Deception in Information Security: Legal Considerations in the Context of German and European Law

Daniel Fraunholz[1], Christoph Lipps[1], Marc Zimmermann[1], Simon Duque Anton[1], Johannes Karl Martin Mueller[2], and Hans Dieter Schotten[1]

[1] Intelligent Networks Research Group
German Research Center for Artificial Intelligence
{firstname}.{lastname}@dfki.de
[2] The Project Group: Constitutionally Compatible Technology Design
University of Kassel
johannes.mueller@uni-kassel.de

Abstract. Deception systems have produced promising results in protecting networks from recent attack campaigns. Their development and operation, however, is regulated by technical and legal circumstances. There are several aspects to be considered when operating a deception system, such as privacy, entrapment and liability. In addition to these general aspects, domain specific law that, for example, applies to research or government, needs to be accounted for. In this work German and European law was investigated with respect to deception systems focusing on the aspects listed above and others. The findings are applied to the design, operation of a Honeypot, as well as the generation and publication of information. We found that it is not forbidden to use deception systems in general but several facets have to be considered in the technical implementation.

Keywords: Information Security, Privacy, Deception, Honeypots, European Law, German Law

1 Introduction

Deception technology and especially Honeypots are an advanced IT-security mechanism to oppose cyber crime. This technology relies on purposely providing false or delayed information, hiding information and misleading ongoing attack campaigns into a course controlled by the operator of the deception systems. Usually the parried attacks are monitored and analyzed to gain further insight of the adversaries intentions and approaches. Questions about privacy, copyright and other legal interests are thrown up by this process. Furthermore, devices connected to the Internet can be reached globally. Still domestic law of every participating party applies which leads to legal uncertainty, especially since the origin of an attacker is not known beforehand. In this work the legal concerns

when employing such technology are investigated. Related work is listed in section 2. We first introduce technical aspects and major peculiarities in chapter 3. In section 4 we introduce relevant parts of German, European and international law. These laws are refined by domain specific law in some cases. In some cases, such as governmental or research applications, domain specific law has to be applied in addition to the general law. Examples of relevant domain specific laws are given in section 5. In section 6, the application of the findings are mapped towards real-world Honeypots and the publication of data. It is motivated by previous works of the authors that faced the legal considerations described in this paper. The findings are concluded in section 7.

2 Related Research

Previous work is mostly focused on American law. Legal concerns in all of the reviewed publications are: Liability, entrapment and privacy [9,11,12]. To the best of our knowledge, there is only one work where the concept of Honeypots is investigated in the light of European law [13]. The law of a specific European country has not been analysed with respect to deception systems yet.

3 Technological Aspects

In this section, the technical intricacies of deception systems are introduced that impact the legal evaluation. Honeypots are the most common kind of deception system. They can be distinguished by their type, the deployment strategy, the level of interaction and the counter attack strategy.

3.1 Honeypot Types

Common types are server-side, client-side and token Honeypots. Server-side Honeypots are considered passive in their context. They wait for incoming connection and respond as a genuine server would. In contrast to that, client-side Honeypots actively connect to servers and pretend to be a genuine client system. Honeytokens, such as Honeyfiles and Honeylinks, are data or information embedded in a context like a *HTML*-document or database. Like server-side Honeypots, tokens are passive and wait for an attacker to illicitly access or misuse them otherwise. In this work, we focus on server-side Honeypots, as there is a broad amount of deception technologies that can hardly be considered in full in one scientific work.

3.2 Deployment and Intention

The deployment is only relevant to server and token Honeypots. Research and production deployment modes are distinguished. Research mode is employed to enable a broad amount of attacks. A common research mode deployment is

connecting a Honeypot directly to the Internet and making it addressable with a public IP address. These systems are employed to investigate large scale campaigns such as botnets and common exploitation techniques to access restricted resources from an external context. On the other hand production mode deployment is a strategy where the Honeypot is within a non public context. Interaction on these Honeypots always indicates breaches in the perimeter thus revealing compromises in early stages.

3.3 Level of Interaction

Honeypots may differ in the depth of emulation of the resource they mimic. Telnet servers, for example, typically prompt for an authentication when connected. A system only registering the connection or emulating the login prompt would be considered as a low-interaction system. More advanced systems grant access and provide the full functionality of the given operating system. These systems are considered as high-interaction systems. In literature, medium-interaction system are not consistently defined but commonly placed between the functional scope of low- and high-interaction systems.

