A. Nieto, R. Rios, and J. Lopez, "Privacy-Aware Digital Forensics", *Security and Privacy for Big Data, Cloud Computing and Applications*, 2019.
NICS Lab. Publications: https://www.nics.uma.es/publications

Chapter 8

Privacy-Aware Digital Forensics

Ana Nieto, Ruben Rios and Javier Lopez[1]

Digital forensics and privacy are two naturally conflicting concepts. While privacy can be defined as the desire of people to decide for themselves when, how and to what extent their personal information is shared with others, digital forensics is aimed at acquiring and analysing relevant data from devices in the scope of digital forensic investigations, following a set of procedures to comply legal proceedings.

Digital forensic investigations are usually carried out after seizing the devices from investigated suspects or third parties, who consequently lose control over the data being accessed by the investigator. Moreover, digital forensic tools are even capable of retrieving information which is apparently no longer present in the device because the user decided to delete it. These tools also have the ability of correlating information from different sources giving rise to new actors in the investigation whose privacy can be affected. Also, the lack of context to determine when and why some of the contents which were intentionally deleted by the users may result in wrong accusations.

All things considered, even when digital investigations are conducted by responsible professionals, the data collected from personal devices may result in dreadful invasions to individual privacy. Inevitably, this leads to a controversial debate on the need for strong privacy guarantees in the context of digital forensics. This chapter aims to shed some light into this imperative and highly demanded debate given the fundamental role that the user and his/her personal data plays in current and future digital investigations.

8.1 Introduction

Digital forensics dates back the early 70's when two experts were capable of recovering a highly fragmented database file that was mistakenly deleted from a computer system [1]. At that time, the variety of devices was rather limited and their capacity was considerably smaller compared to current systems. Also, most electronic devices were not interconnected and when connected, data transfers were minimal because of the low capacity of the networks at that time. Since then, technology has

[1]Network, Information and Computer Security (NICS) Lab, Computer Science Department, University of Málaga, Spain

evolved at a tremendous pace and so did digital forensics in an attempt to keep up with technology changes. In the current picture of digital forensics, the users, their devices and the communication infrastructure are strongly related to each other, more than ever before, resulting in an extremely complex ecosystem (Figure 8.1).

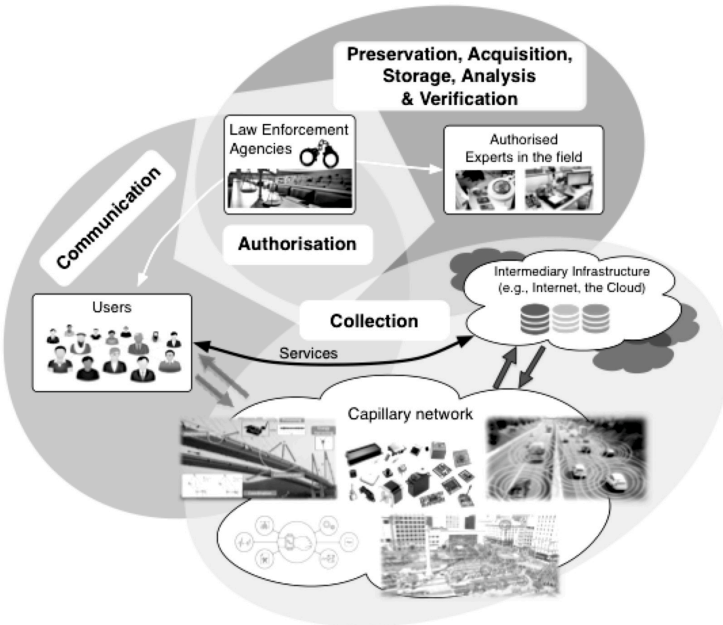

Figure 8.1 Actors in a Digital Forensic Ecosystem

Given the plethora of devices owned by most people and the amount of information stored in them, privacy-preserving digital investigations has been recognised among the key challenges that must be overcome by digital forensics in the near future [2]. While the amount of data collected for an investigation is increasing, usually only a small portion of these data are relevant to the investigation. Moreover, during the process of extracting data for an investigation, some personal data irrelevant to the investigation, may be exposed. These data may be stored in personal devices but also in remote machines, such as Cloud servers, IoT devices, etc. Even more so, given the multi-tenant nature of current computer systems, a single device may not only contain data from the individual being investigated but also from other users not even related to the investigation, thus leading to what is referred to as third-party privacy breach (TPPB) [3].

Some solutions have been devised to prevent violating users' privacy during digital investigations thereby protecting both the users and the investigators from being accused of privacy invasions. Unfortunately, as we will show next, privacy-preserving digital forensics is still in its infancy and much work still needs to be done in this area. In fact, although it has been widely recognised as one of the major

challenges in digital forensics, current tools and methodologies are mostly oblivious to this problem and provide no support for dealing with privacy issues.

In the following sections we will delve into the role of privacy in digital forensics investigations. We start by describing these two disciplines separately to understand their requirements and principles. This will provide the reader with a solid base on how digital forensics and privacy conflict with each other but also shows that there is room for privacy-respecting digital investigations. This is followed by a detailed analysis of the current state of the art in privacy-aware digital forensic approaches. In addition, this chapter gives insight into the social, contextual and technological changes affecting digital investigations.

8.2 Digital Forensics

Digital forensics can be defined as the "*scientific tasks, techniques, and practices used in the investigation of stored or transmitted binary information or data for legal purposes*" (ISO/IEC 27037:2012).

Before the explosion of the Internet and the widespread adoption of social networks, the digital forensics ecosystem was limited to personal computers seized by law enforcement officers in the context of digital investigations. The extraction and analysis of digital evidence could be complex, but the lack of security mechanisms (e.g., secure data erasure or encryption) allowed a large amount of information to be recovered and analysed. In addition, the availability and predominance of certain operating systems and applications facilitated the procedural analysis of data. Thus, the most typical actions in digital forensics were related to data recovery with non-repudiation guarantees and the analysis of digital artefacts (e.g., pictures or other digital files) for cases of fraud or copyright violations. However, the scope of this discipline is broader.

Digital forensics can also help to determine the timeline of events carried out in an entire system to better understand the casuistry or the motivations of an attacker and thus help to prevent future attacks. In fact, there are multiple tools available to forensic professionals for this purpose. The most common ones are software applications aimed to conduct general-purpose digital investigations (e.g., AccessData FTK, MPE+, EnCase, Autopsy) or applications specific to a particular domain (e.g., Volatility for memory analysis). There are also pre-configured environments with tools available to conduct digital investigations (e.g., Kali Liux, Blackarch Linux, SANS DFIR, CAINE). Other utilities are hardware-dependent such as disk copy utilities or specific hardware (e.g., JTAGulator). Finally, some environments and devices, such as smartphones and cars, require the use of manual techniques. An example of this is the chip-off technique which consists of extracting embedded chips for analysis.

The set of tools to be used will be determined by the type of investigation (public or private/administrative) and the specific requirements of the context being analysed (devices, volatility, etc.). All this knowledge requires specific training in the use of methodologies and tools which is not always easy or possible to provide. Some of

the reasons for this may be limitations on hardware resources, on the time to conduct practical cases, or licence requirements.

The volume of data and devices susceptible to analysis is continuously growing. Aside from the operational problems that dealing with vast amounts of data causes to Law Enforcement Agencies (LEA), this has also resulted in the specialisation of certain areas in the field of digital forensics due to the emergence of data from new contexts. These include among other, the Cloud and the Internet of Things (IoT), which in turn have led to Cloud- and IoT-Forensics, respectively. The way in which digital forensics has evolved with the development of new scenarios is analysed in Section 8.2.1.

Cybercrime evolution

According to the Internet Crime Compliant Center (IC3), in 2013-2017 there were a total of 1,420,555 Internet scams affecting victims across the globe, causing around $5.52 billion losses [4]. The motives of the criminals are similar to those that had been years ago (e.g., cyberespionage, financial crime, revenge, cyberterrorism, etc.), but the means to commit crimes, especially the telematic means, are much more powerful [5]. Nowadays, there are multiple cyberweapons or offender resources such as key-loggers, exploit kits or botnet kits, some of which have been developed to affect even the most recent IoT devices. Intrusion detection systems do their best effort to stop these threats but they are not enough.

In recent years, digital forensics has begun to be an area reinforced by different learning courses (either specialised courses or as part of the academic program of various Universities) without a lack of consensus in the training methodologies. One of the reasons is that the specific procedures of a digital forensics professional depend on the law of the country where he/she will practice and the type of investigation he will focus on. The technical background of a digital forensic practitioner should be the same but the legal procedures to commit his/her work can be completely different depending on the context in which the professional is performing his investigation. Another reason that probably affects the training methodologies is the evolution of the digital ecosystem.

8.2.1 Evolution of Digital Forensics

The evolution of digital forensics has been motivated by contextual changes. While many researchers consider digital forensics and computer forensics as equivalent terms, the latter is only a portion the former; as new devices and technologies appear and become more complex, it is more difficult to cover all different aspects of the discipline. Figure 8.3 shows specific areas that can be grouped under the term digital forensics.

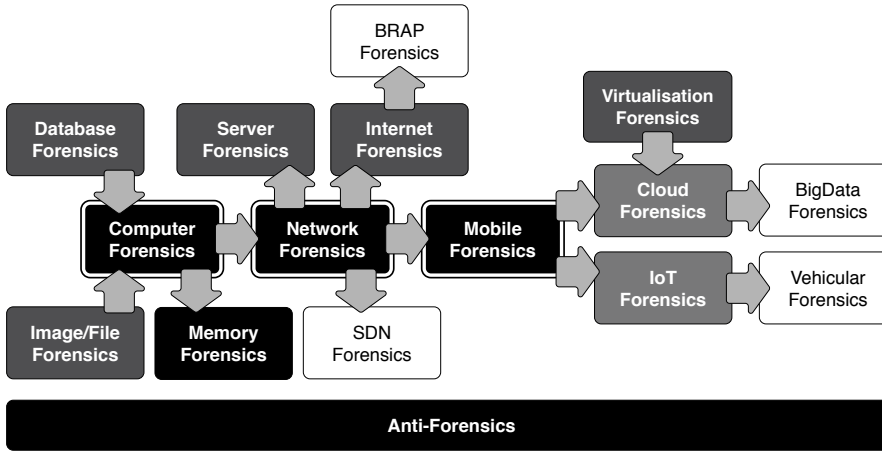

Figure 8.2 Evolution of Topics in Digital Forensics

Computer forensics has its origin in the analysis of evidences that are in digital form, typically found in a computer. This means that it is not necessary to be an expert to perform some of the basic operations considered of forensic nature, such as applying techniques to identify modifications in images or files. Indeed, this is a natural step when analysing digital evidence found in a crime scene or provided by a claimant. This raises the following questions: (i) what makes computer forensics a discipline? and (ii) when is it possible to consider that a certain 'X-Forensics' discipline is *mature*?

There are some indicators that can help answer the previous questions. First, unlike image/file forensics or database forensics, *computer forensics* groups the entire context of a computer and is a term widely used by a group of experts who have defined the problem in books and standards. Moreover, the materialisation of a new discipline typically includes other existing disciplines (now sub-disciplines). For example, memory forensics can be considered as part of computer forensics since, after all, there is memory inside a computer. However, *memory forensics* must be considered as a mature discipline by itself which, in fact, can be applied to other devices and not just computers, for instance, mobile devices [6, 7]. Indeed, there are specific tools developed for memory forensics (e.g., DumpIt, Volatility) and new emerging areas, such as malware analysis, are very dependent on this domain because there are attacks that avoid leaving traces by using only volatile memory. In the same way, *network forensics* started to be analysed because of intrusions and attacks to computer systems that caused great harm in the 80's. Today, there is a wide range of tools and techniques to capture and analyse network traffic, as well as the artefacts and digital evidence generated from network communications [8].

