library of Ethics, Law and Technology 21,
https://doi.org/10.1007/978-3-030-29053-5_15

15.1 Introduction: Dilemmas of Cybersecurity Service Providers

Security software and service providers—usually private companies—play a pivotal role in cybersecurity as they provide the competences and tools for defending the IT infrastructure, devices and data of their customers. Individuals, companies and state agencies put a considerable amount of trust in the tools and services of cybersecurity service providers. Furthermore, those specialised companies often obtain deep insights into the IT infrastructure and information processes of their customers and—more generally—a deep understanding of cyber threats, which provides them with a special responsibility on how to handle such knowledge and customer data. When cybersecurity service providers perform evaluations such as penetration testing, additional responsibilities come into play.

Complicating matters further, new dilemmas have emerged due to partially conflicting regulations, the possibility that some customers may break laws and the fact that state actors are increasingly involved in carrying out cyberattacks. We exemplify these challenges with a dilemma related to threat intelligence and malware detection capabilities of cybersecurity companies.

15.1.1 Example: Dealing with Governmental Malware

The use of malware by governments and Law Enforcement Agencies (LEA) for surveillance and other purposes is a widely accepted fact. In 2014, F-Secure's CRO Mikko Hyppönen said: "If someone had come to me ten years ago and told me that by 2014 it will be commonplace for democratic Western governments to write viruses and actively deploy them against other governments, even friendly governments, I would have thought it was a movie plot. But that's exactly where we are today" (Thomson 2014).

The reasons for governmental use of malware and other details of such operations vary widely, including, for example, terrorist activities investigations, espionage or tracking opposition journalists. One early case is Magic Lantern (Martin 2003), a keystroke logging software developed by the United States' FBI. It could be installed remotely, via an e-mail attachment or by exploiting common operating system vulnerabilities. Already back then there were concerns expressed by antivirus vendors "that FBI software reportedly designed for covert keystroke monitoring could fall into the wrong hands" (Jackson 2001).

In 2011, a well-established group of German hackers accused the German government of releasing a backdoor Trojan into the wild. Security firm F-Secure confirmed that the program included a keylogger and code that could take screenshots and record audio (e.g. for room surveillance; Bott 2011). The group reverse-engineered and analysed the program, which it called "a lawful interception malware program used by German police forces". The malware also offered a remote

control or backdoor functionality for uploading and executing arbitrary other programs. The program's behaviour went well beyond the ability to "observe and intercept Internet based telecommunication" (in other words, wiretapping Internet-based telephony), which is allowed by German courts. In addition, significant design and implementation flaws essentially made all of the functionality available to anyone on the Internet. Figures published recently (2017) revealed that the top five countries originating use of malware and other cyber-attacks were USA, China, Brazil, India and Russia (Ley 2018).

However, improving malware and attack detection capabilities of cybersecurity solutions clearly complicates the development and operation of malware by state agencies. Already in 2001, the public disclosure of the existence of Magic Lantern sparked a debate as to whether anti-virus companies should detect the FBI's keystroke logger or could agree to whitelist it. There were rumours that Network Associates (maker of McAfee anti-virus products at that time) had contacted the FBI following press reports about Magic Lantern to ensure their anti-virus software would not detect the program. Network Associates issued a denial, fuelling speculation as to which anti-virus products might or might not detect government Trojans.

It is likely that the news that North Korea's antivirus software whitelisted malware (Sharwood 2018) did not come as a surprise: "Just why North Korea's government wants software that won't spot some viruses is not hard to guess: a totalitarian dictatorship can only sustain itself with pervasive surveillance and leaving a backdoor that allows viruses in would facilitate just that." There is, however, little evidence on how malware used by Western governments is treated by Western cybersecurity companies. Responding to a transparency plea from leading privacy and security experts in 2013, a number of leading antivirus firms stated that they had strict policies against aiding law enforcement by whitelisting spyware or building backdoors into their software. For example, Symantec stated: "We have a strict policy against whitelisting malware for law enforcement and governments globally. We have never received such a request from law enforcement" (Westervelt 2013). We should note, however, that some of the approached companies failed to respond to the plea (Schwartz 2013). The summary published by Bruce Schneier at that time was: "Understanding that the companies could certainly lie, this is the response so far: no one has admitted to doing so" (Schneier 2013).

In fact, a need to comply with a law or a court order to whitelist governmental malware will clearly leave cybersecurity companies with no choice. Situations can easily be imagined, however, when a difficult choice does exist, because there are no appropriate laws, because a cybersecurity vendor is approached by a LEA from a different state, or due to other reasons. Should the vendor plainly say 'no' to the whitelisting request? Alternatively, should it weigh the consequences of preventing or hampering the LEA operations against privacy of its customers, dangers of a potential use of the backdoor by cybercriminals, and other relevant factors, which may also introduce serious uncertainties in the decision-making?

