COUNCIL OF EUROPE
CONSEIL DE L'EUROPE

Strasbourg, 15 June 2015 T-PD(2015)11

THE CONSULTATIVE COMMITTEE OF THE CONVENTION FOR THE PROTECTION OF INDIVIDUALS WITH REGARD TO AUTOMATIC PROCESSING OF PERSONAL DATA (T-PD)

Passenger Name Records, data mining & data protection:

the need for strong safeguards

Prepared by

Douwe Korff
Emeritus Professor of International Law
London Metropolitan University
Associate, Oxford Martin School, University of Oxford

with advice, comments and review by
Marie Georges
Council of Europe Expert

Directorate General Human Rights and Rule of Law

CONTENTS

Continues overleaf

Contents continued:

page:

* Part V also serves as the Executive Summary of the report

Attachment:

- o – O – o -

Introduction

Much has been said and written about Passenger Name Records (PNR) in the last decade and a half. When we were asked to write a short report for the Consultative Committee about PNR, "in the wider contexts", we therefore thought we could confine ourselves to a relatively straightforward overview of the literature and arguments.

However, the task turned out to be more complex than anticipated. In particular, the context has changed as a result of the Snowden revelations. Much of what was said and written about PNR before his exposés had looked at the issues narrowly, as only related to the "identification" of "known or [clearly 'identified'] suspected terrorists" (and perhaps other major international criminals). However, the most recent details of what US and European authorities are doing, or plan to do, with PNR data show that they are part of the global surveillance operations we now know about.

More specifically, it became clear to us that there is a (partly deliberate?) semantic confusion about this "identification"; that the whole surveillance schemes are not only to do with finding previously-identified individuals, but also (and perhaps even mainly) with "mining" the vast amounts of disparate data to create "profiles" that are used to single out from the vast data stores people "identified" as statistically more likely to be (or even to become?) a terrorist (or other serious criminal), or to be "involved" in some way in terrorism or major crime. That is a different kind of "identification" from the previous one, as we discuss in this report.

We show this relatively recent (although predicted) development with reference to the most recent developments in the USA, which we believe provide the model for what is being planned (or perhaps already begun to be implemented) also in Europe. In the USA, PNR data are now expressly permitted to be added to and combined with other data, to create the kinds of profiles just mentioned – and our analysis of Article 4 of the proposed EU PNR Directive shows that, on a close reading, exactly the same will be allowed in the EU if the proposal is adopted.

Snowden has revealed much. But it is clear that his knowledge about what the "intelligence" agencies of the USA and the UK (and their allies) are really up to was and is still limited. He clearly had an astonishing amount of access to the data collection side of their operations, especially in relation to Internet and e-communications data (much more than any sensible secret service should ever have allowed a relatively junior contractor, although we must all be grateful for that "error"). However, it would appear that he had and has very little knowledge of what was and is being done with the vast data collections he exposed.

Yet it is obvious (indeed, even from the information about PNR use that we describe) that these are used not only to "identify" known terrorists or people identified as suspects in the traditional sense, but that these data mountains are also being "mined" to label people as "suspected terrorist" on the basis of profiles and algorithms. We believe that that in fact is the more insidious aspect of the operations.

This is why this report has become much longer than we had planned, and why it focusses on this wider issue rather than on the narrower concerns about PNR data expressed in most previous reports and studies.

The report is structured as follows. After preliminary remarks about the main topic of the report, PNR data (and related data) (further specified in the Attachment), Part I discusses the wider contexts within which we have analysed the use of PNR data. We look at both the widest context: the change, over the last fifteen years or so, from reactive to "proactive" and "preventive" law enforcement, and the blurring of the lines between law enforcement and "national security" activities (and between the agencies involved), in particular in relation to terrorism (section I.i); and at the historical (immediately post-"9/11") and more recent developments relating to the use of PNR data in data mining/profiling operations the USA, in the "CAPPS" and (now) the "*Secure Flight*" programmes (section I.ii).

In section I.iii, we discuss the limitations and dangers inherent in such data mining and "profiling".

Only then do we turn to PNR and Europe by describing, in Part II. both the links between the EU and the US systems (section II.1), and then the question of "strategic surveillance" in Europe (II.ii).

In Part III, we discuss the law, i.e., the general ECHR standards (I); the ECHR standards applied to surveillance in practice (II, with a chart with an overview of the ECtHR considerations); other summaries of the law by the Venice Commission and the FRA (III); and further relevant case-law (IV).

In Part IV, we first apply the standards to EU-third country PNR agreements (IV.i), with reference to the by-passing of the existing agreements by the USA (IV.ii) and to the spreading of demands for PNR to other countries (IV.iii). We then look at the human rights and data protection-legal issues raised by the proposal for an EU PNR scheme. We conclude that part with a summary of the four core issues identified: purpose-specification and –limitation; the problem with remedies; "respect for human identity"; and the question of whether the processing we identify as our main concern – "dynamic"-algorithm-based data mining and profiling – actually works.

Part V contains a Summary of our findings; our Conclusions (with our overall conclusions set out in a box on p. 109); and tentative, draft Recommendations.

- o – O – o –

Preliminary: what are PNR (and API and SFPD) data?

Passenger Name Records (PNRs) are records, created by airlines and travel agencies, relating to travel bookings. They are concerned with all the aspects of a booking – originally they were not primarily about the passenger or passengers: if a group booking was made, the personal details of the members of the group were often only added later (sometimes as late as the time of boarding). Wikipedia provides the following simple description:[1]

> In the airline and travel industries, a passenger name record (PNR) is a record in the database of a computer reservation system (CRS) that contains the itinerary for a passenger, or a group of passengers travelling together. The concept of a PNR was first introduced by airlines that needed to exchange reservation information in case passengers required flights of multiple airlines to reach their destination ("interlining"). For this purpose, IATA and ATA have defined standards for interline messaging of PNR and other data through the "ATA/IATA Reservations Interline Message Procedures – Passenger" (AIRIMP). There is no general industry standard for the layout and content of a PNR. In practice, each CRS or hosting system has its own proprietary standards, although common industry needs, including the need to map PNR data easily to AIRIMP messages, has resulted in many general similarities in data content and format between all of the major systems.
>
> When a passenger books an itinerary, the travel agent or travel website user will create a PNR in the computer reservation system it uses. This is typically one of the large Global Distribution Systems, such as Amadeus, Sabre, Worldspan or Galileo, but if the booking is made directly with an airline the PNR can also be in the database of the airline's CRS. This PNR is called the Master PNR for the passenger and the associated itinerary. The PNR is identified in the particular database by a record locator.
>
> When portions of the travel are not provided by the holder of the Master PNR, then copies of the PNR information are sent to the CRSes of the airlines that will be providing transportation. These CRSes will open copies of the original PNR in their own database to manage the portion of the itinerary for which they are responsible. Many airlines have their CRS hosted by one of the GDSes, which allows sharing of the PNR.
>
> The record locators of the copied PNRs are communicated back to the CRS that owns the Master PNR, so all records remain tied together. This allows exchanging updates of the PNR when the status of trip changes in any of the CRSes.
>
> Although PNRs were originally introduced for air travel, airlines systems can now also be used for bookings of hotels, car rental, airport transfers, and train trips.

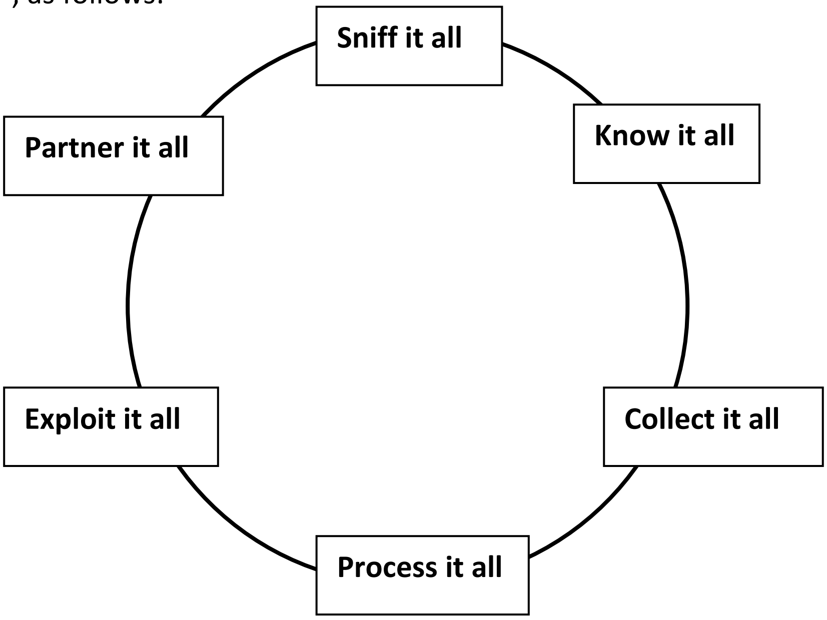
For more formal purposes, there were (and still are) other records, in particular **Advanced Passenger Information (API**, held in the API System, **APIS**) and, in the United States of America, **Secure Flight Passenger data (SFPD)**. These latter records are essentially limited to travel document (passport) information and, in the case of API, basic information about the flights concerned.

[1] See: https://en.wikipedia.org/wiki/Passenger_name_record

By contrast to API and SFPD, PNRs contain extensive information about the whole itinerary of the passenger(s) including hotel and car reservations (if booked with the flights), contact information including addresses, email- and IP-addresses and phone and mobile phone numbers, payment information (credit card details), dietary information (e.g., requests for vegetarian, kosher or hala'l meals), information on disabilities, etc., etc..[2]

For most of the 20th Century, state agencies were not generally interested in PNRs, except perhaps when they thought they might be relevant to ongoing criminal investigations, in which cases access to the records could be sought under the normal criminal procedures, typically with a judicial warrant.

This changed towards the end of the century, when the authorities in a range of countries started to become interested in using information technology more seriously in crime prevention and for more general "social engineering", and started to look at ways of using large collections of data to "identify" "targets" for policy action (see sub-section *III.i*, below). But the main impetus for the collection of large datasets for immigration-, law enforcement and national security purposes came from "9/11". In the USA, in particular, this led to a determination on the part of the authorities to adopt a massively broad approach to data collection, in particular in the fight against terrorism". This "**New Collection Posture**" is described in a slide used in a "top secret presentation [by the US's National Security Agency, NSA] to the 2011 annual conference of the Five Eyes alliance [of the intelligence services of the USA, the UK, Australia and Ne Zealand]", as follows:[3]



We discuss the links between this "new collection posture" – also epitomised in the name of the main early-21st Century US programme "**Total Information Awareness**" – and PNR data in sub-section *III.ii*, below.

[2] See the tables with the data fields required for SFPD, API and PNR in Attachment 1.

[3] The slide is reproduced in Glenn Greenwald, No Place to Hide: Edward Snowden, the NSA and the Surveillance State, 2014, on p. 97.

Here, we should already note that we find later on, in our more detailed discussions of the demands for PNR, in relation to European human rights- and data protection law, that "traditional" passenger information such as API data (or SFPD data in the USA) suffice to meet all the requirements to "identify" "known" people who for some reason are "wanted" or otherwise "looked out for" by the authorities, be that for border control/immigration or normal law enforcement purposes (e.g., because they are wanted convicted criminals who are "on the run", or people formally held to meet the legal requirements of "suspect" under criminal procedure law, or who may be on some other "wanted" or "no-fly" list, perhaps because they are under a court order not to leave the country).

By contrast, we find that the only reason why the authorities – first in the USA, but now also in the EU, and in Russia, Mexico, the United Arab Emirates, South Korea, Brazil, Japan, Argentina and Saudi Arabia – would want full, "bulk" access to *all* the PNR records, on *all* travellers, is because they want to use the additional data for data mining and "profiling" purposes – or as they like to put it, in rather deceptive language,[4] so that they can "identify" "possible" or "probable" or even "potential" miscreants – especially "possible", "probable" or "potential" terrorists, but this is inevitably now being extended to even less-defined "extremists" (and in some of the countries just mentioned is likely to be extended to all manner of dissidents).

In other words, the demands for PNR data are part of the wider demands for suspicionless mass collection-, retention- and analyses of data: of e-communications data, financial transaction data, and now travellers' data and, especially, the linking and combining of those data.

More specifically, the data fields in PNRs with mobile phone information and credit card information obviously allow for easy linking of the PNR data to the other massive "bulk" data collections held by the intelligence agencies, on global e-communications and financial transactions.

The debates about the "proper" and "proportionate" use of PNR data, and about the possible risks and "disproportionate" uses to which they could be put, must therefore take place against these wider contexts: "PNR" is not an isolated issue, but a new symptom of a much wider disease.

This report tries to facilitate that wider debate.

- o – O – o -

[4] We discuss the sometimes deliberately confusing use of the words "identify", "identification" (and "misidentification") in section I.iii.

PART I. The wider contexts

I.i The widest context: anti-terrorist (and wider) surveillance by the police and the secret services, and the blurring of the lines between them

As noted in the Introduction, our central finding is that the various "PNR issues" are part of the wider issues relating to the global surveillance programmes exposed by Edward Snowden. In fact, the link between increased "preventive" data collection generally, and collection of PNR data in particular, and increased surveillance, data mining and profiling, had already been noticed in a study for the UK Information Commissioner in 2004.[5] This concluded as follows:[6]

> From at least the 1970s onward, the presumed general increase in criminality, but more in particular the new threats to society posed by drugs-related and other organised crime, and especially by terrorism, led to a very significant extension of the role of the police from their traditional tasks:
>
> - **the investigation and prosecution of specific criminal offences**; and
> - **the countering of (real and immediate) threats to public order** –
>
> into a further, previously much more marginal area:
>
> - **prevention** of criminal offences being committed, or of threats to public order materialising – or indeed (in line with the trend noted [in the report]), of other activities which are deemed to be socially damaging or unacceptable, even if they are not necessarily criminal.
>
> This relatively new area of police work is typically intelligence-led. In practice, it involves:
>
> ➢ the collecting of personal data on a **wider range of data subject**: i.e. not just on persons (reasonably) suspected of involvement in a criminal offence, or who pose a clear and immediate threat to public order (the targets of "classic" policing), but also on persons who "might" be involved in, or who "might become" involved in, (certain, not always very-well-defined types of) "serious" crime or disturbances and indeed on people who are "in contact with" such already ill-defined targets;
>
> ➢ the increased use of **more intrusive, secret means of data collection** (telephone tapping, "bugging" of homes and offices, the use of informers and undercover agents, etc.) against this wider range of objects of police enquiries;
>
> ➢ **more intrusive means of data processing** including, in particular, ever-wider **"data matching"** and "**profiling**" including the screening of various (not necessarily only police- or public sector-) databases to "filter out" from a general population, individuals who are deemed to merit further police attention of the above kind;

[5] Ian Brown & Douwe Korff, Foundation for Information Policy Research, Privacy & Law Enforcement, study for the UK Information Commissioner, February 2004. The study consisted of five papers (with the first two combined into one), no longer available from the ICO website but still available from:
[ADD SSRN LINK]

[6] Paper No. 5: conclusions & policy implications, pp. 2-3, original emphases.

