FIDIS

Future of Identity in the Information Society

Title:	"D6.7c: Forensic Profiling"
Author:	WP6
Editors:	Zeno Geradts (Netherlands Forensic Institute, The Netherlands),
	Peter Sommer (London School of Economics, UK)
Reviewers:	Mark Gasson (University of Reading, UK),
	Martin Meints (ICPP, Germany)
Identifier:	D6.7c
Type:	Deliverable
Version:	1.0
Date:	Wednesday, 30 April 2008
Status:	[Final]
Class:	[Deliverable]
File:	fidis_wp6_del6.7c_Forensic_Profiling.doc

Summary

This report, on forensic profiling, provides a bridge between forensic science and profiling from technical and legal perspectives.

Conclusions are drawn that new identity systems have their own strengths to detect what was impossible previously. But their weakness is that they can also provide false positives. From the examples it appears that much development is needed in this area before large scale implementation can be used in practice.

It is concluded that the different norms approved at European level remain insufficient. They do not deal with the impact of the widespread use of criminal intelligence, the increased monitoring of the average citizen or the increased linkage of police databases. Significant issues such as how to ensure the transparency and accountability of law enforcement activities, the quality of the data processed, e.g. the differentiation between categories of data subjects, or a strict application of the purpose specification principle remain unanswered.

Each country will thus be called to make the specific balance between the competing interests at stake, in particular to prevent that the increasing use of personal data for risk prediction turns into stigmatisation of parts of the population.

The FIDIS NoE receives research funding from the Community's Sixth Framework Program

Copyright Notice:

Members of the FIDIS consortium

1. Goethe University Frankfurt	Germany
2. Joint Research Centre (JRC)	Spain
3. Vrije Universiteit Brussel	Belgium
4. Unabhängiges Landeszentrum für Datenschutz (ICPP)	Germany
5. Institut Europeen D'Administration Des Affaires (INSEAD)	France
6. University of Reading	United Kingdom
7. Katholieke Universiteit Leuven	Belgium
8. Tilburg University[1]	Netherlands
9. Karlstads University	Sweden
10. Technische Universität Berlin	Germany
11. Technische Universität Dresden	Germany
12. Albert-Ludwig-University Freiburg	Germany
13. Masarykova universita v Brne (MU)	Czech Republic
14. VaF Bratislava	Slovakia
15. London School of Economics and Political Science (LSE)	United Kingdom
16. Budapest University of Technology and Economics (ISTRI)	Hungary
17. IBM Research GmbH	Switzerland
18. Centre Technique de la Gendarmerie Nationale (CTGN)	France
19. Netherlands Forensic Institute (NFI)[2]	Netherlands
20. Virtual Identity and Privacy Research Center (VIP)[3]	Switzerland
21. Europäisches Microsoft Innovations Center GmbH (EMIC)	Germany
22. Institute of Communication and Computer Systems (ICCS)	Greece
23. AXSionics AG	Switzerland
24. SIRRIX AG Security Technologies	Germany

[1] Legal name: Stichting Katholieke Universiteit Brabant
[2] Legal name: Ministerie Van Justitie
[3] Legal name: Berner Fachhochschule

Versions

Version	*Date*	*Description (Editor)*
0.9	01.01.2008	• Draft deliverable 6.7
0.95	17.01.2008	• Integrated all other information
0.96	27.01.2008	• First page and abstract modification requested by reviewers.
0.99	10.04.2008	• Revision after review
1.0	16.04.2008	• Final version

Foreword

FIDIS partners from various disciplines have contributed as authors to this document. The following list names the main contributors for the chapters of this document:

Chapter	*Contributor(s)*
1 Executive Summary	Zeno Geradts (NFI), Peter Sommer (LSE)
2 Introduction	Zeno Geradts (NFI), Peter Sommer (LSE)
3 Definitions	Olivier Ribaux (University of Lausanne, sub-contracted by NFI)
4 Emerging Profiling Technologies	Gerda Edelman, Gert Jacobusse (NFI), Thomas Gloe, Matthias Kirchner (TU Dresden); Olivier Ribaux and Sylvain Ioset, (University of Lausanne, sub-contracted by NFI)
5 Legal Implications of Forensic Profiling	Ekaterina de Vries (Vrije Universiteit Brussel), Fanny Coudert (KU Leuven)
6 Conclusions and Recommendations	All
7 Bibliography	All

Table of Contents

1 Executive Summary

This report, on forensic profiling, provides a bridge between the forensic pre-occupations of FIDIS WP6 and the profiling concerns of WP7. This deliverable is based on earlier workshops on forensic profiling in Amsterdam in 2005 (D7.6a) and in The Hague in 2007 (D7.6b).

The aim of forensic research is to support investigatory and judicial processes by finding traces in otherwise apparently unpromising raw material from which it is possible to build a picture of events and activities. Locard's Principle is at the foundation of what forensic scientists do: "Every contact leaves a trace".

The issue is data collected for one purpose but then used for another – and without there appearing to be any controls on the further use. Data Protection regimes appear to be silent on the topic. The regimes protect personal data, not generalised data which may then be applied as part of a profile to disadvantage an individual. And any tests that may exist in the legislation are also subject to poorly-defined exclusions based on public safety and the needs of criminal intelligence. Within Europe, the issue is further complicated by the difficulties of interpreting the various rules for the exchange of data, including intelligence data, between nation state members.

This deliverable provides some discussion to move these issues forward. Conclusions are drawn that new identity systems have their own strengths to detect what was impossible previously. But their weakness is that they can also provide false positives. He goes on to say that while electronic traces are information among others that are valuable in the context of the criminal justice system and forensic science, the technology itself must be understood within its context of use. Forensic profiling follows various objectives that are related to the interpretation process, to the investigation or for intelligence purposes. The use of these possibilities necessitates structured processes that may provide tools that go beyond technologies in order to discuss opportunities and risks. The significance of a profile is very different depending on the aim of processes being carried out: for instance, a physical description of the offender that corresponds to the profile of a certain proportion of the population may have different relevance from an investigative or court perspective.

In this deliverable we also look in several practical implementations which could be used for profiling, in images, tracking persons, finding relation in transport of drugs and setting up such a centre of expertise for forensic profiling. It appears that much development is needed for having large scale implementations.

Finally the arguments surrounding the proposed Data Protection Framework Decision for data processed in the framework of police and judicial co-operation in criminal matters are discussed.

It is concluded that the different norms approved at European level remain insufficient. They do not deal with the impact of the widespread use of criminal intelligence, the increased monitoring of the average citizen or the increased linkage of police databases. Significant issues such as how to ensure the transparency and accountability of law enforcement activities, the quality of the data processed, e.g. the differentiation between categories of data subjects, or a strict application of the purpose specification principle remain unanswered. The multitude of initiatives creates a complex framework prone to legal loopholes and difficult to comprehend.

2 Introduction

This report, on forensic profiling, provides a bridge between the forensic pre-occupations of FIDIS WP6 and the profiling concerns of WP7. This deliverable is based on earlier workshops on forensic profiling in Amsterdam in 2005 (D7.6a) and in The Hague in 2007 (D7.6b).

The aim of forensic research is to support investigatory and judicial processes by finding traces in otherwise apparently unpromising raw material from which it is possible to build a picture of events and activities. Locard's Principle is at the foundation of what forensic scientists do: "Every contact leaves a trace".

More specifically, the aims of digital forensic research are:

- To identify potential sources of digital evidence These are chiefly unintended artefacts from ICT – not the obvious substantive documents and transaction records but such features as configuration files, temporary files, date-and-time stamps, and deleted but recoverable data
- To examine and analyse them
- To derive, by the use of reverse engineering and testing, rules which describe their behaviour
- To produce convenient tools which enable these findings to be used during investigations

The products of forensic science activity can be aggregated with each other and also with other products of an investigation in order to assist a court in reaching conclusions.

Practitioners in Digital Forensics do not necessarily set out to breach privacy – their aim is to aid law enforcement.

But the effect of their work may be to weaken privacy rights because what can happen is that personal data lawfully acquired may then be subjected to a forensic enquiry which reveals more than was originally anticipated.

Looking specifically and simply at Identity Management Systems: in addition to their obvious role as a means to verify identity and then accord appropriate privileges they create audit trails of activity which can then be used to track the movement of an individual based on when and where an identity was presented for checking. The longer the period over which such audit trails are kept, and the greater the level of detail within them, the greater the potential breach of privacy.

Data matching is the traditional retrospective way of offender profiling, linking individuals with personal identifying data. But there is also the proactive practice of 'data mining' or 'risk profiling', that is, finding patterns and correlations in large databases, inferring a certain profile from these patterns and correlations and subsequently identifying people who fit these computer-generated profiles. People identified in this way may find themselves subject to exclusion, for example, from flying. Moreover whereas in traditional offender profiling an accused has an opportunity to know and test the evidence against him/her, when the techniques are used pro-actively, perhaps against an agenda of public safety, the excluded individual has little opportunity to challenge the profiling. Indeed it may be difficult for anyone to test the profiling algorithms and the quality of the data behind them.

The issue is data collected for one purpose but then used for another – and without there appearing to be any controls on the further use. Data Protection regimes appear to be silent on the topic. The regimes protect personal data, not generalised data which may then be applied as part of a profile to disadvantage an individual. And any tests that may exist in the legislation are also subject to poorly-defined exclusions based on public safety and the needs of criminal intelligence. Within Europe, the issue is further complicated by the difficulties of interpreting the various rules for the exchange of data, including intelligence data, between nation state members.

This deliverable provides some discussion to move these issues forward.

Olivier Ribaux examines the various definitions that are associated with the word "profiling" and then looks at the meanings attributed to "forensic profiling". In order to do so he takes us through the various types of analysis that are used in the investigative and judicial processes. One type consists of reconstruction of events; but another uses statistical information to build a "profile" of a possible perpetrator.

He concludes that new identity systems have their own strengths to detect what was impossible previously. But their weakness is that they can also provide false positives. He goes on to say that while electronic traces are information among others that are valuable in the context of the criminal justice system and forensic science, the technology itself must be understood within its context of use. Forensic profiling follows various objectives that are related to the interpretation process, to the investigation or for intelligence purposes. The use of these possibilities necessitates structured processes that may provide tools that go beyond technologies in order to discuss opportunities and risks. The significance of a profile is very different depending on the aim of processes being carried out: for instance, a physical description of the offender that corresponds to the profile of a certain proportion of the population may have different relevance from an investigative or court perspective.

Having provided a framework of definitions and possible theory, the Report then provides a series of practical instances from emerging technologies. Thomas Gloe and Matthias Kirchner provide an update on digital image forensics. Gerda Edelman describes work carried out by NFI into techniques of co-ordinating and aggregating pictures from multiple different CCTV sources using 3D modelling. Gert Jacobussen describes the work at the Netherlands Forensic Institute to set up a centre of expertise on intelligent data analysis which is deploying a variety of network and data analysis tools. Olivier Ribaux and Sylvain Iose describe an intelligence management system which profiles drug seizures developing patterns based, among other things, on chemical composition and communications patterns between those in narcotics gangs.

Katja de Vries and Fanny Coudert then examine the legal implications of forensic profiling. They investigate the distinctions between "due process" and "due processing". Due process is a fundamental drawn from Article 6 of ECHR. Due processing relates to Data Protection. In relation to forensic profiling is the fact that it does not limit itself to uniquely *individual* information (e.g. the fingerprint of one *particular* individual) but that it makes use of statistical information derived from huge databases (e.g. the profile of the *average* terrorist inferred from a certain pattern of correlations). A second peculiarity of forensic risk profiling, they say, is the fact that it can be used in a *pro-active* and *hypothetical* way. Instead of looking for an individual matching the traces left at a place of crime, forensic data mining can be used to *prevent* a crime. This pro-active or hypothetical character of forensic risk profiling

dissociates it from the investigative process directed at a potential trial. The profile used to detect high risk air-plane passengers is not meant to be used as *evidence* in a criminal trial, but is meant to prevent the high-risk passenger from entering the plane without further screening. The passenger who is told that he cannot enter the plane will often even be unaware of the fact that he was subjected to forensic profiling and simply assume that he apparently looked suspicious.

They then take us through the arguments surrounding the proposed Data Protection Framework Decision for data processed in the framework of police and judicial co-operation in criminal matters. They show that the difficulty of regulating forensic risk profiling in a way which is in accordance with the requirements of a constitutional democracy, is that it is a technique which is almost intrinsically opaque for the data subject who is subjected to it. This makes it hard for the data subject to contest the rightfulness of the processing of his data in court.

They focus particular attention on the attempts to provide regulation of the use of police data. For example, how far is data on criminal convictions "personal"? How do data protection principles interact with Directives designed to reduce crime by promoting cross-border co-operation? What is the position of data collected for one purpose and then re-used for another? What is the position of data collected by private companies and which they are required to retain under some European directive or national law?

They conclude the different norms approved at European level remain insufficient. They do not deal with the impact of the widespread use of criminal intelligence, the increased monitoring of the average citizen or the increased linkage of police databases. Significant issues such as how to ensure the transparency and accountability of law enforcement activities, the quality of the data processed, e.g. the differentiation between categories of data subjects, or a strict application of the purpose specification principle remain unanswered. The multitude of initiatives creates a complex framework prone to legal loopholes and difficult to comprehend.

3 Definitions

3.1 Introduction

In the context of crime or criminal investigation, profiling is often assimilated with *offender profiling*, *psychological profiling* or the use of *investigative psychology*, mostly, although not exclusively, in the context of violent crimes. *DNA-profiling* is a different term that is also familiar to a wide range of the population even if its exact scope remains largely unknown. Another immediate perception of profiling in a forensic context is the application of *data mining techniques* to an important quantity of data collected from crimes and persons in order to recognise patterns that may inform about illegal activities. Less known, but the object of growing interest, is the field of *illicit drug profiling* (systematic extraction and storage of chemical attributes of drugs seized in order to obtain indications on the manufacture and distribution processes, the size and the evolution of the market). There is thus no one single use of the term "profiling" in forensic science and intuitive meanings apparently lead to very different territories. If the psychological viewpoint appears to fascinate and attract many people, DNA, illicit drug profiling and data mining dimensions appear to belong to technical and highly specialised fields, largely inaccessible to the public.

These ambivalent feelings are the result of a distorted perception of all of the dimensions that lead to wrong expectations and fears: common sense vision of forensic science and criminal investigation differs considerably from concrete practice. Moreover, many different communities of researchers participate in the debate by developing similar but loosely connected models and approaches. These are based on different bodies of knowledge mainly borrowed from psychology, sociology, criminology, forensic science, crime analysis and criminal intelligence, or statistic and computer science. Finally, what really works and what does not is not easy to distinguish.

Thus, in the perspective of the FIDIS project, the process of balancing risks for the subjects and opportunities for the data controller is not immediate (Hildebrandt and Gutwirth 2008). For instance, weighting up the risks of being wrongly profiled as a criminal in the course of an investigation, and the opportunity for an investigator, law enforcement agencies or the criminal justice system to be able to neutralise dangerous criminals early, is not straightforward. There is an initial need to find some unity within these scattered pieces of works.

A better formalisation is also essential from a forensic perspective because notions of identity and identification are at the core of the domain and should properly integrate evolutions associated to id-systems and new identities in information society. Moreover, forensic science needs new frameworks in order to make the best use of data mining technology, not only in the treatment of electronic traces, but also to exploit more traditional forensic case data. This convergence between the different fields of forensic science, and particularly what is called forensic Information Technology (forensic IT), with methods for their exploitation such as data mining, seem to constitute one of the biggest challenges for the future.

This is a considerable task, as forensic science is too often considered to be a list of separated and narrow specialities. However, this FIDIS task, connected with results obtained from other FIDIS activities, offers an opportunity to take some steps towards this objective.

Thus, the distinctions that are provided here aim to identify some of profiling-related concepts, inferences and technical methods explicitly or tacitly used, as the object of research

or applied in practice. Reasoning activities that may be assimilated with profiling are pervasive. Of these inference forms, some are identified here in order to support further FIDIS tasks which will integrate them into a more global approach of profiling (Hildebrandt and Gutwirth 2008). This account is not intended to be comprehensive, because relevant dimensions go far beyond what can be explored in the single task of this project.

3.2 Definition of forensic profiling

In other FIDIS tasks, the terms forensic and profiling have been addressed separately.

Forensic: « The term forensic, as used in this report, refers to information that is used in court as evidence » (Geradts and Sommer, 2006) p.10. However, there are various definitions and acceptations of forensic sciences. We will consider here that forensic science is the study of traces resulting from criminal or litigious activities. This is an extension of the definition, according that traces are information not strictly dedicated to the court, but also that may bring knowledge in broader domains linked to security that deal with investigation, intelligence, surveillance, or risk analysis (Ribaux *et al.* 2006).

Profiling: "The process of 'discovering' correlations between data in data bases that can be used to identify and represent a human or nonhuman subject (individual or group), and/or the application of profiles (sets of correlated data) to individuate and represent a subject or to identify a subject as a member of a group or category" (Hildebrandt 2008) p. 41. Data mining technology is generally considered as a mean by which relevant patterns are discovered and profiles are generated from large quantities of data.

We consider that forensic profiling consist in the exploitation of traces in order to draw profiles that must be relevant to the context of supporting various security tasks, mostly in the criminal justice system. A distinction of forms of profiles that are used in this context is necessary before evaluating applications of data mining techniques for forensic profiling.

3.3 Forensic profiling in its context

3.3.1 Linkage blindness and limits of profiling

It may be perceived that the necessary data for forensic profiling is immediately available in a suitable form to the criminal justice system. This is definitely not so. Methods for processing data carefully distinguish a selective collection of traces, the collation of the data coming from different sources, the evaluation of its quality, the analysis of the available information and the timely dissemination of intelligence or knowledge on a need to know and right to know basis (Peterson *et al.* 2000). This decomposition helps to explicit a series of pervasive difficulties when profiling is envisaged.

A broad variety of barriers that go far beyond the inadequate use of technologies (Sheptycki 2004) hamper the fluidity of information. These can lead to a well identified weakness called *linkage blindness* (Egger 1984), an obstacle to the detection of relevant patterns in the information which exist in reality. This incapacity to connect the dots is generally accepted to be at the origin of main intelligence failures (United States 2004). Below are some examples of causes, but other legal, organisational, methodological, technological, human and fundamental (complexity) causes may also lead to linkage blindness:

- Law enforcement data is scattered into different files and in different jurisdictions. For instance DNA and Automatic Fingerprint Identification Systems (AFIS) may be centralised at country level, but both databases are generally treated separately as the result of legal rules. Moreover, databases may also use different classification systems and even preclude extractions of parts of the data, as well as electronic exchanges;
- Beyond police recorded data, administrative data and openly accessible sources, information is generally not directly accessible and available. If we suppose a specific situation, a judicial authority must intervene to authorise the access by the police and to order the possessor to grant access. This may dramatically slow down the whole process. Consequently, this may invalidate the analysis of the data in function of the dynamics of the problem under scrutiny. For example, several months are sometimes needed for obtaining some set of data in the framework of international cooperation agreements;
- Data comes from multiple sources under a broad variety of forms, which can still occasionally be a paper form. Moreover, the whole data treated, even police recorded data, is not prepared for profiling purposes, rather, it is structured for strictly administrative purposes;
- Profiles are hypotheses that are based mainly on imperfect (incomplete and uncertain) information. Thus, profiles may provide irrelevant leads and recovering from wrong investigative directions must be possible through recording assessment of the solidity of the information upon which hypotheses have been drawn.

These difficulties are obstacles to the treatment of data. Whether or not data mining technologies are implemented is not an essential question here. Rather, it appears that collection of data, evaluation of the information and the pre-processing stages for collating different sources of information generally imply a significant effort that must absolutely precede analysis and profiling.

