

Chapter 9

Ethical and Unethical Hacking

David-Olivier Jaquet-Chiffelle and Michele Loi

Abstract The goal of this chapter is to provide a conceptual analysis of ethical hacking, comprising history, common usage and the attempt to provide a systematic classification that is both compatible with common usage and normatively adequate. Subsequently, the article identifies a tension between common usage and a normatively adequate nomenclature. ‘Ethical hackers’ are often identified with hackers that abide to a code of ethics privileging business-friendly values. However, there is no guarantee that respecting such values is always compatible with the all-things-considered morally best act. It is recognised, however, that in terms of assessment, it may be quite difficult to determine who is an ethical hacker in the ‘all things considered’ sense, while society may agree more easily on the determination of who is one in the ‘business-friendly’ limited sense. The article concludes by suggesting a pragmatic best-practice approach for characterising ethical hacking, which reaches beyond business-friendly values and helps in the taking of decisions that are respectful of the hackers’ individual ethics in morally debatable, grey zones.

Keywords Cracker, Black hats · Hacking · Hacktivism · Pentesters, Taxonomy, True hackers, White hats

9.1 Introduction

The goal of this chapter is to provide a conceptual analysis of ethical hacking. The chapter begins (Sect. 9.2) with a historical introduction, describing how the term hacking and different denominations for different varieties of hacking have been

D.-O. Jaquet-Chiffelle (✉)
University of Lausanne, Lausanne, Switzerland
e-mail: david-olivier.jaquet-chiffelle@unil.ch

M. Loi
University of Zurich, Zürich, Switzerland
e-mail: michele.loi@uzh.ch

introduced in everyday, journalistic and technical language. Section 9.3 introduces our proposal of a systematic classification, one that fulfils adequate descriptive purposes and that maps salient moral distinctions into the different denominations of hacker types. It does so by proposing an initial taxonomy (inspired by common usage) and subsequently revising it by adding further nuances, corresponding to further evaluative dimensions. Section 9.4 discusses the concept of ethical hacking, revealing a fundamental ambiguity in the meaning of ‘ethical’ as an attribution to hacking. It presents our main thesis, namely that ‘ethical hacking’ refers to a limited view of ethics which assumes the pre-eminence of business-friendly values and that hacking that is ethical, all things considered, may not be ‘ethical hacking’ according to the common usage of the term. We recognise, however, that in terms of assessment, it may be quite difficult to determine who is an ethical hacker in the ‘all things considered’ sense, while society may agree more easily on the determination of who is one in the ‘business-friendly’ limited sense.

9.2 What Actually Is a ‘Hacker’?

Almost every week mass media communicates about *hackers* having stolen thousands of passwords and other sensitive private information. It is commonplace to read articles about hackers having taken advantage of system vulnerabilities to bypass security barriers in order to fraudulently access private and company networks. The current understanding of the term ‘hacker’ is influenced by the news, and this twists the original definition of what a hacker is (Fig. 9.1).¹

Today’s perception of the term ‘hacker’ tends to be reduced to ‘black hat’ and ‘cyber-criminal’. This has not always been the case, and the term ‘hacker’ conveys a much broader meaning.

9.2.1 *Hackers in the Early Days*

In the 1960s and 1970s, typical hackers were not really driven by malicious intent. They were often supportive of strong (ethical) values, broader than computer security issues, such as democracy or freedom of speech. At the same time, computers, not to mention networks, were still in an early stage of development. The economic weight of computer related business was trifling in comparison to today’s influence of GAFAMs² in the global market. Criminal opportunities were limited. Early

¹ As C.C. Palmer wrote: “Instead of using the more accurate term of ‘computer criminal’, the media began using the term ‘hacker’ to describe individuals who break into computers for fun, revenge or profit. Since calling someone a ‘hacker’ was originally meant as a compliment, computer security professionals prefer to use the term ‘cracker’ or ‘intruder’ for those hackers who turn to the dark side of hacking.” (Palmer 2001: 770)

² The GAFAM acronym stands for Web main players, namely, Google, Apple, Facebook, Amazon and Microsoft.

Fig. 9.2 Shift in the hackers' incentives



insignificant economic consequences at stake seems far away. Hacking has become a business; a very serious one at that.

From the 1960s to the 2010s, we can therefore observe a shift in the nature of hacking incentives: ideological incentives have been replaced by economic ones (Fig. 9.2).

Ethical values at stake have evolved accordingly. In the 1960s, they were essentially described by the so-called hacker ethic. With the development of the Internet, of e-commerce and the increasing economic weight of information, freely shared information as well as many early ideological ethical values entered into conflict with economic-related ethical values, in particular regarding the protection of information ownership.

9.2.3 Modern Hackers

Modern computer hackers are usually defined as skilled programmers and computer experts who focus on software, computer and network vulnerabilities. There is a plethora of terms available to distinguish them: white hats, black hats, grey hats, pen testers, ethical hackers, crackers and hacktivists, to mention the most important ones. Some categories of modern hackers do not even require significant expertise. Indeed, *script kiddies* are non-expert hackers who run programs and scripts developed by other, more expert hackers (Barber 2001). Modern hackers are categorised not only according to their expertise, but also according to the (ethical) values they adhere to or not. Legal values are often implicitly emphasised in this classification (see also Fig. 9.3).

Early hackers were categorised according to their expertise through peer recognition, and were adherent to values described in the hacker ethic. Today, 'hacktivists' still consider IT vulnerabilities as opportunities to promote a cause, a political opinion or an ideology. The group *Anonymous* is a typical heterogeneous group of hacktivists. In her best-seller (Olson 2013), Parmy Olson shows a large variety of profiles and incentives within *Anonymous*. However, most modern hackers use IT vulnerabilities for malicious purposes to commit fraud and make money. Some modern hackers strictly conform to applicable laws, whereas the majority does not really care.

Modern hackers can have a broad spectrum of incentives for their activities. According to Richard Barber, white hats are "[s]ecurity analysts and intrusion detection specialists [...] [who] spend their time—just as police or intelligence analysts do—researching the technologies, methodologies, techniques and practices of hackers, in an effort to defend information assets and also detect, prevent and track hackers" (Barber 2001: 16). White hats do respect applicable laws. In a dichotomic world, they are the good guys. Their incentive is to protect software,

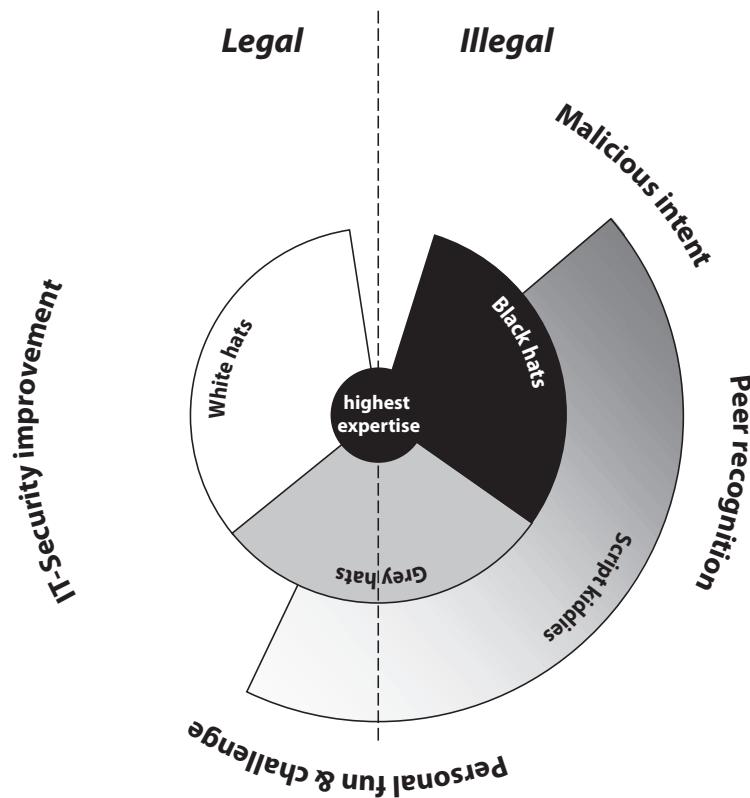


Fig. 9.3 White hats, black hats, grey hats and script kiddies (Note that the outer layer refers to one predominant motivation (not the exclusive one). For example, not only grey hats, but also white hats as well as black hats may have fun in doing their activities or enjoy taking a challenge. White hats might also look for peer recognition)