- **an increased blurring of the distinction between the work of the police and the work of the intelligence services**, on the one hand, **and the work of social and other State services** (such as the NHS, immigration), on the other (the new “full societal alliance” [discussed in the report]);
- **increased centralisation** within countries (including the UK); and
- **increased internationalisation**, especially within the EU but also (more problematically in terms of data protection) with the USA and other Western countries (especially those which are members of, or have special arrangements with, NATO).

In our third paper, we discussed the “**Total Information Awareness**” system in the USA and the related controversy over the transfer of **airline passenger (so-called PNR-) data** from the EU to the USA. We believe that even though the TIA-program has, for now, been suspended, it still represents the ultimate step in moves towards preventive, intelligence-led law enforcement. In a way, it is the natural outcome of the above trends. If the programs being developed under the TIA banner – “next-generation face recognition”, computerised translation of texts in foreign languages, computer-assisted data analysis, etc. – were to be shown to be effective in the fight against terrorism, there would be an unstoppable demand for their introduction in the fight against serious or organised or international crime (which is in any case inseparable from the fight against terrorism).

A particularly problematic aspect of this technology-driven, “intelligence”-based policing-as-part-of-wider-social control (as again well illustrated by TIA), is **the trend to classify people on the basis of supposedly highly-sophisticated pattern-recognition and -re-defining programs**. If computers can reliably classify a person as a “potential terrorist”, they can surely also single out people who are likely to have committed a bank robbery or a rape, or some other heinous crime? Indeed, it would be useful if the system could predict who will rob banks, or will rape people...

We believe we have shown that TIA-type programs of this kind have a long way to go to live up to this promise and that their usefulness even in the fight against terrorism is doubtful. However, we believe that the above trends - unless countered - will nevertheless result in **a wider use of such computer “profiling”, of larger sections of the population, for a range of purposes**, irrespective of such doubts.

There are clear and inherent dangers in the establishment of any secret Government databases or file collections, even of the old-fashioned, primitive kind, as the ECHR-cases of Leander and Rotaru, discussed in our fourth paper, have shown. If such processing is extended and based on supposedly more sophisticated, but at the same time less-controllable computer technology with built-in (but unacknowledged) biases, this will have **a more than just chilling effect on democratic freedoms**. They could lead to the stigmatisation of minorities and ethnic, religious or cultural “out-groups” and can be used to harass political activists and others - with the basis for such stigmatisation and harassment hidden in impenetrable algorithms. Leander was denied a job on the basis of an “error-ridden” secret file; Rotaru was falsely classified as a right-wing extremist in 1949 and half a century later still nearly denied compensation for the persecution he suffered in Communist Romania because of this file. Bigger and more powerful databases are no less susceptible to such errors. In the USA, political activists have already been “flagged” and prevented from travelling,

> without any serious evidence that they were involved in crime (let alone terrorism). In Britain, 30,000 Muslim homes have been raided under anti-terrorist legislation, presumably on the basis of "intelligence", with less than 0.5% of such raids resulting in terrorism-related arrests.

We will look more closely at the "Total Information Awareness" programme and its successors, and the PNR-related programmes linked to them, in the next sub-section, because it has direct lessons for the current proposals on the use of PNR data for various policing and "national security"/"foreign intelligence" purposes.

Suffice it to note here that the above concerns were also already reflected in a 2008 *Issue Paper* of the Council of Europe Commissioner for Human Rights, prepared by one of us.[7] It concluded *inter alia*:[8]

> We are rapidly becoming a "Surveillance Society". This is partly the result of general technical and societal developments, but these trends are strongly reinforced by measures taken in the fight against terrorism.
>
> In the context of the fight against terrorism, this means individuals are at risk of being targeted for being suspected "extremists" or for being suspected of being "opposed to our constitutional legal order", even if they have not (yet) committed any criminal (let alone terrorist) offence.
>
> "Targets" of this kind are moreover increasingly selected through computer "profiles". Even if some may be caught, there will always be relatively large numbers of "false negatives" - real terrorists who are not identified as such, and unacceptably high numbers of "false positives": large numbers of innocent people who are subjected to surveillance, harassment, discrimination, arrest - or worse. Freedom is being given up without gaining security.
>
> In addition, increasing use is made of non-criminal, yet effectively punitive, "administrative" measures against identified suspected "extremists" or new-type "enemies of the State". This robs them of fundamental safeguards, both against the specific measures taken against them and, as groups, against such discrimination. It leads to alienation of the groups in question, and thus actually undermines security.
>
> **In the process, all of us are increasingly placed under general, mass surveillance, with data being captured on all our activities, on-line or in the "real" world. Such general surveillance raises serious democratic problems which are not answered by the repeated assertion that "*those who have nothing to hide have nothing to fear*."**
>
> The response to these developments should be a re-assertion of the basic principles of the Rule of Law, as enshrined, in particular, in the European Convention on Human Rights, and as further elaborated in the case-law of the European Court of Human Rights and the European Court of Justice, as well as in European legal instruments directly or indirectly inspired by the Convention and such case-law, including in particular the still-pre-eminent Council of Europe

[7] Council of Europe Commissioner for Human Rights, Protecting the Right to Privacy in the Fight Against Terrorism, *Issue Paper (2008)3*, prepared by Douwe Korff, available at:
[ADD LINK]

[8] P. 13, emphasis in bold added. The case-law referred to in the last paragraph quoted (overleaf) is discussed further in Part III, below.

> recommendation on data protection in the police sector (Recommendation R(87)15 of the Committee of Ministers).

The Snowden revelations proved that if anything the scale of the global surveillance systems was underestimated in 2004 and 2008: no-one had envisaged quite how ubiquitous and global, and massive, the bulk "hoovering up" of data by the USA's NSA and the UK's GCHQ (and their partners) had become.

The dangers were therefore re-emphasised in a more recent (December 2014) Commissioner for Human Rights *Issue Paper*, written after the Snowden revelations, which stressed, with reference to the Data Retention Judgment of the Court of Justice of the European Union, that:[9]

> European data protection has been further strengthened by a judgment of the Court of Justice of the European Union, which has rejected compulsory, suspicionless, untargeted data retention. In connection with the debate on the practices of intelligence and security services prompted by Edward Snowden's revelations, it is becoming increasingly clear that secret, massive and indiscriminate surveillance programmes are not in conformity with European human rights law and cannot be justified by the fight against terrorism or other important threats to national security. Such interferences can only be accepted if they are strictly necessary and proportionate to a legitimate aim.

Again, we will discuss the specific legal requirements later, in section *VI*. Suffice it to note here that the issue of bulk collection (be that by "push" or "pull" means) of airline passenger data constitutes one important form of "compulsory, suspicionless, untargeted" data collection, and is closely tied in with the wider surveillance systems exposed by Snowden. In the next sub-section, we discuss the models for this, as established in the USA.

I.ii The use of airline passenger data for anti-terrorist screening in the USA: CAPPS I & II and "Secure Flight" – and their links to "Total Information Awareness" and now to the NSA/GCHQ global surveillance programmes

History[10]

The U.S. government's attempts to obtain airline passenger information for travel security purposes pre-dates "9/11" and appears to have started around 1998 with the first "Computer Assisted Passenger Pre-Screening" program, CAPPS (now referred to as CAPPS I). This already relied on profiles, aimed at selecting people who fit the profile for enhanced "secondary security screening". It would appear that these profiles were still rather basic and related to fairly straight-forward, factual data (although details are still hard to come by).

Soon after "9/11" a more advanced version of CAPPS started to be developed, CAPPS II, that was to use much more sophisticated profiles. Again, the precise technical details are not known, except that it is clear that it was **a system designed to profile airline**

[9] Council of Europe Commissioner for Human Rights, The Rule of Law on the Internet and in the wider digital world, *Issue Paper (2014)3*, also prepared by Douwe Korff, available at: [ADD LINK]

[10] In relation to the earlier developments, the text below draws on Douwe Korff, Paper No. 3: TIA & PNR, in: Ian Brown & Douwe Korff, Privacy & Law Enforcement (footnote 5, above).

passengers by reference to *multiple sources*, with the aim of rating each passenger according to the supposed risk he or she poses - with the basis for the assessment being hidden in a computer algorithm.

Such sophisticated data mining and profiling was typical for a series of wider programmes being developed at the time (from "9/11") by the U.S. military and security agencies and –bodies (including the U.S. Defense Advanced Research Projects Agency, DARPA), under the umbrella of, or closely related to, the broad "Total Information Awareness" (TIA) programme (later re-named the Terrorism Information Awareness" programme), apparently orginally conceived by John M Poindexter. The aim of those programmes was to create largely-automated computerised systems for identifying individuals and categorising them in terms of risk, on the basis of "intelligent" computer analyses and data mining, using "self-learning" "artifical intelligence" programmes.[11]

TIA was formally "suspended" in 2003, and CAPPS II formally abandoned in 2004, in particular because of privacy concerns. However, according to a 2012 (i.e., pre-Snowden) New York Times article:[12]

> [When TIA was dismantled] the NSA. was already pursuing its own version of the program, and on a scale that he [Poindexter] had only imagined. A decade later, the legacy of TIA is quietly thriving at the NSA. It is more pervasive than most people think, and it operates with little accountability or restraint.

Furthermore, according to this article:

> After TIA was officially shut down in 2003, the NSA adopted many of Mr. Poindexter's ideas except for two: an application that would "anonymize" data, so that information could be linked to a person only through a court order; and a set of audit logs, which would keep track of whether innocent Americans' communications were getting caught in a digital net.

Of course now, "post-Snowden", we know what programmes the journalist had heard rumours about: the global surveillance programmes run by the NSA (in cooperation with the UK's GCHQ in particular).

Recent developments

CAPPS II also did not die. It was replaced by the "*Secure Flight*" programme, managed by the US Transport Security Administration (TSA). This programme was initiated in August 2004 when, according to the U.S. Government Accountability Office (GAO) – which has reviewed it twice – it simply "match[ed] [passengers] against subsets of the TSDB" – that is, it matched the so-called Secure Flight Passenger Data (SFPD), against the subsets in the general US Terrorist Screening Database (TSDB), maintained by the US Terrorist Screening Center (TSC).[13]

[11] For details, also on the many associated programmes, see Douwe Korff, o.c. (previous footnote), section 3.

[12] Shane Harris, Giving In to the Surveillance State, New York Times, 22 August 2012, at: http://www.nytimes.com/2012/08/23/opinion/whos-watching-the-nsa-watchers.html?_r=0

[13] United States Government Accountability Office Report to Congressional Requesters, Secure Flight – TSA Could Take Additional Steps to Strengthen Privacy Oversight Mechanisms, September 2014, introductory page, under the heading "*Why GAO Did This Study*". The report is hereafter referred to as "the GAO report".

Since 2009, it has become a more integral part of the TSDB, as we shall see. Initially, it seems, it was still basically aimed at matching passengers against lists. However, as we note below, more recently things have changed.

Two issues are of relevant to this report.

First of all, it is interesting that the Secure Flight Passenger Data would appear to be more limited than the data typically included in a full Passenger Name Record. The GOA Report says SFPD:[14]

> includ[es] full name, gender, date of birth, passport information, and certain non-personally identifiable information provided by the airline, such as itinerary information and the unique number associated with a travel record (record number locator).

The word "includes" leaves this rather open-ended, but it would not appear to be as wide as (full) PNRs. In Europe, we also would regard the information listed as "non-personally identifiable" as still "personal data", because they are obviously linked to identified or identifiable natural persons, i.e., the airline passengers in question.

More important are the categories of lists, provided in the latest (September 2014) GAO report. This contains a table setting out **four** (presumably all four) **"high risk" categories of passengers**.[15] The table is reproduced overleaf. Somewhat oddly, the first three categories are described as "subsets" of the TSDB – but the most worrying fourth one is apparently separate (as discussed below).

The first "high risk" list (subset of the TSDB) contains the identities of "individuals who are suspected or known to pose a threat to aviation or national security and [who] are prohibited from boarding an aircraft or entering the sterile area of an airport." Presumably, "suspected" here refers to quite a high level of (concrete) suspicion (see the discussion below).

The second list (subset of the TSDB) covers "individuals who must undergo additional security screening before being permitted to enter the sterile area [i.e., the secure area of the airport] or board an aircraft." Presumably, as long as nothing dangerous is found on the person or in the person's luggage as a result of the "additional security screening", the passenger is allowed to board. Presumably also, there must be some level of suspicion against such persons, and that suspicion must be in some way related to terrorism or terrorist activities – but how high the level of suspicion has to be is unclear. Given that "additional security screening" involves little more than a "pat-down" and having to take one's shoes and belt off,[16] it probably is not a very high level of suspicion (but see again the discussion below).

Table 1: Secure Flight Screening Activities	
Screening Activity	**Description**

[14] *Idem*, p. 7. For further details, see the Attachment.

[15] *Idem*. We have not include the "TSM Prev√^{TM}" lists which list people who have applied for "pre-screening" and who, if succesful, are then subjected to less scrutiny. The table clearly does not contain all subsets, since the GAO report itself mentions a list of "Cleared Persons" (as discussed below, under the heading "**problems and [partial] remedies**") which it says is a subset of the TSDB, but which is not included in the table.

[16] GAO Report, p. 9, footnote 15.