This is particularly evident when dealing with the more fundamental questions of devising models in order to collate data coming from scattered sources. This data is generally available in different formats and must be structured in a suitable form for analysis purposes. Generally, at least three main dimensions of analysis appear relevant when dealing with criminal data for analysis purposes: what are the entities (for instance objects, individual, groups, traces, series, incidents, etc.) and their relations (for instance this person own this car), chronologies (for instance sequence of transactions between bank accounts), and spatio/temporal developments (for instance concentration of activities and their evolutions). It is very doubtful that data mining would be possible without first engaging efforts to collate the data. This is done through models that are based on at least one of those dimensions, depending on what the problem at hand is and what is searched for in the data.

Finally, disseminating obtained results in order to make intelligence products available to an organisation is a critical aspect of the whole methodology. The quality of communication influences the possibility to appropriately use the obtained profiles in the field. The analytical part that entices profiling, at the core of the process, must be thus carefully considered within a broader process.

3.3.2 Data available

Roughly speaking, sets of data available to law enforcement agencies are divided into two categories:

- Nominal data directly designates persons or objects (recidivists, intelligence files and suspect files, stolen vehicles or objects, etc.) and their relations. Nominal data may also be obtained in the framework of specific investigations, for instance a list of calls made with a mobile phone (card and/or phone) that cover a certain period of time, a list of people corresponding to a certain profile, or data obtained through surveillances;
- Crime data consist of traces that result from criminal activities: physical traces, other information collected at the scene, from witness or victims or some electronic traces, as well as reconstructed descriptions of cases (*modus operandi*, time intervals, duration and place) and their relations (links between cases, series).

Nominal data and relations may be abstracted in order to describe the structure of groups of offenders or criminal organisations.

Crime data are ideally also regrouped into abstract descriptions according to recurrent situations that share typical mechanisms. For instance, credit card frauds may be distributed into classes that separate skimming, distraction thefts, other thefts, etc. However, most of the time, data is initially administratively classified according to legal definitions which may mask the real dynamic behind crime problems (Goldstein 1990). This emphasises the necessity to make a distinction between sources of traces (persons or objects), the activity or situation that may explain the traces (the dynamic of the crime: context, immediate environment, victims, offenders) and the offence (legal definition) (Cook *et al.* 1998; Jackson *et al.* 2006).

This separation through crime/criminal data has led to a distinction between the fields of crime analysis, mostly carried out at a regional or local level, and criminal intelligence analysis, mostly the province of central agencies. This duality usually designates two professional communities (Bruce *et al.* 2004)[4]. However, both are obviously linked under many forms, particularly because traces directly result from behaviours of individuals and help provide some kind of description. This is compound by the aim of the investigation to identify, localise, and then provide evidence about the link between a trace and a person, to assume an activity or help determine an offence. In this context, forensic profiling will constitute the process that focuses on the exploitation of traces, but may overlap with criminal intelligence analysis.

3.4 Profiling and the reconstruction process

The *reconstruction* process is central. It starts from the traces (effects) and leads to the management of alternative hypotheses (possible causes) that may explain the existence of the traces. This process can obviously be envisaged through the optic of profiling.

[4] IALEIA: International Association of Law Enforcement Intelligence Analysts; IACA: International Association of Crime Analysts

Each criminal activity is unique, thus the reconstruction process does not aim to extract general principles, but rather to provide an explanation of what occurred in a specific situation with specific persons (what, who, when, where, why and how). However, this activity occurs within a general context (for instance demographic, economical, sociological, criminal, physical, etc.) that also have to be taken into account. For instance, knowing that similar events already occurred in the same region during the same period may be of importance in each phase of the judicial process.

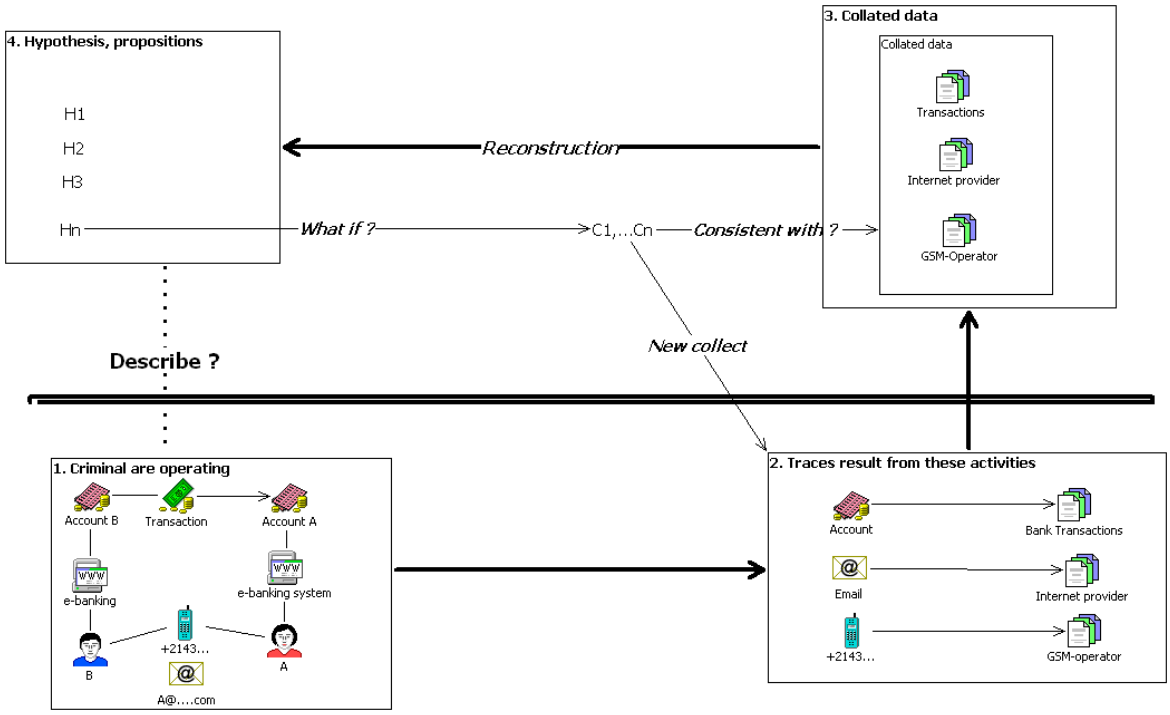

Figure 1: A description of the whole reconstruction process, from the activity to the management of hypotheses

Traces may be visible, latent or even hidden behind irrelevant data. They are frequently fragmentary. For instance, important files on a hard disk may be discretely hidden within directories generally dedicated to other purposes, logically deleted, or even partially replaced by new data. Specific search strategies must be devised for detecting, recognising and collecting relevant traces.

From the traces collected, possible sources, activities and the offence must be described. A modern approach for framing the reconstruction process considers together the offender, the victim, the immediate environment and existing protections (Felson and Clarke 1998; Ribaux and Margot 2007). This so called situational approach emphasises that conditions for a crime to occur are very specific and strongly depend on the motivation and abilities of the offender (perceived risks, expected gain, effort, knowledge or resources), as well as the characteristics of the victims within a specific environment that make her vulnerable and her values attractive (Value, Inertia, Visibility, Accessibility - VIVA). Thus, forensic profiling will not only concentrate on the offender, but also on victims and environment, in order to detect precursors that may lead to repeated victimisation. In fact, this chemistry or opportunities can be helpfully studied in order to look for concentration of crimes in time and place, as well as for explaining the existence and developments of these clusters (Felson and Clarke 1998).

3.4.1 The three chapters of the judicial process

There are many actors of the criminal justice system who search for answers to the what, who, when, where, why, and how questions, depending of their functions, roles, competence and/or knowledge. Each has a different aim that call for specific reasoning forms. The judges and jury will take responsibilities in the definitive decision at the end of the process. The preparation of the file is supported or carried out by the defendant and his lawyer, the police, forensic scientists or a magistrate. The whole reconstruction process may beneficiate from specific knowledge of experts. Witnesses and victims are themselves interviewed to provide insight on offenders and their activities. Finally, criminologists are interested in the mechanisms of the offence in order to test their hypothesis or develop theories of crime. These contributions may considerably differ from one system to another one. In particular, the role of the forensic scientists may range from providing expertise in very restricted situations to directly participating and coordinating part of the investigative process, where the study of traces and a scientific attitude are expected. Profiling then follows different aims that have to be distinguished.

Brodeur (Brodeur 2005) from a general perspective, and Kind (Kind 1994) with a forensic viewpoint, have both argued for the existence of different types of investigations or chapters that have their own logic in the course of the judicial process: the problem to find (identification and localisation of suspects), the problem to prove (how to structure evidence) and the problem of the trial itself where evidence is presented.

3.5 The different forms of forensic profiling in the judicial process

It is crucial to distinguish the different forms forensic profiling may take according to the different aims of each of the three chapters. They are manifold, but unfortunately, they are not all intensively studied. They are often tacitly applied or result from best practices. Here, the distinction between profiling as the result of a well-grounded professional methodology and its tacit counterpart is not always very clear. Moreover, there is a gap between what mythology about police work might suggest and the actual degree of formalisation of profiling methods practically used by law enforcement agencies.

Here is an attempt to capture only some aspects and methods that are applied at different degrees, but this survey does not intend to cover all possible forms. Even, some forms may be variably applied across organisations in terms of formalisation of the processes using databases, as a best practice or tacitly applied at the level of the individual. They will be mostly envisaged from an inferential perspective first; then, they will be considered in relation with the increased availability of electronic traces. Finally, data mining technologies for treating data will be discussed.

We will study profiling along the three chapters paradigm, but in the reverse order than was presented, because forms are limited when presenting evidence to a court and progressively more diverse when dealing with investigation. First, we will sketch the challenge faced when delivering conclusions to a court, then envisage some forms of profiling when structuring evidence, and finally complete the problematic by considering profiling as an investigative aid.

3.5.1 The interpretation process and profiling

Following the hierarchy of propositions coined by (Cook *et al.* 1998) forensic experts are essentially dealing with propositions related to the source (an object or a person) of a trace, alternative activities that may explain a set of traces and explaining how traces may provide other pieces of information aiding the final decision of the judge.

According to traces collected, the expert deals with propositions representing the prosecution and defence positions (Aitken and Taroni 2004). For instance, these propositions may take the form "this electronic trace was generated by this computer versus the alternative possibility that another computer taken in a relevant population caused it". Or at the activity level, according to traces collected: "this person was surfing on the internet during this particular session versus another person among a relevant population did a manipulation that provoked the traces".

In terms of profiles, one may assimilate a trace to an element of description of an object or a person. Then, interpretation necessitates comparisons between profiles built on the basis of traces collected from the scene (profile trace) with the profile of an accused person (profile accused). It often takes the concrete form of comparing traces with reference material taken from the source, for example a finger mark collected from a scene compared with a fingerprint taken from the accused. We will call the result of this comparison a match[5].

The relevant questions are reformulated in a probabilistic form: what is the probability that the profile-trace matches the profile-accused, if the accused person is the source of the trace. Then, from an interpretation perspective, one must consider the alternative possibility: the probability of a same match if another person were at the origin of the trace. Thus the reasoning process must continue by evaluating the proportion of profiles of other persons taken from a relevant population that could have possibly caused traces that match in a similar way.

Population studies (profiles of populations) are of great importance to help provide an informed assessment. Even, evidence is often wrongly considered as absolute: 10-loci DNA profiles, even rare, generally represent more than one single individual on the planet. Frequencies distributions of alleles in different populations constitute necessary background knowledge for assessing evidence. For instance, specific subpopulations and familial relationship may considerably affect allele frequencies.

In the field of electronic evidence, presenting evidence to a court according to this framework is still in its infancy: very few studies help determine on the basis of traces collected on a computer, how the profile built from this information may individualise the user. Many researches have still to be carried out in this direction.

Another form of forensic profiling occurs at this stage because a person must be categorised in function of a criminal rule of law. The virtual person representing the guilty is compared to the profile generated from traces. This is the offence level.

For example it is often asked if a suspect has belonged to or participated in a specific criminal organisation: according to the description of criminal organisation that may differ from one legal system to another one, how is the classification possible on the basis of the profile of the individual? Representation of criminal organisations as defined by the law, classification of a

[5] For simplification purposes, we will not discuss the different types of similarities that may result from this process, but this comparison process is at the heart of interpretation

specific organisation as a criminal organisation and demonstrating the participation of a specific individual in the organisation are frequent and far from obvious questions that have to be answered by the court.

Individual profiling (Jaquet-Chiffelle 2008) is required in this circumstance, as well as when for instance forensic psychiatry supports the assessment of a persons profile in order to evaluate his degree of responsibility that may depend on this kind of diagnosis.

The whole process is synthesised in figure 2. Using profiles in the interpretation process is promising because it rests on definitions of identities and clarifies which attributes are compared.

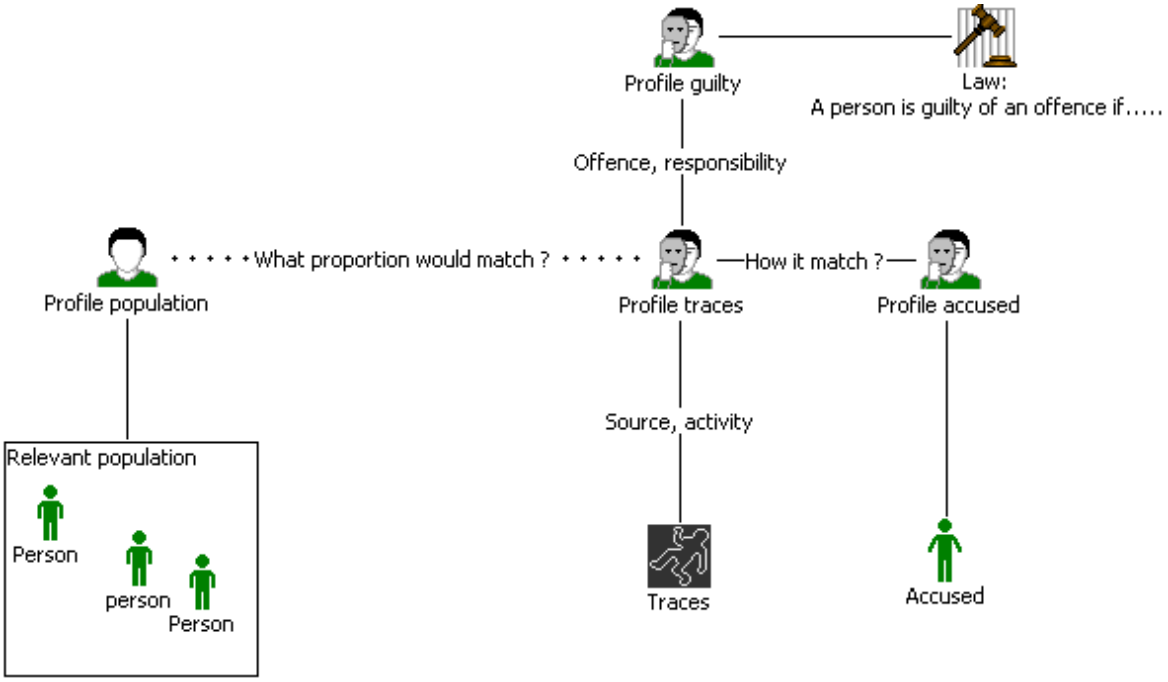

Figure 2: Description of the interpretation process using the notion of profile

3.5.2 Structuring evidence and profiling

When a suspect has been arrested, forensic scientists may advise an authority on how to deal with traces and provide leads on new traces to be collected. At this stage, a lot of activity is dedicated to test the consistency of available information, under the assumption that the suspect is at the source of the traces and the activity. A test of consistency with the hypotheses is not sufficient for going to court (see above), but it may lead to refute the hypotheses if available traces show unexplained discrepancies.

For instance a person who is supposed to have used her credit card at one place could not have used simultaneously her mobile phone at another distant place. In terms of profiles, coherence of the profile of the person under scrutiny has to be tested from various perspectives in order to detect potential contradictions or on the other hand to support hypotheses by demonstrating consistence (it has still to be confirmed how those concordances may occur by coincidence!). For instance, it may be assumed that the use of a mobile phone is part of the *modus operandi* of a serial offender when he is operating. Thus, data related to the localisation of mobile phones should show spatio-temporal coherence with data related to the

crimes themselves. Correlation between different sources of data (traces) may be thus intensively used according to the hypothesis to be tested.

3.5.3 Forensic profiling in an investigative perspective

As stated by many authors (Kind 1987; Wiggett *et al.* 2003; Jackson 2004; Jackson *et al.* 2006; Mennell 2006; Mennell and Shaw 2006), there is the realisation among forensic scientists that their role must extend to the investigation itself. They must be particularly engaged when hypotheses have still not been entirely drawn, in the coordination of the forensic information collected, as well as for proposing new collection of data. In this way, the forensic scientist turns from an evaluator to a more investigative attitude (Jackson *et al.* 2006): who/what is the source of this trace, how can we explain the existence of these traces, what is the offence, what evidence may indicate some possibilities for new data collection, what support and leads to the investigation may be provided, where is the person who committed the crime, etc?

This contribution is based on an entirely different inferential process than for interpreting evidence for the court. Rather than balancing probabilities related to given propositions, it focuses on the development of alternative hypotheses that may explain the existence of traces. Thus, rather than testing the hypothesis of culpability or innocence, we could generally describe the process as starting from the effects (the traces) and imagining possible causes on the basis of general knowledge (abduction and induction). Forms of profiling that arise during the investigative part of the process are manifolds and combine individual profiling with group profiling (Jaquet-Chiffelle 2008). We do not have the pretension of identifying all the possible forms here, only the most typical will be described.

One of the basic operations consists of creating a first profile from the available (collected) information and then searching for all the persons or their relations to objects that correspond to this profile. Profiles are described here as categories that restrict the search within a "selected" population. A person (individual) may generally be described through traces:

- They themselves reflects directly some physical aspects of the sources and have some descriptive capacity, such as fingermarks or DNA profiles extracted from biological marks, a snapshot taken from a camera;
- Traces and where they are found may be used to infer some indications about physical aspects or inform about clothes or accessories: earmarks found at a certain height on a door and the size of shoemarks may indicate (qualitatively) how tall the source is; a snapshot may provide some physical description as well as information about clothes and accessories;
- Traces may indicate the make and model of the printer used to print a recovered document, a bullet collected at the scene of crime may indicate the make and model of the firearm used, while paint marks coming from a car may point to a make and model of the implicated car. These are all types of acquisitions that may indirectly point to a person. Other possibilities include the use of fibre for inferring description of clothes, toolmarks or other marks for obtaining some description of the tools used. In a similar way of thinking, but about persons, DNA profiles indicate the gender (generally not

more about the physical aspect through non-coding DNA sequences chosen for forensic use);

- The activity and behaviour in the immediate environment may be inferred through a global analysis of the spatial (and temporal) distribution of traces, such as a sequence of shoemarks, a sequence of withdrawals with a specific bank card at different ATMs, traces of navigation with an internet browser;
- Circumstances and application of different theories from different bodies of knowledge may help to interpret the situations in order to provide other traits of the person or of his behaviour. For instance, geographical profiling (mostly for serial crimes) aims at providing clues for localising a person (Rossmo 1999), or different theories point out that psychological traits may also be inferred. The person may also be the object of a classification process into different categories (pre-defined classification of computer crime offenders, arson offenders, rapists, etc.).

Each final profile may thus be more or less general. Its attributes are known or unknown, complete or not and mostly uncertain.