Figure 8.3 depicts the evolution of digital forensics by relating consolidated disciplines (i.e., properly defined, accepted by the experts and practitioners, well-documented and with solid tools) with other sub-disciplines (e.g., image/file forensics, database forensics, server forensics) or emergent ones (e.g., Cloud forensics

and IoT-Forensics) that are acceptably described, or, whose challenges are partially identified.

Notably, mobile forensics is quite different from network forensics or memory forensics. Similar to computer forensics, mobile forensics emerged because a new type of device appeared and it was necessary to develop the right tools and methodologies to deal with them. Nowadays, mobile forensics is properly defined but this was an inflexion point. New forensic disciplines (e.g., IoT-Forensics [9]) are also emerging because new types of devices (e.g., sensors nodes) are being developed.

New use cases, scenarios and devices keep emerging, such as autonomous cars and drones [10]. An autonomous vehicle can be stolen or used maliciously by external entities. Therefore, new tools are required to preserve and store digital evidence in these new contexts [11, 12, 13].

It is important to highlight that, unlike computer forensics, new emerging areas are analysing privacy as a requirement from their inception. This is a clear example of how the current social context is changing digital forensics, making new areas in digital forensics incorporate privacy as a requirement.

8.2.2 Digital Forensics Rules

Digital investigations are governed by a set of principles and standards that define the procedures accepted by a broad community of experts in the field. In this section both principles and standards for digital forensics are detailed in order to provide a solid foundation that helps to understand this discipline.

8.2.2.1 Digital Forensics Principles

Digital investigations are conducted following a set of well-defined methodologies and processes accepted by a broad community of experts in the field (see Section 8.2.2.2 for further information). Although there are various methodologies covering different contexts, all of them respect a set of basic principles:

Integrity: the actions carried out during the seizure of digital evidence should not change the evidence itself.

Competence/Expertise: any person accessing the original digital evidence must be forensically skilled and competent.

Availability: the whole digital evidence management process (seizure, access, storage or transfer) must be fully documented, preserved and available for review.

Responsibility: those individuals in possession of digital evidence are responsible for all actions taken with it during the period in which is guarded by them.

Agreement: any agency that is responsible for conducting digital forensic processes must comply with these principles.

Repetitiveness: an independent third party should be able to repeat the entire process applied to digital evidence and achieve the same result.

Most of the previous principles are defined by the *International Organization on Computer Evidence* (IOCE 1999). In particular, the principle of repetitiveness is not defined by the IOCE, but it is assumed that this principle should be ensured through compliance with the rest of principles. There is one exception which justifies not

including such principle among the previous four principles: when volatile data is acquired (e.g., a memory dump during live forensics) it is very difficult, if not impossible, to repeat the process with the same results because the data could be rewritten precisely due to the tools used to make the acquisition of the data in memory.

It should be noted that there are legal and ethical principles that depend on the country where the digital investigation is carried out and the investigator's ethics. In particular, anything not explicitly considered by law (cf. Section 8.3.2.1) is subjective and depends on the interpretation of the investigator. Additional precautions can be taken before the digital investigation starts by establishing contracts (commitment rules and acceptable use policies) between the digital investigator and the client.

8.2.2.2 Digital Forensics Standards

A set of international standards have been developed to define the guidelines for the management of digital evidence and digital forensics processes. One example is the ISO/IEC 27037:2012 standard, which provides guidelines for four basic processes in the management of digital evidence: i) identification, ii) collection, iii) acquisition and iv) preservation. After this, it is assumed that the digital evidence will be analysed in the laboratory. Precisely, the ISO/IEC 27042:2015 standard describes steps for: v) investigation, vi) analysis, vii) interpretation and viii) reporting. Also, other important aspects, such as analytical models to be considered, and the mechanisms and techniques to demonstrate the competence and proficiency are provided.

It is important to highlight that, depending on the model or methodology chosen to conduct the digital investigation, more or less phases/processes are considered. However, six phases are generally considered [14]: planning, identification, collection, preservation, examination, analysis and report. Typically, the planning phase has to be done before the field work - selection of procedures, legal and ethical considerations (e.g., responsibilities, authorisations, etc.), tools, and so on; the identification, collection and preservation occur at the crime scene; and the examination, analysis and report can be done at the laboratory, once the digital evidence has been collected thus preserving the Chain of Custody.

Chain of Custody (ISO/PC 308)

The technical committee (TC) ISO/PC 308 (created in 2016) is currently working on the standardisation of what is referred to as the Chain of Custody (CoC). This is a term applied beyond digital evidence management. In words of the ISO/PC 308 TC, "a chain of custody is a succession of responsibilities for processes as a product moves through each step of the supply chain. Each supply actor has to implement and document a set of measures in order for the chain of custody to function." The goal is to guarantee the traceability and integrity of the product, which in the context of digital forensics is the digital evidence. If the integrity and authenticity of the digital evidence is put in question then, the entire digital investigation can be rendered useless.

The ISO/IEC 27043:2015 and ISO/IEC 30121:2015 standards can be considered horizontal to the previous ones. The former encapsulates "idealised models for common investigation processes across various investigation scenarios", therefore covering all the previous steps or processes of any digital investigation. The latter describes how to conduct digital forensics within an organisation to take legal actions after a security breach or in the case of any other incident in which information technology is a decisive factor.

In addition, the ISO/IEC 27050:2016 - Electronic discovery - standard is closely related to digital investigations. This standard is decomposed in four parts: 1) Overview and concepts, 2) Guidance for governance and management of electronic discovery, 3) Code of practice for electronic discovery, and 4) ICT readiness for electronic discovery. In this standard, *electronic discovery* is defined as the "discovery (3.4) that includes the identification, preservation, collection, processing, review, analysis or production of Electronically Stored Information". These are, in fact, the typical steps or processes already defined for digital evidence management. This standard emphasises the cost associated with the managing electronically stored information (ESI). It considers there are ESI that must be preserved (e.g., logs) while other ESI that can be considered expendable (e.g., deleted data or unallocated space on hard drives). However, this is in conflict with digital forensics.

According to the ISO/IEC 27037:2012 standard, digital evidence is defined as "information or data stored or transmitted in binary form that **may be relied upon as evidence**". This is different from ESI, defined in ISO/IEC 27050-1:2016 as: "data or information of any kind and from any source, whose temporal existence is evidenced by being stored (3.26)" (volatile storage or non-volatile storage) "in or on any electronic medium". Then, a digital evidence is, by nature, the ESI that is *relevant* to a digital investigation.

8.2.3 Digital Forensics Challenges

The challenges in digital forensics have also evolved over the years. A good summary of the history of digital forensics is provided in [1]. According to the author, the history of digital forensics can be divided into three stages, which helps to understand the evolution of the challenges in this field.

Thus, during the first stage, 70's - 90's, digital forensic professionals worked with LEAs "on an ad-hoc, case-by-case basis" and the need to perform digital forensics was rather limited, because as the capacity of the disk was smaller, users saved less data and printed more. In the second stage, 1999-2007, denoted as "the golden age", multiple vendors began to develop specific digital forensic tools that required relatively limited training, allowed to recover deleted files - basic file carving - or to analyse e-mail messages. It was then when new disciplines such as Network and Memory Forensics were born to answer to some specific challenges: obtain data that allows to understand the network events and obtain memory data that would allow us to circumvent the security controls of the computers (c.f. Section 8.5). Furthermore, it was during the *golden age* when the research in digital forensics had rapid growth and the professionalisation of the sector began. This resulted in the accep-

tance on the use of specific tools and procedures to conduct digital investigations by the community of experts.

The third stage considered in [1] is from 2007 to 2010 (year in which the paper was published), but the environmental characteristics and the challenges in digital forensics are basically the same as today (see Table 8.1).

Table 8.1 Digital Forensics Challenges

Characteristic	**Digital Forensic Challenge**
Growing **size** of storage devices	Insufficient **time** to create a forensic image or to process the data
Prevalence of **embedded** flash storage and proliferation of HW interfaces	Storage devices can no longer be easily removed or imaged. **Embedded storage is routinely ignored** during forensic investigations (e.g., persistent memory inside GPUs)
Proliferation of **Operating Systems and file formats**	Increase the requirements, **complexity and cost** of digital forensic **tools**
Multiple devices in a single case	**Correlation** of digital evidence is needed
Pervasive **encryption**	Hinders or **avoids the processing** of data
Cloud for remote processing and storage	Complicates the **identification and acquisition** of digital evidence. Makes **impossible to perform basic forensic methodologies** of data preservation and isolation
Malware not written in persistent storage and capable of using **anti-forensic** techniques	Need for RAM forensics tools which are more difficult to create than disk tools and new systems to **capture the malware for in-depth analysis**
Law & Privacy	Limits the scope of forensic investigations

It is important to emphasise that, regardless of the clear value of technical challenges, the challenges motivated by social changes (e.g., the *need* for privacy) are usually not so prominent in the literature and, nevertheless, play a crucial role given the new areas highlighted in the previous section and summarised in Figure 8.3. For example, one of the major challenges is to make the new areas and methods used (c.f. Section 8.2.2.2) understandable by an audience that is not an expert in the field and that is increasingly involved in digital investigation [15].

In particular, privacy is a major concern in digital forensics, because i) digital forensic tools will be more and more proactive; the inference of user's information will be fundamental to speed up the processing of data, and ii) privacy tools, in general terms, affect the acquisition and analysis of digital evidence, being considered as anti-forensic mechanisms in many cases [16]. Nevertheless, although the confrontation between digital forensics and privacy is intuited, it is unfair to make an assessment of the influence of privacy in digital investigations without first knowing the nature of data privacy. In order to fully understand said relationship, the basic characteristics of digital privacy will be addressed below. Furthermore, Section 8.3.3 will return to the digital forensic challenges but from the point of view of privacy.

8.3 Digital Privacy

Privacy is a difficult to explain concept. There have been many definitions throughout history, each of which consider different aspects and perspectives of this convoluted concept. For example, an extensively used definition of privacy was formulated at the end of the 19th century by Warren and Brandeis [17], who described it as "the right to be let alone". However, this definition of privacy covers only a single dimension of the term and many other jurists, scholars, philosophers and sociologists have considered and introduced new aspects which broadens its scope.

One of the main problems in reaching a satisfactory definition for privacy is the fact that it is a very subjective term. Privacy has to do with the desires and expectations of people. Moreover, desires and expectations evolve and change over time although they are very much conditioned by the past and current situation. For example, a person may be willing to share his religion believes when living in his/her home country but may be reluctant to do so when travelling to a different state or country.

Consequently, privacy is about giving people a feeling of security and confidence. People need to feel in control of what personal information is known to others and want to have the ability to decide how much information and in which circumstances. These arguments lead to another widely-accepted definition of privacy by Westin [18], who describes it as the desire to determine under what circumstances and to what extent personal information is exposed to other entities.

Note that no matter which is the most accurate definition of privacy, what is really important here is the observation that people feel vulnerable and insecure without it. And because of this, privacy has been recognised as an individual right in numerous laws, regulations and treaties all over the world, including the Universal Declaration of Human Rights.

8.3.1 Evolution of Digital Privacy

As previously mentioned, privacy encompasses many different aspects and it very much depends on the context. In the early days when privacy started to be considered as a serious matter, concerns were mostly about physical privacy. That is, the right of people to be free from intrusions into one's physical space, property or solitude.

According to Holvast [19], people could be arrested in England for peeping and eavesdropping, as early as 1361. Also, personal correspondence was protected from reading invasions already in 1624. And, since privacy depends on context, by that time, privacy invasions where mostly perpetrated by acquaintances in close contact with the individual, typically from the same town or village.