No Fly List (high risk)	The No Fly List is a subset of the Terrorist Screening Database (TSDB), the U.S. government's consolidated watchlist of known or suspected terrorists maintained by the Terrorist Screening Center (TSC), a multi-agency organization administered by the Federal Bureau of Investigation. The No Fly List contains records of individuals who are suspected or known to pose a threat to aviation or national security and are prohibited from boarding an aircraft or entering the sterile area of an airport. Secure Flight has matched passengers against the No Fly List since 2009.
Selectee List (high risk)	The Selectee List is a subset of the TSDB containing records of individuals who must undergo additional security screening before being permitted to enter the sterile area or board an aircraft. Secure Flight has matched against the Selectee List since 2009.
Expanded Selectee List (high risk)	The Expanded Selectee List includes terrorist records in the TSDB with a complete name and date of birth that meet the reasonable suspicion standard to be considered a known or suspected terrorist, but that do not meet the criteria to be placed on the No Fly or Selectee Lists. Secure Flight began matching against the Expanded Selectee List in April 2011.
Transportation Security Administration (TSA) rules-based lists (high risk)	The high-risk rules-based lists include two lists of passengers who may not be known or suspected terrorists, but who, according to intelligence-driven, scenario-based rules developed by TSA in consultation with U.S. Customs and Border Protection (CBP), may pose an increased risk to transportation or national security.

Source: U.S. GAO Report

The third list (sub-set of the TSDB) is both puzzling and worrying. It relates to people who "meet the reasonable suspicion standard to be considered a known or suspected terrorist, but [who] do not meet the criteria to be placed on the No Fly or Selectee Lists." This suggests that the "reasonable suspicion" standard referred to is quite low – it is not even high enough to allow the person to be selected for additional screening which, presumably, is not a very high standard.

Before considering the last list, which is expressly said to be of people against whom there is no "reasonable suspicion" (and which is not a TSDB subset), we should note what the GAO report has to say about this standard, i.e.:

> All TSDB-based watchlists utilized by the Secure Flight program contain records determined to have met TSC's reasonable suspicion standard. In general, to meet the reasonable suspicion standard, the agency nominating an individual for inclusion in the TSDB must consider the totality of information available that, taken together with rational inferences from that information, reasonably warrants a determination that an individual is known or suspected to be or have been knowingly engaged in conduct constituting, in preparation for, in aid of, or related to terrorism or terrorist activities. As previously discussed, to be included on the No Fly and Selectee Lists, individuals must meet criteria specific to these lists. The TSDB, which is the U.S. government's consolidated watchlist of known or suspected terrorists, also contains records on additional populations of individuals that do not meet the reasonable suspicion standard articulated above, but that other federal agencies utilize to support their border and immigration screening missions. In addition, according to TSA officials, Secure Flight does not utilize all terrorist records in the TSDB because records with partial data (i.e., without first name, surname, and date of birth) could result in a significant increase in the number of passengers misidentified as being on the watchlist and potentially cause unwarranted delay or inconvenience to travelers.
>
> (Note *b* to the Table)

We cannot quite understand the above, or how it relates to the three subsets; nor can we find any specifics in the report about the "criteria specific to [each list]" – but that can be left aside here.[17]

Rather, for the purpose of our report, the fourth list is the most intriguing, and most worrying.[18] This list is expressly said to *not* be a subset of the TSDB, which has implications in terms of remedies, as we shall see under the next heading. It covers:

> **passengers who may not be known or suspected terrorists, but who, according to intelligence-driven, scenario-based rules developed by TSA in consultation with U.S. Customs and Border Protection (CBP), may pose an increased risk to transportation or national security.**

This is highly revealing. It would appear that there is no "reasonable suspicion" (apparently, of any degree) against the people on the list: they are not "known or suspected terrorists". Rather, they are identified as "high risk" (!) on the basis of **"intelligence-driven, scenario-based rules** developed by TSA in consultation with U.S. Customs and Border Protection (CBP)".

This can only refer to profiles created on the basis of analyses of the data in the various databases used by the TSC. As it is put in the report, this fourth list consists of:

> **passengers who meet intelligence-driven criteria indicating they may pose a greater security risk**

There is no information in the GAO Report on the nature or origin of the intelligence referred to, or on how they "drive" (?) the criteria, i.e., on the nature and origin of the "scenario-based rules" – that is, of the algorithms applied. These core issues are simply ignored.

Presumably, being singled out ("identified")[19] on the basis of such profiles, i.e., on the basis of a supposedly-sophisticated algorithm, does not suffice, in U.S. law, to formally

[17] The American Civil Liberties Union notes in its March 2014 report, U.S. Government Watchlisting: Unfair Process and Devastating Consequences (hereafter "the ACLU Report") that:

> "The TSC [Terrorism Screening Centre] defines a 'reasonably suspected terrorist' as 'an individual who is reasonably suspected to be, or have been, engaged in conduct constituting, in preparation for, in aid of, or related to terrorism and terrorist activities based on articulable and reasonable suspicion.' On its face, this standard is baffling and circular: it essentially defines a suspected terrorist as a suspected terrorist."

It points out that no information is available on the different standards applied to the different lists; or on the interpretation of those standards; or on the interpretation of the term "terrorism". Even more worrying:

> "The permissive standard for labeling someone a terrorist raises serious questions about the reliability of the intelligence underlying government watchlists. That intelligence originates with agencies such as the CIA, NSA, or the Defense Intelligence Agency, but the watchlisting process does not appear to involve rigorous review of the quality or credibility of the intelligence."

Indeed, to this the report adds, in footnote 21 (with reference to an earlier, 2007 GAO Report) that:

> "Neither the TSC nor the National Counterterrorism Center (NCTC), which consolidates terrorism-related intelligence, is positioned to assess the credibility of the intelligence underlying nominations to watchlists. The GAO has reported that both NCTC and the TSC generally treat an agency's designation of a watchlist nominee as presumptively valid."

See:
https://www.aclu.org/sites/default/files/assets/watchlist_briefing_paper_v3.pdf

[18] Regrettably, the ACLU Report mentioned in the previous footnote does not address this list: the report appears to be limited to the TSDB watchlists – although of course many of its findings, e.g., as to the lack of known criteria, the non-verification of agencies' "nominations", or the possibly devastating consequences of being listed, are relevant also to this list.

constitute “reasonable suspicion”. Yet it is still used to categorise the person concerned as “high risk” in a terrorist list.[20]

In section *IV*, we will discuss the dangers inherent in these kinds of automated profiling systems generally. Here, we must note that the *Secure Flight* programme was specifically altered in 2009, so that it no longer merely looks for matches between the Secure Flight Passenger Data and the subsets (lists) in the TSDB. Rather, the SFPD are themselves also fed into that database, and used to enhance the profiles. As it is put in the GAO report:[21]

> Since 2009, Secure Flight has changed from a program that identifies passengers as high risk solely by matching them against subsets of the TSDB, to [a program] that uses PII [read: Secure Flight Passenger Data] **and other information** to assign passengers a risk category: high risk, low risk, or unknown risk.

In other words, the passenger data themselves are “mined”, and linked to the other “Big Data” datasets already used by the TSC, and the NSA, to improve the “intelligence-driven, scenario-based rules”, i.e., the profiling algorithm.[22]

As we shall see below, at IV.iv, this is precisely what is also proposed for the EU PNR scheme.

[19] We will discuss the problems with the word “identified” in section I.iii, below.

[20] According to the table, this category consists of two lists – but the difference between these two lists is not explained.

[21] GAO report, introductory page, under the heading “*Why GAO Did This Study*”, emphasis added.

[22] According to Hasbrouck, this is in reality a very recent change, at least in practice (although, as shown above, the groundwork was clearly already laid in 2009). He dubs this new, “risk-based analysis” of passenger data “CAPPS IV” and wrote on 9 January 2015 that:

> The existence and TSA-mandated implementation of the new so-called “Computer-Assisted Passenger Prescreening System (CAPPS)” was first disclosed publicly in an obscure posting this Monday [i.e., 5 January] on the DHS website and an equally obscure notice published the same day in the *Federal Register*.

See (also for links to these postings):
http://papersplease.org/wp/2015/01/09/capps-iv-tsa-expands-profiling-of-domestic-us-airline-passengers/

Problems and [partial and deficient] remedies

General

As the GAO report explains:[23]

> The Transportation Security Administration's (TSA) Secure Flight program screens approximately 2 million passengers each day, matching passenger-provided personally identifiable information (PII) such as name and date of birth against federal government watchlists and other information to determine if passengers may pose a security risk and to assign them a risk category. By identifying those passengers who may pose security risks, Secure Flight helps protect against potential acts of terrorism that might target the nation's civil aviation system. However, Secure Flight can also have inadvertent and potentially inappropriate impacts on the traveling public, such as when passengers are identified as high risk because they share a similar name and date of birth with an individual listed on a watchlist, and thus experience delays and inconveniences during their travels. In order to minimize such impacts on passengers, the Department of Homeland Security (DHS) Traveler Redress Inquiry Program (TRIP) provides an opportunity for travellers who believe they have been delayed or inconvenienced because they have been incorrectly matched to or wrongly identified as the subject of certain watchlist records to seek redress.

As the above and other passages make clear (at least, on a close read), there are two types of errors that can occur in the screening programme:

- travellers may be "incorrectly matched to or wrongly identified as the subject of certain [TSDB] watchlist records" – i.e., they are erroneously believed to be a person who they are not, e.g., "because they share a similar name and date of birth with an individual listed on a [TSDB] watchlist" ("**mislisted**"); and
- "passengers who may have been misidentified to high-risk, rules-based lists" – i.e., they are the listed person, but they are wrongly assessed and wrongly marked as "high-risk" on a ["rules-based"] list ("**misidentified**").

It is important to note that (as indicated by the square brackets) the remedies that are available depend on the list. Specifically, the DHS TRIP programme provides for redress only with regard to "passengers who may have been incorrectly matched to or listed on high-risk lists based on the Terrorist Screening Database (TSDB)" – i.e., to passengers on the first three categories of lists (No Fly; Selectee; Expanded Selectee), that correspond to relevant TSDB "subsets".[24]

Remedies against "mislisting":

If a person complains to TRIPS, and the authorities (DHS) accept that there was a "mislisting" error, the victim is added to the "TSA Cleared List" and issued with a "redress control number" that can be used if they are subsequently again denied boarding or otherwise "inconvenienced". According to the GAO Report, such mistaken-identity corrections are generally, although not always, effective (even leaving aside the

[23] GAO Report, p. 1.

[24] GAO Report, p. 18.

often long delays, which the GAO Report says are being reduced); the cases of apparent non-effectiveness are ascribed to the fact that:[25]

> Because of the application of other TSA security measures, such as random selection, an individual's presence on the Cleared List may diminish, but will not preclude, the possibility of being selected for enhanced screening.

Moreover, according to the GAO Report:

> As of February 2014, Secure Flight officials were not aware of any passengers who have been misidentified to the CDC Do Not Board List.

However, that does not tell anyone much: if the officials had been aware of any such errors, they would of course have had to correct them of their own motion. In fact, official figures from 2014 suggest that there are considerable numbers of "mislistings":[26]

> DHS TRIP has received and processed more than 185,000 redress requests and inquiries since its establishment in 2007. Once the TRIP review process is complete, and all traveller records have been updated as appropriate, DHS issues a letter to the traveler signalling the completion of the review and closure of the case. **Historically, approximately 98% of the applicants to DHS TRIP are determined to be false positives.** To avoid such instances, DHS TRIP assigns applicants a unique Redress Control Number, which they can use when booking travel.

This suggests that in some seven years, there were approximately 182,000 acknowledged "mislistings", or some 26,000 each year. That may not seem very much against the 2 million" passengers reportedly screened each day – but not only will this number not include many people (especially non-US citizens or residents) who will have chosen not to complain, it also does not cover people who are listed as "high risk" on other lists than the "No Fly" list – they may not have realised that the "additional screening" to which they were subjected was not random or normal. And in any case, of course this will be little comfort for those who were "mislisted" and who may have suffered quite serious consequences (including unwelcome "attention" from security agencies in other countries to which their "high-risk" label was revealed).[27] It underlines the dangers of "false positives" – discussed in section I.iii, below.

More critical sources are thus, unsurprisingly, not nearly as satisfied as the GAO – and suggest higher error rates than admitted by the DHS. As the ACLU notes:[28]

> Well-publicized cases such as that of Rahinah Ibrahim [see box] have confirmed that watchlist entries result from blatant errors. Government audits suggest that these kinds of errors may occur at an alarmingly high rate.

Individual Cases: Rahinah Ibrahim

[25] GAO Report, p. 18, footnote 32.

[26] Written testimony of TSA Office of Intelligence Assistant Administrator Steve Sadler for a House Committee on Homeland Security, Subcommittee on Transportation Security hearing titled "Safeguarding Privacy and Civil Liberties While Keeping our Skies Safe", 18 September 2014, available at: http://www.dhs.gov/news/2014/09/18/written-testimony-tsa-house-homeland-security-subcommittee-transportation-security (emphasis added)

[27] According to the ACLU, the TSDB lists are shared with "at least 22 foreign governments": see the quote from the report at the bottom of page 20.

[28] ACLU Report (footnote 17, above), pp. 5-6, references omitted.

Rahinah Ibrahim, a Stanford PhD student and Malaysian citizen, was prevented from boarding a flight in San Francisco, handcuffed (despite being wheelchair-bound at the time), and held in a detention cell for hours in January 2005 based on what turned out to be a bureaucratic error by the FBI that placed her on the No Fly List. The government fought to avoid correcting the error for years, even invoking the state secrets privilege in an unsuccessful effort to prevent judicial scrutiny. She was permitted to leave the country, but to this day, she has been barred from returning, even though the government admits that she should not have been placed on the No Fly List.[1]

[1] *Ibrahim v. Dep't of Homeland Security, Case No. C06-00545 WHA at 8, 9-11 (N.D. Cal. Feb. 6, 2014).*

- A March 2008 report by the Department of Justice Inspector General described numerous weaknesses in FBI watchlisting procedures and concluded that "the potential exists for the watchlist nominations to be inappropriate, inaccurate, or outdated because watchlist records are not appropriately generated, updated or removed as required by FBI policy."
- A year later, in May 2009, the same Inspector General found that 35 percent of the nominations to the lists were outdated, many people were not removed in a timely manner, and tens of thousands of names were placed on the list without an adequate factual basis.
- A review by the TSC determined that 45 percent of the watchlist records related to redress complaints were inaccurate, incomplete, outdated, or incorrectly included.

When flawed or unreliable information makes its way into the watchlist database, it tends to stay there. Agencies have paid far greater attention to putting people on watchlists than to reviewing or purging of erroneous or outdated information. In short, there is every incentive to place individuals on a watchlist, but little incentive to clear them. And even if bad information is removed from one list, it may remain on other lists to which it was previously exported. As U.S. District Judge William Alsup noted,

> "[o]nce derogatory information is posted to the TSDB, it can propagate extensively through the government's interlocking complex of databases, like a bad credit report that will never go away."