3.4 Counter Measures and Aggressive Honeypots

Recently, more aggressive counter measures such as hacking back the attackers are discussed in the context of self defence [8]. From a technical perspective counter attacks can be classified in the same taxonomies as the initial offensive. They extend from denial of service or resource exhaustion techniques to more specific attacks such as dictionary or brute force attacks against the maintenance protocol (most likely *Secure Shell*) of the attacking server. Mirroring of the attacking technique is promising, particularly against propagating malware or botnets since the system was most likely compromised originally by the applied technique.

4 Legal Concerns with Deception Systems

In this section the impact of German and European law on the above introduced technical intricacies is discussed. We identified five legal key aspects that need to be taken into consideration when operating a deception system. A simplified relation of German national and European law is shown in figure 1.

The Basic Constitutional Law of the Federal Republic of Germany (*GG*) is derived from the German constitution. All domain specific laws have to be compatible to the basic law as well as the constitution. More than that there are two kinds of European legislation: Directives and regulations. Regulations are directly applied in each member country, while directives need to be adapted into national law first. The application of law in Germany therefore consists of the Basic Law, the domain specific laws that derive from the German constitution as well as European directives and European regulations.

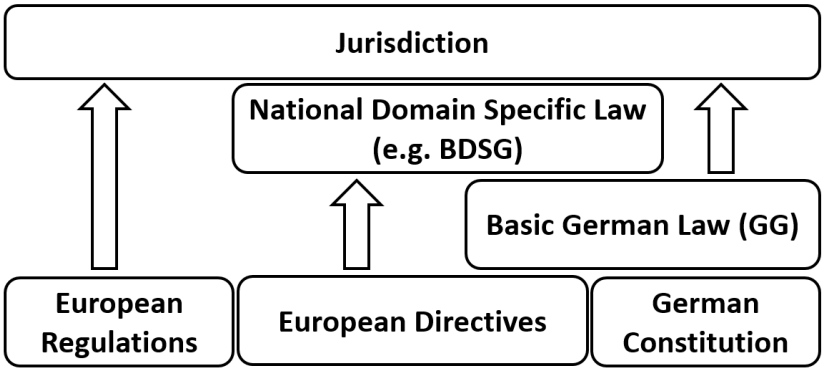

Fig. 1. Application of law

4.1 Privacy

In this section, European regulations and German laws are discussed in the context of privacy. A major task of deception systems is the collection of threat intelligence. The basis of threat intelligence is information collected from the interaction with deception systems. Deception operators need to consider the amount of data they intend to collect with respect to the regulations and laws. §8 of the European Convention of Human Rights provides a right to respect for private and family life in home and correspondence. This convention was published in 1950 and is the foundation of all European and national regulations and laws. It states that persons working with technologies that are capable of collecting or processing personal data where a specific person may be identified are responsible to prevent corresponding violations of privacy.

Fundamental European Rights In 1981, the first appearance of European law containing regulations for the processing of personal data has been published as the European treaty series No. 108 (Convention for the Protection of Individuals with regard to Automatic Processing of Personal Data). This convention specifies a right of privacy for individuals, with regard to autonomic processing of corresponding personal data. In §5 it is clarified that personal data has to be obtained and processed "fairly and lawfully", stored only for specified and legitimate purposes and no longer than required. §5 is only applied to data, that enables identification of a data subject. As a consequence, it is legitimate to store collected data if it is impossible to identify the corresponding person. In §5 an exception for the processing of data in deception systems for scientific research or statistical purposes as well as for law enforcement is given, with respect to the binding to their purpose. The exception requires, that there is no obvious risk of a violation of privacy. However, in §8 every person is granted the right to assert the existence as well as main purpose of collected personal data and the identity

of the data controller. Additionally, every data subject is able to demand the deletion of all relevant data, if the criteria for collection are not fulfilled.

European Directives In 1995, the European Union published the directive *95/46/EC* that focuses on the protection of individuals with regard to the processing of personal data and on the free movement of such data. This directive establishes the fundamental rights in regulations all member states need to regulate in national law. In Art. 7, criteria for the legitimate processing of personal data are named. According to this paragraph, data can only be processed if processing is necessary for

- the data subject unambiguously given consent,
- the performance of a contract,
- compliance with a legal obligation,
- the performance of a task carried out in public interests or vital interests of the data subject or
- legitimate interests as long as they do not interfere with fundamental rights of the data subject.