With the growth in popularity of the newspaper and the more recent invention of photography, the reach and impact of privacy invasions grows. These technologies made it possible to publish information from individuals without their consent and the audience was considerably bigger. Fortunately, this led to an interesting privacy debate that gave rise to the publication of "The Right to Privacy" by Warren

and Brandeis [17], which has been fundamental to the development of privacy laws, mostly in the United States.

As a matter of fact, the emergence of new technologies has been inevitably followed by new ways of invading privacy, which in turn fuelled the privacy debate. For example, the development of telephony led to communication wiretapping and the creation of laws to protect from it, such as the Wiretapping Act in the United States. Other technologies went trough a similar process but it was not until the development of the computer and the widespread use of the Internet that the number and magnitude of privacy invasions reached a whole new dimension.

The ability of computers to collect, store, analyse and disseminate massive amounts of information opened the door to unique opportunities to violate privacy. Information was no longer just local, it could be transmitted and shared with anyone, anywhere in the world almost instantly. Communications exploded in number and size. People started to use their computers, smartphones, and other types of devices to get online. Nowadays, people upload comments, pictures and videos to social networks and the Cloud. All these communications leave traces of what they do, what they like, where they are, and whatnot. These traces can potentially be collected and analysed for different purposes by companies, governments and even criminals.

However, this situation is far from finished. New technologies and paradigms are being developed, such as the Internet of Things [20], which promises to expand the Internet to the physical world by fitting everyday objects with computational, sensing and communication capabilities. In such scenario, the ubiquitous deployment of billions of smart devices will bring countless opportunities to invade privacy. Personal data will be more distributed than ever before, stored in all types of devices, local and remote. People may not even realise of being subject to data collection, who is collecting data and for which purposes. Furthermore, data collection will happen in situations hitherto unsuspected, even in the intimacy of our homes. All this, together with the increase in computing capabilities, advances in data mining and machine learning algorithms (see Figure 8.3) bring unprecedented challenges to privacy protection.

8.3.2 Privacy Protection

There are basically two ways to protect privacy. The first means of protection consists of implementing privacy enhancing technologies (PETs) and privacy-by-design principles. This allows to minimise the collection of personally identifiable information (i.e., data that is linkable to the identity of an individual) and also promote client-side data storage and processing. These approaches aim to anonymise and/or reduce the quality of data before it is released. In this way, the user retains some level of control over the data being offered to third parties. Later in this chapter we will see some PETs applied to digital forensic investigations.

The other means of privacy protection is not technological but legal and regulatory. These are extremely important, especially in situations where the user can be subject to data collection even without taking an active role in the system. This is the case, for example, when a person is in a smart environment surrounded by different types of sensors, cameras, and so on. In situations like these, the user cannot control

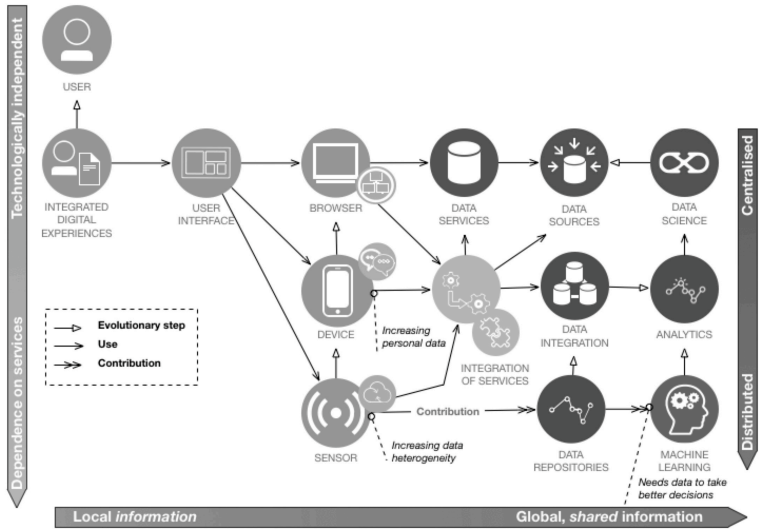

Figure 8.3 Contextual changes for Privacy and Digital Forensics

with technological means the amount of personal information that is being collected. Therefore, to prevent data-hungry entities from invading individual privacy, laws, regulations, audits, and sanctions must be in place.

8.3.2.1 Privacy Laws

The importance of protecting privacy was acknowledged around the globe after a long history of privacy invasions. Privacy is now seen as a fundamental right in the constitution of most countries all over the world [21] and, in 1948, the United Nations recognised privacy in the Universal Declaration of Human Rights [22].

> **Universal Declaration of Human Rights, Article 12:**
> No one shall be subjected to arbitrary interference with his privacy, family, home or correspondence, nor to attacks upon his honour and reputation. Everyone has the right to the protection of the law against such interference or attacks.

Two years later, the Council of Europe created the European Convention on Human Rights [23], which also included the right to privacy. However, there are no universally-accepted privacy laws. Still, there are agreements between countries in relation to the movement of personal data [24], such as the EU-US Privacy Shield and the US-Swiss Safe Harbour Framework.

In most countries, the legal basis for the protection of privacy is defined by constitutional laws. In the United States, the Fourth Amendment of the Constitution protects individuals from unreasonable searches and seizures without a warrant,

which can only be obtained upon probable cause. In other words, individual privacy is protected unless there is sufficient evidence to believe that the person has committed a crime. Like the United States, some countries (e.g., Canada, Germany or Japan) do not explicitly mention the word privacy in their constitutions [21]. In those countries, the courts usually recognise the privacy right as implicit. Other countries like Brazil, South Africa and South Korea directly refer to the inviolability of privacy in their constitutions. Besides constitutional laws, each country has its own specific laws related to privacy protection.

In the United States, privacy provisions in constitutional laws are complemented by statutory privacy laws [25]. The main restrictions to privacy invasions comes from the Electronic Communications Privacy Act (previously the Wiretap Act), the Pen Register Act and the Stored Communications Privacy Act. The first two are related to the protection of all forms of private communications and the meta-data generated from these communications, while the latter regulates access to the information stored by Internet Service Providers. However, it is important to note that when people share information and files with others, they usually lose the reasonable expectation of privacy [26]. Additionally, privacy is considered in tort laws and there is a number of sector-specific laws, such as the Health Insurance and Accountability Act (HIPAA) for the health sector.

On the contrary, in Europe, there is a general framework for the protection of privacy regardless of the sector. The European Data Protection Directive from 1995 (Directive 95/46/EC [27]) was devised as a mechanism to homogenise and unify the privacy laws from different member states. However, being a directive, each member state was free to decide how to transpose its provisions into national laws. In recent years, however, the European Union has been developing a regulation (not a directive) for the protection of personal data. The General Data Protection Regulation 2016/679 (GDPR [28]) has recently superseded Directive 95/46/EC. This regulation not only entered into force in May 2018 in all Member States simultaneously with legal binding but also introduces some notable changes. For example, it introduces bigger fines to organisations not complying the GDPR, which can reach up to 4% of their annual global turnover or 20 Million, whichever is bigger. Another relevant change introduced by the GDPR is that it extends its data protection scope to any organisation processing personal information of European citizens, regardless of its location. Also, the GDPR recognises new rights to data subjects, such as the right to erasure, also known as the right to be forgotten.

Clearly, covering all existing privacy laws is virtually impossible and well beyond the scope of this section. The goal is solely to give a brief overview of some well-known privacy laws. The interested reader is referred to [24] for more details on privacy laws around the world with a special focus on the United States.

8.3.2.2 Privacy Principles

Most privacy laws identify a set of principles that determine the responsibilities of organisations that handle personal data and at the same time shape the rights of individuals. How to successfully comply with these principles depends on each par-

ticular organisation and is not covered by the law but at least these principles define some general guidelines.

The first law to introduce privacy principles was probably the US Privacy Act of 1974 [29]. This law established a set of guidelines to govern the practices of federal agencies in the maintenance, processing and dissemination of personal data records. A few years later, the Organisation for Economic Co-operation and Development (OECD) published 8 principles for the protection of privacy and transborder flows of personal data. These principles, which were revised in 2013 by a group of experts and remained unchanged [30], can be summarised as follows:

Collection limitation: personal data should be collected by lawful means and with the consent of the data subject.
Data quality: Personal data should be accurate, complete and relevant to the purpose for which it was collected.
Purpose specification: The data subject should be aware of the purpose for which personal data is collected not later than at the time of data collection.
Use limitation: Personal data should not be used for purposes other than the ones specified at the time of collection.
Security safeguards: Personal data should be protected against loss, attacks and misuse.
Openness: The data subject must be aware of the policies, procedures and practices of the data holder with regard to personal data.
Individual participation: The data subject must be able to access his/her own data as well as to ask for corrections within reasonable time.
Accountability: The data controller is made responsible for non-compliance with any of the above principles.

Although not all privacy laws consider the same principles, most of them revolve around the same ideas of data minimisation, use limitation, individual participation and access, plus security and accountability. These principles usually change in name or number but not in form. For example, the EU Directive 95/46/EC [27, Chapter II] adopted all these principles and the same happened with the more recent GDPR 2016/679 [28], which references all principles in Article 5 except for individual participation. Notwithstanding, data subject's rights are addressed in part III (articles 15 to 17), which include the rights of access, the right to rectification and the right to be forgotten. All three, can be regarded as sides of the same coin, namely individual participation.

Note that all these laws include some limitations and restrictions to the privacy principles considered in them. These restrictions are included as a mechanism to, among other things, protect national or public security, assure the rights and freedoms of others, or prevent and prosecute criminal activities.

8.3.2.3 Privacy Standards

A technical standard is a document that provides a series of rules, instructions or methods for achieving uniform results across different products or systems. Stan-

dards are mostly developed by standard organisations after a rigorous process involving a number of technical experts in the area.

Some organisations have developed standards with a focus on data protection and privacy in different domains. One of the most prolific standards organisations in this respect are the International Organisation for Standardisation (ISO) and the International Electrotechnical Commission (IEC), with several standards. The devised standards cover several aspects including the establishment of a common framework defining privacy terminology, actors and principles (ISO/IEC 29100:2011), the description of architectural components for systems that process personally identifiable information (ISO/IEC 29101:2013), and the definition of a set of controls for implementing measures to protect personal data (ISO/IEC 27018:2014) in accordance with the privacy principles defined in the aforementioned standard. It is also worth mentioning that a technical committee of experts from the ISO, the ISO/PC 317, is currently working on a new standard (ISO/NP 23485) to ensure the compliance with new regulations during the whole life cycle of products or services.

Other organisations such as the British Standards Institute (BSI), the European Committee for Standardization (CEN) and the National Institute of Standards and Technology (NIST) from U.S. Department of Commerce, have also delved into the protection of personal data and privacy with various standards and recommendations. Some of them are the BSI BS 10012:2009, the CEN CWA 16113:2010, and the NIST 800-122.

8.3.3 Privacy Challenges in Digital Forensics

Considering the main characteristics of the current ecosystem of technologies and the digital forensics challenges summarised in Table 8.1, a list of privacy challenges motivated by digital forensics is provided in Table 8.2.

Some of the papers analysed during this chapter (cf. Sections 8.5-8.7) partially cover some of said challenges. For example, the (unnecessary) privacy exposure of third parties which are not directly related with the digital investigation is widely discussed in Section 8.5.2.3. There are also some challenges that are closely related to each other. For example, the consent for a user does not control that said user stores data about other users; therefore, even with an informed consent, there will be a high risk of third party privacy breach to occur. Besides, informed consents must be simple and clear, and this is increasingly difficult due to the different jurisdictions, laws, standards and the users themselves.

Also, different jurisdictions can affect data privacy. For example, in Cloud computing environments, a privacy-aware digital forensic solution must consider the jurisdiction of the country where the data are stored, and and be able to understand that these data can not be moved to another country that does not meet the same privacy requirements. At the same time, forensic tools in this area should understand these premises.