15.1.2 Dilemmas of Cybersecurity Service Providers

The example above is one of several dilemmas cybersecurity service providers may face in their daily operations, where often no clear legal guidance is provided. This is where ethical considerations come into play to find an optimal solution. In addition to the 'whitelist governmental malware?' question discussed above, the following list provides further examples of such dilemmas:

- *Should questionable customers be protected?* Not all customers of cybersecurity service providers may have good intentions; some may even be labelled 'criminals' with respect to certain legal systems. However, whether certain customers are considered 'questionable' may not in any case be clear, for example if they reside in authoritarian or totalitarian states and are surveyed for political reasons. Furthermore, cybersecurity service providers may also choose to collaborate with such states and thus may become instruments for questionable aims. Although the judicial system in which the cybersecurity vendor is operating (also with respect to export regulations or sanction regimes) may to a certain extent provide a framework for declining certain customers or allowing LEAs to monitor their devices, handling such cases sometimes remains an issue of the company's policy and attitude.
- *Should incidental findings be disclosed?* Cybersecurity service providers have considerable access to the information flows of their customers in order to detect, for example, hostile intrusions. However, what should a cybersecurity vendor do when they find traces of a crime committed by a customer company when monitoring its network? To a certain extent, the legal code of the customer's location may provide answers in such cases, e.g. when the crimes concern child pornography or crimes against humanity. It must also be considered that some legislation may forbid disclosing such information. In-between those legal boundaries, a space for ethical choices remains.
- *Should illegal breaches be profited from?* Some cybersecurity service providers (for example the Italian company 'Hacking Team') provide offensive technology to the worldwide law enforcement and intelligence communities; i.e. tools against which other cybersecurity vendors develop countermeasures. Sometimes, illegal breaches may reveal source-code of such offensive tools, which legally is considered a break of a trade secret. Should an antivirus company analyse the source code in order to improve their tools? Again, a clear legal regulation is not available for such cases and an ethical choice has to be made.
- *Should non-customers be informed about potential risks?* Private cybersecurity service providers operate under constraints to optimise revenue. How should such organisations deal with findings that do not directly lead to an increased revenue? For example, should they inform victims even if these are not their customers? Although most cybersecurity vendors work in a commercial environment, they rely on the work of many volunteers. It is therefore recommended that a commercial security organisation gives something back to the community, be

it information or tools. To what extent are cybersecurity service providers able to contribute to the community and how can such behaviour be incentivised?

In the following, we do not further discuss such dilemmas in detail. Rather, we provide an overview of domains, where cybersecurity service providers should implement policies in order to handle challenging situations in an optimal way.

15.2 Domains for Policy Implementations

15.2.1 Customer Data Handling

Data handling is a fairly heavily regulated domain, especially if personal data is involved (see Chap. 5). Cybersecurity service providers operating under the regime of the General Data Protection Regulation (GDPR) of the EU have to fulfil the principles stated within, in particular transparency. Cybersecurity vendors need to inform their customers of what data they collect, how they process it, for what purposes, etc. To analyse this aspect, we distinguish between data emerging from consumers (e.g. individuals buying an antivirus tool) and companies that usually make use of a broader spectrum of services.

Taking the practices at F-Secure as an example, consumer-related data can be differentiated into the following categories:

- *Client relationship data.* This data is necessary to manage the relationship of the company with its clients, and to market and sell the company services to them or to the legal entity that they represent. Any company on the free market seeking customers will collect such data.
- *Service data.* This data is automatically processed in order to provide the clients with the services that they requested. This also includes the data that the clients actively submit to the vendors when subscribing to their services. Again, any entity on the free market that sells services to customers will collect such data.
- *Security data.* This usually concerns anonymous or pseudonymous data that the company needs to collect to keep the clients secure, for instance, execution details of certain programs on a client device or its networking activities.
- *Analytics data.* This concerns additional anonymous or pseudonymous data that the companies collect to learn when and how their services are found and used, for example, which protection features of a specific security product are enabled by a specific customer or how many infections were detected and blocked in a given customer device.

Whereas data from the first two categories are ‘standard data’ that any company will collect from their customers, the last two categories refer to specific data sets only available to cybersecurity service providers. It is therefore recommended to place these data sets in their own ‘silos’ to ensure that, in particular, security data is processed separately from data of the other types. To defend against a specific

malware, a cybersecurity vendor does not need to know whether a particular user has been infected with that malware. Rather, the company only needs to know that this new form of malware emerged, analyse it, and then provide countermeasures to all of their customers. Analytics data should be processed in pseudonymised form by default, hence enabling the sharing of data among developers without privacy risks to an individual. Service data (i.e. name, email-address and other identifiers) and analytics data are combined only based on specific rules, for instance, to make it possible to send a reminder to a customer which purchased but did not activate certain security service. Via access control policies, cybersecurity service providers ensure that their marketing people do not have access to the analytics data, and all the other departments have no access to the service data.