The dissemination in fact goes well beyond USA government systems: [29]

> **Information from the TSDB is not only shared widely within the federal government and among state and local law enforcement agencies, but also exported to "several non-federal governmental watch lists" (we do not know which jurisdictions or for what purpose) and at least 22 foreign governments.**

Remedies against "misidentification" (i.e., mis-labelling):

If the remedies against "mislisting" are deficient, those against "misidentification" – that is: against the allegedly wrong assessment and labelling of a person as being of "high risk", on the basis of a supposedly sophisticated algorithm[30] – are left essentially

[29] ACLU Report (footnote 17, above), p. 3, references omitted, emphasis added.

[30] On the use of the terms "identified" and "misidentified" (here: "misidentified to [a list]"), see again section I.iii, below.

unspecified, but must be assumed to be even less (probably indeed, much less) effective. According to the GAO Report:[31]

> DHS TRIP is not able to provide redress for passengers who may have been misidentified to high-risk, rules-based lists and subsequently applied to DHS TRIP for redress. However, **according to TSA officials, TSA procedures for using the high-risk, rules-based lists mitigate impacts on passengers who may have been misidentified to these lists.** These officials stated that there is a possibility that a passenger could be misidentified to a rules-based list if their name and date of birth are similar to those of an individual on the list. TSA has established procedures for using the rules-based lists to mitigate impacts on passengers from screening against the lists. These procedures could assist those misidentified as a result of Secure Flight screening and may result in TSA removing passengers from the lists. ...

However, as a footnote clarifies:[32]

> **The details of these procedures are considered sensitive security information.**

The report goes on to say that:

> By removing individuals from rules-based lists, TSA ensures that passengers who are misidentified to those individuals will no longer be identified as a match, and thus delayed or inconvenienced as a result.
>
> In certain circumstances, TSA also reviews questionable matches to the rules-based lists to determine whether individuals on the list should be removed. According to TSA officials, starting in 2012, TSA's Office of Intelligence and Analysis (OIA) began monitoring the number of questionable matches to the list. According to TSA officials, the rate of questionable matches is less than 1 percent of all matches to the list for April 2012 through May 2014. TSA officials stated that the TSA Intelligence Analysis Division manually reviews these questionable matches and removes individuals from the list who have been erroneously included on the list. By removing these individuals from the list, TSA ensures that passengers will no longer be erroneously matched to them, and thus delayed or inconvenienced as a result. However, according to TSA officials, TSA's effort to identify and remove questionable matches does not address all possible misidentifications to the rules-based list. For example, TSA officials stated they do not review some matches because TSA does not have additional information about those passengers—beyond that included in the SFPD—that would be necessary to determine whether the passenger was actually misidentified to the rules-based high-risk list.

However, this whole quote appears to refer only to "mislistings": the erroneous matching of a specific person (a complainant) with a record in which another person is labelled "high risk" – **it does not address the question of how anyone (about whose person there is no confusion) can challenge a "high-risk" label attached to his or her name, or the assessment leading to the label.**

[31] GAO Report, p. 21, emphasis added. The same bland statement about there being "TSA procedures ... to mitigate impacts" are used elsewhere in the report, but as here with little or no useful clarification.

[32] GAO Report, p. 21, footnote 39, emphasis added.

Yet by the very nature of a list created by algorithms applied to inherently ambiguous and subjective "intelligence", such determinations are extremely difficult to challenge – and they become effectively unchallengeable if the underlying "intelligence" and the evaluations of the "intelligence" and the precise algorithm used to weigh the various elements of the "intelligence" cannot be challenged. As of course no victim of such a determination will ever be able to do.

And of course the TSA is about the last body that can "mitigate" against this. It effectively always accepts the "identification" – i.e., in respect to these lists, the labelling – of a person as "high-risk" by one of the agencies that contributes lists (or that contribute to combined lists).[33]

In other words: the secret "mitigating processes" referred to in the GAO Report must be assumed to be meaningless. At most, they can amount to a request from the TSA to an agency that it review its "rule-based" decision to place someone (a complainant) on such a list, but without the TSA having any insight into whether such a review actually took place; whether it was meaningful; or whether any action taken (or not taken) was appropriate. Needless to say, there is no due process attached to these "mitigating processes".

In the next section, section I.iii, we discuss how in any case the "rule-based" lists are **dangerous** because of **inherent limitations on and defects** in such exercises, especially when applied to the search for "[possible] terrorists" or other rare phenomena. This is important, not only in relation to the USA's "national security"/"foreign intelligence" data mining operations – i.e., the mining of the vast bulk data troves exposed by Snowden – but also in relation to similar schemes being proposed, or indeed already being implemented, in Europe.

I.iii The dangers inherent in data mining and profiling

One of us has already discussed the general problems with data mining and profiling in a report presented to the Consultative Committee in 2013.[34] This section reiterates that discussion with some minor edits to relate the discussion more closely to the topic of this report, and with some comments added.

[33] Cf. the finding in the ACLU Report (quoted in footnote 13, above) that "The GAO has reported that both NCTC and the TSC generally treat an agency's designation of a watchlist nominee as presumptively valid." To borrow the wording from that report: Neither the TSC nor the National Counterterrorism Center (NCTC), which consolidates terrorism-related intelligence, is positioned – or indeed allowed – to evaluate or challenge the raw intelligence of the agencies that create the "rule-based lists", or the algorithms those agencies use to determine who will, and who will not be listed on those lists.

[34] Douwe Korff, The use of the Internet & related services, private life & data protection: trends & technologies, threats & implications, presented to the Council of Europe Consultative Committee on Data Protection, March 2013 (T-DP(2013)07), available at:
https://www.coe.int/t/dghl/standardsetting/dataprotection/TPD_documents/KORFF%20-%20T-PD(2013)07Rev_Trends%20report%20-%20March2013.pdf

The aim of data mining

The aim of data mining is to make sense out of very large amounts of "big-but-dumb data": to turn them into "smart data" (the new catchword):[35]

> Systems of decision have to provide relevant, useful, actionable, intuitive, digestible and interactive information to the right person at the right time. The next generation of analytics are systems of decision that can provide the relevant information to every system user, in work context, to make smart business decisions.

Sometimes, the results of these systems may be straightforward and linear: "*If X occurs, do Y because the [big] data shows that this will [always] lead to Z*". But that will be rare. Much more often, indeed increasingly the norm, will be an output that is in reality a probability: "*If X occurs, do Y, because the [big/smart] data analysis shows that this will probably lead to Z*" (or at least, *Z* will be more probable than if you didn't do *Y*). The conclusion ("... *will probably lead to Z*") is based on the automatic analysis of many factors and data from many sources, i.e., on a (possibly dynamic) algorithm; and the conclusion is used to take decisions, including decisions on individuals. In other words, "smart data" analysis – data mining – rests on the creation of "profiles".

Profiling[36]

Profiling is one of the most challenging, and most worrying, developments relating to the use of the Internet, the Internet of Things, and "Big Data", yet is becoming pervasive, and is at the heart of the "rule-" or "scenario-based" listing referred to in the previous section. It means collecting and using pieces of information about individuals (or that can be indirectly linked to individuals) in order to make assumptions about them and their future behaviour.[37]

For example, someone who buys a pram will often also shortly thereafter buy baby clothes and nappies. In more abstract terms, "*people who did X and Y often also did Z. You did X and Y, so we will treat you as if you are likely to do Z*". But that is a very old-fashioned minimal profile, using obvious factors.[38]

In a world of massive "Big Data", innumerable elements can be factored in, and links can be established between factors that no-one would have thought were linked in advance:[39]

[35] See: http://smartdatacollective.com/mfauscette/50705/big-data-smart-data-supporting-critical-business-decisions

[36] This sub-section draws on a section on *Profiling* in a booklet on data protection by EDRi, written by Douwe Korff on behalf of the Foundation for Information Policy Research, UK, available at: http://www.edri.org/files/paper06_datap.pdf

[37] For a more detailed analysis, see http://protectmydata.eu/topics/limitations/ and Douwe Korff, Comments on Selected Topics in the Draft EU Data Protection Regulation (September 18, 2012), available at: http://ssrn.com/abstract=2150145

[38] Apparently, the earliest, minimal analysis of passenger data by the TSA singled out for further screening people who bought a one-way ticket with cash – and that was then about the only "rule".

[39] The quote is from Art Coviello, executive chairman of RSA, the security division of EMC, see: http://www.computerweekly.com/news/2240178641/Embrace-big-data-to-enable-better-security-says-RSA (emphasis added)

> "Big data is not just about lots of data, it is about having the ability to extract meaning; to sort through the masses of data elements to discover **the hidden pattern, the unexpected correlation.**"

Moreover, the logic used in the analyses - the profiling algorithm - can either be determined in advance and left unchanged (static), or, as is increasingly the case, be constantly dynamically re-generated and refined through loops linking back to earlier analyses, in theory constantly improving the outcome. Moreover, the refining is increasingly done by the computer itself, using "artificial intelligence".

Thanks to Snowden, we now know that the NSA has been building exactly these kinds of massive, bulk collections of data on all types of electronic communications, financial transaction, "loyalty cards", "frequent flyer" programmes – and of data on travellers, in the form of their PNRs. It was already clear from the "Total Information Awareness" programmes that the main aim was to "mine" such data troves – and this is undoubtedly exactly what happens under the current TIA-successor programmes. It is also unthinkable that in this the US agencies are not using the most advanced data mining tools and software, including such dynamic algorithms and artificial intelligence, in this data mining.

The problems with profiling

There are however serious problems with profiling. As a UK government study acknowledges, in rather under-stated terms:[40]

> In all cases [of profiling], the challenge will be to be certain that our understanding of human behaviour (both individual and collective), and our capability to capture that understanding in computer code or in sets of rules, is sufficient for the intended use of the model.

There are three main problems with profiling.

The base rate fallacy[41]

The first problem arises when profiles are used to identify rare phenomena, and is referred to in statistical literature as the "base rate fallacy". This phrase is used to refer to the mathematically unavoidable fact that if you are looking for very rare instances in a very large data set, then no matter how well you design your algorithm, you will always end up with either excessive numbers of "false positives" (cases or individuals that are wrongly identified as belonging to the rare class), or "false negatives" (cases or individuals that do fall within in the rare, looked-for category, but that are not identified as such), or both. It is important to stress the mathematical inevitability of this: you

[40] UK Government report on Technology and Innovation Futures: UK Growth Opportunities for the 2020s – 2012 Refresh (meaning the updated version of the 2010 report, issued in 2012), section 2.1(11), at p. 19:
http://www.bis.gov.uk/assets/foresight/docs/horizon-scanning-centre/12-1157-technology-innovation-futuresuk-growth-opportunities-2012-refresh.pdf

[41] From: Douwe Korff, Comments on selected topics in the Draft EU Data Protection Regulation, prepared for EDRi, November 2012, available at:
[ADD]
See also the section on this topic in: Douwe Korff, Technologies for the Use of Images: Automated Processes of Identification, Behavioural Analysis and Risk Detection Control at the Airports (2010), presented at Spanish Data Protection Agency Seminar, Madrid, Spain, 9-11 June 2010. Available at SSRN: http://ssrn.com/abstract=1977874

cannot improve the data set, or the algorithm, to avoid these debilitating results. **You cannot avoid the base rate fallacy.**[42]

Statisticians know this. Epidemiologists know this: they know that it is effective to screen all women over the age of 50 for breast cancer, because in that group there is a sufficiently high incidence of that affliction. But it is not effective to screen all women over the age of, say, 15, because that would throw up enormous numbers of "false positives", which would deplete hospital resources.

Exactly the same applies in anti-terrorist screening based on profiles: there are (thank God) simply not enough terrorists in the general population, or even in smaller populations (say, all Muslims in the UK of Pakistani or Saudi origin), to make the exercise worthwhile. The police and the security services would be chasing thousands of entirely false leads, while a significant number of real terrorists would still slip through the net.

The conclusion must be that profiles should never be used in relation to phenomena that are too rare to make their application reliable, such as trying to "identify"[43] (real, let alone potential) terrorists from a large dataset.

Interestingly, some very recently leaked documents suggest that some NSA analysts have realised for some time that having too much information can actually hamper anti-terrorist data analyses (although they do not seem to have noted the base rate fallacy):[44]

> "THE PROBLEM IS THAT WHEN YOU COLLECT IT ALL, WHEN YOU MONITOR EVERYONE, YOU UNDERSTAND NOTHING."
>
> –EDWARD SNOWDEN
>
> A**N AMUSING PARABLE** circulated at the NSA a few years ago. Two people go to a farm and purchase a truckload of melons for a dollar each. They then sell the melons along a busy road for the same price, a dollar. As they drive back to the farm for another load, they realize they aren't making a profit, so one of them suggests, "Do you think we need a bigger truck?"
>
> The parable was written by an intelligence analyst in a document dated Jan. 23, 2012 that was titled, "Do We Need a Bigger SIGINT Truck?" It expresses, in a lively fashion, a critique of the agency's effort to collect what former NSA Director Keith Alexander referred to as "the whole haystack." The critique goes to the heart of the agency's drive to gather as much of the world's communications as possible:

[42] For a detailed discussion of the analysis of personal characteristics and risk identification, and profiling, see: D. Korff, Technologies for the Use of Images: Automated Processes of Identification, Behavioural Analysis and Risk Detection Control at the Airports (previous footnote). On the baserate fallacy, see in particular also the "security blog" on the issue by Bruce Schneier, referred to in this paper (and in many other papers): Why Data Mining Won't Stop Terror, 3 September 2006, at: http://www.schneier.com/blog/ .

[43] See the discussion of the term later in this section.