In addition, there are clear digital forensics and privacy trade-offs in new areas such as Cloud forensics, IoT-Forenics, or more recently, vehicular/automotive forensics, that must be further explored. For example, renting a car with an on-board computer and synchronising our device with it (e.g. to listen to music or to make a

Table 8.2 Digital Forensics Practices and Privacy Concerns

Digital Forensic Procedures	**Privacy Issues**
Indiscriminate acquisition/collection of digital data	**Third Party Privacy** Breach (TPPB) must be avoided
Full disk images are created and analysed	Deleted files can lead to **false accusations**
Data can be collected from personal devices	Need for **informed consents** complete and understandable by the users
During the investigation, the data to be acquired may be hosted on servers in different countries	Different **jurisdictions** can understand privacy differently
Warrants can be necessary during private investigation.	Matching of **privacy policies** and **warrants** (formally defined) for automated analysis.
Correlation is needed in order to build a timeline	**Multi-device** context (more and more data)
Digital forensics tools and methodologies must be accepted and tested by a broad group of experts in the field	Privacy requirements must be integrated by design in **existing tools and methodologies**
Digital forensic principles must be guaranteed	The **manipulation of data** (e.g., encryption) to protect privacy must be done considering digital forensic principles 8.2.2.1

call), nothing guarantees that these preferences will be erased or that our list of contacts will not be recorded in the car. In contrast, the existence of such mechanisms to secure erasure would eliminate digital evidence that can help establish liability in an accident (e.g., if data from accelerometers are removed) or from other events (e.g., the location of the individual's car at the time of a fine as exculpatory evidence). Moreover, in general, multi-device context and/or multi-tenant architectures implies the access of multiple users to the same platform, compromising the privacy of the individuals sharing data in case a digital investigation is required. Besides, relevant information to the investigation can be divided as pieces of a complex puzzle between multiple personal devices. Therefore, when the digital forensics mechanisms are applied (e.g. based on a warrant) surpassing the security measures, these particular cases must be taken into account.

Considering the previous initial list of privacy-aware digital forensics challenges, during the rest of the chapter, solutions that consider privacy in different digital forensics scenarios are analysed in order to identify the degree of satisfaction of these requirements.

8.4 Law, Privacy and Digital Forensics

One of the expected skills of a digital forensic investigator is to understand the laws and comply with them. Laws help to narrow down the scope of forensic investigations. However, it is possible that due to ignorance or imprudence the forensic

investigator exceeds some limits and ruins all the investigation. For that reason some authors have tried to make these limits more evident by presenting an analysis of privacy laws in different countries.

The authors in [26] concentrate on investigations in the United States. They start by introducing the laws restricting forensic investigations, including the Fourth Amendment and several Acts in the United States Code. But their main contribution is on the presentation of situations where investigators need or need not a court order to conduct the investigation. In general, a search warrant/court order/subpoena is necessary to gather the evidence legally. However, if the investigation does not violate a persons reasonable privacy, does not break the law or falls into an exception of law, then the evidence can be legally obtained without search warrant/court order/subpoena and the evidence will be accepted in court.

Another paper that analyses American law is [25]. The authors describe nine legal areas where more research in digital forensics is necessary. Privacy is only considered in two out of the nine areas (constitutional law and tort law) and unfortunately the authors do not give details on how to approach specific privacy issues. A more privacy-focused analysis is provided in [31]. In this paper, the authors concentrate on the relation of privacy laws with forensic tools. In particular, the analysis is based on the reliability of the tools and how they can protect privacy. The authors claim that one desirable requirement for digital forensic tools is the ability to provide individual accountability through logins. This would help to ascertain who was using the tool during the acquisition or analysis of the data. In addition, the authors propose the inclusion of mechanisms to ensure (by design) that the tool is only used for the purpose the search warrant was granted.

Some authors have looked beyond American and have analysed how laws and regulations in Europe and/or Asia-Pacific countries affect digital investigations. This is done, for example, by the authors in [32]. They also present a survey of three areas of research related to privacy and digital investigations: (1) analysis of privacy policies, (2) modelling of privacy policies and (3) technologies for privacy-respecting investigations.

Some relevant changes introduced by the European General Data Protection Regulation with respect to digital investigations are introduced in [33]. For example, the paper discusses about the influence of the new regulation on the legal proceedings for e-discovery. They observe that with the GDPR, there will be problems in the way cross-border litigations were performed before since data was typically collected on-site and sent to a central e-discovery provider from where lawyers from different countries could access the data. This may no longer be possible with the new data protection regulation. However, the authors conclude that the impact of the GDPR on digital investigations is still unclear.

Finally, the authors in [34] present an state of the art analysis of legal aspects regarding security, privacy and digital forensics in Future Internet scenarios like Smart Cities and the Internet of Things. After reviewing some relevant pieces of legislation in various countries and major cities adopting these novel technologies (Hong Kong, South Korea, Budapest, USA and Europe) the authors observe some of them lack a solid legal framework for data protection. Being the European GDPR the best

structured piece of legislation. The authors also recognise the need for international agreements and cooperation towards a common security and privacy framework, that may be paved by GDPR mandates on protecting transnational flows of data related to European citizens.

8.5 Privacy-Aware Computer Forensics

This section is divided in two broad areas where most of the research on privacy solutions has concentrated over the years, namely database and computer forensics. Very few works consider privacy and memory forensics [35, 36, 6]. These works are mostly devoted to compromising data privacy when it is decrypted in memory and for that reason they are not included here.

8.5.1 Database

As mentioned at the beginning of the chapter, the digital forensic discipline emerged as a problem closely related to data recovery and databases [1]. Privacy problems in the database forensics have been considered from the following points of view: i) the right to protect the access to data, ii) to ensure the secure erasure of data, and iii) to protect the data of honest users during post-mortem investigations towards the definition of specific frameworks for data storage. See Figure 8.4 for a visual summary of privacy-aware digital forensics (PADF) solutions in this area.

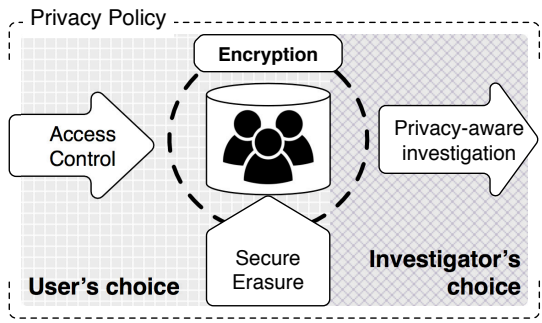

Figure 8.4 PADF Database approaches

In general, most of the solutions in this section use encryption to restrict the access to data, that can be stored in multiple devices or be used for different purposes, as shown in Figure 8.5.

Nowadays, there are database solutions for both servers and mobile devices, and data protection covers not only large servers, but also storage devices that can connect to any computer. In this ecosystem, there are personal devices, used (theoretically) by a single user and devices shared by multiple users or multi-tenant architectures. All these systems needs data to be stored following certain criteria.

In the following sections the contributions in the area of privacy-aware database forensics are analysed. However, it is very important to keep in mind that some of

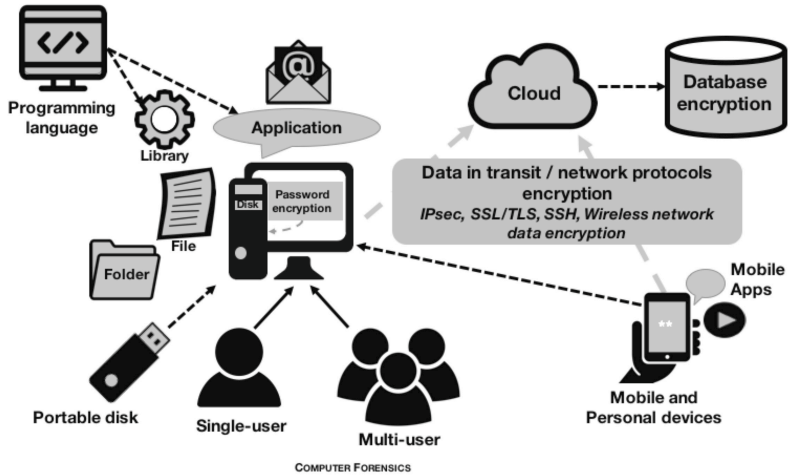

Figure 8.5 Encryption to protect different scenarios

these solutions are also applied to other areas because it is part of the evolution of digital forensics as described in Figure 8.3.

8.5.1.1 Access Control

Access control has been used for two main purposes in digital investigations: as a security mechanism for preventing unauthorised access to data and as a mechanism to control the whole digital forensic investigation. For example, in [37] access control mechanisms are developed to protect individuals' privacy in DNA databases. In case the database is accessed by a non-authorised entity the contents will be unintelligible. Only law agencies officers are capable of accessing this information. The proposed solution uses encryption to ensure that only legitimate queries on the database are allowed. The key to access to the identity of an individual in the database is the result of a set of DNA tests from the specific individual, following a shared secret approach. In this way, only those in possession of the DNA tests can get access to the data.

Note that, in the previous work the contribution is focused on access control as a mechanism to limit access to personal data within a forensics database populated by and belonging to forensic experts. This is different from controlling access to a database containing data stored by a system about individuals (or stored by them). Two solutions [38, 39] can be found to support the latter case. These solutions are aimed at controlling the data that can be accessed by investigators when performing and investigation. In both solutions, data needs to be first categorised in different sensitivity levels.

In [38], the data are later encrypted in such a way that the forensics investigator only gains access to a more privacy-sensitive level once he proves knowledge of data in the level immediately below. In particular, the forensics investigator queries the data controller with a cryptographically blinded hypothesis. The hypothesis is

basically the hash of a predicate regarding the investigation. After checking that the investigator has not exceeded the number of hypothesis test requests for that particular sensitivity level, the data controller returns a message that can be used to unlock the key of the next sensitivity level if the hypothesis of the investigator was correct. However, there are some open issues that should be clarified in future contributions. For example, who is the data controller and when is the categorisation and encryption of data done. This is extremely important from a privacy point of view since the data controller will have access to all the data thus moving the privacy problem from the forensics investigator to the data controller. It is also unclear how these categorisation is done and how accurate the hypotheses need to be.

Similarly, in [39] privacy levels are defined based on a previous classification of data, which is made considering all possible accesses to the system. In this approach both the user and the investigator classify the data. The user chooses between private or not-private and the investigator determines if the data is relevant or not-relevant, based on the goal and the scope of the investigation. Data classified as private and not-relevant is not collected, and data that is private and relevant is subject to the user's choice. In this case, the user can decide whether his/her data will be collected as is or encrypted. The problem of this approach is that by encrypting data, the user alters the digital evidence and therefore could invalidate the digital investigation. Another problem is that the user is in full control and the role of the investigator is extremely limited by the user. This may be in line with privacy principles but not with some of the restrictions included in privacy laws.

8.5.1.2 Secure Erasure

Completely eliminating a data set is not as simple as it may seem. The aim of secure erasure techniques is to make sure that data can not be recovered by any means, and this is the reason why this anti-forensic technique has also been considered as a mechanism to protect privacy. Nowadays it is increasingly common to include secure erasure as a requirement in the security policy of any organisation that manages digital data, even more so after the entry into force of the GDPR.

Deleted data passes through several phases before it is finally removed from the system and making data unrecoverable in the context of database systems is even more challenging. The database system usually makes multiple copies of sensitive data in transaction logs, indexes, etc. that may help to recover data. Based on these findings, the authors in [40] analyse four common database systems, including PostgreSQL and MySQL, and reveal they are vulnerable to some of these data leakage problems. They also propose some changes to the storage manager in MySQL for securely removing deleted data. Basically, the changes consist of calling to the *memset*() operation to overwrite those records which are considered no longer necessary. According to the authors, with careful configuration this imposes insignificant degradation of system performance. On the other hand, they propose to encrypt log records from transaction logs with different keys and simply delete the keys when these log records are no longer necessary.