Corporate customers often use the same products as consumers; hence, the same types of data and policies as described above are relevant. However, corporate customers may in addition use Advanced Threat Protection (ATP), vulnerability scanning and other products that go beyond the standard Endpoint Protection paradigm, as they enhance corporate networks security. Separating the device identity from the security analysis is no longer sufficient for ATP products. Since anti-malware activities have moved from detecting malicious code to detecting malicious behaviour, protecting corporate networks requires more context for analysing device and user behaviour.

The above observation means that anonymisation and data separation are no longer a viable approach. Hence, to safeguard privacy, more focus needs to be put into alternative protection means such as sufficiently granular access control mechanisms, security personnel activity logging or usage guidance. An increasingly typical occurrence is that security and service data of corporate security products are processed jointly, with a pseudonymisation that takes place between corporate customers and a cybersecurity provider, as opposed to pseudonymisation within the provider's systems. In particular, this means that the only way for a security provider to learn actual names of employees of a corporate customer under protection is to ask the customer's Information Security department or management (this may be necessary, e.g., in a security incident investigation). To avoid further complications, analytics data (for product improvement purposes) is usually not collected from corporate security products.

Many cybersecurity vendors publicly state their privacy and data handling principles and practices (one can find examples of such statements at https://www.f-secure.com/en/web/legal/home and https://www.f-secure.com/en/web/legal/privacy). To conclude this section, we would like to list the following—more technical—recommendations to keep in mind when working with different types of data:

- Steps should be taken to ensure that the *telemetry data* collected and stored about security incidents and system configurations *is always anonymous or pseudonymous*.
- *The STRIDE model should be used (*Swiderski and Snyder 2004*), which stands for six categories of security threats:* Spoofing of user identity, Tampering, Repudiation, Information disclosure (privacy breach or data leak), Denial of

Service (DoS), Elevation of privilege. Threat analysis sessions should be conducted when planning new data handling-related functionalities and reviewing their readiness.
- *The quality of pseudonymisation functionality and procedures should be ensured.* The small amount of (e.g. marketing related) telemetry data which contains identifiable data must be pseudonymised before it can be used for analytics purposes, and the ability to reverse the pseudonymisation must be very strictly limited and controlled. Every effort should be made to remove personal identifiers from file paths and file names before they are made available for analytics systems and the pseudonymisation code should be concentrated to a common library as far as possible. Its performance should be reviewed regularly.
- *Unnecessary data should not be collected.* Data collected should always be for a purpose. Review processes should be implemented to regularly check whether telemetry and other data is really needed and to stop collecting it if it is not.
- *Clear data management procedures should be implemented*. Cloud accounts should by default be read-only and extra privileges should be required to be able to make even small changes there. Cloud service account boundaries (e.g. for Amazon Web Services accounts) should be used as a means to isolate more general accounts from data accounts and maintain tight access control on who has access to which data. Encryption contexts should be used to limit the power to decrypt data.

It must be emphasised that such policies and principles would not resolve all problems that may arise in practice. For example, the highly desirable data integrity property may conflict with the rectification and erasure requirements defined by the GDPR. Such a problem can arise if a user asks their cybersecurity provider to remove certain parts of security data collected from her or his machine. Satisfying the user's request will effectively make the collected security data incomplete and possibly useless in incident investigations.

15.2.2 Information About Breaches

A sensitive issue for cybersecurity service providers is when they become a victim of a data breach themselves, as this may have a direct impact on their reputation. There are several incentives to 'hack' a cybersecurity vendor. For example, state actors may be interested in knowing malware detection mechanisms to be able to circumvent them for specific intelligence operations.

According to the GDPR, cybersecurity vendors that are active in the EU have a legal obligation to provide protection against "accidental or unlawful destruction, loss, alteration, unauthorised disclosure of, or access to personal data". Failure to fulfil the obligation may result in authority investigation or administrative fines.

It is therefore necessary that cybersecurity service providers have procedures for managing personal data breaches. This includes criteria for determining whether a

security incident constitutes a personal data breach according to the GDPR and procedures for decision-making and evidence preservation that must be followed when handling such breaches. We list here a number of practical considerations for handling personal data breaches that are relevant for a typical cybersecurity service provider:

- All personal data breaches should be recorded, including suspected ones and confirmed false positives. Timestamped minutes detailing all the facts and assumptions should be included as well as the risk-based reasoning for whether to treat a specific incident as a personal data breach.
- In a team managing a suspected personal data breach case, members of the Chief Information Security Office, Legal Department and Executive Team should be included.
- When performing a risk assessment of a breach, the following factors should be considered: number of impacted individuals; data types; breach types (e.g. accidental, unlawful), data protection mechanisms used for the affected data (e.g. was the data encrypted).
- When making a decision to publish a Personal Data Breach Notification, the following factors should be considered: whether the identities of the affected parties and impact of the breach to them is known; whether the relevant aspects of the security incident and breach are known at a reasonable confidence level; whether taking extra investigation time would benefit the understanding of the breach or help limit its negative effects to the involved parties (data subjects, controller(s), and the cybersecurity service provider), or whether it would instead be likely to aggravate the situation; whether there are specific items that should be omitted when sharing information because disclosing those would aggravate the impact of a breach to the data subjects.
- Unless a personal data breach is unlikely to result in risks to the rights and freedoms of natural persons, the Supervisory Authority should be notified about the breach.