A person whose data has been collected has the rights to get all relevant information about the data and the data controller and has to be notified if someone obtains personal data that has not been directly obtained from that person. According to §17, the controller has the duty to take care of an up-to-date security to protect the data and the processing of the data from accidental or unlawful usage or destruction. EU Directive *2002/58/EC* is an addition to *95/46/EC* and focuses on the processing of personal data and the protection of privacy in the sector of public electronic communications. §5 of this directive contains regulations about the confidentiality of the communication. Communication networks and services shall prevent all kinds of tapping, interception or surveillance of communications without consent. The directive distinguishes data in traffic and location data. Traffic data is used for the transmission of messages and has to be deleted or anonymized when it is no longer required for the purpose of transmission. Traffic data can legally been processed for billing of subscribers and location data only within the necessary duration for transmission or if they are anonymized.

General Data Protection Regulation EU directives *95/47/EC* as well as *2002/58/EC* are the two important directives in the context of collecting data with deception systems. However, due to the fast technical progress the directives are obsolescent and hardly cope with modern communication. The European Union published two new regulations, that are coming into force on May 25th in 2018. EU regulation *2016/679/EC* General Data Protection Regulation (GDPR) annuls directive *95/46/EC* and *2017/0003/EC* (e-Privacy) will annul the prior directive *2002/58/EC*. Both regulate the same domain as their predecessor but cover more details with respect to new processing techniques. In *2016/679/EC*, it is described that a data controller needs to inform a person at the time of

collection but not in case of getting data otherwise and for the purpose of scientific research. Personal data controller have more obligations such as

- securing the data,
- keeping a register about all data and processing steps and
- informing the supervisory authority if processing may violate the privacy of certain persons.

Researchers can process data if they guarantee to respect all rights of privacy and only use a "minimal amount of data". If possible the data needs to be anonymous. Otherwise, it needs to be pseudonymous.

German law The foundation for domain specific law is composed by the *GG*. It establishes several fundamental rights with respect to privacy such as:

- Human Dignity (Art. 1 *GG*),
- Personal Freedoms (Art. 2 *GG*),
- Privacy of correspondence, post and telecommunications (Art. 10 *GG*) and

Article 10 regulates the inviolable privacy of correspondence, mail and telecommunication. However, section 2 state the inviolability as restricted pursuant to other law. Please note, that fundamental rights do not directly apply between natural persons. The Federal Data Protection Act (*BDSG*) is the most important law for privacy in Germany and is currently adapted to the *GDPR*. It focuses on the rights of a person about processing personal data and privacy, thus realizes EU directive *95/46/EC*. Both state that personal data is only collected for a "limited duration of time" and only if "required to provide services" or if the user gave his permission to do so. The *BDSG*, however, allows several exceptions. §4a dictates that the permission has to be given in written form except for valid scientific purposes. Additionally, §4d commits data processing organizations to report to supervisory authorities about automated data processing techniques. German law differentiates between collecting, processing and using data. Collecting data for scientific purposes is not explicitly allowed without the permission of the user, but the processing of received data is allowed. According to §40, personal data, that was collected for scientific purposes can only be used for those purposes and has to be anonymized as soon as possible. The German Telemedia Act (*TMG*) describes the rights and obligations for all electrical information and communication services. The §15 *TMG* state that personal data can only be collected if it is necessary to provide the service or for billing. As in *95/46/EC* the provider has to inform the user about collection and purpose of data and the user has to give his permission. §13 also declares that the data has to be deleted right after the end of the service. The TMG is very restrictive, deception systems with extensive data collection may be illegal in the context of the *TMG*. However, the term *electrical information and communication service* is defined in §1 of the *TMG* and a deception system may not pose a service as required to apply the *TMG*. §13 sec. 7 states that the provider ensures the security of the service and the data. This clause may allow the collection of data of potential threats to the service.