[44] *Inside NSA, Officials Privately Criticize "Collect It All" Surveillance*, The Intercept, 28 May 2015, available at: https://firstlook.org/theintercept/2015/05/28/nsa-officials-privately-criticize-collect-it-all-surveillance/ (slightly redacted in the above quote). The two documents themselves can be found here: https://www.documentcloud.org/documents/2088978-do-we-need-a-bigger-sigint-truck.html https://www.documentcloud.org/documents/2088983-too-many-choices.html

> because it may not find what it needs in a partial haystack of data, the haystack is expanded as much as possible, on the assumption that more data will eventually yield useful information.
>
> Imagine, [another] analyst wrote in a leaked document, that you are standing in a shopping aisle trying to decide between jam, jelly or fruit spread, which size, sugar-free or not, generic or Smucker's. It can be paralyzing.
>
> "We in the agency are at risk of a similar, collective paralysis in the face of a dizzying array of choices every single day," the analyst wrote in 2011. "'Analysis paralysis' isn't only a cute rhyme. It's the term for what happens when you spend so much time analyzing a situation that you ultimately stymie any outcome …. It's what happens in SIGINT [signals intelligence] when we have access to endless possibilities, but we struggle to prioritize, narrow, and exploit the best ones."
>
> [The documents are two] of about a dozen in which NSA intelligence experts express concerns usually heard from the agency's critics: that the U.S. government's "collect it all" strategy can undermine the effort to fight terrorism. The documents, provided to *The Intercept* by NSA whistleblower Edward Snowden, appear to contradict years of statements from senior officials who have claimed that pervasive surveillance of global communications helps the government identify terrorists before they strike or quickly find them after an attack.

There is also recognition of the problem at a higher level. Thus, an authoritative study by the US' National Research Council (the US National Academies) concluded already in 2008 that:

> **Automated identification of terrorists through data mining (or any other known methodology) is neither feasible as an objective nor desirable as a goal of technology development efforts.**[45]

Yet this is still being pursued – also and in particular by means of compulsory bulk PNR data collection, and the matching of these data with other "Big Data" datasets. We will return to this question of the efficacy of profiling in sub-section *V.v*, below.

Discrimination by computer

Apart from the base rate fallacy (which is well-known to statisticians, albeit ignored by too many others, including the two NSA analysts referred to above), the wider implications of algorithm-based decision-making have not been as widely researched as

[45] Protecting Individual Privacy in the Struggle Against Terrorists: A Framework for Program Assessment, study by the United States' National Research Council, 2008, Executive Summary, pp. 3-4, emphasis added. As it is put in somewhat greater detail in the body of the study: "*Automated terrorist identification is not technically feasible because the notion of an anomalous pattern – in the absence fo some well-defined ideas of what might constitute a threatening pattern – is likely to be associated with many more benign activities than terrorist activities. In this situation, the number of false leads is likely to exhaust any reasonable limit on investigative or analytical resources. For these reasons, the desirability of technology development efforts aimed at automated terrorist identification is highly questionable*." (pp. 78-79). For the full assessment, with extensive detail, see Appendix H: *Data Mining and Information Fusion*. The study is available at: http://www.nap.edu/catalog.php?record_id=12452.

they should be. However, the leading research in this area, by Oscar Gandy, shows that (in David Barnard-Wills paraphrase):[46]

> predictive techniques and 'rational discrimination' – statistical techniques used to inform decision making by 'facilitating the identification, classification and comparative assessment of analytically generated groups in terms of their expected value or risk' – perpetuate and enforce social inequality.

This built-in risk - that profiles will perpetuate and reinforce societal inequality and discrimination against "out-groups", including racial, ethnic and religious minorities – is of course especially acute in relation to the screening of passengers by the TSA and the DHS, described above.

Crucially, this can happen even if the algorithms used are in their own terms perfectly "reasonable" and indeed rational. In practice (as Gandy has shown) the results will still reinforce the inequalities and discrimination already perfidiously embedded in our societies. Crucially, this discrimination-by-computer does not rest on the use of overtly discriminatory criteria, such as race, ethnicity or gender (which is why the "anti-discrimination clauses in the EU-US PNR Agreement, and indeed in the proposed EU PNR Directive, are so deficient, as discussed below, under the heading *"Profiling and "sensitive data""*). Rather, discrimination of members of racial, ethnic, national or religious minorities, or of women, creeps into the algorithms in much more insidious ways, generally unintentionally and even unbeknown to the programmers.

But it is no less discriminatory for all that. Specifically, it is important to stress that in international human rights law, the concept of discrimination does not imply some deliberate discriminatory treatment. Rather, in the words of the Human Rights Committee established under the UN Covenant on Civil and Political Rights:[47]

> the term "discrimination" as used in the Covenant should be understood to imply **any distinction, exclusion, restriction or preference** which is based on any ground such as race, colour, sex, language, religion, political or other opinion, national or social origin, property, birth or other status, and which has **the purpose or effect** of nullifying or impairing the recognition, enjoyment or exercise by all persons, on an equal footing, of all rights and freedoms.

Only by constantly evaluating the results of the decisions based on profiles can one avoid these effects. It takes serious effort. As Gandy concludes:[48]

[46] Review of Gandy's main book on the topic, Coming to Terms with Chance: Engaging Rational Discrimination and Cumulative Disadvantage, 2009, in *Surveillance & Society* 8(3): 379-381, at:
http://www.surveillance-and-society.org/ojs/index.php/journal/article/viewDownloadInterstitial/gandy_chance/gandy_chance
For the book itself, see: http://www.ashgate.com/isbn/9780754679615

[47] UN International Covenant on Civil and Political Rights, Human Rights Committee, General Comment No. 18: Non-discrimination, 10 November 1989, para. 7, emphases added, available at:
http://www.unhchr.ch/tbs/doc.nsf/%28Symbol%29/3888b0541f8501c9c12563ed004b8d0e?Opendocument
The HRCtee's definition draws directly on the definitions of discrimination against women, and discrimination on the basis of race, in the major UN Conventions against discrimination against women (CEDAW) and against people on the basis of race (CERD) (and, we might add, in the UN Declaration against discrimination on the basis of religion).

[48] Oscar Gandy, Engaging rational discrimination: exploring reasons for placing regulatory constraints on decision support systems, J Ethics Inf Technol, Vol 12, no. 1, pp. 29-42, 2010, at:

> these systems must be subject to active and continuous assessment and regulation because of the ways in which they are likely to contribute to economic and social inequality. This regulatory constraint must involve limitations on the collection and use of information about individuals and groups.

In Europe, this "regulatory constraint" - this protection against discrimination-by-computer - takes the form of data protection rules (although, regrettably, to date not much action has been taken on this score).[49]

The increasing unchallengeability of profiles - and of decisions based on profiles:

Profiles are becoming increasingly sophisticated and complex. As already noted, these days they tend to be dynamic, in the sense that, in the more developed "artificial intelligence" or "expert" systems, the computers operating the relevant programmes create feedback loops that continuously improve the underlying algorithms - with almost no-one in the end being able to explain the results: the analyses are based on underlying code that cannot be properly understood by many who rely on them, or even expressed in plain language.

This ties in with both earlier topics. First of all, such sophisticated profiles will have been tweeked in the direction of either higher "false positive" or "false negative" rates. Without understanding this, a user can seriously misinterpret the results.[50]

Secondly, it is especially in such dynamic systems that the risk of reinforcing engrained biases is greatest: feedback loops have a tendency to amplify such biases. Yet again, the very complexity of the algorithm tends to mask such effects: many users will not be able to detect such discrimination, or may be uninterested in it as long as the systems work to their benefit.

This danger (of which we warned years ago, unheeded) is beginning to be recognised, especially since algorithms are now used, not just in governance and law enforcement ("identifying" social and criminal deviants), but even in warfare – to "identify" "targets" to kill, by drones or in other ways:[51]

> What we are in the process of building is a vast real-time, 3-D representation of the world. A permanent record of us.
>
> But where does the meaning in all this data come from? For this, one needs ever more complex algorithms, automation, machine learning, and artificial intelligence. Such technologies are powering a wide range of new governance tools that can trace and record movements of people, detect patterns, and

http://academic.research.microsoft.com/Publication/41860489/engaging-rational-discrimination-exploringreasons-for-placing-regulatory-constraints-on-decision

[49] Cf. the (to date, rather deficient) discussion of the proposed rules on profiling in the draft General Data Protection Regulation and in the Council of Europe recommendation on profiling, both of which in principle prohibit the use of sensitive data in profiling (with exceptions), but without noting that discriminatory outcomes can also result from the processing of non-sensitive data in such systems.

[50] See F Kraemer et al., Is there an ethics of algoritims?, Ethics Inf Technol (2011) 13:251–260, at: http://purl.tue.nl/605170089298249

[51] Taylor Owen, The Violence of Algorithms: Why Big Data Is Only as Smart as Those Who Generate It, Foreign Affairs, 25 May 2015, available at: https://www.foreignaffairs.com/articles/2015-05-25/violence-algorithms

> ascribe risk to behaviors outside of programmed norms, to predicting future events.
>
> ...
>
> This is leading us to a place of predictive governance, based on unaccountable and often unknowable algorithms. Although the United States currently has a directive that humans must be a part of any fatal decision in war, this ignores all of the algorithm-based decisions that lead up to this ultimate point. If they are biased, flawed, or based on incorrect data, then the human will be just as wrong as the machine.

The same point is also made in this excellent article:[52]

> Many UK citizens are not concerned about the State having access to their innocuous emails and texts, particularly if this will allow law enforcement agencies to reduce the threat of terrorist attacks.
>
> However, intelligence services collect citizens' online communications data and run profiling algorithms to detect the characteristics of a potentially high-risk person, passenger or consignment. When law enforcement agents use algorithms that reflect unexamined generalizations about what constitutes a high-risk person, these practices may lead to erroneous conclusions that can result in negative outcomes that affect not only the lives of individuals but also of specific sections of society. If, for example, the indicators on which profiling is based relate to religious beliefs, ethnic or national origin, types of websites visited, flights booked to particular destinations, connections to specific groups of people or an individual, or political affiliations or occupation, the net is cast very wide and will include law abiding citizens.
>
> Whilst it is possible to derive profiles that provide valuable insights to intelligence services about suspected terrorists, it is also inevitable that intelligence services will arrive at incorrect conclusions about individuals, or groups of people. Inaccurate profiling can result in a flag being allocated to an individual, which may have repercussions in terms of a law-abiding citizen being subjected to more in-depth surveillance, arrest and detention for a number of days without charge, deportation, limitations on that individual's ability to gain entry to another country, or to secure certain types of employment and other forms of discrimination.
>
> The potential for algorithms to introduce biases and flawed decision-making into data analytics is a concern not only for the intelligence services, but also in both public and private sectors. An algorithm is a computational procedure used to process vast quantities of data. Algorithms are engineered to make decisions, take actions and deliver results speedily and continuously. Data scientists and programmers who write algorithms are effectively translating rules into code and, when those rules impact on citizens' rights, it can be very difficult to determine whether or not laws are being adhered to and rights are sufficiently protected. These issues can be compounded further by automation bias, which is the propensity for humans to assume that automated decision making systems

[52] Rachel O'Connor, Is Cameron proposing to legislate, inadvertently, for a Police State in the UK? Why citizens should urge caution, balance and proportionality, 21 January 2015, on the "TrustElevate" *groovyfuture* blog, available here: http://groovyfuture.com/is-cameron-proposing-to-legislate-inadvertently-for-a-police-state-in-the-uk-why-citizens-should-urge-caution-balance-and-proportionality/

> are infallible and to ignore contradictory information made without automation, even if it is correct.
>
> Many of the most complex algorithms are created by a number of different programmers over many years, which can result alterations to rules guiding the data analytics. The result is that the rules governing the analyses conducted by intelligence services become increasingly opaque over time. If the steps by which a law-abiding citizen is flagged as a potentially high-risk person are both poorly understood and not easily reversed, this not only propagates inaccurate decision making but also impedes legitimate redress.
>
> A democratic state is built, in part, on the premise of accountability however, there is a recognised accountability deficit accompanying the delegation of legislative power to code writers, which is rarely alluded to by politicians advocating the expansion of the remit of the intelligence services. The risks of not being duly cognisant of the potential pitfalls associated with big data analytics are multi-facted and include an inadvertent move from a democratic state to a Police State.

Yet just when algorithms increasingly dictate, or at least inform, policy decisions, the data subjects - the individuals included in or excluded from profile-based selections - are less and less able to challenge those results, at least in their individual cases.

If a company says it will not give you a loan because your income is too low, or you have a history of bad debts; or if an immigration authority refuses you a visa on the basis that you do not earn enough, or you have a criminal record,[53] you can challenge that if the figures or facts the company used are incorrect, or outdated.

But increasingly, a company or state agency will tell you it will not give you a loan, or will not invite you to an interview, or has placed you on a terrorist “no-fly” or “high-risk” list, “***because the computer said so***”: because the computer generated a “score” based on a profile, that exceeded or did not reach some predetermined basic level. If you ask for an explanation (if, that is, you actually find out that such an automated decision has been made on you), the company or agency (or at least the person you are dealing with) is likely to be unable to explain the decision in any meaningful way. They might provide you with examples of some of the information used (age, income level, whatever), but they will not give you the underlying algorithm - partly because the respondent him- or herself does not know or understand that algorithm, which is in any case constantly dynamically changing, and partly because the algorithm is a “national” or “commercial secret”.

It is extremely difficult to provide for serious accountability in relation to, and redress against, algorithm-based decisions generally. As Citron put it already in a 2007 paper:[54]

> Distinct and complementary procedures for adjudications and rulemaking lie at the heart of twentieth-century administrative law. Due process required agencies to provide individuals notice and an opportunity to be heard. Agencies could foreclose policy issues that individuals might otherwise raise in adjudications

[53] In the UK, there are rules that link entry permits for non-EU citizens to minimum income levels.

[54] Danielle Keats Citron, Technological Due Process, University of Maryland Legal Studies Research Paper No. 2007-26; Washington University Law Review, Vol. 85, pp. 1249-1313, 2007. Available at SSRN: http://papers.ssrn.com/sol3/papers.cfm?abstract_id=1012360
The quoted text is from the abstract.

> through public rulemaking. One system allowed focused advocacy; the other featured broad participation. Each procedural regime compensated for the normative limits of the other. Both depended on clear statements of reason.
>
> The dichotomy between these procedural regimes has become outmoded. This century's automated decision-making systems collapse individual adjudications into rulemaking while adhering to the procedural safeguards of neither. Automated systems jeopardize due process norms. Their lack of meaningful notice, and a hearing officer's tendency to presume a computer system's infallibility, devalue hearings. Standard Mathews v. Eldridge cost-benefit analysis is ill-equipped to compare the high fixed cost of deciphering a computer system's logic with the accumulating marginal benefit of correcting myriad inaccurate decisions. Automation also defeats participatory rulemaking. Code, not rules, determines the outcomes of adjudications. Programmers inevitably alter established rules when embedding them into code in ways the public, elected officials and the courts cannot review. Last century's procedures cannot repair these accountability deficits.