15.2.3 Threat Intelligence Activities

Threat intelligence is information that helps understand threats targeting organisations and citizens, in the past, at present and in the future. Thus, the production of threat intelligence is a core activity of cybersecurity providers for preventing, detecting and responding to threats to their customers and general public. Threat intelligence is what cyber threat information becomes once it has been collected, evaluated in an appropriate context (in particular, considering information source and reliability), and analysed through structured techniques, identifying similarities and correlations in vast quantities of data. Threat intelligence production is not an end-to-end-process; rather, it is a circular process whereby requirements are stated; data collection is planned, implemented and evaluated, the results are analysed to

produce intelligence, and the resulting intelligence is disseminated and re-evaluated in the context of new information and consumer feedback. The process is a cycle because it identifies intelligence gaps and unanswered questions, which prompt new collection requirements, thus restarting the intelligence cycle (Intel 2019).

In this circular process, regular access to—potentially sensitive—information is necessary for cybersecurity service providers, which obviously requires defined data management procedures. Ideally, cybersecurity vendors must ensure that they are open about what they collect, that only necessary data is collected, that they use it only for pre-defined and justified purposes and give it out only on a need-to-know basis for legally permissible and ethical use, that they keep it secure and destroy it when they no longer need it, and that sensitive pieces of data are removed or anonymised whenever possible. In reality, however, they often face threat intelligence-related trade-offs and choices with no clearly defined rules.

We focus here on the potential privacy impact of threat intelligence activities. It is typical for cybersecurity vendors to collect publicly accessible data on the Internet, extract metadata from selected files collected on the Internet or received from various feeds (e.g. exchange with other vendors and research groups, see also Sect. 15.2.5), and analyse it all to gather more information for further pivoting and putting context around threats. In the collected data and extracted metadata, important for incident investigations and research, Personally Identifiable Information (PII) may be found and, by connecting pieces of data from multiple online sources and samples, actual identities of persons targeted by the analysed threats may be discovered. For instance, public social media data and profiles, including names, locations, workplaces, etc., may be valuable when only very selected profiles are associated with possible attackers or attack targets. Alternatively, data extracted from decoy documents and malicious emails can be used for pivoting, threat attribution and the identification of attacked organisations, or URL strings and whois data related to threat campaigns may include names, physical and email addresses and organisation names.

It is crucial for cybersecurity service providers to ensure that such data are collected and processed only for specific use cases and the goals of the research or investigation. The collected sensitive data should be assessed for relevance for the use case in question and, if found irrelevant, deleted immediately. Data relevant for the use case should be stored only as long as the use case continues to be relevant and preferably locally in researcher machines. If such data is ultimately stored in external services, appropriate safeguards must be designed and applied. If other organisations are involved in data collection (e.g. business or research partners), the process must be made transparent for all the parties and its privacy impact must be analysed. If collected data or produced intelligence are shared with others (e.g. with law enforcement agencies or cybersecurity research groups), Sect. 15.2.5 discusses the relevant considerations and practices. An important aspect of data sharing to consider is whether data are moved across borders and—if so—to what states.

We conclude this section with a simple piece of advice: Always consider carefully if your threat intelligence goals can be reasonably achieved with less identifiable data.

15.2.4 Vulnerability-Related Information

Finding vulnerabilities in software or system configurations of their customers is among the key activities of cybersecurity service providers. However, if a vulnerability has been found, defined processes are necessary for a proper response. Policies are required regarding documentation, handling, and what kind of and when vulnerability information is shared with other parties or made public. These policies include Security Advisory publishing and communication to ensure the controlled disclosure of security vulnerability information and appropriate balance between (a) letting the customers and partners know enough to protect themselves, and (b) communicating in a way that does not help attackers who want to exploit vulnerabilities. The following considerations should be remembered when preparing a security advisory on a vulnerability:

- Appropriate sensitivity classification of the vulnerability-related information should be assigned and ensured before the public release.
- Providing deep technical details about the vulnerability is typically more useful for persons who want to exploit it than for those who want to protect their systems, so unnecessary details should be avoided.
- All affected products and versions should be listed even if no fix is provided for some of those. There should be openness about the fact that users of discontinued and unsupported versions are running vulnerable software and that the only way to secure their systems is to upgrade to a supported version.
- It is always good to explicitly mention the product groups and versions that are NOT affected by the vulnerability.
- If the vulnerability is found externally, how the reporter wants the credits to be stated should be checked. Some external parties may not want the credits to be stated publicly. Contact information should only be included if approved by the reporter.
- Appropriate national CERTs can be informed before the issue becomes public, in particular, to communicate it to other CERT organisations around the world.
- The correct preparation should be put in place for when the media calls!