One of the main (but not the only) questions of the investigation is the identification of the sources of the traces and how they may be related with the activity. Developing hypotheses about who/what is the source may be straightforward for instances through the use of DNA databases or Automatic Fingerprint Identification Systems (AFIS). Those systems start from the traces that come from a source (data subject as defined in (Hildebrandt 2008)) , transform them into a digital form (attribute of a virtual person (Jaquet-Chiffelle 2008)), compare them with collections of reference material and suggest as output a (list of) possible candidate(s) (or list of virtual persons) that refer to possible data subjects. The result is then interpreted and integrated into the investigation process. When using AFIS databases, a list of candidates is returned by the system, while for DNA databases, usually a single profile[6] is returned. However, with the evolving content of databases and since identical twins have same DNA, occasionally several DNA-profiles may be returned by the database. Moreover, with the extended use of partial DNA or mixtures, putative sources may be multiple.

In order to generalise this process, a useful concept has been stressed by Kind (Kind 1987). He argues for the use of the dual concepts of frame and form. The frame contains the set of entities considered as relevant for the investigation, according to available evidence, while, roughly, the form distinguishes different region of the frame as more or less promising. A list of candidates extracted from an AFIS system constitutes the frame, while scrutinising the content provides as outcome the form. The frame is often constituted of persons or entities that share a common profile. This may also be seen as a non-distributive group profiling approach (Hildebrandt and Backhouse 2005; Hildebrandt 2008; Jaquet-Chiffelle 2008) where a category of individuals is built on the basis of a different set of data and where the decision to insert an individual (or its individual profile) into the frame may depend on features of different natures.

[6] The use of profile for DNA may be confusing in the context of this deliverable. However, a DNA profile may be defined as a description of a person through part of her DNA structures. Even if the parts of the DNA – structure used in a forensic context have been chosen for their polymorphism across the population, the same profile may apply to several persons. A profile thus does not define a single individual, but rather a group.

There are many ways to develop a frame in the course of the investigation, depending on the case and available traces. The direct and simplest way consists in comparing the trace with the collection of reference material (like for DNA or AFIS databases). A similar process consists of comparing images taken from video surveillance systems (CCTV) with collection of photos taken from known persons. The scheme is the same and simple, but obviously the source of data used presents specificities that make the methods routinely applicable, as well as automated profiling possible or not.

Another possibility, when recidivism is known as frequent, is to compare the assumed *modus operandi* of the offender with *modus operandi* used by known recidivists. Here again, when serial crime is considered, a profile extracted from the series of *modus operandi* used by the recidivist (a profile extracted from an already constituted set of information – individual profile) may be used to proceed to the comparison: the burglar usually operated during the night, entered the premises through an open window, and generally selected only credit cards. There may be very different approaches for building such a profile, for instance by expecting that a specific feature occurs in each case or only in the majority of cases, expecting the existence of a specific feature of not, etc. The relevancy of such a profile depends on the expected use of the profile (searching other databases for linking cases, organising specific surveillance, trying to intercept the perpetrators) and thus may take the status of intelligence (see below).

Another important form of profiling is carried out through the application of models and methods used for hypothesising the place where the offender resides, or one of his centres of interest. These methods are known as geographical profiling and may be used in specific situations, for instance when a serial offender operates (Rossmo 1999). With the development of new technologies, data extracted from GSM operators may play an important role in this perspective, for instance by assuming the degree of mobility of a person, where he resides or other spatial dimensions related to his behaviour.

Finally, other possibilities are developed through new id-systems: when a profile of the offender has been developed and some of his activities may be inferred, new frames may be built. For instance if the author was suspected of having used her mobile phone when operating, details of all the calls made during the time of the offence in the region of interest may be asked from the operator, with the hope of detecting the card or the mobile phone used by the offender. If an offender is supposed to have entered a building controlled through id-systems, the list of persons who entered the building may be provided.

All these forms may be used in combination through cross-referencing, for instance when geographical profiles lead to a list of inhabitants, the use of firearms may indicate the relevancy to search among the list of legal possessors, the profile of a car to consider the file of car owners, etc. This data may then be cross-referenced either to build a category of persons corresponding the best to the offender profile, conscious of the fact that the offender may or not appear in these databases. This may, as an outcome, provide a list of relevant identities to be further investigated.

Jaquet-Chiffelle (Jaquet-Chiffelle 2008) stressed that this kind of investigative profiling follows two distinct goals: the first is to identify an individual within a community or infer its habits, behaviour, preferences, knowledge, etc. But the second form is not independent from the first one as it is often not obvious, once identified, to find (ultimately arrest) a person worth being the object of further investigations. Occasionally, the localisation of the person

even leads to his arrest before he is identified. For instance, when a serial burglar operates, his pattern may be detected and used to devise surveillances that may in turn lead to his arrest.

A rich example, well documented, of possibilities to apply such techniques can be found in the review of the investigation of the Yorkshire Ripper during the 70s (Byford 1981). This investigation offers a broad series of inferences and treatment of data typical of complex investigations. Review of the case has led to original ideas about different forms of profiling (Kind 1987). At that time, among other difficulties, the lack of computerisation and possibilities of cross referencing was identified as a severe handicap for the investigation. The ripper was finally arrested through a routine control in the street, because he was circulating with stolen plates. Despite that this arrest was made in isolation from the investigative strategy itself, it was actually also obtained through the use of a systematic control process aided by the databases of stolen plates. Lessons learned from this case have had in particular considerable impact on the development of computerisation for major case management[7] and organisations of incident rooms. It may also be considered as a milestone in the development of analytical capabilities within law enforcement such as geographical profiling or the use of information technologies in the management of serious cases.

3.5.4 Categorical elimination

Once the frame has been built, further techniques or investigations may lead to categorical elimination. This can be performed for instance through DNA screening, when persons are eliminated through comparison of their DNA profile with the profile of the trace.

This filtering process, when applied with DNA, offers some guaranties, but occasionally, the profile may not be accurate and leads to a wrongful elimination. This was the case during the Yorkshire Ripper investigation, with the letters and an audio tape received by the police. The tape contained a message of a person who pretended to be the ripper. He had a strong accent from one specific region. By assuming that the sender of the video was the actual ripper, each person who did not correspond to this profile (in particular the strong accent) was eliminated from the investigation. Actually, the person who sent the letters and tape was identified in 2005 thanks to DNA and had apparently no connection with the ripper. This example illustrates one reason why profiling in the course of criminal investigation is particularly critical and necessitates further formalisation.

3.6 Repetition of crimes as a fertile area for forensic profiling

At this point, the treatment of specific cases has been discussed. However, some forms of criminality, such as high volume crimes, are likely to show a substantial serial component. A series of studies in criminology led to the conclusion that both as small number of persons or group of offenders are responsible for a great quantity of cases (Wolfgang *et al.* 1972; Audit commission 1993), and certain victims and types of victims show different patterns in the way they are repeatedly victimised (Farrell and Pease 2001). Understanding patterns of victimisations and detecting the activity of groups of serial offenders are of crucial interest both from a repressive and preventive perspective, and also from a more strategic perspective, when the structure of some forms of criminality are scrutinised (mobility of offenders,

[7] Development of the HOLMES system (Home Office Large Major Enquiry System)

specialisation, evolution of criminal careers, amplitude of the activity, association between criminals, classification of victims, vulnerabilities, crime phenomena, etc.). This is why modern policing strategies necessitate the implementation of analytical capabilities dedicated to the understanding of the dynamic of crime and the provision of intelligence on which informed decision may be taken.

Forensic profiling has thus a great role to play in this perspective, providing at least that physical material and traces have a demonstrated potential for connecting the dots, or linking cases (Ribaux and Margot 2007).

3.6.1 Repetition of crimes – crime series

Crime series occur when the same offender operates several times. There are no definitive definitions of crime series as group of criminals may operate in different compositions, show different forms of organisation and perpetrate different types of crimes. One may accept that when the same offender operates repetitively, there is a *crime series*, and that individual crimes of that series are *linked*. Several crime series may be connected and form a graph structure which represents the activity of a group of offenders.

Thus, three forms of forensic profiling that rely on police recorded data may occur in this context:

1. Detection of a series of crimes and reconstruction of graph structures that represent the activity of single offenders or group of offenders (group profiling)
2. From a series of crime, extract a profile that best describes the series and its or their author(s) (individual profiling)
3. From the graph structure, extract a profile that best describes the forms of criminality and the group of offenders (group profiling)

Crime series detected through police recorded data are generally incomplete and largely uncertain. Not all crimes are reported by victims (level of reportability may greatly vary according to the type of crime) and the collection of traces at each crime scene may be only partial. Moreover, a series of obstacles that ranges from legal constraints to technical difficulties causes impossibilities to detect links, while they actually exist. This is the well known incapacity of connecting the dots or linkage blindness law enforcement agencies systematically suffer from (Egger 1984; Sheptycki 2004).

From this perspective, basic contribution of forensic science concentrates on crime linking through physical evidence. Detection of links is particularly efficient when the type of crime causes the exchange of physical traces. For instance, DNA databases automatically detect a great number of series, when traces are collected systematically and there is a realisation within the organisation that this detection is useful (Girod *et al.* 2004). A series of arsons, linking through the analysis of illicit drugs seized, or even series of bombing have largely benefited from this approach (Ribaux and Margot 2007)

Other types of traces, such as toolmarks (Geradts *et al.* 1999), earmarks, shoemarks (Girod 2002; Rix 2004), may be used in combination, thus augmenting the chance to detect series of

crimes through at least one of these dimensions. Electronic traces may also be used for detecting the use of the same mobile phone or smart-card for organising the operation, linking cases when surveillance cameras show withdrawal with stolen bank cards perpetrated by the same person, or for linking different withdrawals perpetrated with the same stolen cards, and even linking unsolicited e-mails through technical information found for instance in the header of the message (Birrer *et al.* 2007).

All detected links present some degree of uncertainty: they may exist or may not exist and we are more or less certain of that. However, links inferred from physical material present the advantage of being the result of a direct comparison of measurable data collected at scenes of crime, while other approaches necessitate as a precondition some form of interpretation. For instance, a traditional method for linking crimes (Völlmer 1919) relies upon the comparison of *modus operandi* which are already the result of a reconstruction process and thus necessitate uncertain reasoning to be performed as a precondition.

Other approaches consist of focusing on behavioural aspects of cases which add other levels of inferences and uncertain reasoning. This psychological based methodology has been considered as a breakthrough, as it brought systematic and structure for linking violent crimes. However, a series of drawbacks have now been detected, ranging from difficulties of filling long forms to feed the database to the lack of assessments of the global efficiency (Godwin 2001; Grubin *et al.* 2001; Witzig 2003).

Thus, by using a holistic perspective, the detection of links is envisaged by monitoring and combining all the possible dimensions. However, pragmatic constraints (cost of input and analyse data, time, knowledge, etc.) and other types of constraints limit the practical possibilities to aggregate data and multiply all those perspectives. This inevitably contributes to linkage blindness. This problem ought to be attenuated by choosing the more solid and efficient dimensions through systematic assessment and comparisons of available methods.

Beyond the type of data used, the central mechanism for detecting series rests on proximity measurements: spatio/temporal, *modus operandi* or similarities between traces. There is also space here for developments, seeing that each serial offender shows an individual profile, while automatic methods for comparing cases generally rely on one single metric (see also the example about drug profiling below).

3.6.2 Repetition of crimes – problems and phenomena

Forms of repetition of crimes other than the serial activity of offenders may be relevant to decide which measures (preventive, repressive or a combination) may be appropriately taken. For instance, hot-spots or crime clusters may incite to concentrate police surveillance on specific area during a certain period, independently of the serial activity of single offenders. Concentration of crimes may be also due to the attractiveness of a potential victim, a residence or a specific environment, independently of how many authors are operating. In this perspective, profiling techniques used refer to spatio/temporal crime cluster analysis that aims to alert based on relevant patterns extracted from crime data.

More generally, the concept of the problem analysis triangle helps analysis of situations in terms of a combination of a likely offender who meets a suitable target within a specific environment without the protection of a capable guardian. These approaches belong to a set of very rich theories in criminology, called opportunities theories (Clarke and Eck 2003). In this perspective, crimes (and more generally problems that regroup recurring incidents and

suggest to focus on their causes (Goldstein 1990)) are considered through the lens of situations which may also be the object of profiling (what are the types of situations recurrently observed, how may I interpret this cluster of crimes?).

3.6.3 Repetition of crimes – tactical, operational and strategic analysis

Intelligence-led approaches of policing proposed to estimate the usefulness of a piece of intelligence (knowledge) through its capacity to help in decision at different levels of the organisations. In this perspective, at a tactical level, a profile that results from an investigation or a detected series of crime orients investigation (who is the perpetrator, where can I find him). At an operational level, phenomena, target profiles, repeated vulnerabilities and gaps in the security system are scrutinised in order to organise some specific responses. Crime clusters are good example of profiles that may be used at an operational level. Finally, at a strategic level, understanding of the criminality structure along with its tendencies support the devise of global strategies, and decision to be taken by the management. However, all those levels rest upon a solid detection of series and an understanding of repetitive situations.

For instance, the systematic analysis of e-mails massively sent by scammers (advance fee fraud, 419 scams, phishing, all sorts of counterfeits materials, etc.) in order to find a potential gullible victim provide accurate indications on the criminal mechanisms[8]. This is done from a tactical to a strategic level, and concerns the infrastructures used, the localisation, or the organisation of the fraud (Anderson *et al.* 2006; Birrer *et al.* 2007). This example blurs the line with the wider concept of intelligence and/or surveillance.

3.7 Intelligence, risk analysis, detection and surveillance from a forensic profiling perspective

The analysis of repetitive crime may find an extension in the strongly connected fields of intelligence, risk analysis, detection and surveillance. There are requirements at an European level to set up risk management systems at the level of the countries (see FRONTEX for instance[9]) and initiatives, as well as recurrent recommendations to develop intelligence-led styles of policing[10] (GNIM 2005). Within this modern framework, security and policing are based on a strong analytical capacity that is articulated around risk assessment or problem detection. For instance, certain intelligence-led systems have identified four area of interest

[8] Some typical frauds on the Internet consist of creating a first contact with potential victims or attracting her to a place or a site through an e-mail. The advance fee fraud consists, through massive sending of unsolicited e-mails, of persuading prospective victims to pay modest amounts of money with the promise to receive substantial benefit in return. Various scenarios are generally used, such as an old dictator asking for help in order to move his fortune out of West Africa. Phishing consist of asking the recipient of the e-mail to connect himself onto his e-banking system through a proposed URL link. Actually the link leads to a copy of the veritable e-banking sitewhich is controlled by the scammers. The perpetrators then collect the data that is entered by the victim. A lot of counterfeits materials (watches, bags, etc) are sold on Internet sites. Potential customers are attracted to these sites by e-mails sent massively. Thus, all these e-mails result from criminal activities and can be considered as traces. The analysis of these e-mails has a great potential to help reconstruct some aspects of these mechanisms.

[9] http://www.frontex.europa.eu/ last access 20.10.07

[10] For instance Rec (2001)11 – of the Committee of Ministers to member States concerning guiding principles on the fight against organised crime: "Member states should develop new methods of police work by shifting focus from reactive policing to pro-active policing, including the use of strategic intelligence and crime analysis"

(GNIM 2005): locations (for instance a hot-spot or other patterns), subjects (persons or groups), crimes (crime series) and high risk issues (for instance a forthcoming manifestation).

Thus, forensic profiling may even go beyond repetitive crime analysis. Actually, new types of policing strategies based on intelligence necessitate the development of monitoring processes that help detect various sorts of dangers for the security on the basis of crime data. This goes beyond this framework by exploiting open source or other data that allow signs of danger to be detected. Repetitive crime analysis is only one form of those processes, but they may be better considered in connection with risk analysis and surveillance.

Main changes from this perspective come from the extensive use of ID systems in our societies. Data possessed by credit cards and mobile phone operators, use of automatic plate detection systems, automatic recording of passport and biometric data when controlling people (at the border or for access controls), as well as various other sorts of electronic data that potentially allow what might be considered as dangers to be detected. For instance:

- Surveillance systems (computerised or not) are often focused on the basis of definition of individual profiles, for instance people searched by an authority for any reasons;
- The profile of a criminal activity related to hot-spots area for targeting surveillance.

Of importance is that most of this data, compared to crime data used for repetitive crime analysis, is initially mostly not dedicated to risk management or security (with the exception of explicit identity controls), but generally record traces resulting from a service that rests on an electronic infrastructure (for instance, e-commerce, e-banking, etc.). Generally, aggregation of data for monitoring purposes is not possible as they are possessed and secured by different companies that set up their own pattern analysis system for detecting suspicious activities. However, legislations may range from favouring some limited direct availability of this data to law enforcement agencies in order to open more possibilities: for instance reporting procedures for financial intermediaries to Financial Investigation Units (FIU)[11] might be seen as a central feature for the application of anti-money laundering techniques.

Because data mostly results from a legal activity, most patterns that may be detected are uninteresting from a risk analysis perspective, unless data has previously been pre-processed or filtered on the basis of a well identified selective risk profile. Separation of relevant from irrelevant patterns is thus difficult and may lead to repetitive false alarms. Therefore, rather than the reliability of the technology, the central aspect is the possibilities to devise appropriate risk profiles. This may help detect deviations between what may be inferred as a historical usual activity of a population or of individuals, in comparison with the results of a suspect change of behaviour. For instance a change in the use of a bank account that cannot be explained through standard activities of the whole population may alert and incite to scrutinise in more depth the reasons for these changes. Profiles of criminal activities in GSM fraud may lead to the definition of filters that focus detection on some specific massive calls on specific numbers, in specific conditions. Another example is the detection of paedophilia images that may pass through a channel under surveillance: large amount of known images are collected in databases under the form of their digital “fingerprints” and are compared with the information traversing the channel[12]. Of course, images may be slightly transformed and

[11] Egmont group see http://www.egmontgroup.org last access
[12] see for instance: http://www.netcleantech.com

therefore will not be recognised by the surveillance system. Thus, new challenges consist of finding profiles and measures of similarities rather than detecting the exact image.

Another determinant aspect concerns the integration of a learning process into these techniques: criminality and knowledge about criminality are evolving over time, sometimes rapidly. For instance, a new case integrated into a series of crime may dramatically change the knowledge about the offender (his profile). The detection of new forms of repetitions may lead to the construction of new profiles that, in turn, will influence the detection process, acquisition and classification of new events. For instance, the rise of metal rates has made these more attractive for criminals and has led to an increase for example of copper thefts. Given that specific groups of criminals have focused their activities toward this type of crime, it becomes important from an analytical point of view to devise a new classification dimension for all metal thefts.

The utility of future computerised systems as an aid in an intelligence perspective will greatly depend on their capability to adapt and integrate these learning mechanisms. Moreover, they should not be considered exclusively from a technical perspective, but rather have to be suitably integrated into a complete methodology, at the workplace, taking into account a complex set of pragmatic constraints.

3.7.1 Distribution of tasks

Whether this learning process may be computerised or strongly supervised by human operators is the subject of a subtle balance that depends on the specific problem at hand. This is one of the main difficulties and possibly causes of misunderstanding between the different communities dealing with data mining, crime investigation and intelligence. This is compounded by the fact that intelligence is not only the result of structured treatment of information, but also emerges from intensive tacit communication largely based on confidence. One critical aspect for reducing failures of intelligence resides in how human intelligence can complete signal intelligence. Augmenting the use of new technologies will add very few values if the framework does not integrate these considerations.

Probably less “true” data mining systems are used in practice than is imagined, although promising separated components are providing new forms of aid for investigators, analysts, and deciders. The process of better understanding more adequate repartitions goes through intensive formalisation of the different concepts pervasive in the treatment of data and information within a law enforcement and security context. In particular, usually data is not recorded into databases in a suitable form and necessitates strong pre-processing.

When tasks are distributed, communications between the contributors arise as a critical aspect. This is why visualisations of entities (individuals, objects, events) and their relations, chronologies, as well as spatio/temporal aspects play an important role when different human partners collaboratively solve a specific problem.