She proposes "A new concept of technological due process", which should be:

> a carefully structured inquisitorial model of quality control [that] can partially replace aspects of adversarial justice that automation renders ineffectual [and that] also provides a framework of mechanisms capable of enhancing the transparency, accountability, and accuracy of rules embedded in automated decision-making systems.

In an article entitled *Rage Against the Algorithms*, similarly Diakopolous points out the serious, almost unsurmountable obstacles facing anyone trying to use traditional transparency mechanism (such as data subject access requests or freedom of information requests) as means to "check algorithmic power". He suggests that:[55]

> a new and complementary alternative is emerging. I call it algorithmic accountability reporting. At its core it's really about reverse engineering—articulating the specifications of a system through a rigorous examination drawing on domain knowledge, observation, and deduction to unearth a model of how that system works.

In his article, he refers to a number of journalistic efforts that succeeded in such "reverse-engineering" of commercial algorithms.

However, as O'Connor points out:[56]

> **A key question is how might Citron's and Diakopolous's suggestions be tested and applied in the context of the UK intelligence service?**

We believe that trying to provide answers to that question must be one of the Consultative Committee's top priorities, in relation to commercial-, administrative-, law enforcement- and national security agencies' use of profiles – including in relation to the use of PNR data in such profiles.

[55] Nicholas Diakopoulos, Rage Against the Algorithms: How can we know the biases of a piece of software? By reverse engineering it, of course, The Atlantic, 3 October 2013, available at: http://www.theatlantic.com/technology/archive/2013/10/rage-against-the-algorithms/280255/

[56] Rachel O'Connor, o.c. (footnote 52, above).

It is because of this that we highlighted, in section I.ii, the fourth "high-risk, rules-based" terrorist list used by the US authorities. This is obviously a list that is created by algorithm – but there is no effective possibility to challenge one's inclusion on the list, to fight the algorithm. As we noted, according to the GAO Report, even the "*procedures [aimed at] mitigat[ing] impacts [of these lists] on passengers who may have been misidentified to these lists*" are "*considered sensitive security information*". That leaves those who have been thus "misidentified" without redress.

Even at a higher accountability level, e.g., in relation to parliamentary or judicial or special oversight bodies, it will be effectively impossible to verify the risks inherent in those profiles: i.e., to assess the level of "false positives" and "false negatives", or the possibly discriminatory effect of the profiles on certain groups, without the full, in-depth cooperation of the agency generating the profiles. Yet the latter are likely to be unwilling to be so helpful, unless compelled to do so by law.

Profiling thus really poses a serious threat of a Kafkaesque world in which powerful agencies (like the DHS and the NSA – or in the near future European agencies?) take decisions that significantly affect individuals, without those decision-makers being able or willing to explain the underlying reasoning for those decisions, and in which those subjects are denied any effective individual or collective remedies.

That is how serious the issue of profiling is: it poses a fundamental threat to the most basic principles of the Rule of Law and the relationship between the powerful and the people in a democratic society.

As Taylor Owen puts it in the article, quoted above:[57]

> If algorithms represent a new ungoverned space, a hidden and potentially ever-evolving unknowable public good, then they are an affront to our democratic system, one that requires transparency and accountability in order to function. A node of power that exists outside of these bounds [of transparency and accountability] is a threat to the notion of collective governance itself. This, at its core, is a profoundly undemocratic notion—one that states will have to engage with seriously if they are going to remain relevant and legitimate to their digital citizenry who give them their power.

Before linking the above discussion to the discussions on PNR in Europe (below, at *V*), it is important to clarify three further matters related to data mining and profiling: the misleading use of the words "identify", "identifying" and "identification" (or "misidentification"); the problems with anonymity in relation to large datasets and data mining; and the issue of "sensitive data".

"Identifying" a suspect through data mining/profiling

In section *VI*, in our discussion of the data protection "purpose-specification and –limitation principle", we note that information on airline passengers can be used for a range of rather different purposes, in relation to all of which the terms "identify" or "identification" are used, as follows:

[57] Taylor Owen, o.c. (footnote 51, above).

1. checking the **identity** and credentials (e.g., visa) of an airline passenger for the purpose of verifying whether that person is entitled to enter the country [identity check and immigration control];

   NB: some countries may also have exit requirements, but these are usually related to the purposes listed at (2) and (3), below.

2. **identifying** "known" wanted criminals (for which one should read persons convicted of criminal offences) and persons properly categorised as suspects within the meaning of the relevant national criminal law and criminal procedure law ("known suspects");

   NB: this of course includes such identification of people convicted or formally suspected of terrorism.

3. **identifying** other "known" persons on the basis of specific laws permitting action against the individuals, e.g., preventing a person from leaving a country because he has failed to pay child maintenance;

4. using PNR data to facilitate the *ex post facto* investigation of criminal offences and the *ex post facto* "**identification**" and prosecution of the perpetrators;

5. pro-active "**identification**" of "possible suspects", i.e., the marking of people as a "probable criminal" or "possible criminal", without those people being yet formally categorised as suspects in the criminal law/criminal procedure law sense (i.e., in the absence of any evidence against them that would suffice to properly designate them as formal suspects, in accordance with criminal procedure law); and

6. pro-active "**identification**" of people for "preventive targeting" on national security grounds, in cases in which no action can (yet) be taken against them under the criminal law.

Here, we want to note that in discussing PNR in relation to each of these, it is important to clarify the term "**identify**" in relation to these different purposes:

In (1), (2) and (3), above, the term means "*confirming that a certain person (e.g., a person stopped at a border check) is a specific person named [or otherwise identified] on a list or official record or document*" – i.e., respectively, whether the person is the person to whom a travel document (passport) pertains, or whether the person is a specific person named in an official document such as a court judgment or court order:

> "Mr John Bloggs on the PNR list is the Mr John Bloggs who was convicted of robbery by [a particular court] on [a particular day], and who is wanted as a fugitive."

Here, the question is simply one of matching the data on the convicted criminal (or the person formally classified as a suspect) with the particulars of the traveller. The person can be said to have been "identified" as the wanted criminal or suspect (etc.) if his details correspond to those of the wanted man.

This is a quite different meaning from the word "identify" or "identification" in points (4), (5) and (6). There, the aim is not so much to match a certain person against a pre-existing record. Rather, the aim is to categorise the person.

In (4), the aim is to examine records (*in casu*, PNRs) to see if a person stands out as being linked to a particular crime, *ex post facto*, and if so, to establish whether there is enough evidence against that person to formally classify him as a "suspect" in terms of criminal procedure law, or to charge him, or commit him for trial (or whether he should be "excluded from the investigation").

The use of the word "identify" or "identification" in (5) and (6) is even more ambiguous. Here, the purpose is to indicate that there is a certain *likelihood* that the person thus "identified" will probably (or even possibly) commit a crime, perhaps of a certain nature, e.g., a terrorist offence (point (5)); or even more vaguely that that person "may be a terrorist" (or otherwise "of interest" on grounds of national security or in relation to the collection of "foreign intelligence") (point 6). This is essentially a *label* attached to the person indicating no more than that a person has made this risk assessment, and has reached this conclusion about the specified risk level – or worse, that a computer has made this risk assessment and reached this conclusion.

We believe this semantic difference has seriously hampered the debates on PNR (and on other big data sets). The authorities repeatedly assert that these systems are only used "to identify terrorists" (and/or similar bad people); and most members of the public – and indeed many in positions of authority, such as members of parliament – will understand this term to refer to "identification" in the sense used in (1) – (3), above. But in reality it is increasingly used in the sense used in (4) – (6).

Any honest, transparent debate about the uses of PNR data or other big datasets should be completely clear about the way in which the terms are used. Unfortunately, this is not yet the case. We hope the Consultative Committtee will help to remedy this.

<u>The problems with anonymity in large datasets</u>

In several PNR-related contexts, reference is made to "anonymisation" of data or to "masking" of data, with the suggestion that such "anonymised" or "masked" data can no longer be linked to an identified or identifiable individual. This is done, for instance, in the 2012 EU-US PNR Agreement and in the proposals for an EU PNR Directive (both first noted in Part II, and then further discussed in the light of the law, in Part IV).

However, in reality there are very serious problems with ensuring anonymity in large datasets, especially if the data in the datasets can be linked to data in other large datasets (as we seen is becoming the norm for PNR). In such cases, it becomes almost impossible – indeed, often actually mathematically impossible – to ensure anonymity. This issue is therefore important in relation to PNR. Again, though, it will have to suffice to refer back to earlier summaries of the issues.[58]

[58] What follows is largely taken from a section on *Anonymisation* in the booklet on data protection by EDRi (note 36, above), which in turn drew heavily on advice to a major EU study, provided to the authors of the study (Prof. Douwe Korff and Dr. Ian Brown) by Prof. Ross Anderson, quoted on p. 50 of Working Paper No. 2, produced for that study, and on the FIPR submission to the UK Information Commissioner's Office (the UK Data Protection Authority) on the latter's draft Anonymisation Code of Practice, also drafted by Prof. Anderson. These are available here:

- New Challenges Study, working paper 2:
http://ec.europa.eu/justice/policies/privacy/docs/studies/new_privacy_challenges/final_report_working_paper_2_en.pdf
- FIPR submission to the ICO:
http://www.fipr.org/120823icoanoncop.pdf

Anonymisation means removing or obscuring information from data sources that would allow direct or indirect identification of a person.

One of the big advantages of anonymisation is, for example, to allow research that would otherwise not be possible due to privacy concerns. For instance, using everyone's medical records to find disease patterns could improve health care, but would also seriously infringe on people's privacy. It is claimed that the solution is to remove (or "mask") direct identifiers such as names, birth dates, and addresses, so that the data cannot be traced back to individuals.

Governments, industry and researchers tend to claim that effective anonymisation of personal data is possible and can help society to ensure the availability of rich data resources whilst protecting individuals' privacy.

Unfortunately, this is simply not the case – as scientists have known for a long time. For example, in 1997, researchers were already able to re-identify individual patients from a large set of medical records reduced to post code and date of birth. In 2006, a study found that if you know how a user rated just six films, you can identify 99% of the users in the Netflix (an online video rental service) database.

How is this possible? The main problem is that effective anonymisation does not just depend on stripping away direct identifiers (name, address, national identification number, date of birth) from a data set. Instead, the relevant measure is the size of the "anonymity set" – that is, the set of individuals to whom data might relate. If you're described as "a man" the anonymity set size is three and a half billion, but if you're described as "a middle-aged Dutchman with a beard" it is maybe half a million and if you're described as "a middle-aged Dutchman with a beard who lives near Cambridge" it might be three or four.

Pseudonymisation, that is replacing the name and other direct identifiers with a new identifier, – e.g. "John Smith, 1 High Street" becomes "person 45684231" – does not resolve this problem either, irrespective of whether, or how well, the pseudonym is encrypted. Suppose we gave everyone in the world an ID card with a unique number. What will happen? You start with a single pseudonymous incident, such as a drug prescription: "human no. 45684231 got penicillin on 3 Feb 2009". The anonymity set size just shrunk from seven billion to a few hundred thousand. Then along comes a second incident: "human no. 3,265,679,016 got codeine on 14 May 2009". Now it's down to a few hundred or even a few dozen. A couple more incidents, and the individual is uniquely specified.

As more and more "Big Data" data sets are released, the possibility of identifying people in any single "anonymised" data set by using data from other large data sets increases greatly.[59] With current – and foreseeable future – technology, it is safe to say that anonymisation no longer works when identities are actively sought. This poses major general challenges, in particular in relation to "Big Data", that are insufficiently acknowledged or addressed to date.

[59] There are techniques to limit queries to a specific single database to ensure that re-identification of individuals from that single database is (almost) impossible. This includes in particular "differential privacy", designed by Cynthia Dwork and others. However, this does not work if one can make cross-referenced searches in several large datasets. See: http://research.microsoft.com/en-us/projects/databaseprivacy/ (with references).

The only way to counter these problems is through full transparency regarding the technologies being used, open peer review by security engineering experts and limits on "Big Data" disclosure and linkages. Such procedures will not completely eliminate the problems, but will at least provide early warnings over compromised databases and raise standards.

However, yet again, it is difficult to see how even such (even generally not yet widely implemented and in any case limited) safeguards can be applied to the dataset-linking and data mining operations of intelligence services. Those agencies are notoriously averse to transparency, and to any open peer-reviews of their operations. Moreover, the large commercial companies that currently provide the technical know-how underpinning the agencies' activities greatly benefit from this secrecy since it means no-one is allowed to check the efficacy and effectiveness of their technologies.

For now, we must leave it at the conclusion that the measures supposedly achieving "anonymisation" of data in the "Big Data" data sets mined by the agencies – including the PNR "Big Data" dataset – are meaningless, and serve as little more than fig leaves to hide the actually easy reidentifiability of the data.

Profiling and "sensitive data"

There are similar problems with regard to the supposed safeguards relating to the use of "sensitive data", defined in Article 6 of the Council of Europe Convention on Data Protection (Convention No. 108) as:

> Personal data revealing racial origin, political opinions or religious or other beliefs, as well as personal data concerning health or sexual life [and] personal data relating to criminal convictions.

As we shall see in Part IV, in spite of various references to such data, the provisions in the EU-third country PNR agreements (in particular in the 2012 EU-US PNR Agreement) and in the proposed EU PNR Directive do not really seek to prevent discriminatory outcomes of the uses of the PNR data they claim to regulate. All they do is limit (to a rather limited degree) the overt use of such data.

The problem is that, as already noted, data mining and profiling almost inevitably "perpetuate and [re-]enforce social inequality"[60] – and that this is so, irrespective of any overt limitations on the use of "sensitive data": **you can use entirely "non-sensitive" data in such operations, yet still end up with results that *in effect* discriminate on grounds of race, religion or sexuality etc..**

Given that stigmatisation of "suspect communities" is one of the most serious dangers of any state data mining/profiling operation, no more so than in relation to terrorism, the insufficiency of the safeguards in this respect in the PNR-related instruments is another major issue of concern.

- o – O – o -

[60] See the quote at the bottom of p. 25, above.

PART II. PNR and Europe

II.i The links to the US systems

The Consultative Committee may have begun to wonder why we have dwelled to such an extent on the systems in the USA, on the defects in the US systems, and on the general dangers inherent in data mining and profiling, especially in a terrorist context.