Court Judgments Due to the lack of detail about the content of personal data, there have been several trials. Those trials were mainly about the legitimacy of the collection of personal data and special types of data such as IP addresses. In Germany, the Higher Regional Court of Cologne dealt with those questions in the case *12U16/13* in December 2015. In this case, they had to decide if it is legitimate to hold information about dynamic IP addresses for 4 days. The defendant, a small size service provider, collected data to protect the system from disruptions. From the courts point of view, Denial-of-Service attacks, spam mails and malware can result in disruptions. Therefore, it is allowed to collect and process relevant data as long as needed to get important information about potential disturbances. This case affects provides of communication services but not provides of media services. In October 2016, the European Court of Justice dealt with a case about a litigation between the federal public of Germany and Mr. Breyer that was initialized in October 2014. Breyer stated that it is illegal to collect information about user containing time data and dynamic IP addresses for provider of online media services. The court decided that this data is personal if the provider has the legal means to get the identity of the person behind the address. Furthermore, it is not allowed to store this information for a longer duration than necessary to provide the requested service. Due to this judgment, the German Federal Supreme Court of Justice had to form an opinion about the original case. In May 2017 an announcement was made that no final consideration could be made.

4.2 Entrapment and Accessoryship

Entrapment is not regulated by European law. Members of the European Union are self dependent in this domain. In Germany entrapment is covered by §26 and §30 Penal Code (*StGB*). However, the requirement of both is to dictate somebody to commit a crime. A criminal abusing a honeypot had the criminal intention before. The criminal actively searched for vulnerable systems and exploited the identified vulnerability to take advantage of the honeypot. Even if a client side honeypot is considered, the attacker did set up a server with offensive abilities before. She also implemented a trigger to employ the offensive capacities against a connecting system. Honeypots do not dictate criminal activities to intruders. A conviction for entrapment is unlikely. §27 *StGB* covers accessoryship: "Intentionally rendering aid to another in intentional commission of an unlawful act". This could be the case if a deception system is used to perform attack against third parties. In this case the operator of said deception system could be considered an accessory, since she aids the attacker by implicitly providing resources.

4.3 Liability and other Claims

Any operator of a service can be liable for damage caused by this service. Liability is specifically affecting honeypots as the operator is liable for damage occurring from the honeypot operation as well. As Honeypots are intended to be exploited,

the risk of an attack propagating because of misconfiguration is high compared to well maintained systems. Liability is not regulated in the European Law. In German Law liability is regulated in §823 Civil Code (*BGB*). Liability needs to be considered in several cases:

- Damage of third parties due to an intruders interaction with the honeypot,
- damage due to information published that was collected with the honeypot.

To palliate the probability for a conviction, technical measures need to ensure a proper policy enforcement. Policies need to be defined to mitigate pivoting attempts. Additionally, the collected data needs to be handled with a state of the art security level and publications need to fulfill the requirements of privacy. With adequate technical measures, a reduction of the risk of being held liable to the level of a common service can be achieved. Honeypots do not expose the operator to significant juridical consequences any more than other systems. In addition to liability, it may happen that a honeypot collects personal data from third parties. According to Art. 5, *2016/679/EC*, this kind of collection is illegal as it is not related to a valid cause. Violating the *2016/679/EC* can result in claims against the operators.

4.4 Copyright

High-interaction deception systems enable extensive monitoring of intrusion campaigns. These systems may be able to obtain malicious programs from attackers and botnets. Malicious programs can be obtained in the form of compiled binaries or a sequence of commands. Any person has the rights on her intellectual property, no matter how the program is used. In European law, the directive *2004/48/EC* focuses on the enforcement of intellectual rights. §5 of this directive states that the author of literature or art is determined by a name, indicating the works author. Additionally, in directive *2009/24/EC* on the legal protection of computer programs §1 claims that computer programs are protected by copyright as pieces of literature. In German law, the Act on Copyright and Related Rights (*UrhG*) regulates intellectual rights. This act has the same formulations with regard to computer programs as the European law. According to §§15-22, the author has the right to choose about publication, duplication, spreading and exhibition. If a work is free to use it can be taken as basis to create a new product without notifying the author of the original work. In §§97-99 *UrhG*, the author of any kind of written work can insist on stopping the usage of his intellectual property. After this warning, proceedings against the usage can be initiated to compensatory damages. However, in case of an attacker who creates malicious software it is very unlikely that she will claim damages because she usually has created the software to do illegal actions she can be condemned for.