In this perspective, possibilities to use algorithms for forensic profiling (Anrig *et al.* 2008) have to be carefully studied. This assessment mostly belongs to the research area. A recent conference[13] provides an overview of the adaptation of technologies to this field. Other

[13] NATO Advanced Study Institute on Mining Massive Data Sets for Security, September 10 - 21, 2007, Villa Cagnola - Gazzada – Italy http://mmdss.jrc.it/ (last access 24th October 2007)

authors have also made a global account of the potential of data mining technologies applied to particular sets of forensic case data (Oatley *et al.* 2006; Terrettaz-Zufferey *et al.* 2006). There is a clear trend to consider the whole information available, particularly electronic traces, in the framework of data mining technologies[14].

3.8 Perspectives: virtual persons and forensic profiling

Virtual persons is a model defined in FIDIS which has already been presented (Jaquet-Chiffelle *et al.* 2008) and is the subject of its own deliverable (Jaquet-Chiffelle 2008). It consists of abstract representations that may describe physical persons or groups of persons, computer software, as well as other virtual entities relevant in modern information societies. Associate to virtual persons are their profiles with their attributes, roles, abilities, acquisitions, habits, behaviours, etc.

This concept may be useful to connect crime/criminal intelligence analysis. Unknown offenders may be described as virtual persons that do not necessitate designating a particular physical person to be worked on, or may even point to several different physical individuals changing in the course of the investigation. For instance, virtual persons may be useful when considering serial crimes and the virtual offender(s) imagined through individual profiling (figure 3). Those structures may be scrutinised without necessitating permanent connection with physical persons; this connection may also change in the course of the investigation, according to a hypothetical-deductive reasoning process. From the profile, general knowledge about types of offenders may be used to refine the description of the offender. For instance, from the description of the person and his behaviour and general knowledge about typical patterns of mobility, some useful indications for localisation are inferred. Moreover, this profile may also apply to a particular data subject. For instance, if it is assumed that the individual is living in a particular region, the subject's data concerning habitants are immediately linked.

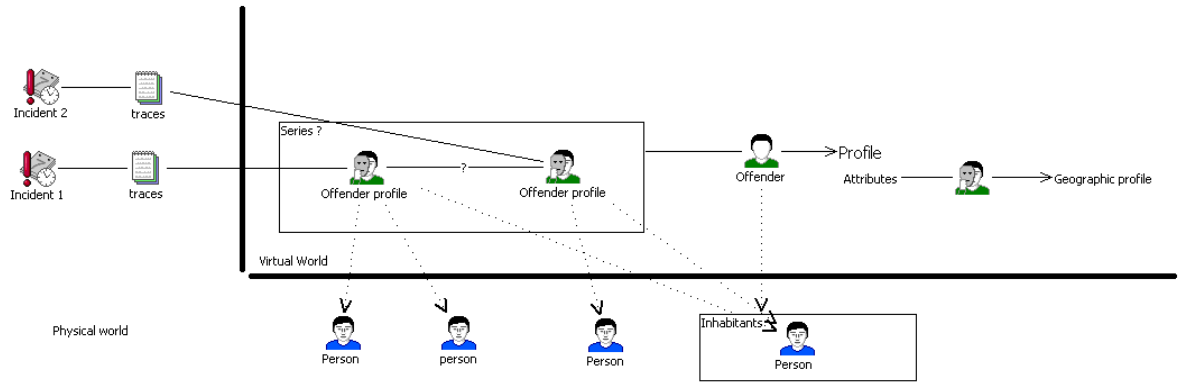

Figure 3: Description of a series of crimes using the concept of virtual persons. The physical and virtual worlds are separated. In the virtual world, perpetrators are not known, but their profiles can be worked on in order to point to a restricted set of potential candidates. In this example, the geographical distribution of the incidents may show a pattern typical to the behaviour of specific types of offenders.

[14] Trends in conferences, see e-Forensics 2008: The 1st International Conference on Forensic Applications and Techniques in Telecommunications, Information and Multimedia http://www.e-Forensics.eu (last access, 16th Novembrer2007)

This formalism may be even more suitable for modelling some confidence games that occur on the internet. Advance fee fraud scenario used by scammers to deceive gullible victims (see above) may be usefully described with virtual persons: persons appearing in the scenario and using different identities to deceive victims, as well as representations of the scammers themselves with their different roles in the fraud. This exercise should be the object of future use cases studies.

However, it must be emphasised that the final objective of the investigation remains proving the existence of a crime, linking the activities to human beings, as well as collecting, preserving, interpreting and presenting evidence. Thus, the connection between virtual and physical worlds concentrates the attention. Using virtual persons within an investigative framework necessitates further research. It is not claimed here to have made a definitive account and even appropriate use of virtual persons, but only provide insight of how this notion may be linked to some aspects of forensic profiling and help provide the material that will help test the adequacy of the model in a forensic context through the use of cases studies.

Forensic profiling follows various objectives related to the interpretation process, investigation or intelligence. Beyond difficulties and limits described above, the number of applications and ideas for potential developments increases rapidly.

The following examples illustrate this variety by some emerging uses of forensic profiling. Some of these approaches also include the search for the application of suitable data mining technologies.

The first example discuss the kind of information digital image may reveal about the used image acquisition device, as well as possible characteristics for interpreting the link between the trace and its putative source. The detection of images alterations belongs also to the whole forensic process presented.

Another original illustration pertains to 3D-modelling of CCTV footage. Through the use of this technique, several parameters useful to reconstruct the activity recorded can be inferred, such as distances, locations of objects and persons, or their speed. In particular, this approach helps to test alternative hypothesis about scenarios.

Emerging profiling and data mining technologies have still to be tested in order to situate better their potential in the analysis of large quantities of crime data, and to help provide timely and accurate intelligence. A centre of expertise on intelligent data analysis set-up in Holland for this purpose is presented.

Finally, the use of forensic profiling for intelligence purposes is illustrated through an actual illicit-drug profiling process. Data mining technologies for detecting patterns in the chemical profile of the illicit substances seized have been tested. They help separate and detect changes in the organisation of the market. The harmonisation of methods used by several forensic laboratories is a condition to extend the use of the approach to an international level.

4 Emerging Profiling Technologies

Forensic science has many areas of expertise. In this chapter we focus on several techniques that could be used for profiling, one on digital image forensics, on tracking persons with video-cameras, on setting up a centre for profiling and on drug trafficking.

4.1 Digital Image Forensics

Thomas Gloe, Matthias Kirchner (TU Dresden)

In our multimedia society, digital images play an important role. The advent of low-cost digital imaging devices as well as powerful and sophisticated editing software gives rise to a wide use of digital images in all areas of our everyday life. Notably, the internet allows a fast distribution of digital image material. In the context of forensic profiling, a digital image may reveal information about the image acquisition device used and, consequently, give a link to a single person or a group of persons who probably took the picture. Furthermore, the question whether a digital image shows an original and unaltered scene or whether it was tampered with subsequent to its generation is of high importance, e.g., when analysing images of surveillance cameras. Both, the question of image source identification and the question of securing an image's integrity, are subsumed by the concept of image authenticity.

Several approaches to address these questions have been proposed, e.g., digital signatures or digital watermarking. Nevertheless, it is important to note that both digital signatures and digital watermarks have to be generated directly in the imaging device since at a later point the image's authenticity can no longer be guaranteed. In contrast, methods with come under the relatively new concept of *digital image forensics* basically rely on particular statistical features, which can be understood as a "natural" and inherent watermark. Consequently, digital image forensics does not require any prior knowledge of the original image.

According to the above-mentioned issues, the area of digital image forensics can be broadly divided into two branches. The first problem linked to digital image forensics is *image source identification*, which is based on specific characteristics of the image acquisition device or technology. The second field of application is to determine whether a specific digital image has undergone *malicious post-processing or tampering*. Forensic algorithms of this type are designed to unveil either characteristic traces of image processing operations, or to verify the integrity of particular features introduced in a typical image acquisition process.

To obtain the goals, digital image forensic techniques exploit either device specific characteristics introduced during the image acquisition process or manipulation artefacts introduced during image processing. Additionally, meta information included in an image file, like the date of exposure or the name of the camera model, could be part of a forensic analysis. In contrast to device specific characteristics or manipulation artefacts, meta information can however be easily modified or deleted with easily available image processing toolboxes.

Figure 4 illustrates the origin of device specific characteristics in the simplified model of a digital camera. Starting with the lens, characteristics like chromatic aberration (Johnson 2006) and radial lens distortion (Choi 2006) are introduced. Chromatic aberration for example is the result of the lens' incapability to refract light of different wavelengths to the same point at the sensor and becomes visible as coloured artefacts especially on edges and straight lines. Further characteristics are introduced by the sensor, namely, sensor defects and sensor noise.

Due to minor inaccuracies during the manufacturing process, these characteristics are unique for each sensor and, therefore, enable not only the separation between different devices but also the determination of a unique imaging sensor. Dependencies between adjacent pixels due to the need of colour interpolation (Popescu 2005) and differences in JPEG compression (Farid 2006) form other typical ingredients for forensic methods. The occurrence of such device specific characteristics in an image under investigation can be estimated to extract information about the source device and, furthermore, can be tested for integrity to detect image manipulations. However, it is also possible to consider the whole image acquisition process as a black box and analyse the device response function (Lin 2005) or macroscopic features of acquired images (Kharazzi 2004).

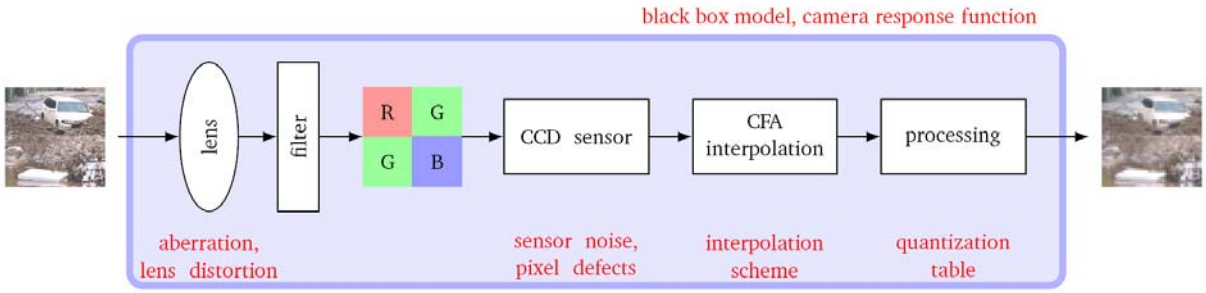

Figure 4: Origin of device specific characteristics in a simplified digital camera model

Contrary to techniques which rely on device specific characteristics, forensic methods based on manipulation artefacts are applicable without knowledge of the digitisation device used. For example, it is possible to reveal pixel dependencies introduced during resizing or rotation of images (Popescu 2005). Other methods used are, for example, statistics of JPEG coefficients to detect recompression (Lukas 2003), inconsistencies in lighting to detect copy forgeries (Johnson 2005), or analysis of phase congruency to detect image splicing (Chen 2007).

It is important to note that in the existing body of literature there is a lack of rigorous discussion of robustness against strategic counterfeiters who anticipate the existence of forensic techniques. As a result, the question of trustworthiness of digital image forensics arises. Forensic methods might benefit from *research on countermeasures* in a similar way as reasoning about attacks in multimedia security in general is useful to improve security. In this sense, attacks on image forensic algorithms can be understood as schemes to systematically mislead the detection methods. In general, such attacks can be assigned to one of the following three objectives: the camouflage of malicious post-processing or tampering of an image, the suppression of correct image origin identification, and furthermore, the forgery of image origin. Initial investigations on attacks against both a source identification scheme (based on sensor noise) and a manipulation detector (based on resampling artefacts) showed that it is in general possible to achieve these goals by spoofing the specific characteristics used (Gloe 2007).

Consequently, current forensic methods, which doubtlessly show very promising results, should be further investigated and extended to be able to cope with, or at least be sensitive to attacks. Although in the future it might be possible to generate digital signatures or digital watermarks directly during the image acquisition inside the device, digital image forensics will still form an important building block for the authentication of digital images as it allows

a highly practicable analysis of digital images with in general no limiting technical constraints.

4.2 Tracking people and cars using 3D modelling and CCTV

Gerda Edelman, NFI

4.2.1 Introduction

In forensic casework, CCTV footage can provide useful information about the crime, perpetrator or witnesses. However, with the growing number of security cameras, the amount of information increases rapidly. The question arises if surveillance images could be used more effectively with the help of 3-dimensional models of the scenes that are visible in the surveillance images.

At the Netherlands Forensic Institute a project is being carried out that aims at the 3D reconstruction of all movements of people and cars before, during and after a big incident from analysis of all available video footage. Methods developed in this project are being applied in forensic casework. A case example is presented in this chapter.

Figure 5: Images of six different security cameras, confiscated by the police

4.2.2 Case example

After a shooting incident, police investigators confiscated footage of six private security cameras in the surrounding area, from each of which a single frame image is shown Figure 5. On these videos many people and cars were visible, all of whom were of potential interest. For scenario testing and the validation of testimonies a 3D reconstruction of all movements was made.

First, a 3D model of the scene was made by aerial stereo photography, in which virtual cameras were created at the positions of the real security cameras (see figure 6 for a top view). The virtual camera parameters (location, focal length, orientation) were adjusted until the virtual camera view was close to the real camera view. An example of a camera view and its virtual imitation is shown in Figure 7. There was some overlap between the different camera views.

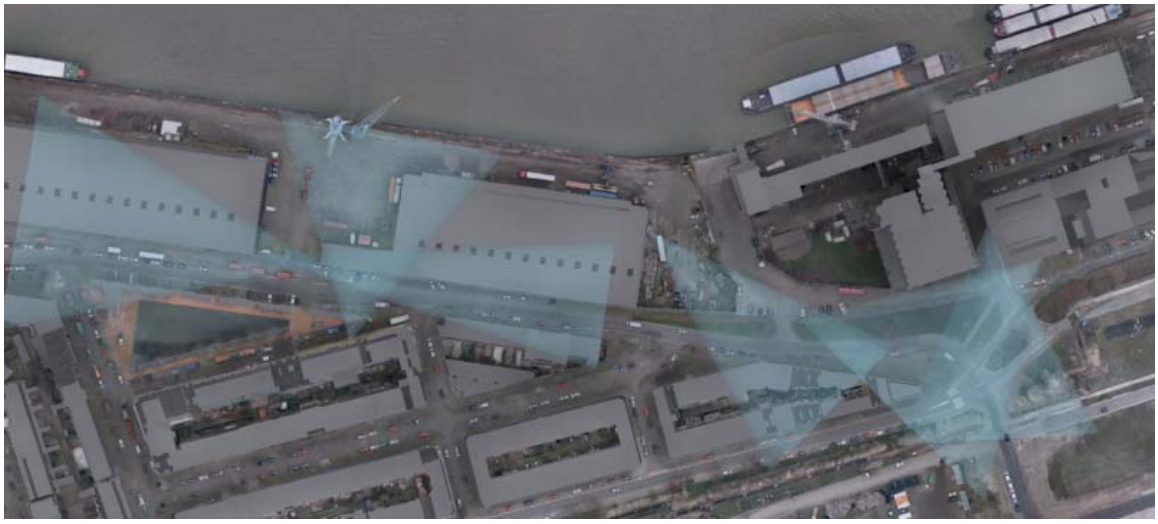

Figure 6: Top view of the 3D model showing the area around the crime scene. Cones demonstrate fields of view of the security cameras.

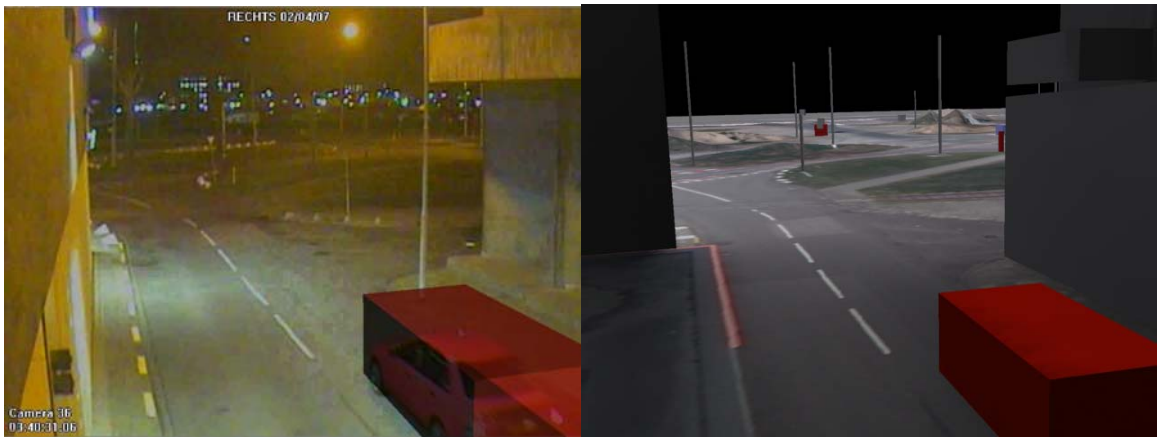

Figure 7: Example of a camera image and its virtual imitation. At the position of a car a box is created

In the 3D model, boxes and cylinders were created at the positions of cars and persons that were visible on the camera images (see Figure 7). When the car/person moved, the box/cylinder was moved to the same location. In this way a 3D animation was created of all movements observed in the area of surveillance, which could be watched from different perspectives.

4.2.3 Discussion

The 3D model allows more insight into the situation because an overview of the scene can be given. In this case, the 3D reconstruction gave insight in the relations between cars seen on different cameras. Because the fields of view of the cameras partly overlapped on the streets, cars could be tracked over six different cameras. However, when a car went out of sight and came back later, it was not possible to identify it again, because of the low quality of the images. On the path areas there was hardly any overlap between different camera views. This, in combination with the low resolution of the images and the size of a person, made it more difficult to track pedestrians on multiple cameras. Nevertheless, more awareness could be created of the locations of people because the scene could be watched from any point of view. This made it possible to verify different scenarios and testimonies.

Another advantage of this method is that it can be combined easily with other spatio-temporal information. Cell sites, for example, can be added to the 3D model, so that evidence concerning telephone calls can be compared with the positions of different people at the same time.

Also, the analysis is objective and can be compared to results of other investigators. Differences in interpretations can be visualized and separated from facts in the animation. This is of high importance in forensic investigations.

4.3 Setting up a centre of expertise on intelligent data analysis

Gert Jacobusse, NFI

Within the NFI (Netherlands Forensic Institute), the knowledge and expertise centre for intelligent data analysis (Kecida) is being set up. The mission of this centre is to advise government agencies in choosing software, methods and techniques for intelligent data analysis. Complementary to that, the centre aims to support the implementation of processes for analysing and merging large amounts of digitally stored data.

Without working together, the scale of individual agencies is often too small to efficiently organise the required capacity and facilities to build knowledge. Also, many efforts are carried out multiple times by separate individual agencies. It is expected that the centre will increase the knowledge about intelligent data analysis within the Dutch government, by taking initiative and providing opportunities to let agencies combine their efforts to build knowledge. The centre already initiated several activities to achieve its aims. Diverse methods, techniques and software products that may help to analyse data in the battle against criminality, fraud and terrorism are closely followed by studying literature, searching the Internet, inviting software suppliers, and consulting experts from universities. To follow up on that, some of the techniques and software products are being tested in “learning, discovery- and evaluation” trajectories together with other government agencies who want to explore the techniques.

One of the first challenges during the actual testing of methods with real data is the need for a safe environment with powerful computers, knowledgeable and trusted employees, and state-of-the-art analysis tools. To avoid interference with the operational environment, we supply a dedicated computer environment for pilot projects.