computers, networks and the IT infrastructures from the bad guys, the so-called black hats or crackers.

According to Sergey Bratus, by contrast, black hats “act for personal gain and without regard for possible damage” (2007: 72). According to Technopedia (n.d.), a black hat is “a person who attempts to find computer security vulnerabilities and exploit them for personal financial gain or other malicious reasons”. They might also have other motivations such as cyber vandalism for example. Their values lead to illegal activities.

Grey hats are hackers whose intentions are not fundamentally malicious, but who accept irregular compliance with the law to reach their objectives, which distinguishes them from white hats. Contrary to black hats, greed is not their typical main incentive.

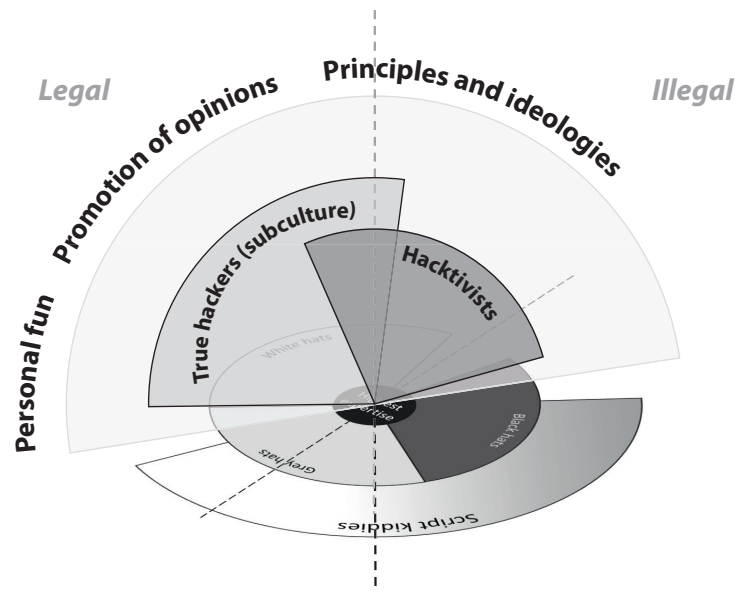


Fig. 9.4 A third dimension to represent true hackers and hacktivists

Grey hats might also share some incentives with white hats and so-called true hackers: personal fun, peer recognition, intellectual challenges, etc. However, they do not really share the original hacker ethic.

To represent true hackers, as well as hacktivists, we need a third perpendicular dimension where the legal perspective only plays a secondary role (Fig. 9.4).

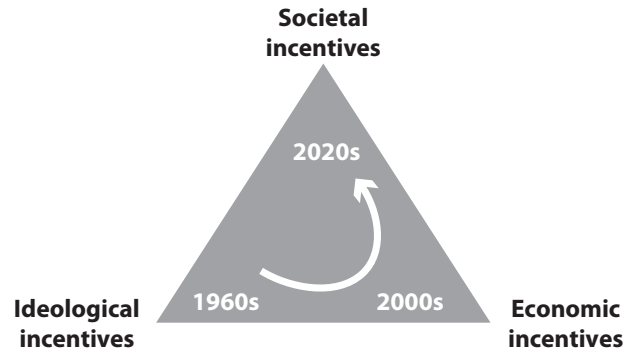
Many different definitions are used for terms categorising modern hackers. These definitions are not always fully compatible. They bring different nuances. There is a need for a more systematic classification.

9.2.4 Today's Hackers

We have already emphasised a shift in hackers' incentives from the 1960s to the 2010s. Since the beginning of the 2000s, information grew as a valuable asset and created new economic incentives for cyber-criminals. In our modern interconnected society, we now observe a new shift: information tends to also increasingly become a societal asset too (Fig. 9.5).

Nowadays, our whole society heavily depends on information and information technologies: transport and communication systems, medical facilities, SCADA control systems, electrical grid, nuclear plants and other critical infrastructures,

Fig. 9.5 A societal dimension in hackers' incentives



government activities and voting systems, commercial exchanges and payment infrastructures, security-oriented surveillance technologies, or even military control systems.

With the advent and the development of smart cars, autonomous drones, smart medical devices and the Internet of Things, our physical world is becoming even more intertwined with the virtual one. To mimic a famous slogan,³ what happens on the Internet does not necessarily stay on the Internet anymore. Lives are at stake. The very functioning of our society now relies on the Internet. A disruption of Internet services and other information infrastructure can paralyse a whole country. This creates a new paradigm and extra incentives for hacking activities. As a direct consequence, we observe the emergence of new categories of hackers: *state-sponsored hackers*, *spy hackers* or even *cyber-terrorists*. The target can be an individual, a company, a facility, an infrastructure or even a state. Whereas black hats foster cyber-crime and cyber-security countermeasures, state-sponsored hackers or cyber-terrorists have given rise to new concepts such as cyber-war, cyber-defence and cyber-peace.

9.3 Towards a More Systematic Hackers' Classification

As pointed out, different meanings of the term 'hacker' coexist in the context of computerised systems. The term seems to have evolved since the 60s and describes very different realities nowadays. True hackers, adept at the so-called hacker ethic, are disappointed by today's mainstream usage of the term 'hacker'. They do not want to be considered in the same category as security breakers and cyber-criminals.

However, in the earliest known appearance of the term 'hacking' in the context of computerised systems (Lichstein 1963)—which appeared in the MIT student newspaper *The Tech* on 20 November 1963—the pejorative connotation is already present.

³What happens in Vegas stays in Vegas!

Traditional dictionaries are of limited assistance in refining the meaning of the term ‘hacker’ in the context of computerised systems. In fact, this word has numerous different meanings in the English language. The Merriam-Webster dictionary provides four definitions for a hacker (“Hacker | Definition of Hacker by Merriam-Webster” n.d.):

1. : one that hacks⁴
2. : a person who is inexperienced or unskilled at a particular activity (a tennis hacker)
3. : an expert at programming and solving problems with a computer
4. : a person who illegally gains access to and sometimes tampers with information in a computer system

Curiously, the second definition seems completely opposite to the typical common understanding as it emphasises the *inexperience* of the hacker at a particular activity.

The last two definitions better capture the main meanings in the context of this chapter. The third one is general and covers most of the modern categories of hackers, whereas the last one is close to what we call a black hat or a cracker.