4.5 Self-defence

Self-defence, for example *hackback*, in the context of cyber crime, describes counter measures against intruders or intrusion attempts. These counter measures can

be criminal acts by §202a-d *StGB*, §206 *StGB*, §263a *StGB* or §303a-b *StGB* depending on the technical design of the counter measure. However, the German law allows active counter measures, such as hackback, under specific circumstances. There are two different cases: Self-defence and necessity. Self-defence is "any defensive action to advert an imminent unlawful attack on oneself or another". It requires that the attacking system is juristic property of the attacker. This cannot be ensured as most attacks stem from infected systems. More than that self-defence also requires the attack to be ongoing. From a technical perspective honeypots are able to trigger instant counter attacks. This requirement can therefore be fulfilled, if the technical implementation is adequate. However, self-defence salvages the risk of a conviction if a third party system is attacked. Necessity means the aversion of an imminent danger, such as a cyber attack, upon a legal interest with any means necessary, given they are proportional and given the legal interest outweighs the effect of the counter measure. In case of necessity, third party actors are also allowed to be attacked. Preventative counter measures are also allowed if a threat is present. Defence motivated by necessity needs to be adequate with respect to the opposed threat. This implies that the kind of attack needs to considered for the counter measure. For example port scans may not legitimate distributed denial of service attacks as response. Counter measures are a legitimate option against attackers. However, the context is significant. Adjacent to the discussed situations, the operator is a major factor. State involved counter attacks may pertain international or martial law.

5 Domain specific Law

In this section we discuss specific domains in which general regulations are supplemented, as, for example, the *StGB* does for criminal law. Table 1 gives an overview of the investigated domains and the corresponding codes of law. The domains law enforcement, research, federal law & public sector and telecommunication providers are the four domains that were identified as relevant for application in the area of deception systems.

Table 1. Domains and Corresponding Laws

Domain	**Corresponding Law**
Law Enforcement	§100g *StPO*, §7 BKAG
Research	§28 S2 no.3 *BDSG*, §40 *BDSG*
Federal Law & Public Sector	§§13,14 *BDSG*
Telecommunication providers	§96 *TKG*, §100 *TKG*

Subsequently an overview of different cases/institutions with specific laws that are able to extend general laws is given.

5.1 Law enforcement

In order to protect personal data in the purpose of the prevention, investigation, detection or prosecution of crime or the execution of criminal penalties, respectively, there is the directive *2016/680 EC*, wherein the fundamental right and freedom of natural persons and their right to the protection of personal data is fixed. As all European directives it needs to be transferred in national law, which is currently in process in Germany. There still exist some sections in German law that allow to collect personal data for law enforcement. In German law there is a fundamental understanding of the commensurability of a governmental intervention. In this context the

- suitability,
- necessity,
- reasonability and appropriateness

of the intervention must be given. Beyond that, as already mentioned above, any legal foundation is required for every intervention. For instance conditions regarding the interception of telecommunication are defined in §100a Criminal Procedure (*StPO*). It is allowed to intercept and record telecommunication, also without the knowledge of the concerned person, if certain facts give rise to the suspicion that a person focuses a serious crime or, in cases where there is criminal liability for attempt, has attempted to commit such a crime or has prepared such a crime.

Due to §100g *StPO* it is also allowed to collect communication traffic data, in terms of §96 *TKG*, if certain facts give rise to the suspicion that a person has committed a crime, in cases where there is criminal liability for attempt, has attempted to commit such a crime or has prepared such a crime or has committed a crime by the means of telecommunication.

Extended permissions are also given to governmental authorities such as the Federal Criminal Police Office. According to §7 Federal Police Law (*BKAG*) they are allowed to collect, process and use privacy data, as well as, for example, time and scene of a crime (§8 *BKAG*), if this is necessary in compliance to their task as central office of law enforcement. Furthermore they are allowed to impose and store other privacy data.

5.2 Research

The German law is very detailed in the context of data security. According to §28 Sentence 2 no. 3 *BDSG* the collection and storage of data shall be admissible, if it is necessary in the interest of a research institute for conducting scientific research. Sentence 6 no. 4 of the same section, extends these permissions in a way that not only the collection but also the processing and the use of special types of personal data is allowed, if necessary, for the purposes of scientific research. But there are also some restrictions, there is a specific section within the *BDSG* especially for research institutes. §40 defines that collected and stored personal data may processed or used only for scientific research purposes. Due to sentence

2 it is also necessary to anonymize personal data if this is possible. Sentence 3 states that a publication of personal data is only allowed if the data subject has consented or if the data is indispensable for the presentation of research findings on contemporary events. Researchers intending to publish results, coercively need to anonymize these.