Finally, the knowledge about the techniques is shared with the government agencies that need it. This is done in different ways. A minimal option is to establish contact between the agency

and a software supplier, or between the agency and an expert from university. Another option is to cooperate with a software supplier to provide demonstrations or software training to people from different agencies at once. Apart from the benefit of more efficient organisation, an additional benefit is that people from different agencies who use the same techniques are brought together.

Knowledge is also shared by applying and advancing it in co-operational pilot projects. The most active option is to directly exploit the knowledge that is built up within the centre. This is done in two ways: first by gathering the knowledge in a database that can be shared with other agencies. Second by inviting people to visit presentations and software demonstrations in which the results of pilot projects are spelled out.

After less than one year it is already clear that the awareness of the need for a knowledge and expertise centre is broadly shared among people from government agencies that work on intelligent data analysis.

4.4 Example of intelligence management system through forensic profiling: drug profiling

Olivier Ribaux and Sylvain Ioset, University of Lausanne

The systematic chemical and physical analysis of illicit drugs seized by law enforcement agencies has greatly developed since the middle of the nineties (Guéniat and Esseiva 2005) (Ioset *et al.* 2005). Illicit substances are seized, transferred to laboratories, and analysed in order to extract a "profile" (list of chemical substances and their quantities). The profiles are then recorded into a data base which is exploited in an intelligence or investigative perspective. For instance the process of linking illicit substance seized in different circumstances may lead to concentrate attention to a specific organised network while they were previously the object of separated investigations. Other indications about cultivation (origin), manufacture processes, or the distribution process of illicit drug trades can be inferred through the systematic analysis of the data base.

The data is organised into a dynamic memory: seizures are not stored individually but are rather collated and grouped into "classes" mainly according to similarity measurements between profiles coming from different seizures (Dujourdy *et al.* 2003; Esseiva *et al.* 2003). Depending on which basis they are formed, these clusters mainly indicate similarities in the traffic at different levels, from the cultivation (origin) to the distribution of the illicit substance.

Beyond standard clustering methods, other original methods for detecting patterns have been tested, particularly through spatio/temporal and graph visualisations. For instance, combinations of cutting agents are often used by drug smuggler before the distribution in the street. The spatio/temporal evolution of these co-occurrences inform on the dynamics of the local market (Terrettaz-Zufferey *et al.* 2007).

However, there is evidence that each drug trafficking network and laboratory develop its own receipts and methods that reflect differently into the intrinsic structure of the chemical profiles (correlations between variables). Thus, there is no suitable universal metric that can be defined, except for those specificities, and can systematically provide the same reliability when measuring proximity between samples. There is a need for a typical learning process as

"classes" or specific groups profiles evolve over time, and show an inherent structure that may in turn influence the classification of new data.

This hypothesis has been tested with data coming from known solved cases. Spectral clustering and its variants have been chosen to train the system and have shown to substantially improve the classification process (Ratle *et al.* 2007). How those ideas may lead to the development of unsupervised methods is now the object of further developments.

However, even if comprehensive European projects have led to some harmonisation and extension in the use of the method, in particular in the field of amphetamines (Aalberg *et al.* 2007a; Aalberg *et al.* 2007b; Andersson *et al.* 2007b; Andersson *et al.* 2007a; Andersson *et al.* 2007c; Lock *et al.* 2007), far from the whole potential of the approach being exploited. In fact, the central question is how to integrate knowledge extracted from drug profiling data bases with the analysis of other (traditional) sources of information (geopolitical, coming from investigations, etc.). Full aggregation of data, even theoretically ideal, can now be difficult to imagine as organisations that deal with the set of data are different (mostly forensic laboratories and the police), cover different countries and are based on different specialities. A more pragmatic model consists in the development of communication channels between partners organised as a network. For instance, chemical links can be systematically provided to the police and used in the investigative process. Conversely, investigative hypothesis can be tested through chemical profiling (Ioset *et al.* 2005). This integration process must attract much more attention than the lack of communication between the organisations actually allows in practice (police, forensic laboratories and Universities).

5 Legal implications of forensic profiling: of good old data protection legislation and novel legal safeguards for due processing.

Fanny Coudert (ICRI, Katholieke Universiteit Leuven) and Katja de Vries (Vrije Universiteit Brussel), Jacek Kowalewski (Katholieke Universiteit Leuven)

Human rights are not simple timeless pre-givens. A famous example hereof is the right to privacy – defined as the right "to be let alone" – which emerged in the late nineteenth century in the light of new technologies like "instantaneous photographs and newspaper enterprise" (Warren and Brandeis, 1890, p. 195). Thus, sometimes new technologies may create a need for the protection of new rights.

In the last decades, computing power has increased enormously. At the same time several structural changes have occurred within the law enforcement field: organised crime has become an issue of international concern; police does not hold only data about the people they suspect of having committed a criminal offence but also about the average citizen; criminal intelligence files have evolved so as to take a significant place in law enforcement policies; and finally, databases are interconnected so the information and intelligence does not remain separate from the criminal records anymore (Second evaluation report of the relevance of recommendation N° R (87) 15, Council of Europe, 1998). Thus, forensic profiling technologies are today no longer limited to the classical retrospective linkage of one particular trace (e.g. a fingerprint) with one particular suspect – as it has become possible on the one hand to *interconnect and compare* extensive databases in order to find evidence about crimes and, on the other hand, to apply *data mining and risk profiling techniques* in order to detect abnormal patterns with *crime prevention* purposes. As a result of the interconnection of databases both the classical *retrospective* (solving crimes which have already taken place) as well as the new *prospective* outlook of forensic profiling (preventing crimes which might take place) have seriously expanded their scope: pointing out to persons who otherwise might have been completely unsuspected.

Data protection laws ensure a series of safeguards to limit the expansion of police powers with regard to the processing of personal data which could cause harmful consequences on the rights and freedoms of individuals. However, the aforementioned changes in the main structure of law enforcement techniques and tools have highlighted insufficiencies in the protection brought by data protection legislations and forces present day lawyers and politicians to rethink if these laws are sufficiently resilient to answer these new technological challenges. In that sense, Commissioner Frattini recalled that "The protection of fundamental human rights such as privacy and data protection stands side-by-side public safety and security. This situation is not static. It changes, and both values are able to progress in step with technological advances. But it also means that there must be lines which cannot be crossed, to protect people's privacy" (Franco Frattini, 20 November 2007).

This chapter will examine the new threats posed by the use of profiling techniques to fundamental freedoms and the current answers provided by data protection laws, pointing out their insufficiencies. The first two sections provide a general overview of forensic profiling (section 5.1) and the existing legislation with respect to data protection (section 5.2). After these introductory sections the problems for constitutional democracy posed by two new aspects of forensic profiling technology – the interconnection of databases and risk profiling –

are analysed more in-depth. Section 5.3 focuses on the *interconnection* of databases and its legal implications, while the following two sections provide an analysis of *risk* profiling: section 5.4 exploring the insufficiencies in the existing legal instruments with respect to risk profiling and section 5.5 presenting some suggestions for alternative legal safeguards. It is shown how these alternative data protection safeguards ('due processing' and adequate remedies) could play a twofold role in protecting individuals' freedoms: before the decision-making takes place, the integrity of the data and the strict respect of the purpose principle will guarantee the accuracy and legitimacy of the processing; once the decision has been made on the basis of the profile, the data subject should be protected from harmful consequences through a strict application of the principle of transparency and adequate mechanisms of redress. To conclude with, some concluding remarks are made about data protection with respect to the interconnectivity of databases and risk profiling.

5.1 Forensic profiling the old and the new way

Forensic profiling is a practice in the field of law enforcement which could be described as "the process of inferring a set of characteristics (typically behavioural) about an individual person or collective entity and then treating that person/entity (or other persons/entities) in the light of these characteristics" (Bygrave, 2002). This broad definition encompasses both the classical practice of 'data matching' and the more recent practice of 'data mining' or 'risk profiling'.

Data matching is the traditional *retrospective* way of offender profiling, *linking individuals with personal identifying data* (Citron, 2007; Steinbock, 2005, p. 3-4), e.g., identifying a suspect by a fingerprint or determining the likelihood that two fingerprints were produced by the same individual. The profiling goal of data matching is "to help investigators examine evidence from crime scenes and victim and witness reports to develop an offender description. The description can include personality traits and behaviour patterns, as well as age, race or geographic location. Investigators might use profiling to narrow down a field of suspects or figure out how to interrogate a suspect already in custody" (Psychology and Law enforcement - Criminal profiling, 2004).

However, since it has become possible to *interconnect* databases the practice of data matching is not so "classical" anymore – raising an abundance of new questions on, e.g., the quality of the data processed, referring in particular to their accuracy and legitimacy, and how the purpose limitation principle applied to the information standing the different databases should be respected. In this respect it is good to keep in mind that "forensic information can be extracted from many electronic devices to be used in Court. However, in the examination process, it is important to consider the likely integrity of the data, i.e. how failsafe the retrieval system is, since this will undoubtedly have an impact on the identity of the real person involved as a suspect" (De Hert, presentation 4). These issues surrounding the interconnection of databases (influencing both classical data matching as well as the new form of risk profiling described hereunder) will be discussed in more detail in section 5.3.

Apart from 'data matching' forensic profiling also encompasses the *proactive* practice of 'data mining' or 'risk profiling', i.e., finding *patterns* and *correlations* in large databases, inferring a certain algorithmic profile from these patterns and correlations and subsequently identifying people who fit these computer-generated profiles (Citron, 2007; Steinbock, 2005, p. 3-4). A notable example resides in the surveillance of flight passengers where "the profile

constitutes the basis for decisions on fly/no-fly, arrest, detain for questioning and so on" (Steinbock, 2005, p. 3-4; see also e.g.: http://www.frontex.europa.eu). In this way risk assessment techniques are used for the purpose of crime prevention. Thus, in the case of risk profiling, the problems in terms of data protection do not only reside in the interconnections of databases but also on how to protect individuals against arbitrary decisions, information asymmetry and abusive and overall surveillance. In order to prevent the negatives to overweight the positives, adequate safeguards should be installed. The specific dangers arising from risk profiling will be discussed in section 5.4. However, first we will turn our attention to the existing data protection instruments.

5.2 Scope of application of the different data protection instruments

In the last decade there has been a lot of debate on the processing of personal data processed in the framework of police and judicial cooperation in criminal matters. The Council of Europe recommended already in 1998 to adopt an additional instrument with regard to data collected and processed for the purpose of suppressing criminal offences. Since then several legislative developments have been taking place within the European Union. A Council Framework Decision on the exchange of information and intelligence between law enforcement authorities of the Member States was approved in 2006 (Council Framework Decision 2006/960/JHA of 18 December 2006 on simplifying the exchange of information and intelligence between law enforcement authorities of the Member States of the European Union). In 2007, two political agreements were reached on the integration of the Prüm Treaty[15] to the European Union framework (Initiative of the Federal Republic of Germany on the stepping up of cross-border cooperation, 9 November 2007) and on a general framework on data protection in the third pillar (Proposal for a Council Framework Decision on the protection of personal data processed in the third pillar, 11 December 2007). However, the multitude of initiatives increases the level of complexity of the legislation and the risk of creating loopholes in the protection. The different level and nature of crime in the different countries as well as the varying pressing social needs result in substantial differences in European countries in the dividing lines between data protection, criminal procedure and rules organising the police (Stepping up of cross-border cooperation, procedure file on the Prüm Treaty, 2007) hindering the approval of uniform rules at international level. In the remainder of this section we will take a closer look at the legislative instruments for data protection in the third pillar.

5.2.1 Criminal data as personal data

This chapter focuses on criminal data, i.e., data which are collected and processed for the purpose of suppressing criminal offences. This includes not only the data gathered in the course of a criminal investigation where there are reasonable grounds for suspicion against an individual but also data collected for purposes of criminal intelligence.

[15] Convention between the Kingdom of Belgium, the federal Republic of Germany, the Kingdom of Spain, the French Republic, the grand Duchy of Luxemburg, the Kingdom of the Netherlands and the Republic of Austria on the stepping up of cross-border cooperation, particularly in combating terrorism, cross-border crime and illegal migration, signed by the contracting parties in Prüm (Germany) on the 27 May 2005.

Data protection legislation will only apply to processing that involves the analysis of personal data. This concept refers to any information relating to an identified or identifiable natural person (the so-called, "data subject"). According to the EC Directive on Data Protection (Directive 95/46/EC) an identifiable person is one who can be identified, directly or indirectly, in particular by reference to an identification number or to one or more factors specific to his physical; physiological, mental, economic, cultural or social identity. To determine whether a person is identifiable, account should be taken of all the means likely reasonably to be used either by the controller or by any other person to identify the said person (Recital 26). It follows that 'what is of legal importance is the capability or potentiality of identification rather than the actual achievement of identification.'

Especially in the case of forensic data matching techniques it might quite frequently be the case that it is unclear if the data involved could be qualified as personal data. Such retrospective profiling will often involve evidence not related to individuals from the crime scene which could consist in digital footprints, photographs of the scene, etc. However, according to the interpretation of the Working Party 29, when the data processing only makes sense if it allows the identification of specific individuals and treatment of them in a certain way, it should then qualify as processing of personal data. All forensic profiling would thus fall under the scope of application of data protection legislations.

5.2.2 Data protection instruments applicable to forensic profiling

Directive 95/46/EC, the so-called "data protection directive", exclusively applies to Community Law, i.e. it excludes processing operations concerning public safety, defence, State security (including the economic well-being of the State when the processing operation relates to State security matters) and the activities of the State in areas of criminal law. It follows that police files will be excluded from its scope of application. However, criminal data processed by private companies or other public authorities will be included. This will be particularly the case when legal provisions mandates the retention of specific data for law enforcement purposes, e.g. the retention obligation of communication data by Internet Service Providers made by the Data Retention Directive (Directive 2006/24/EC) before they are being transferred to law enforcement authorities.

Police files are not however left without any protection as long as both Article 8 of the European Convention of Human Rights (hereinafter referred to as of "ECHR") and the Council of Europe Convention for the Protection of Individuals with regard to Automatic Processing of Personal Data, ETS n°108 hereinafter termed 'CoE Convention') remain fully applicable. The legitimacy of intrusions in the right to privacy will be assessed against the criteria set up by Article 8 ECHR and the jurisprudence of the European Court of Human Rights. Such intrusion should be in accordance with the law and necessary in a democratic society in the interests of, amongst others, national security, public safety or for the prevention of disorder or crime. This CoE Convention establishes a series of data protection principles which have inspired most of European data protection systems, including the Data Protection Directive. The convention gives clear and precise indications on the purpose to be achieved by each principle, but leaves to each Party the definition of the best way to implement its principles into national law.

In application of this Convention, the Council of Europe has adopted a specific recommendation on the use of police data (Council of Europe, Recommendation nº R (87) 15,

17 September 1987) which translates the general principles set up by the Convention into more specific guidelines applicable to processing with law enforcement purposes. Both norms are referred to by other instruments regulating law enforcement activities and constitute the frame of reference.

Finally, three Council Framework Decisions more specifically related to the exchange of personal data within EU member States should be mentioned:

- The Council framework Decision 2006/960/JHA of 18 of December 2006 aims at regulating the exchange of information and intelligence between law enforcement authorities of the Member States. It covers any type of information or data which is held by law enforcement authorities and by public authorities or by private entities and which is available to law enforcement authorities without taking coercive measures. It however does not contain specific provisions on data protection but relies on the existing national legislation, transposing the CoE Convention and the Recommendation R (87) 15 and the principle of mutual recognition.

- The Council reached a political agreement in June 2006 on a Decision on the stepping up of cross-border cooperation, particularly in combating terrorism and cross-border crime. This Decision contains provisions based on the essential parts of the Prüm Treaty and is designed to improve the exchange of information between authorities responsible for the prevention and investigation of criminal offences. To this end, the Decision contains rules in the following areas:

  - On the conditions and procedure for the automated transfer of DNA profiles, dactyloscopic data and certain national vehicle registration data.
  - On the conditions for the supply of data in connection with major events with a cross-border dimension;
  - On the conditions for the supply of information in order to prevent terrorist offences, and on the conditions and procedure for stepping up cross-border police cooperation through various measures (Stepping up of cross-border cooperation, procedure file on the Prüm Treaty).

The decisions refer to the CoE Convention and the Recommendation R (87) 15 and complement this framework by sector specific rules.

- The draft framework decision on data protection in the third pillar (Proposal for a Council Framework Decision on the protection of personal data processed in the third pillar, 11 December 2007) intends to regulate the exchange of personal data between European law enforcement authorities, setting up a series of data protection principles applicable to third pillar activities. However, it fails to provide a comprehensive set of rules applicable to third pillar activities in the same way as the Data protection directive did in the first pillar and contains major derogations to data protection principles. Furthermore, the Working Party on Police and Justice recently stressed “that for certain aspects the current text of the proposal does not provide for the same level of protection as defined in Convention 108. This certainly

seems to be the case with the provision on the further use of data received from a Member State (Articles 3 and 12) and the right of access (Article 17)"[16].

5.3 Forensic profiling and the interconnection of police databases

5.3.1 Accuracy of the information processed

The specific nature of law enforcement activities calls for the processing of different kinds of personal data whose accuracy and reliability is not always guaranteed.

First, depending on their trustworthiness, these personal data could be divided into "hard data" and "soft data". "Hard data" are data flowing from a well established source. "Soft data" are very vague indications about somebody's possible involvement with serious crime. They can stem from an anonymous source, resulting in complete uncertainty about its trustworthiness (Second evaluation report of the relevance of recommendation N° R (87) 15, Council of Europe, 1998).

Second, depending on to whom they are referring, data can be classified into "about persons suspected of having committed a specific crime" or "about persons about whom there are indications that they are involved in committing or preparing a serious crime, either as part of an organisation or alone" (Second evaluation report of the relevance of recommendation N° R (87) 15, Council of Europe, 1998).

Both "soft data" and the data about persons not being the subject of founded suspicions as defined by the Codes of criminal procedures are proper to criminal intelligence activities where lower level of safeguards are tolerated. As highlighted by the Council of Europe, "as police and judicial powers in most Codes of criminal procedures are limited to cases where there is a suspicion against a person with regard to a specific criminal offence, new information technology is increasingly used to store data about criminals as person as such, without relation to specific criminal offences. (…) The data are used to solve any crime, either already committed or expected to be committed in the future. Their use is not limited to the investigation of, or use as evidence in, a specific criminal offence." General data protection principles apply to such processing.

Article 5 of the CoE Convention (1981) states that personal data undergoing automatic processing shall be obtained and processed fairly and lawfully. This mainly raises the question with regard to criminal intelligence as of who can be a data-subject as part of criminal intelligence. The Council of Europe however does not get into such evaluations and leaves to each Member State the task to define the criteria for identifying the targets that can be subject of criminal intelligence.

An interesting example resides in the UK DNA database. This database is the largest in the world covering details about 4.5 million people including information on every person arrested, convicted or not, and on 900,000 children. "The high rate of inaccuracies, e.g. incorrect dates, spelling mistakes and duplications established by the Data Quality and Integrity Team Unit, could lead to innocent people being accused of crimes and wrongly arrested" (House of Lords, "Surveillance and data collection", 14 November 2007).

[16] Comments with respect to the proposal for a Council Framework Decision on the protection of personal data processed in the framework of police and judicial cooperation in criminal matters, The Working Party on Police and Justice, 7 November 2007, available online at: http://www.statewatch.org/eu-dp.htm.