The American Heritage dictionary gives similar definitions for a hacker (“American Heritage Dictionary Entry: Hacker” n.d.):

1. (a) One who is proficient at using or programming a computer; a computer buff.
(b) One who uses programming skills to gain illegal access to a computer network or file.
2. One who demonstrates poor or mediocre ability, especially in a sport: *a weekend tennis hacker*.

Those definitions only describe large categories of hackers. We need to delve deeper into subtle differences to distinguish between the many terms used nowadays to characterise hackers in the context of computerised systems and eventually to precisely define what an ethical hacker is.

A more systematic classification requires, as a first step, a *taxonomy*, i.e. the creation and definition of classes with clear identities. A second stage of classification is *ascription*, i.e. placing each hacker into its class. Ascription corresponds to the identification of a hacker as belonging to a specific class. Identification itself is a “decision process attempting to establish sufficient confidence that some identity-related information describes a specific entity in a given context, at a certain time” (Pollitt et al. 2018: 7). When the entity is a person, i.e. for people identification, the identification process relies on authentication technologies in order to corroborate

⁴The verb ‘to hack’ has numerous meanings. According to the Merriam-Webster dictionary, the first definition is “*to cut or sever with repeated irregular or unskillful blows*” which has nothing to do with computer hacking.

(or to exclude) the fact that the given identity-related information describes this person in the given context, at the time of reference, with sufficient confidence.

Authentication technologies are classified themselves into four categories, namely:

- Something you know
- Something you are
- Something you do
- Something you have

A key aim of this paper is to develop a classification of (modern) hackers, related to categories of authentication technologies.

9.3.1 A First Taxonomy

In order to reach a new systematic classification of (modern) hackers, different perspectives can be chosen. A first approach consists in defining classes according to hacker's expertise (its scope and its level) and to hacker's values (his/her objectives and moral principles). Expertise can be seen as a collection of internal resources—something that the hacker *knows*—while values followed by the hacker can be seen as an internal attitude—something that the hacker *is*. Those classes are defined in compliance with the first two categories of authentication technologies (Table 9.1).

Hacker's expertise is defined by both its scope and its level. It corresponds to what the hacker knows and is able to do. The scope considers the expertise environments (OS, protocols, network, etc.), the objects covered by this expertise—those being physical (computers, phones, medical devices, smart cars, drones, etc.) or virtual (websites)—as well as the tools and programming languages mastered. The level of expertise appears to be a decisive criterion within hackers' communities to grant access to peer recognition. Next to their technical skills, some hackers might possess social engineering expertise. This might appear to be useful for black hats in order to bypass physical or logical security measures.⁵ Social engineering can be

Table 9.1 A first classification based on expertise and legal goals

	High expertise	Low expertise
Legal goals	White hats	–
Illegal goals	Black hats	Script kiddies
Unlegal ^a goals	Grey hats	
	True hackers	
	Hacktivists	

^a*Unlegal* qualifies a value that is neither legal nor illegal

⁵Social skills may also be useful for white hats, when testing again the possibility of black hat hackers' intrusions.

used to gain a first internal access into a company computer network, for example. However, social engineering requires significant social skills, and not all hackers are social engineering experts. Hackers can be geeks. In his book (Marshall 2008: 1), Angus Marshall humourously defines a geek as “a nerd with social skills, and an extrovert geek looks at *your* shoes when he/she is talking to you.” Conversely, most social engineering experts are not hackers. However, they can work together, typically under the direction of the same entity, a conductor.

Hacker’s values encompass both his/her objectives and his/her moral principles. Hacker’s objectives can be noble: make the digital realm a better and more secure place; they can be ideological: promote political views and ethical values (freedom of speech, democracy); they can be self-oriented (fun, personal intellectual challenge, peer recognition); and they can be malicious (information theft, money extortion, vandalism). Hacker’s moral principles define the limits, if any, that they respect while trying to reach their objectives. These limits can be legal and/or ethical. They can also be personal or related to a particular community.

To give an example based on this first classification, we only consider both the expertise level (high or low) and the legal nature of hacker’s goals. We use *illegal* to qualify a goal which is *not legal*—typically a value related to malicious intentions—and *unlegal* to qualify a goal which is neither legal, nor illegal in nature, for example ‘to have fun’ or ‘to make the world a better place’.

9.3.2 A Second Taxonomy

We can extend the first taxonomy to develop a finer classification (Table 9.2). In our attempt to determine a more systematic classification of modern hackers, a second approach consists in considering not only the internal resources (expertise) and the internal attitude (values), but also external attitudes, as well as the external resources hackers have access to. Following the analogy with authentication technologies, the external attitude corresponds to something the hacker does and the external resources to something that he or she has.

The external attitude describes the *modus operandi*. Hackers’ *modi operandi* are numerous. Actions can be potential or actual. Some hackers will act according to what they are able to do, as long as this is compatible with their goals. Others will stop as soon as their actions could become illegal or incompatible with some moral principles. Hackers’ targets belong either to the physical world (smart objects, computers, networks, critical infrastructures, banks) or to the virtual one (e-commerce,

Table 9.2 Analogy between authentication technologies and criteria to classify hackers

	Resources	Attitude
Internal	<i>Something you know</i>	<i>Something you are</i>
	Expertise	Values
External	<i>Something you have</i>	<i>Something you do</i>
	Tools	Modus operandi

e-banking, websites, crypto-currencies). These targets span from individual properties, to companies or even to country-level assets. Hackers can work alone, in (criminal) networks or in state-sponsored groups. They can work for themselves or as mercenaries on behalf of a conductor.

In the economic paradigm, hackers can be classified according to three categories, namely what they know (their expertise, i.e. their internal resources), what they are (their values, i.e. their internal attitude) and what they do (their *modi operandi*, i.e. their external attitude). In the societal paradigm, hackers are also characterised by what they have (their tools), i.e. the external resources they have access to. Indeed, state-sponsored hackers can have access to classified information and weaponised zero-days, to sneaking, eavesdropping or deep packet inspection tools. More traditional hackers usually do not have access to these resources. Some state-sponsored hackers might even have privileged access to specific locations: Internet backbone or other key physical IT-infrastructures. State-sponsored hackers can work directly for a government, e.g. if they belong to a government agency. Alternatively, they might work for official companies selling hacking products and services to governments. Eventually, they might also belong to mercenary groups selling their services to governmental or non-governmental organisations.

In this second taxonomy (see also Fig. 9.6), a *white hat* is a skilled programmer and computer expert who looks for vulnerabilities in software, protocols, OS, computers and servers, in other physical or virtual devices, and in network systems in order to improve the IT-security of a system. As a principle, he or she abides by applicable laws. He or she will stop any action as soon as it has the possibility of becoming illegal. A white hat might work alone and disclose vulnerabilities to the legitimate owner of the targeted system, with or without a financial compensation. Most of the time, white hats are professional hackers employed by IT-security companies, the clients of whom are other companies that need their own IT-security to be assessed. *Pen testers* are white hats specialised in penetration tests using the client's IT-infrastructure. All pen testers are white hats, but not all white hats are pen testers. Indeed, a white hat might decide to analyse the code of some specific open source software without being mandated by its developer or by any third party.