Some other contributions are focused on the analysis of different tools for secure erasure more closely related to the operating system and thus they are analysed in the following section, which concentrates on computer forensics.

8.5.2 Computer

Typically the terms digital forensics and computer forensics have been used interchangeably. Perhaps one of the main reasons is that, for a long time, computer forensics was the area that covered everything necessary in this discipline. As detailed in Section 8.2, numerous contributions have been developed in this area, and, as it could not be otherwise, also contributions regarding privacy. Being computer forensics the propeller of digital forensics, this explains to some extent why most of the privacy-aware contributions analysed have been made in the area of computer forensics. The topics in this area are quite similar to privacy-aware database forensics but tends to be more complex due the heterogeneity of computer architectures, file systems and operating systems available.

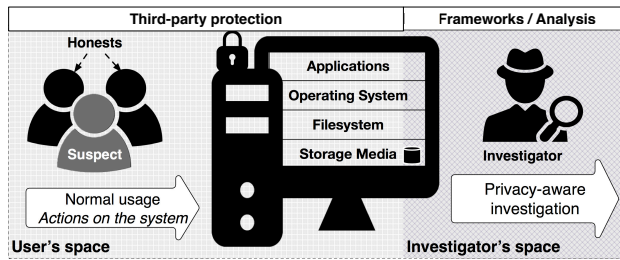

Figure 8.6 PADF Computer approaches

This is the reason why some contributions in this area focus on particular technologies, contexts or platforms. Also some authors have surveyed the field of privacy-aware computer forensics [41, 42, 43]. Although they usually provide different ways of classifying existing solutions, they generally reach similar conclusions.

8.5.2.1 Frameworks and Policies

It is extremely complex to control all the factors that may affect privacy during a digital forensic investigation without turning to a framework that guides the investigator throughout the process and gives recommendations on the different issues. For that reason, the PET (Privacy Enhancing Technology) framework is proposed in [44] to protect the privacy of honest users during a post-mortem digital forensic investigation. Similarly to [37], the solution controls the requests to the databases. The paper is extended in [3], where the concept of *Third Party Privacy Breach* (TPPB) is properly defined in the context of computer forensics. This will be further analysed in Section 8.5.2.3.

The high-level framework proposed in [45] allows enterprises to effectively conduct digital forensic investigations of privacy incidents. The authors extend a general forensic framework to incorporate privacy-related components in the auditing

and monitoring of business processes. In particular, the privacy-specific business processes introduced in the framework are borrowed from the *Generally Accepted Privacy Practices* (GAPP) standard. Similarly, the privacy-specific business policies are based on the *Fair Information Privacy Principles* (FIPS). It is important to highlight that inside an organisation the type of investigations are private, which may be less restrictive than public ones.

The framework presented in [46] is based on three modules: i) expert system, ii) evidence extraction and iii) ranking. The module for evidence extraction will collect the digital evidence and will be then processed by the expert system. This module decides whether the digital evidence is relevant or not to the case so that it can be processed by the investigator. The decision of the expert system is based on previous investigations that are considered to be similar to the current one. Note that, the solution is very dependent on the learning phase that, as highlighted in previous papers, is critical since i) it depends on the quantity and quality of the previous investigations and ii) it represents a privacy issue by itself because it needs access to the case data.

Finally, some authors consider policy-based solutions for preserving privacy in computer forensics. Privacy policies are used to guide the data treatment process. In this area, [47] proposes a set of privacy policies for guiding the investigator throughout the various phases of a forensic investigation (identification, collection, preservation, analysis and presentation).

8.5.2.2 Secure Erasure

In [48] six counter-forensic privacy tools for Windows operating systems are analysed. The authors identify a number of limitations on these tools, which fail to provide a sufficient level of protection to users that wanted to delete sensitive information from their computers. According to the authors, the main problem with these tools is that it is extremely difficult for them to keep up with the number of ways different applications manage data and interact with the operating system. In case the user needs stronger privacy guarantees, the authors point to alternatives to the analysed counter-forensic tools, namely, disk encryption and booting from a CD with all disks removed from the computer.

Similar conclusions are reached in [49]. In this case, six privacy software packages are evaluated based on the amount of information recoverable after they are used. Some of these solutions are intended for secure erasure, either using wipe or deleting specific files. The results show that there are tools that don't purge the unallocated space (in case of wipe a disk) and also that these tools tend to leave some trace of the deleted files - stored by the operating system. Besides, they also evaluate the effectiveness of the tools in erasing targeted user and operating system. As in the previous case, the operating system and the applications create files that are obviated by the evaluated software.

To conclude, in [50] the authors highlight the need to consider different types of evidence since that most of the works on privacy protection in digital forensics concentrate on emails or documents. Moreover, they criticise the incompatibility of the proposed solutions with existing software forensic tools. This is an important

issue, given that new solutions are more difficult to be accepted in court without a clear acceptance (and training) by the experts in the field.

8.5.2.3 Third Party Privacy Breach

A *third party* is usually understood as an entity that is not the main actor (or interested party) in the scene/context/protocol. In the contributions analysed next, a third party is equivalent to the concept of *honest user*. Unfortunately, honest users are very difficult to distinguish from malicious users. The protection of honest users is also a field of study in network forensics (see Section 8.6).

In the area of computer forensics, the problem of protecting honest third parties is defined as a *third party privacy breach* (TPPB) [44]: "the event that a third party, not culpable in the actions leading to the investigation, may be investigated". This is possible, for example, when pattern matching techniques return data about multiple users that must be analysed by the investigator, although they are finally discarded if they are not relevant.

To prevent the aforementioned problem, the same authors [51] propose a solution that forces the investigator to issue focused (i.e., specific) queries in order to get results from the system. The solution consists of various components, one of which is in charge of categorising documents based on their similarity, using n-grams and distance measures. Then, a response filter component is responsible for determining whether the result of a query may lead to a privacy violation based on the distance of results (i.e, being very different). In that case, the query is logged and the investigator is suggested to be provide a more specific query. This type of solutions do not necessarily assume a dishonest investigator. They prevent this threat but also reduce the risk of an honest investigator violating the privacy of third parties unwittingly. Moreover, they can be useful for reducing the amount of irrelevant data that needs to be processed by the investigation thus significantly reducing the time to conclude an investigation.

The authors in [52] focus on the problem of discerning honest users from inside attackers. Their goal is to determine the type of data that helps to catch this type of attackers while complying with data protection laws. They argue that the analysis of data available from physical security systems (e.g., biometrical access control systems) can be crucial to delimit the investigation. Thus, they suggest the data from these systems to be anonymised and only when applying anomaly detection analysis these data can be partially de-anonymised. This would allow to track down insider attacks without revealing the identity of honest users. The problem is that once data is partially de-anonymised it may be possible to identify some particular users.

8.6 Privacy-Aware Network Forensics

This section considers solutions for the protection of privacy in three main domains related to networked systems: servers, networks, and browsers. Note that some of the solutions analysed here could fall into some other categories. For example, privacy solutions for server forensics could fall into the category of computer forensics since a servers is, in essence, a computer.

8.6.1 Server

A servers is a specific-purpose computer optimised to provide a set of services to remote users through the network. Most papers in this area consider web and e-mail servers as case studies. More specifically, web data and e-mails are used to test numerous forensic tools used for pattern recognition (e.g., using keywords). There are two main type of contributions in this area: i) revocable anonymity and ii) searchable encryption. In these cases, the digital investigation is focused on server data, as shown in Figure 8.7.

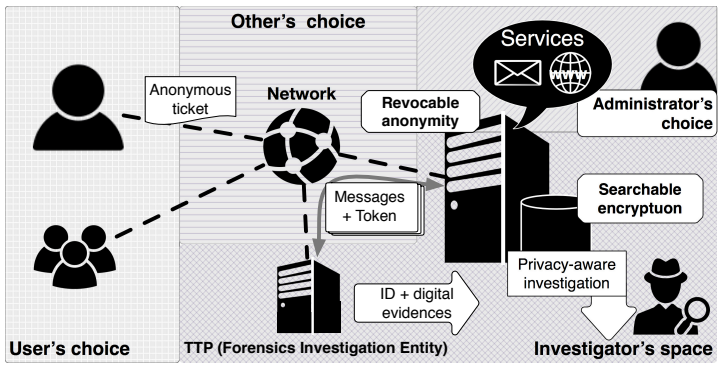

Figure 8.7 PADF Server approaches

8.6.1.1 Revocable anonymity

Revocable anonymity refers to the process of allowing users to operate or receive services anonymously unless they misbehave, in which case they can be re-identified, typically with the help of a trusted third party.

This is the type of approach followed by PPINA (Protect Private Information, Not Abuser) [53]. The idea is to allow users to connect to the server through an anonymous communication network but before doing so, they have to generate a public/private key pair and an access token. The access token is cryptographically linked to the public key of the user. Then, the token is sent to a trusted third party (the Forensics Investigation Entity) that verifies the validity of the token and stores it with the identity of the user. The server also receives a copy of this token signed by the TTP. The server verifies the token without knowing the identity of user, who establishes a connection to it through an anonymity network. The server also stores the token and all messages signed with the key corresponding to the token. In case the user misbehaves, the server sends all the messages and the token to the TTP for it to decide whether an attack occurred or not and if positive reveal the identity of the user.

A similar approach is followed by the ERPINA protocol [54]. This protocol is intended to respect both the desire of the user to remain anonymous while accessing a server and the right of the server to know the actual identity of the user if he misbehaves. The main difference with the previous solution is that, in this case,

the anonymous ticket obtained by the user embeds a policy of use. The ticket is sent by the user through an anonymous communication network and has to be validated by the server before granting access to the service. If the server considers the anonymous user is not following the rules defined by the policy of use, it sends the ticket together with the policy to a trusted third party, which after reviewing the case, decides whether to reveal the identity of the user or not.

8.6.1.2 Searchable encryption

Searchable encryption is a cryptograhpic technique that allows to issue queries to an encrypted database. This is a promising technique to protect user privacy while conducting digital investigations.

A searchable encryption scheme is used in [55] to allow the forensics investigator to query for data matching some specific keywords in the context of e-mail servers. First, the disk image is analysed and an index file matching keywords to files (or sectors in the disk image) is generated. The index file and the image are encrypted at this point. Then, the forensics investigator generates a list of keywords which are relevant to the investigation and passes the list to the data owner. The data owner can then generate a trapdoor, which is a data structure that allows the investigator to search the encrypted index file for a given keyword or a set thereof. Then the investigator can ask the data owner for that specific location of file in the disk interactively. A notable limitation of this scheme is that the data owner can potentially hide information to the investigator since he/she is in charge of creating the index file, the trapdoor and decrypting the files. Another limitation is that the data owner gains sensitive information about the case from the keywords provided by the investigator to obtain the trapdoor.

The solution proposed in [56] aims to prevent the server administrator from learning what is the investigator looking for. To that end, the investigator generates a public/private key pair and shares the public key with the administrator. Then, the administrator divides the documents into keywords and encrypts them with the public key provided. The investigator encrypts its n keywords with the same public key and transforms them into a polynomial of degree n. After that, the investigator sends the administrator the coefficients of that polynomial to hide the actual encrypted keywords. Finally, the server administrator makes a similar transformation of the files into coefficients and using the coefficients from the investigator, the administrator can determine which files contained keywords of interest to the investigator. An important point that is not sufficiently discussed in the paper is how can the administrator prevent the investigator from asking for keywords which are irrelevant to the investigation. Also, it seems that if the number of keywords of interest to the investigator is small, the administrator con more easily brute-force the coefficients and obtain the keywords. This is in contradiction with user privacy preservation since the investigator should ask only for the minimum amount of information that allows him or her to close the case.