One instrument to gain information about vulnerabilities in own software are bug bounty programs. Using such programs, individuals can receive recognition and compensation for reporting bugs in software, especially those pertaining to exploitable vulnerabilities. Such programs allow the developers to discover and resolve bugs before malicious actors become aware of them, preventing abuse. Many organisations and companies have implemented bug bounty programs—and cybersecurity vendors use such programs, too. Again, policies are necessary if cybersecurity vendors launch bug bounty programs to identify vulnerabilities in their own software.[1] In particular:

[1] An example of a bug bounty program of F-Secure can be found here: https://www.f-secure.com/en/web/labs_global/vulnerability-reward-program

- A group of experts has to be established for reviewing cases reported to a bug bounty program.
- Rules for setting amounts of money to be paid for reported vulnerabilities need to be defined.
- The up-to-date program conditions and details should be clearly presented on the vendor's public website.
- A procedure for communicating with the bug reporters has to be defined and followed.
- Metrics for measuring the effectiveness and efficiency of the program are important for its business viability.

At first sight, bug bounty programs are beneficial, because they are a means to decrease the number of vulnerabilities while ensuring that independent analysts are compensated for their effort. However, such programs come with their own set of problems. High-profile companies have to dedicate sufficient manpower to the program in order to triage the incoming reports. Moreover, bug bounty programs can be criticised because they create harsh working conditions due to high fluctuations in pay and a work model that is entirely driven by results.

15.2.5 Data Sharing with Peers

Threat intelligence and attack-related objects (so-called samples, which are primarily malicious or unwanted programs, documents and other files, or URLs) that result from the activities of cybersecurity service providers can be shared with (usually only) a limited number of reputable and vetted partners in the cybersecurity domain to improve global cyberattack resilience. This enables faster and more accurate protection for customers of cybersecurity vendors and high-impact cybersecurity research. In this section, we present some simple principles and considerations in establishing exchange partnerships which are followed by many organisations and groups in the cybersecurity domain.

Very informal agreements often suffice between reputable partners in the cybersecurity industry and research, in particular if they meet the following criteria:

- They are members of well-established groups, such as the Anti-Malware Testing Standards Organization (https://www.amtso.org/) and the Association of Anti-Virus Asia Researchers (https://aavar.org/), or they are represented in the Computer Antivirus Research Organization (http://www.caro.org/index.html).
- They have no known misdemeanour in their track record, including activities that cast doubts over their ethics as an organisation.
- They are involved in activities within cybersecurity, such as antivirus, security research, data protection and suchlike.

Depending on the sensitivity and type of information being exchanged and the partners' background, a written signed agreement may also be concluded prior to

any exchanges. In any case, the following points are usually stated explicitly to avoid any doubts:

- Samples and URLs must be handled in a safe manner.
- Samples and URLs that are exchanged must not be re-shared as such without a clear additional contribution.
- Samples and URLs must not be redistributed to untrusted parties.
- Shared samples and URLs are free from compensation.
- Each party is responsible for its own use of the exchanged material.
- Each party is responsible for having the necessary rights and authorisations for this activity.

If a potential partner does not meet the criteria mentioned above, an extensive background check is usually carried out and the community is contacted for feedback regarding that organisation. A written agreement is always required in such cases prior to any exchanges to confirm the partner's commitment to comply with the established terms and rules of sample and URL exchange.

On-demand sample and URL sharing is normally allowed with trusted individuals in the cybersecurity community without formal agreements in place as long as they comply with best practices in the safe handling of samples and URLs. In this scenario, PGP encryption is always the preferred option to avoid the risk of unintended recipients being able to open exchanged packages.

As a part of such best practices, exchanged samples and URLs are never decrypted nor packaged on production machines, only in special safe environments. Furthermore, files inside packages should not have their original file extensions and package names must clearly describe their contents to prevent accidental execution. Samples and URLs marked as confidential, marked as illegal or containing potentially private data (e.g. email addresses, usernames and passwords within a URL) are excluded from sharing by all responsible organisations and groups. Whenever feasible, the origin of exchanged objects and data is anonymised.

A good example of a formal intelligence sharing partnership is presented by Cyber Threat Alliance (CTA, https://www.cyberthreatalliance.org/), a not-for-profit organisation working to improve the global digital ecosystem cybersecurity by enabling near real-time, high-quality cyber threat information sharing among its members. Members are granted access to the CTA's automated platform for sharing timely, actionable, contextualised and campaign-based cyber threat intelligence which can be used to improve their products and services to protect their customers, and regularly share insights and best practices. All potential members undergo a thorough vetting process, and the CTA also considers their potential value to the Alliance along with any possible security risks.