Black hats are skilled programmers and computer experts who look for vulnerabilities in software, protocols, OS, computers and servers, in other physical or virtual devices, and in network systems in order to support their malicious intentions. They do not abide by ethical values and do not respect laws. Black hats typically use bugs and exploits to gain unauthorised access to a computer system or an IT-infrastructure with both malicious intent and, typically, illegal means. They aim to steal sensitive information, and personal or corporate data. They attempt to trick users or companies in order to get money transferred to accounts they have access to. They might work alone, belong to professional criminal networks or act as mercenaries by selling their services to such networks or a conductor (crime-as-a-service). All black hats are cyber-criminals, but not all cyber-criminals are black hats. Indeed, many cyber-criminals do not have much expertise. They are not hackers themselves; rather, they buy and use tools or services developed by black hats.

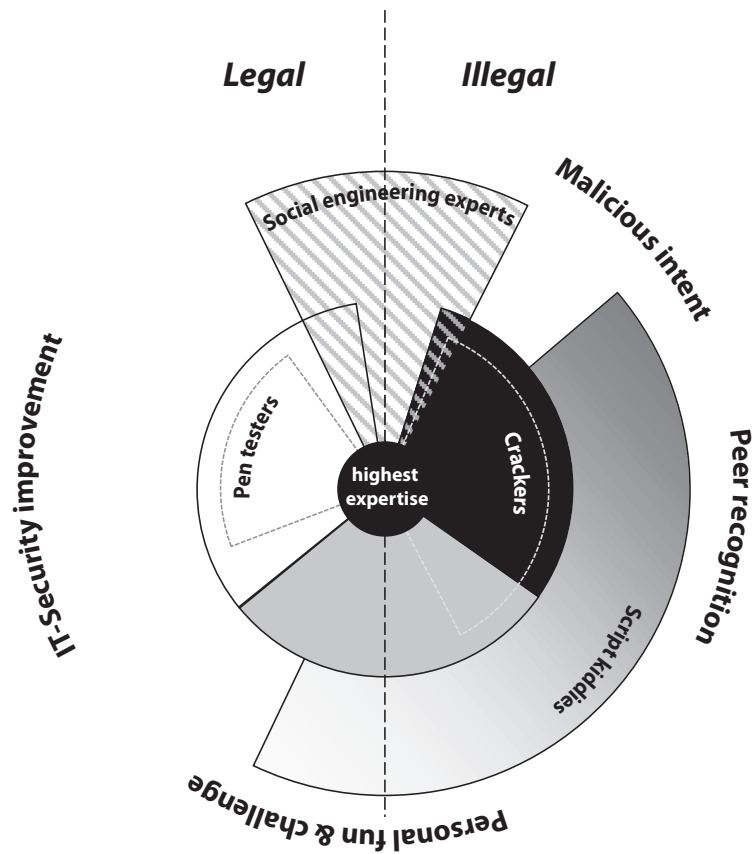


Fig. 9.6 Crackers, pen testers and social engineering experts

Grey hats are skilled programmers and computer experts who look for vulnerabilities in software, protocols, OS, computers and servers, in other physical or virtual devices, and in network systems in order to have fun, to play around, to solve a challenge, to be granted peer recognition, or to improve the IT-security of a system. Usually their intentions are not malicious and financial gain is not their main incentive. They might comply with their own moral principles that can differ from the original hacker ethic. They do not necessarily respect applicable laws, which distinguishes them from white hats.

Below we select the level of abstraction to describe the intentions and voluntary constraints of the different types of hackers at the right level of abstraction in order to distinguish them more analytically. For example, a hacktivist may share attributes with a black hat or a grey hat if he/she breaks the law, while pursuing ideological objectives (not personal gain). Grey hat hackers may also pursue apparently malicious goals, ideological or personal objectives (e.g. fun, etc.) while disregarding law altogether, but who, unlike black hats, do not aim at committing crimes. One possible way to distinguish white, grey and black hats is in terms of their relation to the law and organisations or individuals:

- A white hat acts legally and tries to be trustworthy for companies or other organisations that (may) purchase his or her services.
- A black hat acts both illegally and maliciously, e.g. against a victim (a company or another organisation or an individual), either alone or within a criminal network.
- A grey hat does not attempt to be trustworthy for companies or organisations; he or she may act illegally when required to pursue his or her goal. However, he or she does not act maliciously and attempts to minimise harm and avoid unnecessary harm.

For example, a grey hacker motivated by ideological goals (e.g. the love of justice) may illegally break the security system of a political party to highlight inadequate privacy protections, but refrains from downloading data, publishing them and causing (serious) harm. Nonetheless, he acts illegally (in most jurisdictions) because he lacks the consent of the attacked party and may also cause some harm (e.g. reputational harm for the party), which is ‘offset’ by the broader benefit for the party members’ deriving from the awareness of the vulnerability, so the act could be seen as being prevalently benevolent.

*Crackers*⁶ are black or grey hats who perform computer and system break-ins without permission. As a consequence, their activities are illegal. *Phreakers* are phone crackers.

Note that such descriptions correspond to hackers described as *personae*, or social roles, not to flesh and bone individuals. It is logically possible for the same individual to sometimes act as a white hat and sometimes as a grey hat hacker *in incognito*. However, such an individual would have to keep those identities—corresponding to the different persona, the white and the grey hat—completely separated for the public eye. Indeed, the reputation as a grey hat hacker undermines all grounds for trustworthiness that are essential to being employed as a white hat hacker. Of course, it is also theoretically possible for an individual to transact from one persona to another one: e.g. from being a black hat to becoming a white hat hacker. To be credible, however, such role changes would have to be understood as a ‘full conversion’ by others—a change in the overall motivational set of the individual. Moreover, the conversion may not be sufficient to make the individual trustworthy. Indeed, many security companies would not hire a former black hat. For example, at least until 2001, IBM had a policy to “not hire ex-[black hat]-hackers” (Palmer 2001: 772).⁷ The television series ‘Mr Robot’ (Mr. Robot n.d.) tells the story of an individual who routinely switches between the roles of a white-, grey- and even black-hat hacker, even in the course of the same day. However, the character has an unstable personality and is schizophrenic.

⁶Some authors consider black hats and crackers as equivalent terms. We introduce here some distinctions. In particular, we consider that crackers might be grey hats acting for fun with no malicious intent.

⁷This may have been the case up to 2001; the authors were not able to determine if a change of policy occurred since then.

9.3.3 *Ethical Hacking*

*Ethical hackers*⁸ are white hats mandated by clients (companies) who want their own IT-security to be assessed. They abide by a formal set of rules that protect the client, in particular its commercial assets. All pen testers are ethical hackers, but ethical hackers do not limit themselves to penetration tests. They can use other tools or even social engineering skills to stress and evaluate their client's IT-security (see also Fig. 9.7).

An ethical hacker will try to act similarly to a black hat but without causing any tort to the company. He will look for vulnerabilities that could be exploited by malicious hackers, both in the physical world and in the virtual one. In ethical hacking, the conductor of the attack is the target itself or, more precisely, the target's representative who mandated the ethical hacker to stress and assess the target's IT-security. In comparison, the conductor of a black hat's attack is never the target itself, but either the black hat or a third party—different from the target—if the black hat acts as a mercenary.

Ethical hackers adopt a strict code of conduct that protects their relationship with their clients and their client's interests. Such a code of conduct sets a frame for their attitude. It describes rules that the ethical hacker must abide by. These rules prevent the ethical hacker from taking any personal advantage of his relationship with his client. This fosters the creation of a trusted relationship similar to the special relationship between a medical doctor and his or her patients, or between a lawyer and his or her clients. The client's trust is of utmost importance in order for the ethical hacker to get the contract and to be granted permission to maybe successfully penetrate the system. Indeed, during the course of such an attack, the ethical hacker might discover trade secrets or other very sensitive data about his or her client's activities, as well as personal data about employees. The company needs to trust that the ethical hacker will not misuse his or her potential privileged access into its IT-infrastructure in order to introduce backdoors or to infringe privacy, neither during the mandate, nor after the contract is fulfilled.