In European jurisdiction there is a restriction for the storage of privacy data for the purpose of research. According to §6 section 1e *2000/31 EC*, data shall not be stored longer as necessary for the purpose and it shall not be possible to identify participating entities.

5.3 Federal Law and Public Sector

Public bodies of the Federation are regulated in a specific section within the *BDSG*. In the sections §§12 - 14 *BDSG* it is defined which actions are explicitly authorized. The collection of personal data shall be admissible if the knowledge of them is needed to perform the duties of the bodies collecting them (Section 13 sentence 1 *BDSG*). If it is necessary for the performance of the duties of the controller of the filing system and if it serves the purpose for which the data was collected, it is allowed to store, modify or use personal data (Section 14 sentence 1 *BDSG*).

The German Federal Office for Information Security is responsible for the security within information technology. In order to defend against risks for critical infrastructure, they are empowered to collect and evaluate data, especially information about security vulnerabilities, malware, happened or attempted attacks aligned on information security and also the exact proceeding of the attackers (§8b S.2 Nr. 1 - *BSIG*).

5.4 Telecommunication Providers

Service providers are committed to safeguard the secrecy of telecommunication according to §88 Telecommunications Act (*TKG*). Furthermore they are obligated to protect the personal data of communication participants (§91 ff. *TKG*). If there is any justifying purpose, however, they are granted certain permissions, like collecting communication traffic data, especially phone numbers, and connection meta data, such as time and duration (§96 *TKG*). In case of disturbance of their infrastructure, they are allowed, according to §100 *TKG*, to collect additional data. To define the case of disturbance there is a legal decision of the Higher Regional Court of Cologne (*I-12 U 16/13 OLG Koeln*) in which cyber attacks are defined as such a disturbance. According to §98 *TKG*, they are also allowed to collect location data if they provide additional services relying on this data, which however needs to be anonymised. A specific restriction for the storage of this data is the duration. Service providers need to retain data only for a period of ten weeks, location data only for four weeks (§113b TKG).

6 Honeypot Design and Threat Intelligence

In this section, we apply the findings from the previous sections to the design and threat intelligence of Honeypot systems. Regulations that restrict the collection and usage of data are considered, as well as possibilities to still gather and publish information. Furthermore, lessons learned during the research for this work and the operation of Honeypots are explained.

6.1 Application

The design, operation and threat intelligence can be split into four categories. First, in the deployment, the kind of Honeypot to operate is chosen. After that, the possible operation modes are introduced. The data processing is evaluated after that. Finally, the possibilities for publishing information and insights are discussed.

Deployment Due to the specific nature of production Honeypots they need to follow significantly less restrictions than research Honeypots. As they are placed within the perimeter an attacker has to have breached the network security, having committed a crime already. The authors conclude that entrapment is not an issues. As the operator is always liable, measures have to be taken to ensure the attacker does not influence third party systems. This can be achieved firewall rules and policy enforcement. A best practice is to block or limit outgoing traffic [4]. Blocking can be problematic as no interaction is possible, for example with the Command and Control (*C&C*)- or download-server, hindering further analytics. Sometimes, the inability to create outbound traffic leads to the deletion or abort of infection routine of the malware, a common anti-forensics mechanism [3]. In contrast to research Honeypots, production Honeypots are usually placed in productive environments. If activity can be detected, an attacker has obviously gained access to this environment, potentially endangering the production facility. Counter measures as means of defending the production infrastructure can therefore be considered as self-defence.

Research honeypots, on the other hand, are usually publicly accessible and not connected to productive systems. Therefore, any client can access them and no damage to one entity's assets is imminent. This makes entrapment possible, as it offers obvious vulnerabilities. Entrapment, according to German and European law, requires more than just offering an opportunity, however. It is necessary to actively try to get a victim to perform an illegal action, which cannot be seen here, neither for client-side, nor for server-side Honeypots, according to the authors. Liability lies with the operator, as she has to make sure that no outbound traffic can harm any third-party system. The best practices are the same as described above.