Finally, in [57] it is provided a searchable encryption scheme with the following features: (a) keyword search is non-interactive, meaning that the forensic investigator does not need to contact the data owner every time it wants to issue a query, and

(b) the data owner remains oblivious to the queries (keywords) of interest to the investigator. The data owner determines which are the keywords that can be queried for and establish a threshold of t keywords that need to be present in the file for disclosing it. Very basic keywords, such as pronouns, can be excluded to prevent trivial attacks. The approach is based on Shamirs secret sharing scheme, that is, the key for decrypting a file is constructed based on the keywords in that file and t of such keywords reveal the key. The authors acknowledge two main limitations to the scheme. First, the need to have exact keyword matches for searching. Second, there is the possibility that the investigator performs brute-force/dictionary attacks on keywords to reveal the key. In addition to these limitations, there is the problem of a potentially malicious data owner who wants to limit the ability of the investigator retrieving data. This would be as simple as blacklisting some keywords.

8.6.2 Networks

Unlike previous sections, the contributions analysed here concentrate mainly on how to protect data in transit that must be monitored - either by the network equipment or by the investigator - and lately analysed by the investigator.

Interestingly, the first papers in this area were focused on advising practitioners on the best way of performing monitoring actions so as not to have legal repercussions [58], while the most articles have a wider awareness of privacy. Therefore, there is tendency towards recognising user's rights and the potential impact on the privacy of honest third parties using the same communication channel.

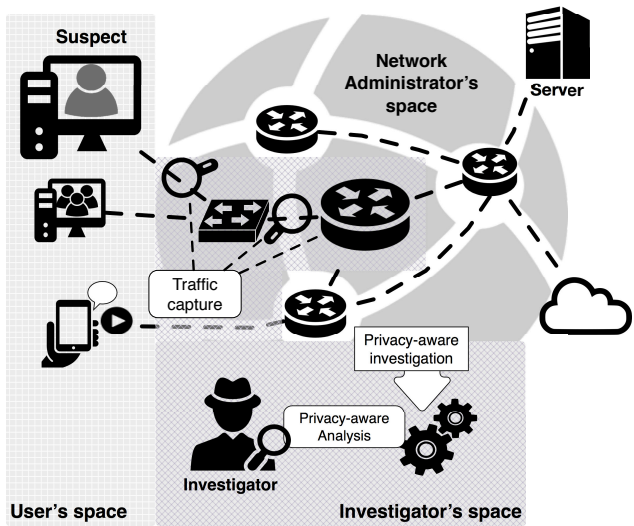

Figure 8.8 Privacy-aware Network Forensic approaches

Being the network a shared medium where multiple individuals exchange information, it becomes a great source of evidence and at the same time a great problem

for privacy. Figure 8.8 describes the scope of this section. It includes privacy solutions for both the process of traffic capture and traffic analysis.

8.6.2.1 Traffic Capture

A solution to network flow recording with partial privacy guarantees is presented in [59]. The authors propose to divide captured network traffic into files containing 5 minutes of network flows from a single IP. Each of these files are read separately and statistics are extracted from them and imported into a database. After this, each file is encrypted with a random AES key and all these keys are IBE-encrypted with the IP-timestamp as public key encryption. Plain text files are deleted form the system. The result is short encrypted files (5 minutes of activity) so that when law enforcement requires data they can be provided with all files pertaining to a specific period rather than all the data. Moreover, the authors propose to use a secret sharing approach to divide the IBE key used for decryption of the files into several shares so that files can only be decrypted if all key holders agree to do so.

The authors in [59] also take care when populating the statistical database to prevent sensitive information from being revealed when querying it. This is enforced by replying to queries only when the result satisfies some privacy conditions. For example, requiring a minimum number of bytes being transmitted to prevent website fingerprinting (i.e., recognising the website being accessed by a user based on the bytes transferred).

Tools are an important part in traffic capture. In this respect, Carnivore was a packet sniffer that allowed the use of filters to capture traffic only from a particular individual, instead collecting all network traffic. This was a tool after which there was controversy because it was secretly used in secret by the FBI from 1999 to 2005. Carnivore was made public in 2000 and several failures were discovered. This caused it to be replaced by commercial products. This tool allowed restricting the capture of traffic, so it was considered that when properly used could help to protect honest users' privacy [31]. Nowadays, most traffic capture tools enable the use of filters.

8.6.2.2 Traffic Analysis

The first approach we are going to discuss here is based in searchable encryption, which has also been used in other contexts (cf. Section 8.6.1.2). In this case, the authors propose a scheme [60] based on bilinear pairings that is intended to allow the investigator to collect evidences from network traffic under the assumption that an attack have been perpetrated. The idea is to analyse the traffic without revealing the identity of the potential attacker until there is sufficient evidence. The paper focuses on evaluating the efficiency of searchable encryption scheme and it is not clear how it can be applied to a real digital forensic investigation.

A network-layer capability named "privacy-preserving forensic attribution" is proposed in [61] to protect the privacy of users while ensuring data traceability. Using a packet-level cyrptographic signature mechanism, a packet is self-identifying and linked to the physical machine that sent it. Privacy is considered as a requirement to avoid non-authorised entities from examining the packets. Another requirement is that packet signatures must be non-identifying; two packets sent by the same source

must carry different signatures. The solution uses a hardware chip that is considered a trusted third party if a single third party is required but considers a secret sharing approach otherwise.

8.6.3 Browser and Applications

The actions a user performs on the Internet leaves a data trace in the user's device, which can be analysed in the context of Internet Forensics (cf. Section 8.2). In particular, the trail left by the user when using a web browser has been analysed in various articles [62, 63, 64]. Also, as applications become more Internet-dependant to operate, new challenges arise.

Applications (the browsers among them) have changed the communication habits of the user in the network. They introduce new ways of interacting with other users. The volume of data stored in user devices increases as the applications become more dependent on users' data. Data is usually privacy sensitive as it is related to user relationships, and may include locations, photos, chats or other types of data. Also, the way the user interacts with the application may reveal sensitive information, including the identity, mood, etcetera. This led to emergence of the term BRAP (Browser and Application) forensics [65].

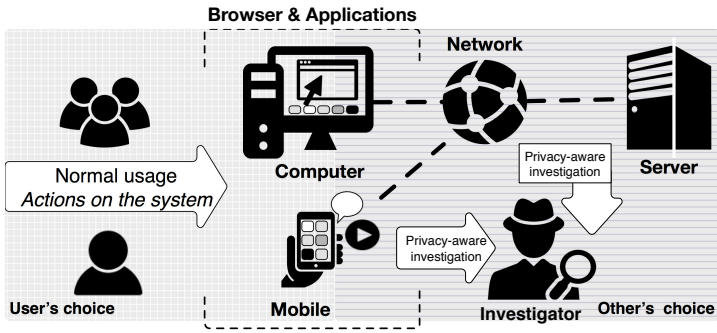

Figure 8.9 BRAP Approaches

Due to the fact that BRAP-related problems are usually analysed as part of computer forensics, the contribution in the area of BRAP forensics is rather limited or dispersed among different areas, especially in the context of mobile forensics [66]. Figure 8.9 shows the focus of BRAP forensics and its relationship with other forensic disciplines.

Most of the approaches that consider privacy in BRAP forensics are closely related to data storage. For example, [67] presents a solution to identify the keywords that enable the inference of topics which are relevant to the investigation. After performing some data mining experiments on web data and e-mails, the forensic investigators could unintentionally extract sensitive data if their tools fail to associate the keywords to the topics of interest. This problem grows with the amount of data to be processed and will be exacerbated when the data is shared among different organisations.

The challenges of BRAP forensics are detailed in [65], where privacy issues are highlighted. The reason for this is that the concept of BRAP forensics is beyond the analysis of logs. The purpose of novel applications is to learn as much information about the user as possible to be much more functional and adapt to user needs. How and where these data are stored depend on each software developer. BRAP forensics will be affected by this and will differ from application to application.

Intuitively, the widespread use of the novel applications raises numerous privacy issues. Therefore, it is worth trying to achieve the same functionality without storing too much personal information. Also, as stated in [65] it is more effective to protect privacy by not storing data than by encrypting them. However, even without storing data our information can be deduced based on our relationships with other individuals (e.g., appearing with a friend in a picture). Doubtlessly, this is a challenging problem.

Unfortunately, there are currently no comprehensive privacy-respecting forencis solutions for BRAP, beyond the generic frameworks that are not directly applicable to specific applications.

8.7 Beyond Computer and Network Forensics

This section describes the contributions in areas other than computer and network forensics. Unlike previous sections, the following analysis shows an evolution from the traditional concept of computer forensics. Starting with mobile forensics, which required the development of specialised tools, the same happened with the Cloud or the Internet of Things.

Basically, these are new contexts that are being analysed by the scientific community and for which specific privacy challenges are envisioned.

8.7.1 Mobile

Mobile forensics is probably the most mature discipline considered in this section. Intrusions to privacy started to receive attention in this context since mobile phones became an integral part of our lives. However, this area considers also other types of mobile personal devices. Indeed, personal devices have motivated various papers in this field and have also strengthened other areas, such as BRAP forensics (cf. Section 8.6.3).

A common approach in this area is to display a banner to inform the user about privacy expectations and garnet their consent. This is a solution followed in DroidWatch [68] and Digital Witness [69] for mobile phones. However, this approach does not protect privacy, it only informs the user about potential privacy problems in some situations.

In [66] the authors analyse existing solutions and methodologies to perform privacy assessments of mobile applications based on how data is stored by mobile applications. The authors follow a forensic methodology to check the information stored by the applications on Android devices. They use the Android Device Bridge (adb)

for data acquisition and typical unix command-line tools (sqlite3, hexdump, tree and strings) for data analysis.

The authors in [70] also analyse Android devices but in this case they focus on the possibility of recovering authentication credentials from volatile memory. One of the observations made by the authors is that password managers can be compromised if the attacker has physical access to the device. Although this problem is independent of the user, in some cases privacy problems are due to user behaviour. This is precisely the goal of the research conducted in [71], to show how user decisions affect his/her own privacy in the context of a mobile platform. In this case, the authors use the mobile forensic tools to help users understand how their behaviour affect their privacy.

Finally, it is worth noting that the area of mobile computing is closely related the Cloud and the Internet of Things. Mobile devices generate huge amounts of data and due to memory limitations has to be outsourced to the Cloud. Moreover, mobile devices can serve as gateways or user interfaces to IoT devices and the data produced by them can be also relayed to the Cloud. In addition, the Cloud is used as intermediary in the communication among different devices.

8.7.2 Cloud

To the best of our knowledge, there are not many papers on the topic of privacy-aware cloud forensics. At the time of writing there are basically two solutions both of which provide cryptographic techniques to limit access to data and resources in the Cloud.

A scheme based on secret sharing and message authentication codes is proposed in [72] to provide a robust logging of Cloud events for forensic investigations. According to the authors, in Cloud environments there is one (or more) logging servers that collect logs from all attached servers. The data to be written to the log file is accompanied by a message authentication code creating a chain to prevent an attacker from deleting events without being detected. To further complicate the task of the attacker, each event of the log is divided into n shares and distributed into random computers of the Cloud. Finally, they also propose to record these events in an immutable database, meaning a database that not even the system administrator can modify. Thus, making it easier for the investigator retrieve evidence.

The paper [73] aims to provide a cloud-forensic solution that minimises the number of virtual machines to be investigated. The proposed method is based on a set of inputs which define the historical activity data for the virtual machine (e.g., network logs, CPU usage) and an array of characteristics of the investigation. The virtual machines that do not match the previous requirements are removed from the search space of the investigation. The data is protected using anti-forensic techniques (both memory and storage are encrypted) but if a security breach occurs, it is still feasible to conduct an investigation using statistical techniques.