The typical content of such a code of conduct contains rules which guarantee that the ethical hacker:

- will get *written permission* prior to stressing and assessing his or her client's IT-security
- will act *honestly* and stay within the scope of his or her *client's expectations*
- will *respect* his or her client's as well as its employees' *privacy*
- will use *scientific*, state-of-the-art and *documented processes*
- will *transparently communicate* to his or her client all the *findings* as well as a transcript of all his or her *actions*

⁸Some authors consider white hats, pen testers and ethical hackers as equivalent terms. In this chapter, we introduce some slight distinctions.

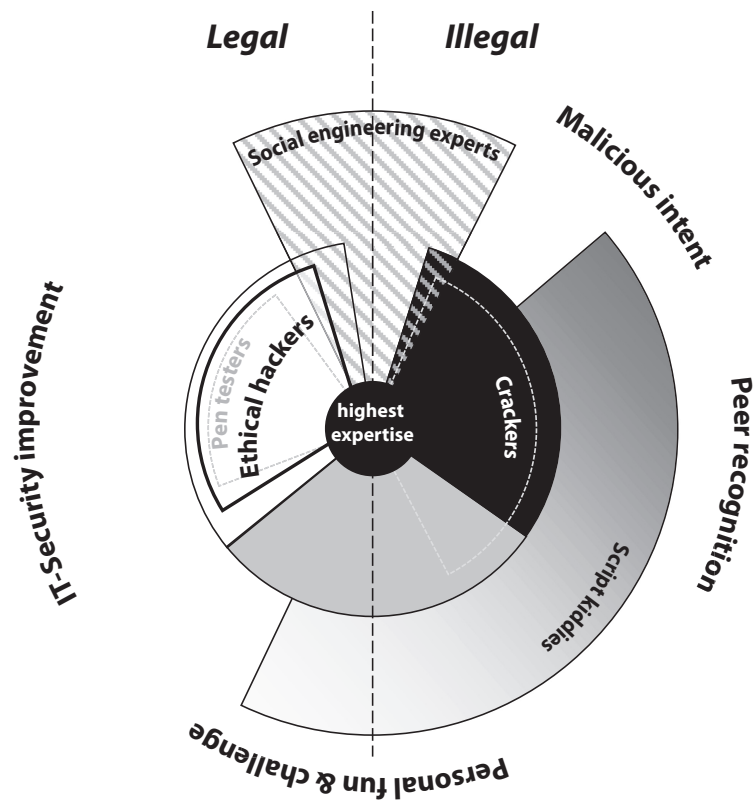


Fig. 9.7 Ethical hackers

- will remove his or her traces and will *not introduce* or keep any *backdoor* in the system
- will *inform* software and hardware vendors about *found vulnerabilities* in their products

These rules also aim at protecting the ethical hacker and making his or her work legal de facto. Different curricula even propose training and certifications in order for a hacker to become a certified ethical hacker (CEH).

9.4 Is ‘Ethical Hacking’ Ethical?

Ethical issues are evaluated according to a collection of ethical values and moral principles in regards to objectives and behaviours in a specific context.

9.4.1 *Inethical, Unethical and Ethical Hacking*

Inethical hacking can be defined as hacking that does not abide by any ethical value. Inethical hacking does not imply *unethical* behaviour, but removes ethical barriers and in doing so increases the risk of actual unethical behaviour. Greed is not an ethical value or a moral principle. Black hats typically perform inethical hacking that leads to unethical behaviour. However, what is *ethical hacking* fundamentally? Is it hacking that respects at least an ethical value? Certainly not, as such a hacking might infringe other fundamental ethical values. Indeed, intuitively, in order for hacking to be deemed ethical it should respect at least the most important ethical values at stake, balanced in a reasonable way. Therefore, non-inethical hacking is not necessarily ethical.

Precisely defining ‘ethical hacking’ in a fundamental, context-independent way is not a trivial matter, if even possible. We could start to define *prima facie unethical hacking* as hacking that infringes at least one ethical value or moral principle in an actual context. *Prima facie* means that the hacking seems unethical, although it may cease to appear so after a thorough examination of the issue. By contrast, the *ultima facie* ethical or unethical choice considers all relevant reasons, also those pulling in opposite directions, and tries to determine what is best *all things considered*. The ‘all things considered’ best act is the choice that is supported by most reasons, or by the strongest ‘undefeated’ reason, including all moral reasons, if any, bearing on the matter (Scanlon 1998). Under this logic, *non-prima facie unethical hacking* would be hacking that respects all ethical values and moral principles in that context. It makes sense to consider that any non-*prima facie* unethical hacking is *ethical*. However, should we require hacking to be non- *prima facie* unethical in order to be deemed ethical? This would lead to an overly restrictive definition. Indeed, with such a restrictive definition of ethical hacking, almost no hacking could be deemed ethical. In practice, we often face competing ethical values. Not all ethical values can be respected simultaneously; they need to be prioritised in regards to objectives and behaviors in a specific context. Therefore, a general concept of *ethical hacking* should not be reduced to non-*prima facie* unethical hacking as it would lead to a useless definition.

The *prima facie* unethical category can be further sub-divided into three categories:

1. Morally problematic: when at least one value is violated; however, the action may be justified ‘all things considered’.
2. Non (ethically) optimal (*weakly* unethical): when the action is not the best one, considering all ethical reasons bearing on the issue.
3. Ethically impermissible (*strongly* unethical): when there is a strong moral reason not to perform the action; e.g. the action violates an important moral duty (what Immanuel Kant refers to as a ‘perfect duty’), e.g. the duty corresponding to another person’s moral right.⁹

⁹ An imperfect moral duty is a duty like the duty to do charity. Wheres—Kant maintained—we all

This distinction is mirrored in terms of a normative moral psychology, specifying the emotions that a morally decent person should feel in correspondence to each category of cases: hacking that is morally wrong in the strong sense (i.e. impermissible) should evoke feelings of blameworthiness by others and moral guilt by the moral agent. Morally problematic hacking may not even be unethical *ultima facie*, and may reasonably lead to no moral blame and no feelings of moral remorse; however, some have argued that it may lead to some kind of moral regret (Williams 1981, 27–28). Non-ethically optimal hacking is unethical (*ultima facie*) but in a *weaker* sense compared to ethically impermissible hacking; it may then justifiably lead to moral remorse and regret.

We have mentioned the idea of the *all things considered* (morally) best choice. Note that in a case of value conflict, a pluralist society may not agree with a single way of balancing and resolving trade-offs between values in practice. As an example of disagreement on balancing, consider *supporting trust in cybersecurity* vs. *achieving justice*. Both values could be in conflict when a white hat hacker discovers proof of unethical behaviour, or possible signs of crimes by a company during pen testing. In order to be trustworthy, the hacker should not act in any way against the interest of the company and cannot, for example, blackmail the company, in order to induce it to stop a *weakly* unethical practice. Moreover, a white hat should avoid any investigation—even pursuing the signs of a possible crime—which is out of the scope of his or her mandate. Moreover, such an investigation might lead to discoveries that further reinforce the conflict between promoting justice and being trustworthy, e.g. the discovery of a *strongly* unethical practice by the company. We can assume that companies would have a counter-incentive to hire the services of penetration testers unless they trust them to promote their own interests in any circumstance, creating a trusted relationship similar to the relationship between a medical doctor and a patient, or between a lawyer and her client. We might also claim that widespread and protected trust in the services of white hat hackers is necessary to achieve good levels of cybersecurity for society at large, which is ethically desirable, in utilitarian terms.