The authors operate several research Honeypots, whose deployment has been described in previous works [5]. As no outbound traffic is possible on our systems, we take no risk of liability for damage on third-party systems. It is, however, not inconceivable that an attacker compromises the underlying operating system, for example via a previously unknown vulnerability, to execute attacks against third parties. In this case, liability seems improbable, as the code was created according to best practices for secure programming.

Operation Honeypots can be operated in different fashions. One of the most important distinguishing feature is their ability to counterattack. Especially since malware usually runs on host systems that are not owned by malicious adversaries, hackbacks can only be executed under certain conditions, as described in section 4. To verify these conditions beforehand is infeasible, making hackbacks risky. On the other hand, many botnets have only been defeated by law agencies because traffic was infiltrated into *C&C* communication [1]. The authors conclude, that law enforcement agencies have a higher tolerance for actively counterattacking malicious adversaries than research institutes. The Honeynet we previously introduced [5] is not capable of hacking back.

Data Processing Honeypots are created and operated to gather data. Typical kinds of collected data are *IP*-addresses, timestamp, location of the attacker and payload, such as credentials or command sequences. IP addresses will supposedly be considered personal data according to German and European law [2, 6]. Timestamp, location and other metadata can be personal data if they are able to identify an individual. If the conjunction of different kinds of data with metadata allows someone to identify an individual, metadata is classified as personal data. The collection of metadata therefore poses a threat to the operator of Honeypots, since it can be personal data to which strict regulations apply. Payloads of attacks can be copyright protected data, especially if an attacker exploits a vulnerability that was unknown beforehand. Such exploits are sold for sums of several thousand dollars up to $1.5 million [15]. The simple collection of copyright protected data, however, does not pose a problem to the operator of the Honeypot.

According to German and European law, personal data has to be deleted as soon as it is no longer required for technical reasons, for example offering web services. The purposes of research and law enforcement create exceptions. Law enforcement agencies are allowed to store personal data as long as the investigation is ongoing and it is allowed by a judge. Research institutes are allowed to store personal data for the duration of their research. It needs to be noted that all assessments in this section only account for storing data, not distributing or utilising it. The author's Honeypots store all this information in a secure way to comply with German privacy protection regulations.

Publication Publication and distribution of collected data is desirable for different reasons.

Law enforcement agencies exchange data in the context of cross-border prosecution of crime. Internet architecture makes it easy for an attacker to avoid the jurisdiction of the country she is attacking in by originating her attack in a different country. This needs for international cooperation in crime fighting.

Research institutes publish scientific results and findings. The publication of personal data, however, is strictly forbidden, unless it is necessary to describe the situation or unless it is a person of public interest. That means personal data has to be anonymised to avoid legal consequences. Another possibility of making use of the data and publishing results is statistical analysis and release of the findings. This allows for full usage of the data without compromising personal data and has, among others, been employed by the authors in previous works [5]. Another common practice, as practiced by *Google Analytics* [7] for example, is deleting one octet of an IP address. This way, some information, such as local area, can be obtained without disclosing identities.

Telecommunication provider Usually, personal data, such as IP and access times, are stored by the telecommunication providers only for a short duration of several days. This data can be shared with law enforcement agencies for prosecution of crimes, as well as with other providers for misuse prevention.

Copyright protected data, as described in section 6.1, must not be distributed or copied in an unauthorized manner. However, since the creator of this data has to claim his rights, she will inevitably admit for having commited cyber crimes. This makes in unlikely for the operator of a Honeypot to be held accountable for distribution of copyright protected data, even though she is technically transgressing the law.

6.2 Lessons Learned

During the operation of our Honeynet, we found that the legal foundations can change in a way that influences our research. This leads to the insight that constant evaluation of the legal landscape is necessary for operators of deception systems in order to prevent being prosecuted. Sometimes, the laws change for the better from a research institutes perspective, as with the *General Data Protection Regulation*, presented in section 4. It explicitly states exceptions for storage of personal data for research institutes. Beforehand, this was tolerated but not regulated by German law. Despite the release of laws, a factor of uncertainty always lies in the jurisdiction. The relatively new topic of cyber crime and its defence has not often been dealt with in court, making the outcome uncertain as there are not many test cases. Furthermore, a challenge lies in the origin of attacks. According to common law, the jurisdiction of the country of origin is applied, making it infeasible for the operator of a Honeypot to check all possible laws. The authors

of this work alone have monitored attacks from 174 countries in 222 days and analysed access from 95 countries monitored during 111 days in previous works [5].