CTA intelligence sharing is grounded in five guiding principles, as stated on their website (available at: https://www.cyberthreatalliance.org/who-we-are/; last access July 7, 2019):

1. *"For the greater good*. We protect customers, strengthen critical infrastructure, and defend the digital ecosystem. Our automated platform empowers members

to share, validate and deploy actionable intelligence to customers in near-real time.
2. *Time is of the essence*. We prevent, identify and disrupt malicious activities by rapidly sharing timely, actionable intelligence and reducing the effectiveness of malicious actors' tools and infrastructure.
3. *Context rules*. We reward context sharing to identify indicators of compromise and provide useful information about those.
4. *Radical transparency*. We attribute intelligence to the member who submits it, but anonymise any and all victim and sensitive data.
5. *You must give to receive*. We require all members to share a minimum amount of intelligence with the alliance to prevent the free-rider problem."

It is noteworthy that one of the key CTA's rules states: "Affected entity's data in shared intelligence must be anonymised".

15.3 Special Considerations for Penetration Testing

Building systems that are absolutely secure from the beginning is virtually impossible due to high complexity and limited resources. Security-relevant mistakes can also be made during deployment and operations. A common approach to reduce the resulting risk is to perform regular security assessments (see Chap. 2). Penetration tests are one type of such assessment (Bishop 2007). In a penetration test, white-hat hackers (also called ethical hackers; see Chap. 9) target productively used systems under realistic conditions in a systematic fashion. Testers use similar or the same tools and techniques as malicious actors. Penetration tests are usually carried out by specialised cybersecurity service providers. The result of a penetration test is a report with a list of security-relevant findings and their severity, often including the steps to reproduce (exploit) them.

Conducting a penetration test involves numerous ethical dilemmas. In the following, we survey selected challenges. We also provide guidance for ethical decision-making based on experience obtained in various engagements.

15.3.1 Order Initiation

Clients that request the services of a penetration testing provider may not be upfront with their intentions. For instance, a client might request a test of a particular version of a product which is not in use in their organisation at all. However, it could very well be the case that the client secretly knows that this version is in use at another organisation, for instance a competitor. The client could also be a nation state agency involved in law enforcement or political espionage. Many penetration testers aspire to 'improving security' by improving defence measures and closing

vulnerabilities. They would not be willing to help clients gain technical expertise for offensive activities. However, it is quite difficult to identify such cases. For instance, clients could pretend that they do not use the to-be-tested product because they are in the middle of a buying decision and this particular product is one of the promising candidates.

Accepting such a job and delivering an in-depth report with details about found vulnerabilities could make the penetration testers complicit in conducting morally questionable activities. Depending on the financial situation at the penetration testing provider, rejecting such ambiguous assignments may not be an option. However, due to the information asymmetry between testers and clients, the testers can still influence the outcome and the potential harm resulting from their findings. For instance, they can refrain from creating and providing a working exploit and omit technical details to make it more difficult to write such an exploit based on their report. The ethical dilemma arises from the fact that it is now the testers who are not completely honest with their client, which is another facet of professional integrity.

Some penetration testers may research particular products independently (i.e. without a mandate by the vendor and without receiving payment) in their spare time. Such activities are often used to advertise the competencies of a penetration testing service provider. Typically, the testers will follow responsible disclosure procedures, i.e. notify the vendor about any security vulnerabilities (see Chap. 2), before a report with the results is published.

Several ethical dilemmas can be encountered during this process. Firstly, vendors of high-profile products know that many penetration testing providers will scrutinise their products immediately upon release because of the comparatively large advertisement value of finding vulnerabilities in them. Effectively, this means that these vendors get high-quality penetration testing without having to pay for it. Should penetration testers react to this form of financial exploitation and refrain from testing the products of vendors that implement this approach to create an incentive for vendors to perform security testing on their own? Alternatively, should it be considered that consciously refraining from independent tests of particular vendors introduces a systematic disadvantage for the (large numbers of) customers of these vendors?

Secondly, during the responsible disclosure process, vendors may try to prevent the penetration testers from publishing a report of their findings at the end, which might result in a loss of reputation for the vendor. This can either be done by threatening the penetration testers with legal means or by offering them a well-paid engagement, which—of course—comes with a non-disclosure agreement that prevents the testers from reporting the findings publicly. Again, it is necessary to establish a balance between professional integrity and generating a steady stream of revenue.

15.3.2 Execution

During the execution of penetration tests, testers make many decisions with significant technical and ethical implications. One of the most straightforward questions is: How aggressively should the test be conducted? More aggressive tests may uncover more vulnerabilities but may also create more harm (e.g. downtime). It is also interesting to ask how thoroughly a discovered vulnerability should be demonstrated. Is it sufficient to state that a vulnerability gives full access to a database with sensitive information of all employees or should the tester actually 'go the extra mile' and retrieve records from this database and include them in the final report to make it 'juicier'? Although this may actually be required to stress the severity of a vulnerability, it can also be seen as an avoidable violation of privacy.