It could be argued that this ‘favouring trust between white hat hackers and companies’ should include companies that do not have a perfectly blank sheet in terms of ethics and legal behaviour. This is in conflict with another strong value: the goal of achieving immediate justice and of protecting possible victims of a crime or of a strongly unethical treatment. Therefore, it is not clear if a penetration tester should always reveal strongly unethical behaviour or clues of crimes to the public, or if he or she should at least threaten to do it, in order to give the company an incentive to address the problem.

The way the term ‘ethical hacking’ is used appears to presuppose a clear and unilateral solution to the problem of value balancing: the solution that gives the high-

have a duty to charity, the duty is not perfect in the sense that we have discretion concerning when, how, and to whom we act charitably. Act-utilitarianism rejects the distinction between perfect and imperfect duties, because according to act-utilitarianism the acts that maximise aggregate utility are both right and dutiful and all other acts are wrong and impermissible in the context.

est priority to (a) refraining from acting against the interests of the company hiring the services of the hacker, (b) only acting within boundaries that have been explicitly consented to, and (c) fulfilling the expectations of the client in a way that preserves the white hat hacker's reputation for trustworthiness.¹⁰ It seems that these three conditions do not conflict in practice. A so-called 'ethical hacker' enjoys the contractual freedom to act in ways that would be illegal if they had taken place without the consent of the party hiring his or her services. He/she acts in a trustworthy way because, in addition to that, he or she acts conscientiously towards the party placing trust in him or her (Becker 1996). We may add to this 'respecting the law'; respecting all law in the pertinent jurisdictions, not only the law of private property.

As mentioned above, an 'ethical' hacker could face situations involving a trade-off between, on the one hand, preserving trust in himself or herself and white hat hackers in general and, on the other hand, achieving justice or other ethical values directly, in the short term. Note that the trade-off between trustworthiness and other ethical values could be solved differently depending on the legal framework in which the white hat hacker operates. Suppose that the hacker operates in a jurisdiction with a law that mandates the white hacker to violate a confidentiality agreement should he or she establish proof of serious crimes. In this case, the individual choice of the hacker to act against the interest of the company hiring him or her, e.g. by revealing proof of strongly unethical behaviour (which happens to also be illegal), would not in itself undermine trust. Indeed, trust relies on rational expectations and we could claim that a company could not rationally expect a hacker to protect its interests when this is explicitly prohibited by the law. Note, however, that the legal framework itself would make some companies less likely to *rely* on white hat hackers to enhance their cybersecurity, since some companies may prefer to run cybersecurity risks rather than giving others legal opportunities to reveal their illegal and/or strongly unethical activities.

To maximise the incentive to rely on white hat hackers, society could pass laws allowing and requiring them, like lawyers, priests and medical doctors, to maintain confidentiality about all behaviours, including crimes, discovered in the course of their professional activities. In such a context, a hacker would undermine trust by revealing clues, or even proof of illegal activities by firms. Note, however, that this is not the same as acting *strongly unethically*: the severity of the unethical behaviour discovered could make it the case that *all things considered*, the choice involving a breach of trust is the most ethical (ethically optimal), or even the *only* ethical (morally required) choice. Nothing guarantees that the (most, or only) ethical way to act is always the legal way to act.

It should also be noted that in choosing between these two legal frameworks, society, or its elected representatives, have to choose a trade-off point between different, equally legitimate, social values. The choice involves a balance between, on the one hand, maximising incentives to rely on white hat hackers or, on the other hand, discovering some serious crimes in the short term. Societies may make this choice based on their understanding of where the utilitarian optimum lies, but some

¹⁰For the link between trust, trustworthiness and reputation see (Pettit 1995).

societies may also adopt legislation reflecting non-utilitarian considerations. For example, the public discussion of a case in which a white hat hacker had a legal *duty* to keep an ugly crime confidential may turn public opinion against confidentiality protection, irrespective of whether it is the utility-maximising solution. A society may be moved by moral indignation to adopt legislation less protective of companies, even if the rationally expected result is that unethical companies will not hire ethical hackers and thus expose their clients to more risks.

In the previous section, we presented the well-established concept of ethical hackers (white hats mandated by clients who want their own IT-security to be assessed, and who abide by a formal set of rules that protect the client, in particular its commercial assets.) Ethical assessment in this context prioritises honesty towards the client, as well as legal and commercially-oriented values. However, other ethical values could interfere with these prioritised values. If the company which IT-security is assessed has some *ultima facie* (weakly or strongly) unethical activities, is it ethical to reinforce its IT-security? What about if its core business is deemed to be *ultima facie* unethical, in the strong sense (morally impermissible)? This shows the limit of an automated analysis of ethical behaviour based on a standard set of rules. So-called ethical hackers might perform ethical hacking in the context of their trusted relationships with their clients, while this same ethical hacking appears unethical (weakly or strongly) if we take a broader perspective.

This ethical problem cannot be solved by simply prescribing absolute respect of the law of a country. As highlighted above, nothing in the world guarantees that the ‘all things considered’ best act is always compatible with the laws of the country in which the ethical hacker operates.

Legislation might prioritise trust relations between hackers and companies above all other values.¹¹ However, it is possible—at least logically—that considerations of trust and trustworthiness do not override, or defeat, any other consideration in every context.¹² Hence, the ‘all things considered’ best act may sacrifice trust and trustworthiness.¹³ Therefore, a hacker who is ethical—in the sense of doing the best ‘all things considered’ act—is not necessarily an ‘ethical hacker’ according to the ordinary definition, which presupposes both actions to be lawful and acting in a way that proves trustworthiness *to mandating firms*.

Actually, the well-established concept of an ‘ethical hacker’ is misleading. In some ways, it is a misappropriation of the term ‘ethical’. The expression ‘trustworthy

¹¹ Maybe, it (correctly) identifies this policy as the one promoting the utilitarian optimum—maximum aggregate utility—in the long term.

¹² Even if preserving trustworthiness maximises long-term utility, for it may even be the case that the best moral view is not utilitarian.

¹³ If the ultimately *right* morality is *not* utilitarian morality, the morally right act can be one that violates a policy that has a rule-utilitarian justification (the policy that would optimise utility in the long run). It is even conceivable that the morally best/right act for *social* morality (the morality behind laws and public policies) and for *individual* morality are *different* acts, because the two moralities differ, due to constraints (e.g. of impartiality, objectivity, inter-subjectivity, integrity) that apply with different force in the two cases. If this unfortunate moral hypothesis is correct, individuals in high-stake roles are condemned to face hard-to-solve moral dilemmas occasionally. See Sect. 4.2.

for business and lawful hacker' would fit better. Indeed, the rules that the ethical hacker has to abide by are fundamentally business-oriented. They foster economic-compliant ethical behaviour,¹⁴ and they create a clear trust-enabling distinction between ethical hackers and black hats. They also protect ethical hackers in making their activities legal de facto. However, these rules do not consider the possibility of ethical issues competing with the need of a trusted relationship and a protection of economic interests. Often, ethical hackers essentially agree to stay faithful to their client whatever the client's activity is. This creates an inviolable trusted relationship similar to the relationship between a lawyer and his or her client, or between a priest and his faithful. Is it ethical to keep secret (and protect) the illegal activities of a client? In utilitarian terms, it depends on the existence or not of a greater public interest to improve companies' IT-security even at the cost of covering critical non-ethical behaviours. Even if it were not a matter of public interest, covering critical non-ethical behaviour may simply be irreconcilable with reasonable individual moralities (e.g. of a more deontological type). Some ethical hacking companies introduce a provision allowing them to report observed illegal activities, at least if questioned by the police in the course of an investigation.