Finally, [74] describes some basic security and privacy problems in edge and fog computing. The authors argue that new laws on data protection, and in particular the European GDPR, may require service providers to delete data they collected and are no longer necessary. However, how to secure erase data and the challenges

associated to do so in a highly distributed environment like the one envisioned by edge and fog computing is not described in the paper.

8.7.3 Internet of Things

Although IoT-Forensics is becoming a promising area of research (cf. Section 8.2), there are very few contributions that consider privacy in this context.

One of the most recent contributions in privacy-aware IoT-Forensics is [15]. This paper is based on a previous contribution by the same authors, in which the Privacy-aware IoT-Forensic (PRoFIT) model is proposed [75]. In addition, the authors define a methodology to integrate privacy properties in accordance with ISO/IEC 29100:2011 throughout the phases of a digital forensic model adapted to the IoT. Unlike previous approaches, the methodology encourages users to collaborate as witnesses in a digital investigation by sharing their digital evidence. The methodology allows the users to collaborate voluntarily with full control on the data they provide to the investigator.

Similarly, the concept of anonymous digital witnessing in IoT environments is defined in [76]. This type of solution is interesting to promote the cooperative approaches in the context of a digital investigation in the IoT. However, the entire process should be part of a reference model or framework in order to ensure consistency, facilitate the traceability of evidence, documentation and identify possible mistakes. These are some of the reasons why PRoFIT is used in [15] to adapt an already defined IoT-Forensic solution, the digital witness, to be respectful with privacy. In this article, two case studies were provided to help understand the convergence between privacy and digital forensics in new, challenging scenarios.

In the first scenario, the user in possession of a PRoFIT-compliant digital witness (named Bob) initiates a digital investigation using his device during a dinner in a smart restaurant equipped with IoT devices owned by the restaurant (e.g., smart oven, smart windows) and by clients (e.g., smart watch). Both personal and non-personal IoT devices coexist in the restaurant. To conduct the digital investigation Bob's device must inform Bob about the digital forensics procedures that will be necessary to acquire its own digital evidences from the environment. Also, Bob's device must ask for collaboration to the rest of IoT devices in the restaurant that are handled by a responsible (the Maître), who must agree with the requests. Furthermore, the Maître will will have the right to know the status of the data that has been provided to the digital investigator and can request these rights to be withdrawn at any time. In this case of study, it is possible to determine the source of malicious software that could have affected more clients in the restaurant. The person in charge of the restaurant, the Maître, collaborates in the investigation and this allows to identify the origin of the problem and also prevents future incidents. In the second scenario, the user who initiates a digital investigation is a police officer (called Max). Max has to search a warehouse where there are several IoT devices (e.g., cameras and temperature sensors). In this case, Max's device contains a search warrant and is able to conduct the digital investigation without asking for the voluntary cooperation of other personal IoT devices. This resembles current proceedings, when there is a reasonable ground to suppose that a charge of criminal conduct is well-founded

and thus the approach does not consider privacy preferences or policies. As a matter of fact, it is important to understand the context in where the digital investigation is conducted and the actors involved in it. Otherwise, it will be very difficult to promote solutions flexible enough to be adopted by the different experts.

In addition, there are other solutions that could be considered within the context of privacy-aware IoT-forensics. For example, Themis is an architecture aimed to acquire data from sensors in smartphones taking into account privacy requirements [77]. The solution is focused on the platform - the mobile phone - from which the data will be collected. The user is notified about the collection of digital evidence but this does not protect the privacy of the possible third-party data contained in the device.

In general, current IoT-Forensics solutions do not consider third-party privacy notifications because most of the solutions are not cooperative. This is not necessarily wrong, it just means that they were designed for another purpose. For example, the authors in [78] argue that privacy in the Home IoT context "may not necessarily equate to expectations of privacy in social networks". However, the cooperation of individuals and their personal devices may be increasingly necessary in digital investigations. This can be pretty similar to a social network in the sense that many individuals can be involved in the same digital investigation.

Finally, smart vehicles can also be considered part of the Internet of Things (cf. Section 8.2). In this area, [12] provides a high-level description of the implementation of a mobile app aimed to give the driver control on the parameters collected by the car's event data recorder. Basically, the app is allowed to connect to the car's internal network (after authentication) and collect event data as a backup mechanism. Consequently, the user has more control on the data collected by the car and its status. The data collected by the app can also be backed up in the Cloud. During a case, the forensics investigator has the option to retrieve the data either from the user, the car or the cloud. Thus, the investigator can check the consistency of the data and the user can at least know which data is being accessed by the investigator.

8.8 Conclusions and Final Remarks

Digital forensics is an evolutionary discipline but with solid legal and ethical principles. As part of this evolution, privacy has taken an increasingly relevant role. Initially, privacy mechanisms (e.g., encryption, secure erasure) were considered anti-forensic. Over the last few years, the new digital forensics disciplines (e.g., IoT-Forensics) tend to consider the need for a symbiosis between both disciplines. Probably because the contexts are increasingly user-centric.

Table 8.3 summarises the results of the analysis made in this chapter. The synthesis of contributions in the area of digital forensics and privacy trade-offs shows that, although privacy solutions have been devised for almost all the topics shown in Figure 8.3, it is not broadly considered in digital forensic scenarios. For example, it is common to examine privacy as a legal requirement and make use of access control solutions, sometimes using cryptographic techniques, to protect access to data. Note that techniques such as revocable anonymity or searchable encryption are used in

unrestricted environments but IoT-based approaches are not considering (yet) these solutions. In part due to the current hardware limitation of devices but also because the challenges in privacy-aware IoT-forensics are different [15], being more focused on the deployment of cooperative approaches in common frameworks and methodologies.

Table 8.3 Level of accomplishment of Privacy-aware Digital Forensics

Context	**Privacy-aware Digital Forensics solutions**					
	Privacy levels, policies and filters	Frameworks, models	Revocable anonymity	Searchable encryption	Secret sharing	User consent
Database	*A*	*NI*	*NI*	*NI*	*N*	*P*
Computer	*A*	*A*	*PA*	*NI*	*N*	*P*
Server	*NI*	*NI*	*NI*	*A*	*N*	*NI*
Networks	*PA*	*A*	*N*	*A*	*A*	*N*
Applications	*N*	*B*	*N*	*NI*	*N*	*A*
Mobile	*N*	*N*	*N*	*N*	*N*	*A*
Cloud	*PA*	*N*	*NI*	*NI*	*A*	*PA*
IoT	*PA*	*A*	*N*	*N*	*N*	*PA*

A:Addressed, *P*:Partially addressed, *B*:Barely addressed, *N*:Not addressed, *NI*:Not addressed but can be learned based on the experience of another context.

It is also possible to check in Table 8.3 which solutions or techniques can be re-used or adapted to other digital forensic areas (e.g., it can be understood that solutions for computer forensics can be adapted adapted to servers). This adaptation is not always possible or necessary, though. Intuitively, this is either because of the lack of resources or because the new environments have other requirements. Besides, it is possible that open challenges will require very specific solutions for that area. For example, typical privacy user consents will not be directly applicable to Cloud computing because there are additional issues regarding the jurisdictions that must be considered (cf. Section 8.2.3).

Furthermore, note that, in the current ecosystem, not only personal devices but also intermediary platforms, such as the Cloud, have become core elements of forensic investigations. Not all user data is kept in a single location, instead they are stored on third party platforms and shared with other users around the globe. This dispersion of data also leads to possible problems because a common legal framework is lacking for all countries. For example, Europe has the GPDR but the United States has its own laws. Therefore, privacy-aware digital forensic mechanisms for Europe will be presumably different from those developed in USA.

Doubtlessly, information systems will continue to store increasing amounts of information about users. Nevertheless, we should not fall into the mistake of simplifying privacy-aware digital forensics issues to something that governments, organisations or users must resolve by themselves. This is a convoluted problem that can

only be addressed if all actors have a common understanding of the problem and they are willing to pay the price of the change.

References

[1] Garfinkel SL. Digital forensics research: The next 10 years. Digital Investigation. 2010;7:S64 – S73. The Proceedings of the Tenth Annual DFRWS Conference. Available from: http://www.sciencedirect.com/science/article/pii/S1742287610000368.

[2] Caviglione L, Wendzel S, Mazurczyk W. The Future of Digital Forensics: Challenges and the Road Ahead. IEEE Security & Privacy. 2017 November/December;15(6):12–17. Available from: doi.ieeecomputersociety.org/10.1109/MSP.2017.4251117.

[3] van Staden WJ. An investigation into reducing third party privacy breaches during the investigation of cybercrime. In: 2014 Information Security for South Africa; 2014. p. 1–6.

[4] Internet Crime Complaint Center. IC3 2017 Internet crime report; 2018. Available from: https://pdf.ic3.gov/2017_IC3Report.pdf.

[5] Chon KH. Cybercrime precursors: Towards a model of offender resources. The Australian National University; 2016. PhD Thesis.

[6] Thing VL, Ng KY, Chang EC. Live memory forensics of mobile phones. digital investigation. 2010;7:S74–S82.

[7] Willassen S. Forensic analysis of mobile phone internal memory. In: IFIP International Conference on Digital Forensics. Springer; 2005. p. 191–204.

[8] Meghanathan N, Allam SR, Moore LA. Tools and techniques for network forensics. International Journal of Network Security Its Applications (IJNSA). 2009;1(1).

[9] Oriwoh E, Jazani D, Epiphaniou G, et al. Internet of things forensics: Challenges and approaches. In: Collaborative Computing: Networking, Applications and Worksharing (Collaboratecom), 2013 9th International Conference Conference on. IEEE; 2013. p. 608–615.

[10] Horsman G. Unmanned aerial vehicles: A preliminary analysis of forensic challenges. Digital Investigation. 2016;16:1–11.

[11] Huang C, Lu R, Choo KR. Vehicular Fog Computing: Architecture, Use Case, and Security and Forensic Challenges. IEEE Communications Magazine. 2017 NOVEMBER;55(11):105–111.

[12] Mansor H, Markantonakis K, Akram RN, et al. Log your car: The non-invasive vehicle forensics. In: Trustcom/BigDataSE/I SPA, 2016 IEEE. IEEE; 2016. p. 974–982.

[13] de Fuentes JM, González-Manzano L, Gonzalez-Tablas AI, et al. WEVAN–A mechanism for evidence creation and verification in VANETs. Journal of Systems Architecture. 2013;59(10):985–995.

[14] Rogers MK, Goldman J, Mislan R, et al. Computer forensics field triage process model. Journal of Digital Forensics, Security and Law. 2006;1(2):2.

[15] Nieto A, Rios R, Lopez J. IoT-forensics meets privacy: towards cooperative digital investigations. Sensors. 2018;18(2):492.

[16] Conlan K, Baggili I, Breitinger F. Anti-forensics: Furthering digital forensic science through a new extended, granular taxonomy. Digital investigation. 2016;18:S66–S75.

[17] Warren SD, Brandeis LD. The Right to Privacy. Harward Law Review. 1890 December;4(5):193–220.

[18] Westin AF. Privacy and Freedom. New York: Atheneum; 1967.

[19] Holvast J. 27 - History of privacy. In: Leeuw KD, Bergstra J, editors. The History of Information Security. Amsterdam: Elsevier Science B.V.; 2007. p. 737 – 769. Available from: http://www.sciencedirect.com/science/article/pii/B9780444516084500286.

[20] Gubbi J, Buyya R, Marusic S, et al. Internet of Things (IoT): A vision, architectural elements, and future directions. Future Generation Computer Systems. 2013;29(7):1645 – 1660. Available from: http://www.sciencedirect.com/science/article/pii/S0167739X13000241.