There are two reasons for this. First of all, as noted below, at IV.i, we believe the latest (2012) EU-US PNR Agreement can be read as allowing the passing on of PNR data by TSA/DHS to other US agencies, including the NSA, for the purpose of "identifying" "possible" terrorists through general data mining and profiling of the kind just discussed. This may not have been in the minds of the EU negotiators, but we believe it is certain to be thus read by the US agencies.

In that context, "Europe" must also examine the highly credible claims by Edward Hasbrouck (reported in section IV.ii) that the USA has been systematically violating previous agreements, and is still systematically by-passing European data protection law, by accessing the CRSs used in global airline reservation systems hosted in the USA to obtain full PNR data on most flights, including most European flights (including even entirely intra-European ones), outside of any international agreements.

We also note, at IV.iii, that, unsurprisingly, more and more countries are now demanding that airlines hand over their PNR data in bulk to them. A recent EU Council note, submitted by Spain, mentions Russia, Mexico, United Arab Emirates, South Korea, Brazil, Japan, Argentina and Saudi Arabia.[61] Will they also use these records for data mining? Link them to other bulk data? Will the EU address this in PNR transfer agreements with these countries? Would such agreements be effective in prohibiting, or at least limiting that, given that the EU-US PNR Agreement does not stop the USA from such activities?

Our second, and perhaps main reason for our extensive discussion of the US systems and the dangers inherent in them, is that we believe that the current proposals for an EU PNR system in reality seek to introduce similar schemes in Europe. In particular, as we will show at *V.v*, below, the "purpose-specification" provisions in the proposed EU PNR Directive, if read closely, appear to be specifically geared to this end.

Before addressing these specific matters, we first, in the next section, II.ii, refer to descriptions of "strategic surveillance" in Europe, which we believe directly correspond to the "rule-based" analyses we described earlier for the USA. We then, in Part III, set out the European human rights- and data protection standards that should be applied to the above trends and proposals, i.e., to the existing PNR transfer agreements; to any future "horizontal" regulation of such transfers; and to the proposed EU PNR/ "strategic surveillance"-facilitation scheme.

We then go on, in Part IV, to summarise the legal criticism current and proposed European PNR measures, raised by various bodies, with comments from us on the issues we find of most concern. In section IV.iv, we summarise what we believe to be the core issues in terms of the law.

[61] Spanish delegation Note to the EU Council, 5 March 2015, available at: http://www.statewatch.org/news/2015/mar/eu-council-pnr-mexico-argentina-6857-15.pdf

The above exposés and summaries of necessity had to be kept brief. It would also have been neither useful nor necessary for us to repeat the full analyses of others. Rather, we provide extensive references and links to the full texts.

II.ii “Strategic surveillance” in Europe

In case anyone doubted it, it is now becoming increasingly clear that the kind of data mining/profiling operations we described earlier with reference to the USA, are also being developed in Europe. They are briefly described, in what we feel are somewhat euphemistic terms, in an important recent report of the Council of Europe Commission for Democracy through Law (the Venice Commission), which updates some earlier reports.[62] Although this update is specifically aimed at the question of accountability of state security agencies, it actually also addresses some wider issues, including the new(ish) phenomenon of “**strategic surveillance**” (and accountability for that).

This is exactly the algorithm-based mining of bulk datasets that we have identified as our main concern, i.e., as it is put in the report, situations in which: [63]

> **the material actually examined is obtained by searching the bulk material acquired by means of computer algorithms (selectors).**

Elsewhere, the report describes this in more detail (with reference to communications content and “metadata”), as follows:[64]

> Strategic surveillance involves access both to internet and telecommunications content and to metadata. It begins with a task being given to the signals intelligence agency to gather intelligence on a phenomenon or a particular person or group. Very large quantities of content data, and metadata, are then filtered and collected in a variety of different ways.[65] The bulk content is subjected to computer analysis with the help of “selectors”.[66] These can relate to language, persons, key words concerning content (e.g. industrial products), communication paths and other technical data or all of these. This is one of the important stages for balancing personal integrity concerns against other interests. In practice,

[62] European Commission for Democracy through Law (Venice Commission), Update of the 2007 Report on The Democratic Oversight of the Security Services and Report on the Democratic Oversight of Signals Intelligence Agencies, Study No. 719/2013 (CDL-AD(2015)006), Strasbourg, 7 April 2015 (Based on comments by Mr Iain Cameron, Member, Sweden), available at:
[ADD]

[63] Para. 96, emphasis added.

[64] Paras. 46 – 48 and 50 – 51, emphasis in bold added.

[65] For technical details, see M. Cayford, C. van Gulijk & P.H.A.J.M. van Gelder, *All swept up: An initial classification of NSA surveillance technology*, in Nowakowski et al. (Eds), Safety and Reliability: Methodology and Applications, Taylor and Francis, 2015, part of the SURVEILLE research project. An explanation of the SIGINT process as a whole can be found in *chapter 2* of the report of the National Research Council of the National Academies, Bulk Collection of Signals Intelligence: Technical Options, National Academy Press, 2015 (hereinafter: “National Research Council”). [original footnote 9]
Note that the National Research Council is the research council of the US Academy of Sciences. For critical comments on the NRC report, see Bruce Schneier’s blog of 9 February 2015, at:
https://www.schneier.com/blog/archives/2015/02/national_academ.html [added]

[66] The National Research Council use “discriminant” to refer to terms employed to filter collection; as the collection process occurs in real time, the terms must of necessity be simpler than those used to search the bulk collected data (“selectors”). A “query” directed to collected data can combine several “selectors” (ibid., p. 38-9). For the sake of simplicity, “selector” is used for both terms in the present report. [original footnote 10]

> whether this process adequately limits unnecessary intrusion into innocent personal communications depends on both the relevance and specificity of the selector used and the quality of the computer algorithm employed to sort for relevant data within the parameters chosen ... (however, see also para 62 below).
>
> The bulk metadata is analysed to identify communication patterns. This usually takes the form of checking whether previously identified suspect telephone numbers (X) are in contact with other numbers (Y) and then whether Y is in contact with other numbers (Z) (so-called "contact chaining"). Contact chaining by means of metadata analysis also used for internal security and law enforcement investigations, but, as shown [in another section of the report], there are (or can be) differences, both as regards the scope and quantity of the chaining and as regards the applicable safeguards for privacy.
>
> After the initial computerized searching, and deletion/refining, human analysts subject the data which is left to further analysis, deleting irrelevant material (often called "minimization"). This is another important stage for balancing privacy concerns against other interests. The material left is further refined and added to with other intelligence material, to produce a final product which is then stored for future use, disseminated etc.
>
> ...
>
> **The process of devising and refining selectors is dynamic.** The signals intelligence agency continually tests search methods, communication channels etc. anticipating and dealing with actual or potential counter-measures by the target. In the course of such testing, useful intelligence may also be obtained.
>
> **Strategic surveillance** thus differs in a number of ways from surveillance in law enforcement or more traditional internal security operations. It does not necessarily start with a suspicion against a particular person or persons. It **can instead be proactive: finding a danger rather than investigating a known danger.** Herein lay both the value it can have for security operations, and the risks it can pose for individual rights. Prosecution is not the main purpose of gathering intelligence. The intelligence is, however, stored and used in a number of ways which can affect human rights. ...

The report acknowledges that the above "strategic surveillance" of electronic communications data is just one example of such surveillance:[67]

> One can argue, broadening the perspective, that strategic surveillance is only one part of an overarching trend towards more proactive surveillance of the population; gathering data on a large segment of the population, retaining it for a period of years and making it available for searches. Other such examples are legal requirements on companies to retain and make available airline passenger name records (PNR) data, telephony and internet metadata and financial transactions.

The main point we want to make here is that what is described above, if applied to PNR data, is exactly what is happening in the USA with the use of SFPD and PNR data in the Terrorist Screening Database. The profound implications for human rights noted by the Venice Commission and others (such as the EU Fundamental Rights Agency) with regard

[67] Para. 61.

to compulsory suspicionless retention and bulk handing over of communications data therefore also apply to compulsory suspicionless bulk handing over of PNR data.

In that regard, **the report in our opinion seriously underplays two matters** (in ways similar to what was argued in the USA).

First of all, as the words emphasised in bold above make clear, strategic surveillance (read: demands for the bulk provision of often intrusive personal data) is generally intended to be used "proactively", i.e., as a means of "*finding an [unknown] danger rather than investigating a known danger.*" But the words used are disingenuous – at least the Americans were more open in this respect: it is one of the most important, perhaps most important, aims of "strategic [=bulk, suspicionless] surveillance" to "find", not abstract "dangers" but *people* who are dangerous. In the terms used in the USA, the aim is to "identify" people who "pose a risk" – or to be more precise, to "rate" people on a risk scale (e.g., "high risk") on the basis of the "strategic surveillance" analyses. **The words "finding a danger" are as misleading as the term "identifying" in relation to people classified as "high risk", discussed earlier.**

Secondly, as can be seen above, the Venice Commission report confirms that "*the process of devising and refining selectors is dynamic*" – but it does not explain what that entails in terms of human rights implications or –safeguards. The words that follow this acknowledgment merely obscure: the point is not whether "*the signals intelligence agenc[ies] continually tests search methods*"; or that dynamic algorithms are aimed at "*anticipating and dealing with actual or potential counter-measures by the target*". The point about dynamically-"improved" search algorithms is that they lose the link with the originally simple (or at least relatively simple) "selectors". Rather, in "dynamic data mining systems, the software itself looks for "unknown correlations", not even thought about by analysts in advance, and "enhances" the results, as discussed in section IV, above.

Suffice it to note that it is clear, also from the Venice Report, that the trend in Europe too is to establish "dynamic algorithm"-based data mining and profiling operations on the exact same lines as those clearly already established in the USA (and indeed likely to be linked to those). This clearly has important implications in terms of human rights- and data protection law, as we will discuss in Part IV, after first setting out the relevant standards, in Part III, below.

- o – O – o -

PART III. THE LAW

The case-law of the European Court of Human Rights has, to date, almost entirely had to do with **targeted surveillance** (although several cases submitted in the wake of the Snowden revelations and related to "general surveillance" are now pending).[68] One of us has already summarised this case-law in a 2013 submission to the European Parliament committee's inquiry into the surveillance systems exposed by Edward Snowden. It may suffice to reprint that summary here, at I and II, with references to some other such summaries, and to some more recent cases, including the important CJEU *Data Retention Judgment* and a number of ECHR cases, added at III and IV.

I. General ECHR standards[69]

Since the 1978 case of Klass v. Germany, the ECtHR has consistently held that interception of telephone communications by State bodies, including national security agencies (NSAs), constitutes an "interference" with the right to private and family life, home and correspondence, that is guaranteed by Article 8 of the Convention. There is no doubt that the same applies equally to other forms of electronic communications surveillance (Cf. Liberty and Others, para. 56). Indeed:

> **the mere existence of legislation which allows a system for the secret monitoring of communications entails a threat of surveillance for all those to whom the legislation may be applied.** This threat necessarily strikes at freedom of communication between users of the telecommunications services and thereby amounts in itself to an interference with the exercise of [individuals'] rights under Article 8, irrespective of any measures actually taken against them. (Weber and Saravia, para. 78, emphasis added)

The Court is also particularly concerned that if intercept data are destroyed and the persons concerned are not notified of the fact that they were under surveillance, "this may serve to conceal monitoring measures which have been carried out by the authorities" (*idem*, para. 79). Such surveillance (also by [national security agencies]) must therefore be "in accordance with law", serve a "legitimate aim in a democratic society", and OF Ebe "necessary" and "proportionate" in relation to that aim.

The first of these requirements is crucial. In particular, the Court accepts that safeguarding national security, preventing disorder and preventing and fighting crime are of course "legitimate aims" of a democratic State (Klass, para. 46, cf. Weber and Saravia, para. 104) - although it is notable that in the latter case the Court did not repeat the reference to "the economic well-being of the country" that was mentioned as a further aim of the relevant surveillance law by the German Government (see para. 103).

[68] See in particular Application no. 58170/13, *Big Brother Watch and Others v. the UK*, lodged on 4 September 2013.

[69] The summary in the text is based especially on an analysis of two important decisions by the European Court of Human Rights: the inadmissibility decision in Weber and Saravia v. Germany (2006) and the judgment in Liberty and Others v. the UK (2008), that build on earlier case-law, including in particular Klass v. Germany (1978), Malone v. the UK (1984), Leander v. Sweden (1987) and S. and Marper v. the UK (2008).

Moreover, while the Court grants States "**a fairly wide margin of appreciation** in choosing the means for achieving the legitimate aim of protecting national security", it adds that:

> Nevertheless, in view of the risk that a system of secret surveillance for the protection of national security may undermine or even destroy democracy under the cloak of defending it, the Court must be satisfied that there exist **adequate and effective guarantees against abuse**. This assessment depends on all the circumstances of the case, such as the nature, scope and duration of the possible measures, the grounds required for ordering them, the authorities competent to authorise, carry out and supervise them, and the kind of **remedy** provided by the national law. (Weber and Saravia, para. 106, with reference to Klass, Leander, Malone and other cases; emphases added.)

In other words, in judging whether secret surveillance is "necessary" and "proportionate", the Court looks mainly at the **nature and quality of the law** in question, and at the **available safeguards** against abuse. I will now look at those more closely

In accordance with law

On the point of whether surveillance is "in accordance with law", the Court has developed a number of "minimum safeguards", which we shall examine below. First, however, it should be noted that the Court says that "these safeguards should be set out in **statute law**" (Weber and Saravia, para. 95). In other words, these matters are so fundamental that they may not be left to subsidiary rules or –legislation. This reflects the German constitutional concept of *Gesetzesvorbehalt*, according to which certain restrictions on fundamental rights may only be imposed by statute law, i.e., by a formal law adopted by the democratic representatives of the people. It goes beyond the normal Convention requirement that interferences with fundamental rights must be based on legal rules that are "accessible" to those (potentially) affected (cf. the fourth bullet-point, below).

Minimum safeguards

The "**minimum safeguard**s that should be set out in statute law in order to avoid abuses of power"relate to the following:

- the nature of the offences in relation to which electronic surveillance may be ordered;
- the definition of the categories of people who are liable to be placed under surveillance;
- the limits on the duration of the surveillance;
- the procedure to be followed for ordering the examination, use and storage of the data obtained; these "should be set out in a form which is open to public scrutiny and knowledge";
- the precautions to be taken when communicating the data to other parties; and
- the circumstances in which the intercept data may or must be erased or destroyed.