These specificities have motivated the formulation of specific rules in Recommendation R (87) 15 (Council of Europe, 17 September 1987) which "advocates for police bodies, as far as possible, to distinguish data according to their degree of accuracy and in particular between data based on facts and data based on opinion. The same distinction is advocated before any data transfer to a third party where the degree of accuracy of the information should be indicated, as far as possible, to the recipient" (Kosta et al., 2007). However, "the initial distinction between different categories of data according to their degree of accuracy and reliability and between categories of data subjects (criminal, suspect, victim, etc.), which was included in the Commission proposal to address this concern, has been omitted from the later versions of the draft decision" (Kosta et al., 2007). In this respect, it is worth noting that the latest Proposal for a Framework Decision on data protection in the third pillar (Proposal for a Council Framework Decision on the protection of personal data processed in the third pillar, 11 December 2007) stipulates that the receiving body cannot use the inaccuracy of the data supplied as grounds to evade its liability vis-à-vis the injured party under national law. If damages are awarded against the receiving body because of its use of inaccurate transfer data, the body which supplied the data shall refund the amount paid in damages to the receiving body in full (Article 19).

5.3.2 Re-use of personal data

The interconnection of databases raises the problem of the re-use of personal data originally collected for one purpose for a different one. The personal data can be obtained from other police databases either internal or external to the law enforcement bodies carrying out the investigation, or from private databases held by private companies (financial information, communication data, etc.).

Principle 2 of the Recommendation R (87)15 states that the collection of personal data for police purposes should be limited to such an extent as is necessary for the prevention of a real danger or the suppression of a specific criminal offence. Furthermore, in accordance with Article 5 of the CoE personal data undergoing automatic processing shall be stored for specified and legitimate purposes and not used in a way incompatible with those purposes. In case the second purpose cannot be acknowledged as compatible, it will form a different processing from the original ones and should be grounded on a different, explicit and legitimate purpose. Derogations are foreseen when it constitutes a necessary measure in a democratic society in the interests of protecting State Security, public safety or the suppression of criminal offences (Article 9). These principles will apply in a different way depending on the source of the data and the purpose of the processing.

Personal data obtained from other internal police databases

The question which immediately pops up when discussing the re-use of personal data is whether personal data collected within one criminal investigation could be used in another. When, from the data collected within an investigation, there are enough indications to base a reasonable suspicion to investigate a new unrelated offence, the processing of such data will be deemed compatible (Second evaluation report of the relevance of recommendation N° R (87) 15, Council of Europe, 1998).

However, the question of the compatibility of further use of such personal data for the solution of possible future offences is more complicated and will depend highly on the

balance made in each national legislation. When the personal data, e.g. fingerprints or photographs, relate to a suspect or a person who is convicted afterwards their conservation for possible future offences may be considered as compatible. The Council of Europe however notes that "there is however divergence with regard to the necessity of deleting such data in cases of acquittal by lack of evidence though the suspicion remains. But it is less questionable when somebody's innocence has been established. The conservation of personal data related to persons other than the suspect or the convicted person should be deleted, their further use would be deemed incompatible, unless a legal basis exists to ground such conservation or processing."

In that sense, the Council of Europe has recommended that "any power to perform a general data surveillance check or matching for the purposes of the suppression of crime on the basis of police data gathered in the course of criminal investigations on the basis of vast amounts of persons possibly completely unrelated to any crime, be limited to specific cases described in the Code of Criminal Procedure and be granted on the basis of a specific mandate of the judiciary."

Personal data obtained from police databases held by law enforcement authorities from other Member States

A different problem arises when the personal data are obtained from another law enforcement authority. The CoE Recommendation R (87) 15 already has set up rules for the communication of personal data between law enforcement authorities. However, in 2007, three different Council framework decisions regulating this topic have been approved or agreed on.

The Framework decision on personal data in the third pillar first limits further processing of the data to compatible purposes provided that the law enforcement authority is entitled by a legal provisions to carry out such processing and finally that this processing is necessary and proportionate (Article 3). However, in the words of the European Data Protection Supervisor, "article 3 is far too broad and does not cover an appropriate limitation of the purposes for storage, also required by Article 5(b) of Convention 108, mentioned above. The general reference to the purposes of Title VI of the EU-Treaty can not be seen as specified and legitimate purposes. The purpose of police and judicial cooperation is not by nature legitimate, and certainly not specified. Article 3 does not contain any derogation as would be possible pursuant to Article 9 of Convention 108. However, Article 12 of the proposal lays down a very broad and not clearly defined series of derogations to the purpose limitation principle in the context of personal data received from or made available by another Member State. In particular, the condition that derogations shall be necessary is not explicitly laid down in the article." Therefore, the EDPS stresses "that this broad and open derogation does not fulfil the basic requirements of adequate data protection and even contradicts the basic principles of Convention 108."

With regard to criminal intelligence data, the Council Framework Decision 2006/960/JHA of 18 December 2006, despite referring to the aforementioned Recommendation, specifies that information and intelligence provided under its provisions may be used by the competent law enforcement authorities of the Member State to which it has been provided solely for the purposes for which it has been supplied or for preventing an immediate and serious threat to public security. Processing for other purposes should be permitted solely with the prior authorisation of the communicating Member State and subject to the national law of the

receiving Member State. Finally, when providing information and intelligence, the providing competent law enforcement authority may pursuant to its national law impose conditions on the use of the information and intelligence by the receiving competent law enforcement authority. The Decision on the stepping up of cross-border cooperation, particularly in combating terrorism and cross-border crime implements indirect access to the personal data through reference data. It contains similar provisions to the Decision 2006/960/JHA. Article 26(1) 1) allows processing for other purposes only if this is permitted under the national law of both the supplying and the receiving Member State.

Personal data obtained from private companies

As pointed out by the European Data Protection Supervisor, "there is now a trend to impose cooperation for law enforcement purposes on private actors on a systematic basis" (Opinion on the draft Proposal for a Council Framework Decision, EDPS, 20 December 2007). Several Directives foresee mandatory retention of specific data collected by the private sector in the course of its commercial activity to be used for law enforcement purposes. One of those directives, for instance, is the EC Directive 91/308/EEC, of 10 June 1991 on prevention of the use of the financial system for the purpose of preventing criminal offences, whose article 6 compels to the retention of certain financial data. These data are collected for the suppression of a specific category of crime, irrespective of the fact if the persons are subject of a reasonable suspicion of committing these crimes. These data should not be used for other purposes, unless explicitly permitted by the law. As highlighted by the Council of Europe, "for a specific area there is thus general data surveillance of the population for the purpose of the suppression of a specific form of crime according to specific criteria. The question to be answered is whether and to what extent the police have access to the data gathered" (Second evaluation report of the relevance of recommendation N° R (87) 15, Council of Europe, 1998).

In that sense, the Council of Europe recommends these processing to be explicitly covered by legal provisions defining the criteria and the purposes (Second evaluation report of the relevance of recommendation N° R (87) 15, Council of Europe, 1998).

5.3.3 Storage of personal data

Now that we have explored the accuracy (section 5.3.1) and the re-use (section 5.3.2) of personal data, we turn to a final crucial issue with regard to the interconnection of police databases: the *storage* of personal data.

Personal data should be preserved in a form which permits the identification of the data subjects for no longer than is required for the purpose for which those data are stored. This obligation is difficult to apply in the specific case of criminal intelligence. An example which exemplifies this difficulty is to be found in the EU draft Proposal for a Council Framework Decision on the use of Passenger Name Record (PNR) data for law enforcement purposes (Proposal for a Council Framework Decision on the use of Passenger Name Record (PNR) data for law enforcement purposes, European Commission, 22 October 2007) which foresees the retention of the PNR data for a period of 13 years. This period is judged excessive by the European Data Protection Supervisor (Opinion on the draft Proposal for a Council Framework Decision, 20 December 2007).

The definition of the period of storage will depend on the purpose of the processing as well as on the status of the data subject (i.e., whether the data subject is a victim, a suspect, a convicted, a witness, etc.). The Council of Europe advocates for the deletion to occur after some years after the last time any relevant data has been added to the record. After this period a periodic review could be realised (as done in article 112 of the Schengen Agreement) and, if there are no reasonable grounds to justify further storage then deletion should be the rule.

Another important issue resides in the relevance of the personal data processed. The Council of Europe pointed out the fact that "sometimes, the police, in order to do their work properly have to collect vast amounts of data either by downloading computers during searches in premises, by intercepting communications or by searching the emails of criminals. The storage can only be justified for the time needed to find out that they are really unrelated, unless other compatible use or other use explicitly permitted by law come in view" (Second evaluation report of the relevance of recommendation N° R (87) 15, Council of Europe, 1998). In the Campbell case, the European Court of Human Rights judged that 'the existence of facts or information (should) satisfy an objective observer' that there is reasonable cause to use such data for the purpose of combating crime (Campbell v. United Kingdom, European Court of Human Rights, 1992).

5.4 Risk profiling

5.4.1 Risk profiling – acting proactively on information inferred from aggregative data

After our exploration of the interconnection of police databases in section 5.3, we now turn to another new development in forensic profiling: so called *risk profiling*. There are at least *two* aspects of forensic risk profiling which make it profoundly different from other methods of criminal investigation (i.e. classical forms of *data matching*: even from very advanced data matching like, e.g., searching in several interconnected databases for a certain shoe print). Two central characteristics of forensic risk profiling are that it involves (a) hypothetical information derived from aggregative data and (b) its pro-active character. Two other issues that are sometimes at stake in forensic risk profiling are (c) automated decision making, and (d) the opacity of the reasoning involved.

(a) Hypothetical constructions derived from aggregative data

The *first* aspect which makes forensic risk profiling stand out against more classical methods of investigation is the fact that it does not limit itself to uniquely *individual* information (e.g. the fingerprint of one *particular* individual) but that it makes use of statistical information derived from huge databases (e.g. the profile of the *average* terrorist inferred from a certain pattern of correlations). Of course forensic risk profiling might still aim at the identification of one *unique* individual, but the data (that is, the *patterns* of data) to which a particular individual is compared are not a *unique* marker (e.g. *his* fingerprint) but a *hypothetical construction* derived from information from various people (i.e., a *database*). The fact that forensic risk profiling uses a *construction* derived from the data gathered from *more than one individual* makes it likely that the data mining algorithm will get wrapped up in *opaqueness*: not only could disclosure of the algorithm lead to an infringement of the privacy of those people whose personal data were used in the construction of the algorithm, but disclosure

could also make the algorithm valueless (screening passengers for certain criteria might not be so useful if everybody knows exactly which characteristics are sought for) or possibly be a breach of the intellectual rights of the body which constructed the algorithm – sometimes at high costs. A forensic profile might sometimes turn out to be as secret as a 'secret recipe', surrounded by technologically inspired terminology like "data *mining*", "data *harvesting*", "*crunching* raw data", "data *processing*" and "DNA *banking*".

(b) The pro-active character of forensic risk profiling – keeping it outside the investigative process

A *second* peculiarity of forensic risk profiling is the fact that it can be used in a *pro-active* and *hypothetical* way. Instead of looking for an individual matching the traces left at a place of crime, forensic data mining can be used to *prevent* a crime (e.g. not admitting a potential terrorist to an airplane) or to raise a *hypothesis* about the characteristics to look for (e.g. "the robber is likely to drive a red car, because a significant majority of the robbers in our database did"). This pro-active or hypothetical character of forensic risk profiling dissociates it from the investigative process directed at a potential trial. The profile used to detect high risk air-plane passengers is not meant to be used as *evidence* in a criminal trial, but is meant to prevent the high-risk passenger from entering the plane without further screening. The passenger who is told that he cannot enter the plane will often even be unaware of the fact that he was subjected to forensic profiling and simply assume that he apparently looked suspicious. This means that forensic risk profiling will often stay outside, far from the scrutiny and transparency of a trial, because it is not so much part of a criminal *process* as it is a *hypothetical result in itself*: "Data matching and data mining give no process as the law understands that term. There is no notice, no opportunity to be heard, no confrontation with evidence, no reason given – only a result. Under any theory of due process, decisions based solely and irrevocably on the results of data matching or data mining are deficient, where they affect substantial interests" (Steinbock, 2005, p. 45).

(c) Automated individual decisions

Sometimes risk profiling is also embedded in an automated decision system. A notable example resides in the surveillance of flight passengers where "the profile constitutes the basis for decisions on fly/no-fly, arrest, detain for questioning and so on" (Steinbock, 2005; see also e.g., the European Commission Proposal for a Council Framework Decision on the use of Passenger Name Record (PNR) for law enforcement purposes, 6 November 2007.). The problem of decisions with substantial effects on individuals based on automated software has already been debated within the debate relative to the draft of the Directive. This issue has been revived by the use of profiling techniques for law enforcement purposes (i.e., third pillar activities)

The Framework decision on data protection in the third pillar of the EU (i.e., concerning Police and Judicial Co-operation in Criminal Matters) dedicates an article to the taking of automated individual decisions by law enforcement authorities. Article 8 of the decision stipulates that a decision which produces an adverse legal effect for the data subject or seriously affects him and which is based solely on automated data processing for the purposes of assessing individual aspects of the data subject shall be permitted only when the legitimate interests of the data subject are safeguarded by law.

This article echoes article 15 of Data Protection Directive 95/46/EC (made within the first or 'Community' pillar) which had tried to tackle the issue of increased automation in the decision-making process, mainly with regard to organizational decisions. Article 15 aims at protecting "the interest of the data subject in participating in the making of decisions which are of importance to him." This use of extensive data profiles of individuals by powerful public and private institutions risk to "deprive the individual of the capacity to influence decision-making processes within those institutions, should decisions be taken on the sole basis of his "data shadows"" (Bygrave, 2002). The problem of the lack of transparency (see above) was already at the centre of the debate.

A second fear which was expressed in the debates surrounding Data Protection Directive 95/46/EC was that the "automatic acceptance of the validity of the decisions reached and a concomitant reduction in the investigatory and decisional responsibilities of humans" (Bygrave, 2002). According to the European Commission: "the result produced by the machine, using more and more sophisticated software, and even expert systems, has an apparently objective and incontrovertible character to which a human decision-maker may attach too much weight, thus abdicating his own responsibilities" (COM(90)314 final - SYN 287, 13 September 1990, p. 29). As highlighted by Bygrave (Bygrave, 2002) "the increasing automatisation of decision-making processes engenders automatic acceptance of the validity of the decisions reached and a concomitant reduction in the investigatory and decisional responsibilities of humans. Thus, there is an implicit assumption that the reasoning linking the premises and conclusions for these predictive judgments will be grounded in reality."

(d) Opacity of the reasoning

A related issue which is sometimes at stake in forensic risk profiling is the opacity of the reasoning process. The problem arises with regard to the opacity of the algorithms used which are not necessarily connected to truisms about human behaviour. Furthermore, as stressed by Steinbock, contrary to human judgement, "computer analysis has no way to evaluate the probable accuracy of the data on which it relies" (Steinbock, 2005). Computer reasoning is finally also more difficult to evaluate than human assessment (Steinbock, 2005). Whereas the wording of article 15 of the Data protection Directive enables the data subject to be informed of the logic underlying the processing, article 8 of the Framework Decision does not even make a reference to this difficult issue. Once again, the needs of law enforcement activities will need to be balanced with individuals' right and alternative solutions may have to be found when such information should be kept secret. Prior safeguards, such as a strict assessment of the conditions required for the legitimacy of such processing, may be needed. Intervention of independent authorities may be required as well.

5.4.2 Keeping Risk Profiling fair: Due Process?

Due Process: a transparent trial giving the individual a fair opportunity for defence.

The fears about the present day avalanche of data technologies are mainly framed in terms of *loss of privacy* (Lyon, 2007; Travis, 2007). Terms like "Big Brother" and "Surveillance Society" abound and even sometimes risk becoming worn out mantras (Zarsky, 2002-2003, p. 3; for a clarification of some of the fuzziness surrounding the notion of privacy see e.g. Hildebrandt *et al.*, 2005, p. 35-42). Yet, the consequences of the collection, storage and

processing of personal data are not *limited* to a simple dissolution of the respect for private and family life (as described in e.g. art. 8 of the European Convention on Human Rights), nor do they *necessarily* imply a loss of privacy at all (Hildebrandt, 2008, p. 324-326).

What is sometimes overlooked in the debates on data technologies which focus solely on their *intrusiveness* in the private sphere is the *information asymmetry* (Hildebrandt, 2008, p. 324-326) which it might bring. Thus, although a citizen will probably *not* perceive the fact that the government knows his year of birth *as an intrusion to his private sphere*, the use of this piece of information in a profiling practice of whose existence he is not aware, nor accessible to him if he would be aware of it (e.g. a classified algorithm predicting the likelihood that somebody who was born in 1979 is a terrorist) puts him in a situation of information asymmetry. Such information asymmetry can undermine some of the principles on which constitutional democracies are built – i.e., undermine the empowerment of the individual citizen against the force of the State. Especially within the domain of criminal investigation and adjudication, one of the big achievements of Western constitutional thought is to come up with safeguards which give the individual a fair chance to *contest* the allegations made by a State which tends to have a larger amount of resources than the individual and therefore be more powerful.

It is especially the notion of *due process* which is at the core of this empowerment of the *individual* citizen towards the State (in particular expressed by the idea of *equality of arms*): it is seen as *the* adequate shield against over intrusive State power. In Europe, *due process* is identified with the right to a *fair trial* such as described in art. 6 ECHR (which encompasses, e.g., the right to a public hearing before an independent and impartial tribunal within a reasonable time, the right to adversarial process and the presumption of innocence). In the US (Steinbock, 2005, p. 6-7) the principle of 'due process' is mainly to be found in two provisions of the Bill of Rights, i.e. in the 5th and 14th amendments to the US Constitution and, in the context of criminal investigation or prosecution, the 4th amendment (the right to be secure against unreasonable searches and seizures). Even though the words wherein the notion of due process is framed differ, both in Europe and the US due process is a right which should prevent citizens ending up in Kafkaesque opaque situations – where a suspect is not told what the accusation is, where he has no legal assistance or any other means of protecting himself, where the rules of the process are unclear, where the process is held in a language he does not understand, where the judges are bribed or seem to act completely irrational and arbitrary, etc. Central to due process is thus the idea of a criminal process culminating in a *trial* where all the cards are put on the table – a *transparent* situation – in order to give the *individual* citizen a fair opportunity to refute the allegations made against him.

Why Due Process is at odds with Risk Profiling

The pivotal ideas of due process (an investigative process aiming at the production of evidence at a *trial*, in a *transparent* way, where the *individual* can have his say about it) are endangered when forensic risk technologies are applied which are *not* at all aiming for a trial, transparency or to address one specific individual; *not* providing individuals notice and an opportunity to be heard (Citron, 2007). Thus, with respect to forensic risk profiling technologies, due process might not turn out to be sufficiently protective anymore because they could be considered as either *jeopardising due process* or simply just falling *outside the reach of the right to due process*.

However, it is quite difficult to realise that the venerable notion of due process might fall short in the case of some of the modern forensic data technologies.

This can be illustrated by looking at one of the most well known applications of forensic risk profiling: the use of Passenger Name Record (PNR) data in combination with risk assessment profiles for the surveillance of flight passengers. As highlighted by the European Data Protection Supervisor, "suspected persons could be selected according to concrete elements of suspicion included their PNR data as well as on the basis of "patterns" or an abstract profile. The main concern of the EDPS relates to the fact that decisions on individuals will be taken on the basis of patterns and criteria established using the data of passengers in general. Thus decisions on one individual might be taken, using as a reference (at least partially), patterns derived from the data of other individuals. It is thus in relation to an abstract context that decisions will be taken, which can greatly affect data subjects. It is extremely difficult for individuals to defend themselves against such decisions" (Opinion on the draft Proposal for a Council Framework Decision, EDPS, 20 December 2007). In order to make this technology – which is so highly-intrusive nature into individuals' privacy – legitimate the processing of PNR data should be in accordance with the criteria developed by the jurisprudence of the European Court of Human Right. To comply with the general data protection principles, the processing should be transparent to the data subject and adequate legal remedies should be implemented to protect the data subject against arbitrary decisions.