Any practical definition of ethical hacking should incorporate the existence of possible competing ethical values, even within a fixed context (see also Chap. 3). In other words, hacking could be deemed ethical when it sufficiently respects ethical values and moral principles at stake in regards to objectives and behaviours in a specific context. This provides a practical definition of *ethical hacking*. We are not suggesting that this definition should replace the ordinary one. The most important purpose fulfilled by having a new definition is to distinguish both concepts. One possibility would be to use 'trustworthy for business and lawful hacker' and 'ethical hacker' to distinguish both of them. An alternative would be to use 'ethical hacker' in the usual (business-oriented) way and invent some other label for the sufficiently 'all things considered' ethical hacker instead. This new definition—as well as ethical assessment actually—is intrinsically vague, subject to interpretation and context-dependent. This emphasises the fact that ethical evaluation cannot be reduced to an *a priori* assumption that business-oriented values should take priority, and the qualification of ethical should not be limited to a narrow definition of professional ethics.

9.4.2 Competing Ethical Values

Ethical evaluation, like any evaluation process, produces values that can be fed into a decision process (Pollitt et al. 2018: 8). The values resulting from an evaluation process are not restricted to numbers. They can be impressions, feelings, opinions or judgments. In her axiological sociology essay (Heinich 2017), Nathalie Heinich

¹⁴This behavior may, or may not, be optimal in utilitarian terms (it is often very difficult to determine what maximises utility in the long term and some economic behavior may be harmful, all things considered). Even if it is optimal in utilitarian terms, it may not be ethical, if, as many people think, utilitarianism is not the right ethical theory.

identifies three ways to attribute a value: measurement, attachment, judgement. An ethical evaluation is typically of the third kind: some form of judgement. The decision process following an ethical evaluation usually allows or does not allow an action, an activity or a behaviour to be pursued.

A priori, the ethical assessment of relevant ethical values related to hacking could perform an ethical evaluation of all four criteria used to classify hackers (see also Table 9.2):

- hacker’s expertise
- hacker’s tools
- hacker’s values
- hacker’s *modus operandi*

However, a hacker’s expertise is knowledge. It is ethically neutral and does not carry out direct ethical issues. Tools available to the hacker are not relevant from an ethical standpoint either. This does not mean that hacking tools do not create ethical issues. Indeed, *the creation or not* of some hacking tools, e.g. weaponised zero-days, leads to important ethical issues at a societal level: on the one-hand, weaponised zero-days allow countries to develop cyber-weapons to dissuade potential enemies, on the other hand, unpatched vulnerabilities—if discovered by or made available to black hats—can endanger large scale IT-systems. The WannaCry worldwide ransomware attack that shut down UK hospitals and numerous systems in May 2017 shows the impact of such a weaponised zero-day falling into criminal hands (Mohurle and Patil 2017).

Eventually, only the hacker’s values and *modus operandi* need to be ethically assessed by the evaluator. Note that the evaluator can be either the hacker or another person.

The result of an ethical evaluation depends on the evaluator’s expertise, on the available information, and on his or her way of handling and processing this information, as well as on his or her own criteria and values’ prioritisation and interpretation. State-sponsored hackers, for example, might be deemed ethical if the evaluator prioritises values of the sponsoring state, whereas these same hackers might be considered simultaneously unethical by evaluators living in the targeted country. The interpretation of the facts (state-sponsored actors do not necessarily follow traditional white hats’ rules; they typically try to introduce and keep backdoors in the targeted system; they might use zero-days and not divulge them to the developers) really depends on the evaluator’s perspective, interpretation and prioritised values.

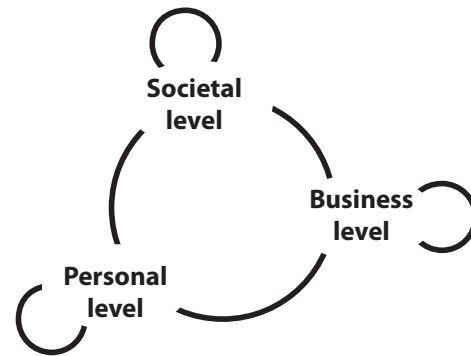
Ethical evaluation parameters also present similarities with the four classes of authentication technologies (Table 9.3).

The evaluator’s level of expertise allows a distinction to be made between an ethical opinion and an ethical expert evaluation (Heinich 2017). The information available to the evaluator might change over time, possibly resulting in new conclusions. This is in particular true when a so-called ethical hacker penetrates his or her client’s infrastructure and discovers ethically sensitive new information. The way the evaluator processes the information relates to quality procedures and best practices; it influences the confidence in the conclusion. The core of the evaluation resides in the evaluator’s own prioritisation of (competing) values at stake.

Table 9.3 Similarities between authentication technologies and ethical evaluation parameters

	Resources	Attitude
Internal	<i>Something you know</i>	<i>Something you are</i>
	Expertise	Values prioritisation
External	<i>Something you have</i>	<i>Something you do</i>
	Available information	Information processing

Fig. 9.8 Potential conflicts between collections of possibly competing ethical values



When addressing ethical hacking, we should consider at least three collections of possibly competing ethical values (see also Fig. 9.8): one at a personal level (hacker's own perspective), one at a business level (company's perspective) and one at a societal level (global perspective). Ethical conflicts can happen within one of these collections or between some of them.

So-called ethical hackers can ethically evaluate their own attitude, i.e. their values and their modus operandi, and they probably will because they chose not to use their expertise for malicious purpose. The code of conduct that ethical hackers have to abide by strongly focuses on the collection of values at a business level. Therefore, these values must belong to the own hacker's ethical values and moral principles. Already at this stage, competing ethical values can appear if, for example, protecting an employee's privacy (whose emails reveal that he is blackmailed by a competitor's board member) conflicts with transparently communicating all the findings to the mandating client. Generally speaking, it will be easier to assess if a hacker is ethical in the narrow (and usual) sense of the term, which assumes the priority of business-oriented moral values.

Ethical hackers also have their own values and moral principles at a personal level. They might share some of the original hacker ethic. If their ethical values conflict with those at a business level, their ethical evaluation of the situation will depend on the prioritisation of the values. A strong personal ethical value or a well-established important societal value might prevail on any other business-related value and lead to breaking the code of conduct. This is in particular true if the ethical hacker unveils critical non-ethical behaviours within the company. In this case, the evaluation of whether the hacker is ethical will be significantly more complex. It is likely to achieve reasonable disagreement, even between equally well-informed persons, concerning what is the ethically optimal act in a given context. There might be no pre-established harmony between values—e.g. no way to maximise fairness

and aggregate well-being at the same time—(Berlin 1991; Nagel 1991; Raz 1986). Moreover, even individuals who rely on monistic moral views (e.g. utilitarianism, which recognises only utility, understood as well-being) and single-rule based moralities (e.g. again utilitarianism: maximise aggregate well-being in the long term) may disagree on what the actual best choice turns out to be (see also Chap. 4 for a discussion of ethical frameworks in cybersecurity).