[21] Solove DJ. Understanding Privacy. Harvard University Press; 2008. ISBN 9780674035072.

[22] The Universal Declaration of Human Rights; 1948. Available from: https://www.ohchr.org/EN/UDHR.

[23] European Convention on Human Rights; 1950. Available from: https://www.echr.coe.int/Documents/Convention_ENG.pdf.

[24] Solove DJ, Schwartz PM. Privacy Law Fundamentals 2017. International Association of Privacy Professionals; 2017.

[25] Nance K, Ryan DJ. Legal aspects of digital forensics: a research agenda. In: System Sciences (HICSS), 2011 44th Hawaii International Conference on. IEEE; 2011. p. 1–6.

[26] Huang J, Ling Z, Xiang T, et al. When digital forensic research meets laws. In: Distributed Computing Systems Workshops (ICDCSW), 2012 32nd International Conference on. IEEE; 2012. p. 542–551.

[27] The European Parliament and the Council of the European Union. DIRECTIVE 95/46/EC on the protection of individuals with regard to the processing of personal data and on the free movement of such data; 1995. Available from: http://data.europa.eu/eli/dir/1995/46/oj.

[28] The European Parliament and the Council of the European Union. REGULATION (EU) 2016/679 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC (General Data Protection Regulation); 2016. Available from: http://data.europa.eu/eli/reg/2016/679/oj.

[29] The U S Department of Justice. Privacy Act of 1974; 1974. Available from: https://www.gpo.gov/fdsys/pkg/STATUTE-88/pdf/STATUTE-88-Pg1896.pdf.

[30] Organisation for Economic Co-Operation and Development (OECD). The OECD Privacy Framework; 2013. http://www.oecd.org/sti/ieconomy/oecd_privacy_framework.pdf.

[31] Adams CW. Legal issues pertaining to the development of digital forensic tools. In: Systematic Approaches to Digital Forensic Engineering, 2008. SADFE'08. Third International Workshop on. IEEE; 2008. p. 123–132.

[32] Dehghantanha A, Franke K. Privacy-respecting digital investigation. In: Privacy, Security and Trust (PST), 2014 Twelfth Annual International Conference on. IEEE; 2014. p. 129–138.

[33] Ryz L, Grest L. A new era in data protection. Computer Fraud & Security. 2016;2016(3):18–20.

[34] Losavio MM, Chow KP, Koltay A, et al. The Internet of Things and the Smart City: Legal challenges with digital forensics, privacy, and security. Security and Privacy. 2018;p. e23:1–11. DOI: 10.1002/spy2.23.

[35] Ghafarian A, Seno SAH. Analysis of privacy of private browsing mode through memory forensics. International Journal of Computer Applications. 2015;132(16):27–34.

[36] Aljaedi A, Lindskog D, Zavarsky P, et al. Comparative analysis of volatile memory forensics: live response vs. memory imaging. In: Privacy, Security, Risk and Trust (PASSAT) and 2011 IEEE Third Inernational Conference on Social Computing (SocialCom), 2011 IEEE Third International Conference on. IEEE; 2011. p. 1253–1258.

[37] Bohannon P, Jakobsson M, Srikwan S. Cryptographic approaches to privacy in forensic DNA databases. In: International Workshop on Public Key Cryptography. Springer; 2000. p. 373–390.

[38] Croft NJ, Olivier MS. Sequenced release of privacy-accurate information in a forensic investigation. Digital Investigation. 2010;7(1-2):95–101.

[39] Halboob W, Mahmod R, Udzir NI, et al. Privacy Levels for Computer Forensics: Toward a More Efficient Privacy-preserving Investigation. Procedia Computer Science. 2015;56:370–375.

[40] Stahlberg P, Miklau G, Levine BN. Threats to privacy in the forensic analysis of database systems. In: Proceedings of the 2007 ACM SIGMOD international conference on Management of data. ACM; 2007. p. 91–102.

[41] Verma R, Govindaraj J, Gupta G. Data Privacy Perceptions About Digital Forensic Investigations in India. In: Peterson G, Shenoi S, editors. Advances in Digital Forensics XII. Cham: Springer International Publishing; 2016. p. 25–45.

[42] Aminnezhad A, Dehghantanha A. A survey on privacy issues in digital forensics. International Journal of Cyber-Security and Digital Forensics (IJCSDF). 2014;3(4):183–199. Available from: http://usir.salford.ac.uk/34016/.

[43] Ledbetter EM. Facing the Challenge: Protecting Data Privacy During Forensic Investigations. Utica College; 2017. Master Thesis.

[44] Van Staden W. Third party privacy and the investigation of cybercrime. Orlando, Florida, USA: Springer; 2013.

[45] Reddy K, Venter H. A forensic framework for handling information privacy incidents. In: IFIP International Conference on Digital Forensics. Springer; 2009. p. 143–155.

[46] Gupta A. Privacy preserving efficient digital forensic investigation framework. In: 2013 Sixth International Conference on Contemporary Computing (IC3); 2013. p. 387–392.

[47] Halboob W, Mahmod R, Udzir NI, et al. Privacy policies for computer forensics. Computer Fraud & Security. 2015;2015(8):9 – 13. Available from: http://www.sciencedirect.com/science/article/pii/S1361372315300750.

[48] Geiger M, Cranor LF. Scrubbing stubborn data: An evaluation of counter-forensic privacy tools. IEEE Security & Privacy. 2006;4(5):16–25.

[49] Hou S, Uehara T, Yiu SM, et al. Privacy Preserving Confidential Forensic Investigation for Shared or Remote Servers. In: 2011 Seventh International Conference on Intelligent Information Hiding and Multimedia Signal Processing; 2011. p. 378–383.

[50] Afifah K, Perdana RS. Development of search on encrypted data tools for privacy preserving in digital forensic. In: 2016 International Conference on Data and Software Engineering (ICoDSE); 2016. p. 1–6.

[51] van Staden W. Protecting Third Party Privacy in Digital Forensic Investigations. In: Peterson G, Shenoi S, editors. Advances in Digital Forensics IX. Berlin, Heidelberg: Springer Berlin Heidelberg; 2013. p. 19–31.

[52] Zimmer E, Lindemann J, Herrmann D, et al. Catching Inside Attackers: Balancing Forensic Detectability and Privacy of Employees. In: International Workshop on Open Problems in Network Security. Springer; 2015. p. 43–55.

[53] Antoniou G, Wilson C, Geneiatakis D. PPINA–a forensic investigation protocol for privacy enhancing technologies. In: IFIP International Conference on Communications and Multimedia Security. Springer; 2006. p. 185–195.

[54] Antoniou G, Sterling L, Gritzalis S, et al. Privacy and forensics investigation process: The ERPINA protocol. Computer Standards & Interfaces. 2008;30(4):229–236.

[55] Law FY, Chan PP, Yiu SM, et al. Protecting digital data privacy in computer forensic examination. In: Systematic Approaches to Digital Forensic Engineering (SADFE), 2011 IEEE Sixth International Workshop on. IEEE; 2011. p. 1–6.

[56] Hou S, Uehara T, Yiu S, et al. Privacy preserving multiple keyword search for confidential investigation of remote forensics. In: Multimedia Information Networking and Security (MINES), 2011 Third International Conference on. IEEE; 2011. p. 595–599.

[57] Armknecht F, Dewald A. Privacy-preserving email forensics. Digital Investigation. 2015;14:S127–S136.

[58] Yasinsac A, Manzano Y. Policies to Enhance Computer and Network Forensics. In: Proceedings of the 2001 IEEE Workshop on Information Assurance and Security; 2001. p. 289–295.

[59] Shebaro B, Crandall JR. Privacy-preserving network flow recording. digital investigation. 2011;8:S90–S100.

[60] Lin X, Lu R, Foxton K, et al. An efficient searchable encryption scheme and its application in network forensics. In: International Conference on

Forensics in Telecommunications, Information, and Multimedia. Springer; 2010. p. 66–78.

[61] Afanasyev M, Kohno T, Ma J, et al. Privacy-preserving network forensics. Communications of the ACM. 2011;54(5):78–87.

[62] Marrington A, Baggili I, Al Ismail T, et al. Portable web browser forensics: A forensic examination of the privacy benefits of portable web browsers. In: Computer systems and industrial informatics (iccsii), 2012 international conference on. IEEE; 2012. p. 1–6.

[63] Satvat K, Forshaw M, Hao F, et al. On the privacy of private browsing–a forensic approach. In: Data Privacy Management and Autonomous Spontaneous Security. Springer; 2014. p. 380–389.

[64] Dharan D DG, Nagoor Meeran AR. Forensic evidence collection by reconstruction of artifacts in portable web browser. International Journal of Computer Applications. 2014;91(4):32–35.

[65] Berghel H. BRAP forensics. Communications of the ACM. 2008;51(6):15–20.

[66] Stirparo P, Kounelis I. The mobileak project: Forensics methodology for mobile application privacy assessment. In: Internet Technology And Secured Transactions, 2012 International Conference for. IEEE; 2012. p. 297–303.

[67] Chow R, Golle P, Staddon J. Detecting privacy leaks using corpus-based association rules. In: Proceedings of the 14th ACM SIGKDD international conference on Knowledge discovery and data mining. ACM; 2008. p. 893–901.

[68] Grover J. Android forensics: Automated data collection and reporting from a mobile device. Digital Investigation. 2013;10:S12–S20.

[69] Nieto A, Roman R, Lopez J. Digital witness: Safeguarding digital evidence by using secure architectures in personal devices. IEEE Network. 2016;30(6):34–41.

[70] Ntantogian C, Apostolopoulos D, Marinakis G, et al. Evaluating the privacy of Android mobile applications under forensic analysis. Computers & Security. 2014;42:66–76.

[71] Keng JCJ, Wee TK, Jiang L, et al. The Case for Mobile Forensics of Private Data Leaks: Towards Large-scale User-oriented Privacy Protection. In: Proceedings of the 4th Asia-Pacific Workshop on Systems. APSys '13. New York, NY, USA: ACM; 2013. p. 6:1–6:7. Available from: http://doi.acm.org/10.1145/2500727.2500733.

[72] Weir G, Aßmuth A, Whittington M, et al. Cloud accounting systems, the audit trail, forensics and the EU GDPR: how hard can it be? In: British Accounting & Finance Association (BAFA) Annual Conference 2017; 2017. .

[73] Odebade A, Welsh T, Mthunzi S, et al. Mitigating anti-forensics in the Cloud via resource-based privacy preserving activity attribution. In: Software Defined Systems (SDS), 2017 Fourth International Conference on. IEEE; 2017. p. 143–149.

[74] Esposito C, Castiglione A, Pop F, et al. Challenges of connecting edge and cloud computing: a security and forensic perspective. IEEE Cloud Computing. 2017;(2):13–17.

[75] Nieto A, Rios R, Lopez J. A Methodology for Privacy-Aware IoT-Forensics. In: Proceedings of the 2017 IEEE Conference on Trustcom/BigDataSE/ICESS, Sydney, NSW, Australia; 2017. p. 1–4.

[76] Nieto A, Rios R, Lopez J. Digital witness and privacy in IoT: Anonymous witnessing approach. In: Proceedings of the 2017 IEEE Conference on Trustcom/BigDataSE/ICESS, Sydney, NSW, Australia; 2017. p. 1–4.

[77] Mylonas A, Meletiadis V, Mitrou L, et al. Smartphone sensor data as digital evidence. Computers & Security. 2013;38:51–75.

[78] Oriwoh E, Sant P. The forensics edge management system: A concept and design. In: Ubiquitous Intelligence and Computing, 2013 IEEE 10th International Conference on and 10th International Conference on Autonomic and Trusted Computing (UIC/ATC). IEEE; 2013. p. 544–550.