Many penetration tests involve 'social engineering' (Mitnick and Simon 2005). Here, employees of the penetration testing provider deceive employees of the client in order to assess their security awareness and compliance. Social engineering ranges from sending tailored spear-phishing mails or giving them calls, for instance "from the IT department that needs their password", to entering the premises under the pretext of gaining physical access.

These interactions are problematic for both parties, client and service provider, and the ethics of social engineering are quite involving (Hatfield 2019). In contrast to searching for bugs in software, social engineering uncovers unprofessional behaviour in humans. It is all too easy to jump to the conclusion that the 'problem' can be resolved by replacing a particular employee who made a mistake. It may be difficult for a client to accept the more inconvenient explanation: a successful attempt at social engineering means, firstly, that the organisation lacks training procedures, and, secondly, that it lacks technical means such as strict access control and segmentation that limit the power that attackers can gain via social engineering. Therefore, engagements involving social engineering should always be implemented in a way that ensures that the client's employees do not face personal consequences based on the results.

Social engineering is also challenging for testers. Not only does it involve lying to other humans, it means strategically deceiving and tricking them, with the explicit objective of making them fail. Testers might have to go to great lengths to build up sufficient trust with their 'victims' to be successful. Social engineering providers should establish clear boundaries of acceptable behaviour for such engagements and should offer dedicated training to their employees. Moreover, they should give employees the freedom to reject social engineering assignments.

The final stage of a penetration test is the creation of a report. Clients may have a political agenda, asking for a 'green' report that plays down the severity of the findings. Other clients may ask for the opposite: a 'red' report, for instance, as the basis to request more funding within their organisation or to discredit the work of colleagues or other units. Such negotiated and intentional modifications of the original assessment conflict with professional integrity and are seldomly justifiable. However, the original message may also be modified unintentionally, for instance,

when clients ask for a management summary without too many technical details. Such a ‘dumbed down’ report can be easily misinterpreted.

15.3.3 Supervision and Governance

In the previous sections, we discussed ethical dilemmas in the realm of penetration testing. How can service providers ensure that they act in a responsible fashion?

First, penetration testing providers and their employees are in a powerful position because of a significant information asymmetry. Only the testers know the full truth. Their employer and the client have to trust them that they report all findings accurately and completely. Although it may be possible to reproduce the actual findings, it is difficult to verify their severity judgement, and virtually impossible to determine whether anything has been omitted or forgotten.

There are two ways of mitigating this information asymmetry between clients and providers. Firstly, clients can choose to engage different penetration testing providers to correlate their findings. Secondly, before the penetration test, clients can intentionally introduce vulnerabilities into their systems to test the abilities of the testers. A historic example of this approach is reported by Karger and Schell (2002) for the Multics system, who inserted malicious code in the system that was not uncovered even after the testers had been informed of its existence.

Even in the absence of such incentives, penetration testing service providers should work towards establishing a sense of work ethics and procedures that ensure high-quality work. For instance, it may make sense to pair up highly skilled junior security analysts with more experienced personnel to channel the curiousness and drive of the young and to avoid them overshooting the mark when they are ‘in the flow’. This is especially important because, as stated in Sect. 15.3.2, it may not be desirable to do everything that is technically possible during an attack, even though it may be very enticing.

Moreover, penetration testing providers should encourage their employees to reflect on the effects of their work on their clients and society at large as well as the way how they conduct it (i.e. whether the end result justifies the means). In addition, formal training, an open culture with informal meetings and discussions as well as opportunities for engaging with the community of security professionals may help penetration testers keep track of the right course.

Penetration testing providers should also consider institutionalising ethical decision making. One approach which seems to work well in practice consists of establishing an Ethics Board that is given authority to decide on the course of action in all morally questionable cases, ranging from whether or not to take an ambiguous engagement to operational questions during testing. Ideally, members of the Ethics Board should be elected by the workforce. Executive stakeholders may be members of the Ethics Board; however, they should be in the minority. Decisions of the Ethics Board should be binding for the company.