Note that our argument does not rely on a rejection of ethical realism or cognitivism. Realism is entailed by the view that the question concerning ‘the all things considered best choice’ can be objective, because it is determined by moral objective facts existing independently of mental states (beliefs, attitudes, emotions) about the choice in question. Cognitivism is entailed by the view that these objective moral reasons, or facts, are not facts about what (all, or the majority) of people actually *want* to be the case. The key point is that, even conceding that morality is grounded in objective facts independent of will of any agent, it may be *in fact* extremely difficult to determine what the *morally best* choice is.

9.4.3 A Pragmatic Best Practice Approach

Pen-test companies and other IT-security hiring white hats face a competing values dilemma (see also Chap. 15). On the one hand, they need to create a trusted relationship with their clients. On the other hand, they need to respond and even anticipate their employees’ ethical expectations. There is certainly no perfect solution to solve this dilemma, as ethical evaluation has an intrinsic personal component, is subject to interpretation and is context-dependent.

As explained above, companies hiring ethical hackers develop a code of conduct that reinforces the business-related ethical behavior of their employees, guarantees that their hacking activities are compliant with applicable laws and fosters a trusted relationship with their clients.

As already mentioned, some ethical hacking companies have introduced a provision allowing them to report observed illegal activities, at least if questioned by the police in the course of an investigation.

To minimise the inherent risks related to the competing values dilemma, an active European pen-test company with about 40 employees created an internal ethical committee. This ethical committee is composed of three employees, freely elected by all employees. Company board members are not allowed to be elected in order to avoid business-related biases in the ethical evaluation. Any employee can submit his or her ethical concerns about an upcoming project if this employee fears that participating in such a project could create a conflict with his or her own values or moral principles, or with other societal ethical values. Members of the ethical committee are in a position to make an independent ethical evaluation. Their decision is binding and cannot be challenged, neither by the direction nor by the other employees. If the committee decides to block a project, the company will stop it independently from having financial consequences.

This example illustrates a possibility to anticipate potential competing ethical values in order to avoid employees breaking their code of conduct or leaving the company. Such an approach enriches and strengthens the concept of ethical hacking and goes beyond a rule-based definition. It promotes an ethical evaluation that is not reduced to an automated process or a checklist, and allows a fine interpretation of the context and a more subtle ethical evaluation, as well as context-dependent decisions.

9.5 Conclusion

The term ‘hacker’ has many different meanings, even within the context of computerised systems. It should not be amalgamated with that of a cybercriminal only. In this chapter, in order to capture a much broader perception of the term and to describe its nuances more faithfully, we developed a new systematic and neutral classification based on four categories: the hacker’s expertise (his or her internal resources), the hacker’s own values and moral principles (his or her internal attitude), the hacker’s *modus operandi* (his or her external attitude), and the tools and information that he or she has access to (his or her external resources). These four categories can be related to the four categories of authentication technologies: something that the hacker knows, something that the hacker is, something that the hacker does, and something that the hacker has.

The term ‘ethical hacker’ in its wide acceptance appears to be misleading and a misappropriation of the term ‘ethical’. Particular pluralist societies, those that recognise that different ethical values are valid and there is no single simple way of measuring or ranking them, are likely to disagree on what is the morally best behaviour for a hacker to adopt in every given circumstance. The expression ‘business-oriented ethical hacker’ would fit better. Moreover, it gives the false impression that it is sufficient for hacking activities to abide by a list of fixed rules in order to be deemed ethical. Ethical evaluation *cannot* and *should not* be reduced to a checklist of rules to abide by those rules that are legal and/or ethical. This is especially true in contexts where at-the-edge hacking opportunities are sometimes in a grey zone which is not covered by current laws, e.g. for spy and state-sponsored hacking activities.

The creation of a code-of-conduct with rules to abide by is a welcome and necessary step in order to support ethical hacking. However, it is not sufficient. Other mechanisms—such as internal ethical committees—have to be created within the pen-test companies or the Gov-CERT units in order to allow a finer interpretation of each context, a more subtle ethical evaluation, and context-dependent decisions.

Acknowledgments The authors would like to thank their colleagues Eoghan Casey and Olivier Ribaux, who reviewed a draft version of this document, for their fruitful comments. 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.

References

- American Heritage Dictionary Entry: Hacker (n.d.) <https://www.ahdictionary.com/word/search.html?q=hacker>. Last access 7 July 2019
- Barber R (2001) Hackers profiled—who are they and what are their motivations? *Comput Fraud Secur* 2001(2):14–17
- Becker LC (1996) Trust as noncognitive security about motives. *Ethics* 107(1):43–61
- Berlin I (1991) The crooked timber of humanity: chapters in the history of ideas. In: Hardy H (ed) Knopf: distributed by Random House, New York
- Bratus S (2007) What hackers learn that the rest of us don't: notes on hacker curriculum. *IEEE Secur Priv* 5(4):72–75
- Heinich N (2017) *Des Valeurs. Une Approche Sociologique*. Editions Gallimard, Paris
- Lichstein H (1963) Telephone hackers active. *The Tech*, MIT. <http://tech.mit.edu/V83/PDF/V83-N24.pdf>. Last access 7 July 2019
- Marshall AK (2008) *Digital forensics: digital evidence in criminal investigations*. Wiley-Blackwell, London
- Mohurle S, Manisha Patil (2017) A brief study of Wannacry threat: Ransomware At-tack 2017. *Int J Adv Res Com Sci Udaipur* 8(5). <https://search.proquest.com/docview/1912631307/abstract/DEF9AE2FF2924E35PQ/1>. Last access 7 July 2019
- Mr. Robot (n.d.). <http://www.imdb.com/title/tt4158110/>. Last access 7 July 2019
- Nagel T (1991) *Mortal questions*. Cambridge University Press, Cambridge
- Olson P (2013) *We are anonymous: inside the hacker world of LulzSec, anonymous, and the global cyber insurgency*. Back Bay Books, New York
- Palmer CC (2001) Ethical hacking. *IBM Syst J* 40(3):769–780. <https://doi.org/10.1147/sj.403.0769>
- Pettit P (1995) The cunning of trust. *Philos Public Aff* 24(3):202–225. <https://doi.org/10.1111/j.1088-4963.1995.tb00029.x>
- Pollitt M, Casey E, Jaquet-Chiffelle D-O, Gladyshev P (2018) A framework for harmonizing forensic science practices and digital/multimedia evidence. OSAC.TS.0002. OSAC Task Group on Digital/Multimedia Science. OSAC/NIST. <https://doi.org/10.29325/OSAC.TS.0002>
- Raz J (1986) *The morality of freedom*. Oxford University Press, Oxford
- Scanlon T (1998) *What we owe to each other*. Belknap Press of Harvard University Press, Cambridge, MA
- Techopedia.Com (n.d.) What is a black hat hacker? – definition from Techopedia. <https://www.techopedia.com/definition/26342/black-hat-hacker>. Last access 7 July 2019
- Williams B (1981) *Moral luck: philosophical papers 1973–1980*, 1st edn. Cambridge University Press, Cambridge