Thus, in this vain the first pillar Data protection Directive foresees several mechanisms to ensure transparent processing. They intend to empower the citizens and give them the possibility to control the processing carried out on them by a first obligation of prior information on the processing and by granting him with rights of access, rectification and deletion. In particular, when exercising his right to access with regard to automated decisions, the controller of the processing should inform the data subject about the logic of the processing, apart from their source and the person to whom they have been communicated.

However, transposing this first pillar logic into the third pillar by simply demanding that forensic risk profiling is made transparent would be contrary to the large part of law enforcement. Risk assessment of flight passengers requires a certain *opacity* – and the same goes for many third pillar law activities. Transparency could make the privacy of the individuals within a database at stake, creating the possibility that 'contestation' by a conscious data subject might turn out to be distortion in disguise and would risk making an algorithm valueless as profiling techniques are frequently applied for investigation and prevention (e.g. surveillance, investigative stops and frisks, searches) which can (or even must) be done without informing the profiled person. So what about *due process* then? The present situation is that forensic risk profiling is a technique which is clearly outside the reach of the classical notion of due process, as well as it is outside the reach of any other form of democratic and constitutional control.

Is it possible to create alternative safeguards which protect and empower the individual citizen without the need to dissolve the opaque and aggregative nature of forensic profiling? In order to answer that question we need to take a closer look at the existing legislation and especially the proposed Frame work decisions.

5.4.3 Critical analysis of the existing safeguards with regard to risk profiling

As was shown above a major issue arising from the use of profiling techniques resides in the (lack of) transparency of the processing. The European Data Protection Supervisor recalls the jurisprudence of the European Court of Human Rights, according to which domestic law may be sufficiently precise to indicate to citizens in what circumstances and on what terms the public authorities are empowered to file information on their private life and make use of it. The information "should be accessible to the person concerned and foreseeable as to its effects, a rule being "foreseeable" if it is formulated with sufficient precision to enable any individual – if need be with appropriate advice – to regulate his conduct" (Rotaru vs. Romania, nº 28341/95, par. 50, 52 and 55).

However, the Framework decision on data protection – which is supposed to clarify further when transparency is required – has met with a considerable amount of critique since it was first put forward by the Commission on 4 October 2005. The European Data Protection Supervisor published two critical Opinions on the proposal (First Opinion of the EDPS on the proposed Framework Decision; Second Opinion of the EDPS on the proposed Framework Decision) in which he voiced his concerns that (Third Opinion of the EDPS on the proposed Framework Decision)

"…developments in the negotiations were leading towards a level of protection of personal data not only below the standards laid down in Directive 95/46/EC, but also incompatible with the more generally formulated Council of Europe Convention No 108".

After the proposed Council Framework Decision was revised by the German Presidency of the European Parliament on 13 March 2007 (Council document 7315/07 of 13 March 2007) the European Data Protection Supervisor published a third Opinion on the 23rd of June 2007 wherein he noted his appreciation for the German attempt but also stated that he was disappointed about the content, which has become according to him:

"…a lowest common denominator approach that would hinder the fundamental rights of EU citizens as well as hamper the efficiency of law enforcement" (Third Opinion of the EDPS on the proposed Framework Decision)

Part of the reason for why the proposal of 13 March 2007 failed to fulfil the expectations was probably the decision-making procedure in the Council which asks for unanimity and leads to this "lowest common denominator". Another likely reason is the political climate ("war on terrorism") which puts more stress on crime control than due process.

However, another cause may have been the fact that the drafters of the first Framework Decision draft of 4 October 2005 seemed to have overlooked the fact that the classical legal protection such as offered by the idea of *due process* (presuming an investigative process aiming at the production of evidence at a *trial*, in a *transparent* way, where the *individual* can have his say about it) will offer no protection in the case of forensic risk profiling. Although *due process* is a right to be cherished in many a context it seems to be inadequate in protecting the individual citizen against the information asymmetries arising from forensic risk profiling technologies.

The extent in which the proposed Framework Decision assumes transparency and the possibility for the individual to stand up for his rights varies slightly between the different

drafts (first draft 4 October 2005; revised draft 13 March 2007; revised second draft 23 October 2007; latest draft 11 December 2007).

The paradox contained within the text of the first draft is the fact that on the one hand it acknowledges the existence of profiling technologies of which the data subject is unaware and which are opaque to him (e.g. art 20 (2) "*Right of information where the data have not been obtained from the data subject or have been obtained from him without his knowledge*"[17]) but on the other hand it still seems to assume an informed data subject standing up for its rights. But how can a data subject *without knowing* that he was profiled know if its rights were infringed? And even *if* the data subject is aware of the fact that he was subject to a forensic risk profiling practice - what use of knowing that your data were processed if you do not know *how*? Who is going to determine if the grounds for an exception to the right of information (*art 19 (2) and art 20 (2)*[18]) are present if nobody who is affected by the profiling practice is aware of the existence of the practice? Who is going to claim the right to be informed about data collection and processing if there is no awareness about its existence?

In the revised German draft and in the latest draft the paradoxical articles 19 and 20 have disappeared: instead of solving the underlying paradox the new texts apparently try to avoid it by reducing the right to information of the data subject. Article 16 (*information for the data subject*) of the latest draft of the proposed Framework Decision (Third Opinion of the EDPS on the proposed Framework Decision, p. 1 ff) is a very much stripped version of the earlier articles:

Article 16: Information for the data subject

1. Member States shall ensure that the data subject is informed regarding the collection or processing of personal data by their competent authorities, in accordance with national law.

2. When personal data have been transmitted or made available between Member States, each Member State may, in accordance with the provisions of its national law referred to in paragraph 1, ask that the other Member State does not inform the data subject. In such case the latter Member state shall not inform the data subject without the prior consent of the other Member State.

---

[17] Article 20(2) *Right of information where the data have not been obtained from the data subject or have been obtained from him without his knowledge*
The information laid down in paragraph 1 shall not be provided if necessary
(a) to enable the controller to fulfil its lawful duties properly,
(b) to avoid prejudicing of ongoing investigations, inquiries or proceedings or the fulfilment of the lawful duties of the competent authorities,
(c) to protect public security and public order in a Member State,
(d) to protect the rights and freedoms of third parties, except where such considerations are overridden by the need to protect the interests or fundamental rights of the data subject.

[18] See also the previous footnote.
Article 19(2) *Right of information in cases of collection of data from the data subject with his knowledge*
The provision of the information laid down in paragraph 1 shall be refused or restricted only if necessary
(a) to enable the controller to fulfil its lawful duties properly,
(b) to avoid prejudicing of ongoing investigations, inquiries or proceedings or the fulfilment of the lawful duties of the competent authorities,
(c) to protect public security and public order in a Member State,
(d) to protect the rights and freedoms of third parties, except where such considerations are overridden by the need to protect the interests or fundamental rights of the data subject.

In article 17 of the same draft of December 2007 we find that the data subject has a right of access '*on request*' – but how can one make requests about something one is not aware about? If the responsible authority considers that there are legitimate grounds to deny the data subject access to the data which were processed, the data subject shall be advised – according to art. 17 (3) – "*that he may appeal to the competent national supervisory authority, a judicial authority or to a court*". Thus, the legality of a profiling practice will only be considered by a competent authority or court if the data subject – who will be normally unaware of the existence of the practice – places a request. Viewed from this perspective article 20 of the latest draft ("*…, the data subject must have the right to seek judicial remedy for any breach of the rights guaranteed to him by the applicable national law*") risks becoming a right that is difficult to exercise in practice.

5.5 Alternative legal safeguards for Risk Profiling: Adequate Remedies and Due Processing.

In this section the twofold role which alternative data protection safeguards (see for other suggestions: Dinant, 2008) could play in protecting individuals' freedoms are presented.

In the first place it will explore what could be done in those cases where the damage has already been done: once the decision has been made on the basis of the profile, the data subject should be protected from harmful consequences through a strict application of the principle of transparency and adequate mechanisms of redress.

In the second place it will present a safeguard for the protection of the freedom of individual citizens before the decision-making takes place: not by due process but by due processing. This would involve an independent board that the processing is done in accordance with the rights of individual citizens: e.g., the integrity of the data and the strict respect of the purpose principle to guarantee the accuracy, accountability and legitimacy of the processing.

5.5.1 Adequate remedies

The aforementioned article 20 of the Framework decision on data protection in the third pillar stipulates that the data subject must have the right to seek judicial remedy for any breach of the rights guaranteed to him by the applicable national law. The data subject should be compensated for the damages resulting from an unlawful processing operation or any act incompatible with the national provisions adopted pursuant to the framework decision.

The definition of appropriate remedies and compensation are let to the domestic legislations. Such remedies should complement the protection granted to the data subject against automated decisions with harmful consequences, such as the possibility to contest the decision, i.e., its result, its logic, the accuracy of the data used for the processing. Indeed, even if this specific safeguards introduced by article 15 of Directive 95/46/EC is not translated to article 8 of the Framework Decision, it participates from the principle of transparency, empowering the data subject to exercise his scrutiny upon the processing of personal data, principle at the core of any data protection legislation. It thus appears important to define the ways how these guarantees could be implemented in the specific field of risk profiling.

Adequate judicial remedies may be of particular importance against abusive use of Passenger Name Record (PNR) for law enforcement purposes as foreseen by the Proposal for a Council

Framework Decision. The use of the PNR is not meant for the identification of individuals but to "contribute to carrying out risk assessment of persons, obtaining intelligence and making associations between known and unknown people" (Proposal for a Council Framework Decision on the use of Passenger Name Record (PNR) data for law enforcement purposes, European Commission, 22 October 2007). The purpose is "to identify persons who are or may be involved in a terrorist or organised crime offence, as well as their associates" (article 3(5)). Recital 9 of the proposal states explicitly that data must be kept for a sufficiently long period as to fulfil the purpose of developing risk indicators and establishing patterns of travel and behaviour

This proposal excludes the possibility of enforcement actions taken by the automated processing of PNR data or by reason of a person's race or ethnic origin, religious or philosophical belief, political opinion or sexual orientation. However, it does not exclude the automated filtering of individuals according to standard profiles, nor does it prevent the automated constitution of lists of suspected persons and the taking of measures such as extended surveillance (Opinion on the draft Proposal for a Council Framework Decision, EDPS, 20 December 2007).

It will thus be necessary to implement adequate and efficient legal remedies to prevent data subjects becoming the victim of mistakes or abuses in such situations. In that sense, Steinbock argues for an elaborate system which would keep a balance between opaqueness and transparency. Instruments in such a system would be not only independent oversight but also e.g. summary hearings, post-deprivation correction rights and compensatory damages (Steinbock, 2005).

5.5.2 Instead of Due Process: Due Processing

As was made clear above the difficulty of regulating forensic risk profiling in a way which is in accordance with the requirements of a constitution democracy, is that it is a technique which is almost intrinsically *opaque* (or even completely invisible) *for the data subject* who is subjected to it. This makes it hard for the data subject to contest the rightfulness of the processing of his data in court.

An alternative approach might be to acknowledge the specific character of forensic risk profiling – a practice somewhere *in between* 'a regulative policy' and 'a step within an investigative process subjected to control by the judicial system'(Citron, 2007) –and the impossibility to address it in a classical due process way because it requires an active, knowing citizen and the possibility of a transparent trial. Next to rights aimed at the individual (*due process*) legislation could be made in order to have some democratic control ("independent oversight of the validity of forensic profiling techniques" Steinbock, 2005) on those techniques while they are constructed and applied (*due processing*):

"…information technology review boards that provide opportunities for stakeholders and public at large to comment on a system's design and testing. One might imagine information technology consultants working on behalf of advocacy groups who would ensure that testing and audit trails employed by contractors comported with best practices. Such boards also could check the accuracy of information stored in databases [….]. Although finding the ideal makeup and duties of such boards would require some experimentation, they would secure

opportunities for interested groups to comment on the construction of automated systems that will have an enormous impact on their communities once operational." (Citron, 2007)

Such legislation on due "processing"[19] would provide a democratic control mechanism for forensic risk profiling, without destroying the opaqueness which is needed by the technique to function properly. Independent controlling bodies concerned with such *due processing* could potentially also stand up for the rights of individual data subject – making *due process* with respect to forensic risk profiling indirectly a realistic possibility again.

5.5.3 Values guiding the Due Processing: Legitimacy and Proportionality

Post 9/11 has seen greater interest in preventing crime, in contrast to the traditional practice of deterrence by reacting to past acts of antisocial behaviour through the criminal process or otherwise. Some data matching or forensic risk profiling results are now being used not only as a reason to begin or intensify investigation but also as the sole basis for decision.

The dangers for the individual citizen stem from the fact that "risk is an invention based on imagined fears and on imaginative technologies for dealing with them. (…) In risk society, policing is not just a matter of repressive, punitive, deterrent measures to control those who are morally wrong. It is also a matter of surveillance, producing knowledge of populations that is useful for administrating them. The focus is on knowledge that allows selection of thresholds that define acceptable risks and on forms of inclusion and exclusion based on this knowledge. (…) Everyone and everything is to be made knowable through surveillance mechanisms. Everyone is presumed guilty until the risk profile proves otherwise" (The United Kingdom Parliament, Home Affairs, Third report, 24 May 2007).

With the help of an independent board such dangers could possibly be lessened. Such a board would need to overlook that the processing is done in accordance with the rights of individual citizens: e.g., the integrity of the data and the strict respect of the purpose principle to guarantee the accuracy, accountability and legitimacy of the processing. However, the evaluation process of such a board could also involve more normative control to assess the legitimacy of the processing.

The principle of proportionality is a common and constant requirement to ground the validity of any measure restrictive of fundamental rights. To assess the legitimacy of the processing against fundamental data protection principles, the processing should pass the proportionality test: it should be adequate to achieve the goal foreseen (*adequacy test*), not being possibly replaced by other less intrusive means at least equally efficient (*necessity test*) and finally to provide sufficient benefits to overcome the negative impact it has on fundamental rights (*proportionality test stricto sensu*). With regard to this last requirement, as indicated by P. Breyer, "the positive and the negative effects of the measure on individuals and society as a whole must be balanced against each other. This cannot be achieved by means of general considerations on the interests and rights in question, since it is impossible to establish an

[19] See for the definition of *processing* art 2 of the latest draft of the proposed Framework Decision:
"processing" shall mean any operation or set of operations which is performed upon personal data, whether or not by automatic means, such as collection, recording, organisation, storage, adaptation or alteration, retrieval, consultation, use, disclosure by transmission, dissemination or otherwise making available, alignment or combination, blocking, erasure or destruction;

absolute order or ranking of interests and rights. Instead, it is necessary to determine how useful the measure will actually be, and what harmful effects it will actually have" (Breyer, 2005). The more severe the infringement of privacy, the more important the legitimate objective in each case will need to be. In most cases, the interference will be judged against whether it meets a pressing social need, and the extent to which an alternative, less intrusive interference would achieve the same result.

The legitimacy of risk forensic profiling practices and uses should be clearly defined by a law, as established by article 8 of the Framework Decision. However, this article does provide more indications regarding the safeguards to be implemented to ensure such legitimacy. It lets the difficult tasks to Member States to balance the interests at stake. This may result in important disparities in the protection granted to individuals.

However, additional safeguards may be put in place in order to ensure an effective application of the principle of proportionality, as it is already taking place in other fields where traditional data protection safeguards seems to struggle to ensure an efficient protection. A possibility thus consists in increasing the role of Data Protection Authorities in the assessment of the legitimacy of the processing as it is already proposed or even installed in other fields, namely processing involving the use of biometric data or originating by video surveillance techniques. Data Protection Authorities may be associated to the implementation of the risk profiling processing controlling its legitimacy, i.e. its conformity with data protection principles, in particular its compliance with the principle of proportionality. Such procedure has already been put in place in Italy with regards to biometrics and video surveillance processing via a voluntary procedure of prior checking.

6 Conclusions and Recommendations

Profiling in forensic science is still inchoate as we can see from the examples, although there is much research in this area. As with searches in databases, one should be aware of false interpretations of hits. False hits can be caused by the size of the database, by the techniques used, and since databases are often not very 'clean'. The persons that interpret the information from profiling should be very aware of the limitations of the methods. In the example of the camera surveillance, one should be aware that artefacts which are used for identification can also be changed. This should always be considered in forensic evidence, and should be included in the chain of evidence.

New ID systems with strengths to detect what was impossible previously, but weaknesses when they provide false positives, still offer new opportunities for improving and consolidating security. Indeed, electronic traces are information among others that are valuable in the context of the criminal justice system and forensic science.

In the light of new technological advances in the field of forensic profiling, i.e. the interconnectivity databases and risk profiling, the existing data protection instruments are not always effective anymore. As commissioner Frattini recalled "the protection of fundamental human rights such as privacy and data protection stands side-by-side public safety and security. This situation is not static. It changes, and both values are able to progress in step with technological advances. But it also means that there must be lines which cannot be crossed, to protect people's privacy" (Franco Frattini, 20 November 2007). However, as pointed out by the European Data Protection Supervisor, the different instruments adopted at European level "have in common that they enable a global monitoring of movements of individuals, even if from different perspectives. The way in which they can already contribute to the fight against forms of crimes, including terrorism, should be subject to in-depth and comprehensive analysis."

In that sense, the European Parliament pointed out that "Governments and EU institutions have often responded to terrorist attacks by adopting laws that have not been sufficiently discussed and some times in violation of basic human rights such as right to privacy or to a fair trial. Members call for further scrutiny of intelligence operations and for more proportionate and evidence-based legislation in the future."

In fact, the different norms approved at European level remain insufficient as they do not deal with the fundamental issues at stake before the widespread use of criminal intelligence, the increased monitoring of the average citizen or the increased linkage of police databases. Such instruments, fruit of difficult political consensus, implement principles broadly formulated and containing important derogations to the general data protection principles. Significant issues such as how to ensure the transparency and accountability of law enforcement activities, the quality of the data processed, e.g. the differentiation between categories of data subjects, or a strict application of the purpose specification principle remain unanswered. Moreover the comments of the European Commission, the European Data Protection Supervisor and the European Parliament are often not taken into account. At the level of the Council of Europe, the principles formulated in the eighties remain broad and subject to interpretation by Member countries.

Another complication is that the multitude of initiative creates a complex framework prone to legal loopholes and difficult to comprehend. The draft Framework decision on data protection in the third pillar has been limited to the exchange of personal data between law enforcement

authorities and fails to provide the third pillar with a comprehensive and strong data protection framework. Furthermore, the European Data Protection Supervisor stressed that for certain aspects the current text of the proposal does not provide for the same level of protection as defined in Convention 108. This certainly seems to be the case with the provision on the further use of data received from a Member State (Articles 3 and 12) and the right of access (Article 17)."

All these factors create legal uncertainty and should lead each Member State to face individually the challenges of ensuring that the new activities developed within the law enforcement field are subject to the principles of "scrutiny", "accountability" and "transparency", in a context of increased international activity and exchanges of criminal data. Each country will thus be called to make the specific balance between the competing interests at stake, in particular to prevent that the increasing use of personal data for risk prediction turns into stigmatisation of parts of the population.

It is, however, too soon to evaluate how the European Commission will implement the required safeguards and balance the different needs at stake. It suffices to say that the proposal for a Framework Decision for data protection in the third pillar constitutes a first laboratory where the aforementioned safeguards will have to be implemented.

7 Bibliography

Aalberg L., K. Andersson, C. Bertler, H. Borén, M. D. Cole, J. Dahlén, Y. Finnon, H. Huizer, K. Jalava, E. Kaa, E. Lock, A. Lopes, A. P.-v. d. Meer and E. Sippola (2007a). "Development of a harmonised method for the profiling of amphetamines I. Synthesis of standards and compilation of analytical data." Forensic Science International 169: 219-229

Aalberg L., K. Andersson, C. Bertler, H. Borén, M. D. Cole, Y. Finnon, H. Huizer, K. Jalava, E. Kaa, E. Lock, A. Lopes, A. P.-v. d. Meer, E. Sippola and J. Dahlén (2007b). "Development of a harmonised method for the profiling of amphetamines II. Stability of impurities in organic solvents." Forensic Science International 169: 231-241

Aitken C. C. G. and F. Taroni (2004). Statistics and the Evaluation of Evidence for Forensic Scientists. John Wiley & Sons, London.