15.4 Conclusion: Improving Ethical Decision-Making

Policies, practices and recommendations explained in this chapter are an important instrument for cybersecurity service providers in order to ensure that the handling of sensitive data or operations is ethical and secure. However, not all dilemmas exemplified in the beginning of this chapter can be resolved by policies and guidelines. It remains important that cybersecurity vendors also develop some competences in ethics (e.g. by providing their personnel with training in ethics) and an 'ethical culture' that supports the handling of unexpected situations. In several domains such as health care and business ethics, decision heuristics have been developed for an ethical assessment of problems (for example in computer science, see Linderman and Grillo 2006). Those decision heuristics usually involve a stepwise procedure for analysing problems where no clear guidelines are available:

1. Ethical sensing: As a starting point, an attempt should be made to answer the following question: What provokes ethical debate? Feelings of outrage, shame, guilt or bad conscience could be indicators of an ethical challenge.
2. Gather the facts and legal framework of the case. This also includes identifying the relevant stakeholders and contextual information. It is desirable to include all perceptual perspectives of the participating experts to avoid bias.
3. Identify the moral question and own positions/values. The goal of this step is to identify the 'ethical core' of the problem. Disclosing the personal values of the involved experts should also be integrated here—again to avoid (normative) bias.
4. Analyse the arguments by using ethics frameworks. Examples for frameworks to be used are provided in Chap. 4 of this book.
5. Develop options and decide. Here, developing several courses of action that lead out of an 'either-or' can be helpful. When evaluating and weighing arguments, one-sidedness in the argumentation patterns should also be identified and discussed.
6. Implement the solution. This includes assessing the possibilities of implementing the decision and taking measures for successful implementation. Communicative aspects (how is the decision communicated to whom?) should be considered as well. Finally, possible criteria for reassessment should be identified and examined to learn from the decision.

In addition to such decision heuristics, cybersecurity service providers should also consider implementing procedures for whistleblowing. Company insiders should have ways to raise their concerns without risking losing their employment or becoming the object of other sanctions. We have many great examples of high ethical standards of technology-minded people and we strongly encourage cybersecurity vendors to support the ethical awareness and culture of their employees.

Finally, intensifying education and training about ethical decision-making is undoubtedly important for vendors of security products and services. However, it is equally important that ethics and security are much more deeply integrated into the education of software engineers, system designers and operators. After all, these

roles make much more significant decisions with critical consequences for security and morality.

Acknowledgments The chapter was created with funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 700540 and the Swiss State Secretariat for Education, Research and Innovation (SERI) under contract number 16.0052-1. We would like to thank Karmina Aquino, Mikko Hyppönen and Hannes Saarinen for their comments and helpful discussions.

References

Bishop M (2007) About penetration testing. IEEE Secur Priv 5(6):84–87

Bott E (2011) German government accused of spying on citizens with state-sponsored Trojan. ZDNet. https://www.zdnet.com/article/german-government-accused-of-spying-on-citizens-with-state-sponsored-trojan/. Last access 7 July 2019

Hatfield JM (2019) Virtuous human hacking: the ethics of social engineering in penetration-testing. Comput Secur 83:354–366

Intel & Analysis Working Group (2019) What is cyber threat intelligence? https://www.cisecurity.org/blog/what-is-cyber-threat-intelligence/. Last access 7 July 2019

Jackson W (2001) Antivirus vendors are wary of FBI's Magic Lantern. The Uppernet. https://archive.is/20120910214651/http://www.gcn.com/online/vol1_no1/17572-1.html. Last access 7 July 2019

Karger P, Schell R (2002) Thirty years later: lessons from the multics security evaluation proceedings of the annual computer security conference

Ley J (2018) State-sponsored hacking out of the shadows and into a business near you. Ivanti. https://www.ivanti.com.au/blog/state-sponsored-hacking-shadows-business-near. Last access 7 July 2019

Linderman J, Grillo J (2006) Ethical decision making and information technology: an introduction with cases, 2nd edn. McGraw-Hill Higher Education

Martin RS (2003) Watch what you type: as the FBI records your keystrokes, the fourth amendment develops carpal tunnel syndrome. Am Crim Law Rev 40:1271

Mitnick KD, Simon WL (2005) The art of intrusion: the real stories behind the exploits of hackers, intruders and deceivers. Wiley Publishing, Indianapolis

Schneier B (2013) How antivirus companies handle state-sponsored malware. Schneier on security blog. https://www.schneier.com/blog/archives/2013/12/how_antivirus_c.html. Last access 7 July 2019

Schwartz MJ (2013) Do antivirus companies whitelist NSA malware? Dark reading. https://www.darkreading.com/vulnerabilities-and-threats/do-antivirus-companies-whitelist-nsa-malware/a/d-id/1112911. Last access 7 July 2019

Sharwood S (2018) North Korea's antivirus software whitelisted mystery malware. The Register. https://www.theregister.co.uk/2018/05/02/north_korea_silivaccine_av_software_analysis/. Last access 7 July 2019

Swiderski F, Snyder W (2004) Thread modeling (Microsoft professional). Microsoft Press

Thomson I (2014) Government-built malware running out of control, F-Secure claims. The Register. https://www.theregister.co.uk/2014/02/28/governmentbuilt_malware_running_out_of_control_fsecure_tells_trustycon/. Last access 7 July 2019

Westervelt R (2013) Antivirus firms: whitelisting malware for law enforcement against policy. CRN. https://www.crn.com/news/security/240159502/antivirus-firms-whitelisting-malware-for-law-enforcement-against-policy.htm. Last access 7 July 2019