In previous works, we published statistical analysis of the attacks to avoid legal issues. The published data is not suitable to identify an individual. For research projects the authors are currently working on, the generation of attack signatures is required. This could create a conflict due to the implicit publication of copyright protected content and needs to be checked thoroughly. German institutes, such as the *Deutsche Telekom AG* with their *DTAG Honeynet* [14], only publish the country of origin, the timestamp and the content of any attack. This information without an *IP*-address is not able to identify individuals. American institutes, such as the *Norse Corp.* with their *NorseMap* [10], publish the IP addresses as well. This is due to the significantly different laws on personal data in the United States of America. These were out of scope in this work, but highlight the drastic differences in jurisdiction between different countries.

Art. 5, *2016/679/EC* demands several conditions to be fulfilled for processing of data to be legal. Many of them, for example the right of deletion of data or the necessity of a valid cause are inherently incompatible with the idea of deception systems. A legally sound implementation is a challenging task, as many restrictions on anonymisation are to be met. This will be discussed in future works.

7 Conclusion

There are several European directives that take all relevant aspects of the legitimacy of deception systems into consideration, as well as regulations, such as the *GDPR* and the *e-Privacy* regulation, that will be applied by national jurisdiction. They contain regulations on privacy, entrapment, liability, copyright and self-defence. Additionally, most of them distinguish between domains such as law enforcement, research, public sector and telecommunication providers. European directives have to be implemented into acts by every member state. Therefore, operators of honeypots have to consider which domain they represent and which law is relevant for them. Especially in case of collecting and processing personal data for research purposes, there are restrictive regulations regarding privacy implemented in European law. For further processing and publishing results, personal data has at least to be pseudomised or anonymised so that it is not possible to identify a certain individual. Several research and Honeypot projects show that it is possible to operate a Honeypot and publish the results in a legally conform way.

Acknowledgment

This work has been supported by the Federal Ministry of Education and Research of the Federal Republic of Germany (Foerderkennzeichen KIS4ITS0001, IUNO). The authors alone are responsible for the content of the paper.

References

1. Andriesse, D., Rossow, C., Stone-Gross, B., Plohmann, D., Bos, H.: Highly resilient peer-to-peer botnets are here: An analysis of gameover zeus. International Conference on Malicious and Unwanted Software 8, 116–123 (2013)
2. Bundesgerichtshof: Bundesgerichtshof zur zulässigkeit der speicherung von dynamischen ip-adressen (2017)
3. Edwards, S., Profetis, I.: Hajime: Analysis of a decentralized worm for iot devices
4. Fraunholz, D., Pohl, F.: Towards basic design principles for high- and medium-interaction honeypots. European Conference on Cyber Warfare and Security 16 (2017)
5. Fraunholz, D., Zimmermann, M., Duque Anton, S., Schneider, J., Schotten, H.D.: Distributed and highly-scalable wan network attack sensing and sophisticated analysing framework based on honeypot technology. International Conference on Cloud Computing, Data Science & Engineering 7 (2017)
6. Gerichtshof der Europäischen Union: Urteil in der rechtssache c-582/14 (2016)
7. Google Inc.: Google analytics (2017), analytics.google.com
8. Koch, A.: Die rechtlichen rahmenbedingungen von hackback (2008)
9. Mokube, I., Adams, M.: Honeypots: Concepts, approaches, and challenges. Proceedings of the 45th Annual Southeast Regional Conference (2007)
10. Norse Corp.: Norse attack map (2017), http://map.norsecorp.com/#/
11. Radcliffe, J.: Cyberlaw 101: A primer on us laws related to honeypot deployments. Information Security Reading Room (2007)
12. Scottberg, B., Yurcik, W., Doss, D.: Internet honeypots: Protection or entrapment? International Symposium on Technology and Society (2002)
13. Sokol, P., Misek, J., Husak, M.: Honeypots and honeynets: issues of privacy. EURASIP Journal on Information Security (2017)
14. Telekom DTAG: Fruhwarnsystem, sicherheitstacho (2017), http://www.sicherheitstacho.eu/
15. ZERODIUM: Zerodium exploit acquisition program (2017), zerodium.com