Anderson D. S., C. Fleizach, S. Savage and G. M. Voelker (2006). Spamscatter: Characterizing Internet Scam Hosting infrastructure. Proceedings of the USENIX Security Symposium, Boston, MA.

Andersson K., E. Lock, K. Jalava, H. Huizer, S. Jonson, E. Kaa, A. Lopes, A. P.-v. d. Meer, E. Sippola, L. Dujourdy and J. Dahlén (2007c). "Development of a harmonised method for the profiling of amphetamines VI. Evaluation of methods for comparison of amphetamine." Forensic Science International 169: 86-99

Andersson K., K. Jalava, E. Lock, Y. Finnon, H. Huizer, E. Kaa, A. Lopes, A. P.-v. d. Meer, M. D. Cole, J. Dahlén and E. Sippola (2007a). "Development of a harmonised method for the profiling of amphetamines III. Development of the gas chromatographic method." Forensic Science International 169: 50-63

Andersson K., K. Jalava, E. Lock, Y. Finnon, H. Huizer, E. Kaa, A. Lopes, A. P.-v. d. Meer, M. D. Cole, J. Dahlén and E. Sippola (2007b). "Development of a harmonised method for the profiling of amphetamines IV. Optimisation of sample preparation." Forensic Science International 169: 64-76

Anrig B., W. Browne and M. Gasson (2008). The Role of Algorithms in Profiling. Profiling the European Citizen: Cross Disciplinary Perspective. M. Hildebrandt and S. Gutwirth. Springer. 39-50.

Audit commission (1993). Helping with Enquiries. London, HMSO: Police paper no. 12.
Birrer S., O. Ribaux, J. Cartier, Q. Rossy, S. Capt and M. Zufferey (2007). "Exploratory study for the detection and analysis of links between prospective advance fee fraud e-mails in an intelligence perspective." IALEA Journal 17: 11-21

Breyer, P. (2005). Telecommunications Data Retention and Human Rights: The Compatibility of Blanket Traffic Data Retention with the ECHR. European Law Journal 11, pp. 365-375.
Brodeur J.-P. (2005). "L'enquête criminelle." Criminologie 38(2): 39-64

Bruce C. W., Hick S. R., and Cooper J. P. (2004). Exploring Crime Analysis: Readings on Essential Skills. International Association of Crime Analysts, Overland Park.

Byford, L. (1981). The Yorkshire Ripper Case: Review of the Police Investigation of the Case. H. M. s. I. o. Constabulary, Home Office.

Bygrave, L. A. (2002). "Data Protection Law: approaching its rationale, logic and limits," Kluwer Law international.

Campbell v. United Kingdom (1992), European Court of Human Rights, 15 EHRR 137.

Chen W, Y. Q. Shi, and W. Su. Image splicing detection using 2-d phase congruency and statistical moments of characteristic function. In E. J. Delp and P. W. Wong, editors, Proceedings of SPIE: Security and Watermarking of Multimedia Content IX, volume 6505, page 65050R, 2007.

Choi, K.S; E. Y. Lam, and K. K. Y. Wong. Automatic source camera identification using intrinsic lens radial distortion. Optics Express, 14(24):11551–11565, 2006.

Citron, D. K. (2007). Technological Due Process. University of Maryland Legal Studies Research Paper No. 2007-26 Available at SSRN: http://ssrn.com/abstract=1012360.

Clarke R. V. and J. Eck (2003). Become a Problem Solving Crime Analyst in 55 Small Steps. Jill Dando Institute of Crime Science, University College London.

COM(90) 314 final - SYN 287, Brussels, 13 September 1990. Commission of the European Communities. Proposal for a Council Directive Concerning the Protection of Individuals in Relation to the Processing of Personal Data.

Convention for the Protection of Individuals with regards to Automatic Processing of Personal Data. European Treaty Series, no. 108, Strasbourg, 28 January 1981; International Legal Materials, 1981, I: 422".

Cook R., I. W. Evett, G. Jackson, P. J. Jones and J. A. Lambert (1998). “A hierarchy of propositions: deciding which level to address in casework.” Science & Justice 38: 103-111

Council Framework Decision 2006/960/JHA of 18 December 2006 on simplifying the exchange of information and intelligence between law enforcement authorities of the Member States of the European Union. OJEU of 29 December 2006, L386/89.

Council of Europe, Recommendation nº R (87) 15, regulating the use of personal data in the police sector, 17 September 1987.

De Hert, P. Presentation 4: Profiling issues and due process. Fidis Deliverable D.6.5/6.6: Second thematic workshop forensic implications combined with the workshop on forensic profiling: Crime control and due process.

Dinant, J.-M., Lazaro, C., Poullet, Y, Lefever, N., Rouvroy, A. (2008). Application of Convention 108 to the profiling mechanism. Some ideas for the future work of the consultative committee (T-PD) Available at: http://www.coe.int/t/e/legal_affairs/legal_cooperation/data_protection/Documents/Reports_and_studies_by_Experts/CRID_Profiling_2008_en.pdf

Directive 2006/24/EC of the European Parliament and of the Council on the retention of data generated or processed in connection with the provision of publicly available electronic communications services or of public communications networks and amending Directive 2002/58/EC. Official Journal L105, 15 March 2006, pp. 54-63.

Directive 91/308/EEC, on prevention of the use of the financial system for the purpose of preventing criminal offences, of 10 June 1991. Official Journal L 166 of 28.06.1991.

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. Official Journal L 281, 23 November 1995.

Dujourdy L., G. Barbati, F. Taroni, O. Guéniat, P. Esseiva, F. Anglada and P. Margot (2003). "Evaluation of links in heroin seizures." Forensic Science International 131: 171-183

Egger S. A. (1984). "A Working Definition of Serial Murder and the Reduction of Linkage Blindness." Journal of Police Science and Administration 12(3): 348-355

Esseiva P., L. Dujourdy, F. Anglada, F. Taroni and P. Margot (2003). "A Methodology for Illicit Heroin Seizures Comparison in a Drug Intelligence Perspective Using Large Databases." Forensic Science International 132: 139-152

Farid, H. Digital image ballistics from JPEG quantization. Technical Report TR2006-583, Department of Computer Science, Dartmouth College, Hanover, NH, USA, 2006.

Farrell G. and K. Pease (2001). "Repeat Victimization." Crime Prevention Studies 12: special issue

Felson M. and R. V. Clarke (1998). Opportunity Makes the Thief: Practical theory for crime prevention. Police Research Series. London, Home Office, Research, Development and Statistics Directorate, Policing and Reducing Crime Unit: 98.

First Opinion of the EDPS on the proposed Framework Decision, OJ C 47, 25.2.2006, p. 27.
Franco Frattini, European Commissioner responsible for Justice, Freedom and Security, Closing speech on Public Security (20 November 2007). Speech /07/728. Privacy and Technology Conference on Public Security, Privacy and Technology", Brussels, Charlemagne building.

Geradts Z., J. Keijer and I. Keereweer (1999). "A new Approach to Automatic Comparison of Striation Marks." Journal of Forensic Sciences 39(4): 974-980

Girod A. (2002). Exploitation et gestion systématiques des traces de souliers: une approche complémentaire pour l'investigation criminelle des cambriolages. Ecole des Sciences Criminelles, Université de Lausanne, PhD Thesis.

Girod A., O. Ribaux, S. J. Walsh and P. Margot (2004). "Bases de données ADN: un potentiel peu exploité de mise en relation d'événements criminels." Revue Internationale de Criminologie et de Police Technique et Scientifique(2): 131-147

Gloe, T; M. Kirchner, A. Winkler, and R. Böhme. Can we trust digital image forensics? In MULTIMEDIA '07: Proceedings of the 15th international conference on Multimedia, September 24–29, 2007, Augsburg, Germany, pages 78–86, New York, NY, USA, 2007. ACM Press.

GNIM (2005). Guidance on the National Intelligence Model, ACPO, http://police.homeoffice.gov.uk/news-and-publications/publication/police-reform/Interactive_NIM_1_.pdfhttp://police.homeoffice.gov.uk/news-and-publications/publication/police-reform/Interactive_NIM_1_.pdf (last access December 14th, 2007).

Godwin M. (2001). Weakness in Computerized Linking Databases. Criminal Psychology and Forensic Technology. M. Godwin. CRC Press. London.

Goldstein H. (1990). Problem Oriented Policing. Temple University Press, Philadelphia.
Grubin D., P. Kelly and C. Brunsdone (2001). Linking Serious Sexual Assaults through Behaviour, Home Office, Development and Statistics Directorate: Research Study 215.
Guéniat O. and P. Esseiva (2005). Le profilage de l'héroïne et de la cocaïne. Presses Polytechniques et Universitaires Romandes, Lausanne.

Hildebrandt M. (2008). Defining profiling: a new type of knowledge ? Profiling the European Citizen: Cross Disciplinary Perspective. M. Hildebrandt and S. Gutwirth. Springer. 39-50.

Hildebrandt, M. (2008). Profiling and the Identity of the European Citizen. In "Profiling the European Citizen: Cross-disciplinary Perspectives" (M. Hildebrandt and S. Gutwirth, eds.), pp. 320-360. Springer.

Hildebrandt, M. (2008). Profiling and the Identity of the European Citizen. Profiling the European Citizen: Cross-disciplinary Perspectives. M. Hildebrandt and S. Gutwirth, Springer: 320-360.

Hildebrandt, M., Gutwirth, S., and De Hert, P. (2005). "Deliverable 7.4. Implications of profiling practices on democracy and rule of law.." FIDIS (Future of Identity in the Information Society).
Hildebrandt, M., S. Gutwirth, et al. (2005). Deliverable 7.4. Implications of profiling practices on democracy and rule of law., FIDIS (Future of Identity in the Information Society).

House of Lords, Mr. R. Thomas, Mr. David Smith and Mr. J. Bamford, "Surveillance and data collection", Minutes of the Evidence of hearing taken before the select Committee on the Constitution, 14 November 2007.

Initiative of the Federal Republic of Germany, with a view to adopting a Council Decision on the implementation of Decision 2007/…/JHA on the stepping up of cross-border cooperation, particularly in combating terrorism and cross-border crime, 9 November 2007, OJ [2007] C267, p. 4.

Ioset S., P. Esseiva, O. Ribaux, C. Weyermann, F. Anglada, S. Lociciro, P. Hayoz, I. Baer, L. Gasté, Anne-LaureTerrettaz-Zufferey, C. Delaporte and P. Margot (2005). "Establishment of an operational system for drug profiling: a Swiss experience." Bulletin of Narcotics 57 (1-2): 121-146

Jackson G. (2004). The Nature of Forensic Science Opinion – a Possible Framework to Guide Thinking and Practice in Investigations and in Court Proceedings. Forensic science society autumn meeting, Wyboston, UK.

Jackson G., C. Champod, I. W. Evett and S. McCrossan (2006). "Investigator/Evaluator - a Possible Framework to Guide Thinking and Practice for Forensic Scientist." Science & Justice 46(1): 33-45

Jaquet-Chiffelle D.-O. (2008). Reply: Direct and Indirect Profiling in the Light of Virtual Persons. Profiling the European Citizen: Cross Disciplinary Perspective. M. Hildebrandt and S. Gutwirth. Springer. 55-63.

Johnson M.K. and H. Farid. Exposing digital forgeries by detecting inconsistencies in lighting. In MM&Sec'05, Proceedings of the Multimedia and Security Workshop 2005, August 1-2, 2005, New York, NY, USA, pages 1–10, 2005.

Johnson M.K. and H. Farid. Exposing digital forgeries through chromatic aberration. In MM&Sec'06, Proceedings of the Multimedia and Security Workshop 2006, September 26-27, 2006, Geneva, Switzerland, pages 48–55, 2006.

Kharrazi, M.; H. T. Sencar, and N. Memon. Blind source camera identification. In Proceedings of the 2004 IEEE International Conference on Image Processing (ICIP 2004), pages 709–712, 2004.

Kind S. S. (1987). The Scientific Investigation of Crime. Forensic Science Services Ltd, Harrogate.

Kind S. S. (1994). "Crime investigation and the criminal trial: a three chapter paradigm of evidence." Journal of the Forensic Science Society 34(3): 155-164

Kosta, E., Coudert, F., and Dumortier, J. (2007). Data protection in the third pillar: in the aftermath of the ECJ decision on PNR data and the data retention directive. International Review of Law, Computers and Technology 21, 343-358.

Lin, Z; R. Wang, X. Tang, and H.-Y. Shum. Detecting doctored images using camera response normality and consistency. In Proceedings of the 2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR 2005), volume 1, pages 1087–1092, 2005.

Lock E., L. Aalberg, K. Andersson, J. Dahlén, M. D. Cole, Y. Finnon, H. Huizer, K. Jalava, E. Kaa, A. Lopes, A. P.-v. d. Meer and E. Sippola (2007). "Development of a harmonised method for the profiling of amphetamines V. Determination of the variability of the optimised method." Forensic Science International 169(77-85)

Lukáš J. and J. Fridrich. Estimation of primary quantization matrix in double compressed JPEG images. In Digital Forensic Research Workshop, Cleveland, 2003.

Lyon, D. (2007). Surveillance Studies. An Overview. Cambridge, Polity Press.
Mennell J. (2006). "The Future of Forensic and Crime Scene Science Part II, A UK Perspective on Forensic Science Education." Forensic Science International 157(Supplement 1): S13-S20

Mennell J. and I. Shaw (2006). "The Future of Forensic and Crime Scene Science Part I - A UK Forensic Science User and Provider Perspective." Forensic science international 157(Supplement 1): S7-S12.

Oatley G., B. Ewart and J. Zeleznikow (2006). "Decision support systems for police: Lessons form the application of data mining techniques to "soft" forensic evidence." Artificial Intelligence and Law 14: 35-100.

Opinion on the draft Proposal for a Council Framework Decision on the use of Passenger Name Record (PNR) data for law enforcement purposes, 20 December 2007. European Data Protection Supervisor.

Peterson M., B. Morehouse and R. Wright (2000). Intelligence 2000: Revising the Basic Elements. Law Enforcement Intelligence Unit (L.E.I.U.) et International Association of Law Enforcement Intelligence Analysts (IALEIA), Sacramento, Lawrenceville.

Popescu, A.C. and H. Farid. Exposing digital forgeries by detecting traces of re-sampling. IEEE Transactions on Signal Processing, 53(2):758–767, 2005.

Popescu, A.C. and H. Farid. Exposing digital forgeries in color filter array interpolated images. IEEE Transactions on Signal Processing, 53(10):3948–3959, 2005.

Proposal for a Council Framework Decision on the protection of personal data processed in the framework of police and judicial cooperation in criminal matters (produced by the German Council Presidency), EU doc no: 7315/07 of 13 March 2007, available at: http://register.consilium.europa.eu/

Proposal for a Council Framework Decision on the protection of personal data processed in the framework of police and judicial cooperation in criminal matters, EU doc no: 16069/07 of 11 December 2007, available at: http://register.consilium.europa.eu/

Proposal for a Council Framework Decision on the use of Passenger Name Record (PNR) data for law enforcement purposes, European Commission, COM(2007) 654, 22 October 2007.

Proposal for a Council Framework Decision on the use of Passenger Name Record (PNR) for law enforcement purposes, COM(2007) 654 final, 6 November 2007.

Psychology and Law enforcement - Criminal profiling. APA Monitor on Psychology, 35, July/August 2004, No. 7, available at: http://www.apahelpcenter.org/articles/pdf.php?id=64.

Ratle F., C. Gagné, Anne-LaureTerrettaz-Zufferey, M. Khanevski, P. Esseiva and O. Ribaux (2007). "Advanced Clustering Methods for Mining Chemical Databases in Forensic Science." Chemometrics and Intelligent laboratory Systems in press

Ribaux O. and P. Margot (2007). La trace comme vecteur d'information au service du renseignement. Traité de sécurité intérieure. M. Cusson, B. Dupont and F. Lemieux. Hurtubise HMH. Montréal: 300-321.

Ribaux O., S. J. Walsh and P. Margot (2006). "The Contribution of Forensic Science to Crime Analysis and Investigation: Forensic Intelligence"." Forensic Science International(156): 171-181

Rix B. (2004). "The contribution of shoemark data to police intelligence, crime detection and prosecution." Findings, Home Office, Research, Development and Statistics Directorate(236)

Rossmo K. (1999). Geographical Profiling. CRC Press.

Rotaru vs. Romania, nº 28341/95. Vol. Official Journal L 166 of 28.06.1991.

Second evaluation report of the relevance of recommendation N° R (87) 15 regulating the use of personal data in the police sector, Council of Europe, 1998.

Second Opinion of the EDPS on the proposed Framework Decision, available on the EDPS website: www.edps.europa.eu.

Sheptycki J. (2004). "Organizational Pathologies in Police Intelligence: Some Contributions to the Lexicon of Intelligence-led Policing." European Journal of Criminology 1(3): 307-332

Steinbock, D. (2005). "Data Matching, Data Mining, and Due Process." Georgia Law Review 40(1): 1-84.

Steinbock, D. (2005). Data Matching, Data Mining, and Due Process. Georgia Law Review 40, 1-84.

Stepping up of cross-border cooperation, particularly in combating terrorism and cross-border crime, European Parliament, the Legislative Observatory, 2007, Procedure file on the Prüm

Treaty, available online at: http://www.europarl.europa.eu/oeil/file.jsp?id=5456232.

Terrettaz-Zufferey A.-L., F. Ratle, O. Ribaux, P. Esseiva and M. Kanevski (2006).

"Assessment of Data Mining Methods for Forensic Case Data Analysis." Journal of Criminal Justice and Security (Varstovoslovje) Special issue(3-4): 350-355

Terrettaz-Zufferey A.-L., F. Ratle, O. Ribaux, P. Esseiva and M. Khanevski (2007). "Pattern Detection in Forensic Case Data Using Graph-Theory: Application to Heroin Cutting Agents." Forensic Science International 167: 242-246

The United Kingdom Parliament, Home Affairs, Third report, 24 May 2007. Available at: http://www.publications.parliament.uk/pa/cm200607/cmselect/cmhaff/76/7602.htm

Third Opinion of the EDPS on the proposed Framework Decision, OJ C 50, 23.6.2007, p. 1.
Travis, A. (2007). New powers vital to avert surveillance society, says watchdog. The Guardian, 1 May 2007.

Travis, A. (2007). New powers vital to avert surveillance society, says watchdog. The Guardian.

United States (2004). The 9/11 Commission Report.

Völlmer A. (1919). "Revision of the Atcherley Modus operandi System." Journal of the American Institute of Criminal Law and Criminology 10: 229-274

Warren, S. D., and Brandeis, L. D. (1890). The Right to Privacy. Harvard Law Review 4, 193-220.

Wiggett A., A. Walters, L. O'Hanlon and F. Ritchie (2003). "Forensic Science Society Spring Meeting 2002: Intelligence." Science & Justice 43(2): 109-118

Witzig E. W. (2003). "The New VICAP." FBI Law Enforcement Bulletin(June): 1-7

Wolfgang M., R. Figlio and T. Sellin, Eds. (1972). Delinquency in a Birth Cohort. University of Chicago Press.

Zarsky, T. Z. (2002-2003). ""Mine Your own Business!": Making the Case for the Implications of the Data Mining of Personal Information in the Forum of Public Opinion." Yale Journal of Law & Technology 5: 1-56