These principles, which were first listed in this way in Weber and Saravia (para. 95, with references to earlier case-law), apply not just to "strategic monitoring" of communications based on "catchwords", but to all interceptions of and surveillance over (e-)communications (Liberty and Others, para. 63; the quote in the fourth bullet-point is from para. 67).

II. The ECHR standards applied to surveillance in practice

It is very instructive to contrast the findings in relation to these tests in Weber and Saravia v. Germany on the one hand, with those in Liberty and Others v. the UK on the other hand.

In Weber and Saravia, the Court found that the German surveillance law (the "amended G 10 Act"), as further restricted by the German Constitutional Court:

- "**defined the offences**" which could give rise to an interception order **"in a clear and precise manner**". (para. 96);
- **indicated which categories of persons** were liable to have their telephone tapped with **sufficient precision** (para. 97);
- limited interception orders to a period of **three months** (renewable as long as the statutory conditions for the order were met) (para. 98);
- set out **strict procedures** for the imposition of surveillance (in particular, for automated "strategic monitoring" through "catchwords"), including **prior authorisation** from an **independent commission** (the G10 Commission) that is appointed by **Parliament** (in consultation with the Government);
- contained sufficient "**safeguards against abuse**", including **strict purpose- (use-) limitation**-, **data disclosure- and data destruction rules** , and close oversight over surveillance by a Parliamentary Board and by the G10 Commission (cf. paras. 116, 120ff, and *passim*); and
- "effectively ensured that **the persons monitored were notified** in cases where notification could be carried out without jeopardising the purpose of the restriction of the secrecy of telecommunications." (para. 136).

In its judgment in Liberty and Others v. the UK, the Court held that surveillance in the UK, too, had a basis in domestic law, i.e., in the Interception of Communications Act 1985 (ICA) and the Regulation of Investigatory Powers Act 2000 (RIPA). However, in contrast to the case of Weber and Saravia, above, the Court held that in the UK the law:

- "allowed the executive an **extremely broad discretion** in respect of the interception of communications passing between the United Kingdom and an external receiver ... The **legal discretion** granted to the executive for the physical capture of external communications was ... **virtually unfettered**;
- the detailed "arrangements" for surveillance were contained in "**internal regulations, manuals and instructions**" that were **not contained in legislation or otherwise made available to the public**;
- the **supervision** provided by the Interception of Communications Commissioner (further discussed below), **did not contribute towards the accessibility and clarity** of the scheme, since he was not able to reveal what the "arrangements" were; consequently, the procedures to be followed for examining, using and storing

intercepted material were not "set out in a form which is open to public scrutiny and knowledge"; and

- the fact that "extensive extracts" from the Code of Practice on surveillance had belatedly been made public "suggests that it is possible for a State to make public certain details about the operation of a scheme of external surveillance without compromising national security."

The Court concluded that:

> the domestic law at the relevant time [did not indicate] with sufficient clarity, so as to provide adequate protection against abuse of power, the scope or manner of exercise of the very wide discretion conferred on the State to intercept and examine external communications. In particular, it did not, as required by the Court's case-law, set out in a form accessible to the public any indication of the procedure to be followed for selecting for examination, sharing, storing and destroying intercepted material. The interference with the applicants' rights under Article 8 was not, therefore, "in accordance with the law".
>
> It follows that there has been a violation of Article 8 in this case.
>
> (Liberty and Others, paras. 69-70)

The European Court of Human Rights considerations and minimum requirements relating to **targeted surveillance**, adduced in Sections I and II, are summarised overleaf.

III. Other summaries

The ECHR and "strategic surveillance": the Venice Commission report

The Venice Commission report, already discussed in sub-section *V.ii*, above, in relation to the description of "strategic surveillance", also contains a section on *The ECHR and strategic surveillance generally*.[70] On the more general issues, the report notes basically the same general considerations to be taken into account with regard to surveillance as are listed above, at I and II.[71]

continued on p. 47

[70] Venice Commission report (footnote 62, above), section VI.B, para. 90ff.

[71] The report also focusses on the *Weber and Saravia* and *Liberty* cases (para. 92). In relation to those, it says in that paragraph that:

> The Court has so far only looked at two cases relating to strategic surveillance, Weber and Saravia v. Germany and Liberty v. UK. The latter case concerned only the issue of "accordance with the law". The former case was an admissibility decision, albeit an unusually detailed and well-reasoned decision. But the issues of "necessity in a democratic society"/proportionality and remedies have not yet been extensively discussed by the ECtHR. Nor can it be said that the standards set out in Weber and Saravia judgment, which concern the German model, are necessarily wholly applicable to national legislation which is constructed in a different way.

It is of course true that the standards set out in *Weber and Saravia* cannot just simply be transposed to any other country or legal system. However, in the summaries provided at I and II, above, an attempt was made to still distill from the case at least the main issues and considerations to be borne in mind.

ECtHR CONSIDERATIONS & MINIMUM REQUIREMENTS RELATING TO TARGETED SURVEILLANCE:

The case-law of the ECtHR shows the following considerations and requirements of European human rights law relating to surveillance:

- A system of secret surveillance for the protection of national security may undermine or even destroy democracy under the cloak of defending it.
- The mere existence of legislation which allows a system for the secret monitoring of communications entails a threat of surveillance for all those to whom the legislation may be applied.
- In view of these risks, there must be adequate and effective guarantees against abuse.
- The first of these is that such systems must be set out in statute law, rather than in subsidiary rules, orders or manuals. The rules must moreover be in a form which is open to public scrutiny and knowledge. Secret, unpublished rules in this context are fundamentally contrary to the Rule of Law; surveillance on such a basis would *ipso facto* violate the Convention.

The following are the “minimum safeguards” that should be enshrined in such (published) statute:

- the offences and activities in relation to which surveillance may be ordered should be spelled out in a clear and precise manner;
- the law should clearly indicate which categories of people may be subjected to surveillance;
- there must be strict limits on the duration of any ordered surveillance;
- there must be strict procedures to be followed for ordering the examination, use and storage of the data obtained through surveillance;
- there must be strong safeguards against abuse of surveillance powers, including strict purpose/use-limitations (e.g., preventing the too-easy disclosure of intelligence data for criminal law purposes) and strict limitations and rules on when data can be disclosed by NSAs to LEAs, etc.;
- there must be strict rules on the destruction/erasure of surveillance data to prevent surveillance from remaining hidden after the fact;
- persons who have been subjected to surveillance should be informed of this as soon as this is possible without endangering national security or criminal investigations, so that they can exercise their right to an effective remedy at least *ex post facto*; and
- the bodies charged with supervising the use of surveillance powers should be independent and responsible to, and be appointed by, Parliament rather than the Executive.

Under the ECHR, these principles must be applied to anyone who is affected by surveillance measures taken by any Council of Europe Member State.

In addition, European States have a “positive obligation” to protect their citizens from surveillance contrary to the above, perpetrated by any other State. *A fortiori*, they are under a legal obligation not to actively support, participate or collude in such surveillance by a non-European State.

On the need to base strategic surveillance on primary legislation, it adds the following:[72]

> the Court has stressed the need for *statute* law to govern the *main elements* of secret surveillance. Case law, even where it lays down detailed standards and comes from the supreme, or constitutional court, is in itself not sufficient to regulate the area[73] and nor is subordinate legislation. The purpose of defining powers with precision is to reduce the scope for misuse of, or overuse of, power. Where a power is framed in wide terms in a statute, and oversight is limited to checking if an agency remains within its statutory mandate, then the oversight is of limited use.[74] Moreover, other things being equal, the more the power in question interferes with privacy, the greater the potential damage to privacy if the power is misused or overused. Precision focuses the minds of everyone involved in the investigation and authorization process on their responsibilities, which are ultimately backed up by the criminal offence of misuse of office. However, the main issue here is what parts of the system can be subject to internal, i.e. secret, regulation (...).[75] This involves looking behind the idea of statutory law to identify its underlying values. These could be said to be three: foreseeability/stability, democratic legitimacy and institutional competence. A statutory regulation is more stable and more transparent than regulation by means of subordinate legislation. As regards the second of these, little need be said: suffice it to say that it is for the representatives of the people to draw balances between competing interests in an area so important as this. The third value relates to the time and expertise which the parliament has at its disposal to devise appropriate general rules, and the completeness of the debate (taking into account all the relevant factors) which accompanies, or should accompany, discussion of legislative proposals. In any event, the ECtHR dismissed the UK government's arguments in the Liberty case that the accessibility requirements should be lower.[76] The Court stated that it "does not consider that there is any ground to apply different principles concerning the accessibility and clarity of the rules governing the interception of individual communications, on the one hand, and more general programmes of surveillance, on the other."

In the Executive Summary, the report notes the following in this respect:[77]

> Form of the mandate. Most democratic states have placed at least part of the mandate of the signals intelligence function in primary legislation, as required by the ECHR. More detailed norms or guidelines are normally set out in subordinate legislation promulgated either by the executive (and made public) or by the Head

[72] Para. 98, original italics.

[73] See Heglas v. Czech Republic, No. 5935/02, 1 March 2007, para. 74. [original footnote 74]

[74] E.g. the narrowness of the review performed by the IPT (see below para. 100) can be criticized. See, e.g. Justice: Freedom from suspicion: surveillance reform for a digital age (2011) p. 133-153, Leigh, I., *A view from across the channel: intelligence oversight in the UK*, in van Laethem, W. and Vanderborght, J. (eds), Regards sur le control, Intersentia, 2013. [original footnote 75]

[75] We have omitted the reference in brackets, which is to paragraph 113, but that paragraph does not deal with the issue of secret regulations.

[76] The Court's emphasis of the accessibility requirements in this case are probably due to the wide, indeed "virtually unfettered" (para. 64) discretion the British legislation gave to the authorizing body. [original footnote 76]

[77] Executive Summary, point 7. The details to be listed in primary legislation are summarised in points 14ff. They basically correspond to those listed at II, above, and in the one-page summary on p. 44, above.

> of the relevant agency (and kept secret). There may be issues of quality of the law (foreseeability etc) in this respect.

We would put it considerably more strongly: such secret rules are incompatible with the rule of law and with the case-law of the European Court of Human Rights.

More specifically, the report notes that state activity aimed at protecting "national security" goes beyond the investigation of specific criminal offences, i.e., that the gathering of "foreign intelligence" "for the economic well-being of the country, public safety or for the prevention of disorder or crime" can also be covered – but usefully adds that the national legislator should be more precise in the stipulations of the mandates of the relevant agencies.[78] It says correctly that the Strasbourg case-law does not distinguish between surveillance by reference to the means used (cable or radio-borne);[79] and notes that strategic surveillance can interfere not just with article 8 and 13 of the Convention (i.e., private life and the right to a remedy), but also with freedom of expression and information.[80]

The report then continues as follows:[81]

> Fourthly, the ECtHR has clarified that strategic surveillance involves *multiple* interferences in personal integrity. The first interference is when there is an authorization to intercept telecommunications, i.e. when the law specifies that telecommunications companies must allow access in some way to the signals intelligence agency to all, or given categories of these communications, or the signals intelligence agency is give a legal power to acquire all or given categories of these communications. As explained [earlier in the report], for strategic surveillance of content, the material actually examined is obtained by searching the bulk material acquired by means of computer algorithms (selectors). Thus, **the implication of the ECtHR's approach is that there must be legal authority for issuing selectors as regards the content of the data, and as regards metadata, for issuing instructions for contact-chaining and otherwise analyzing this data.**
>
> The second interference is after the bulk data has been processed and analysed, at the point that it is transmitted to, and used, by authorities other than the signals intelligence agency.
>
> Thirdly and finally, the ECtHR considers that an interference with private life occurs in so far as the rules provide for the destruction of the data obtained and for the refusal to notify the persons concerned of surveillance measures taken.[82]
>
> This means that specific statutory authority – accessible and otherwise fulfilling the ECtHR's case law on quality of law – must exist for each of these interferences.

We feel there is a problem with regard to the issue we highlighted: he provision of "legal authority for issuing selectors" for the mining of bulk datasets. That would indeed

[78] Para. 93.

[79] Para. 94.

[80] Para. 95.

[81] Paras. 96 – 97, emphasis added.

[82] [*Weber and Saravia v. Germany*], para. 79. In Liberty and others, the Court contented itself with stating that it "considers that the existence of these powers, particularly those permitting the examination, use and storage of intercepted communications constituted an interference with the Article 8 rights of the applicants, since they were persons to whom these powers might have been applied" (para. 57). [original footnote 73]

be possible for straight-forward selectors, e.g., anyone on the same flight as a "targeted" person, or (say) anyone who has flown to Pakistan and asked for a hala'l meal and whose credit card is listed in another (linked) database (but without that card itself being directly linked to some serious crime). The German "G10" law actually includes provisions on the need for approval of "selectors".

However, this cannot work with regard to "dynamically-improved" algorithms. The whole point of such algorithms is that they try to "identify" "unknown correlations". This is precisely one of our main concerns.

The conclusion must be that either "dynamically-improved" algorithms should be regarded as intrinsically contrary to the ECHR, because they cannot be properly controlled; or that actually effective means of controlling them must be found, e.g., to check on how reliable the application of the algorithms is: how many "false positives" and how many "false negatives" did they generate? And were the results (unintentionally) discriminatory? As already noted, that is a much bigger challenge than is acknowledged in the above.

The ECHR and EU PNR: the FRA Opinion

Finally, in this sub-section, we must note the general summaries of the requirements of the Convention in relation to surveillance, listed by the EU Fundamental Rights Agency in its opinion on the proposed EU PNR scheme.[83]

The opinion starts with a review of issues of discrimination relating to the scheme, to which we will return later. Here, it will suffice to note the general ECHR and, more in particular, EU Charter of Fundamental Rights (CFR or "the Charter") requirements listed in the opinion.

Given that it is generally accepted, also by the EU Commission when it proposed the scheme, that the compulsory bulk handing over and use of PNR data interferes with the right tot private life and with the right to data protection,[84] the opinion focusses on the justifications for such interferences. It notes that:[85]

> Article 52 (1) of the Charter establishes conditions for the limitation of the exercise of fundamental rights and freedoms recognised in the Charter and states that any limitation must be "provided for by law" and "respect the essence of those rights and freedoms". Article 52 also stresses that, subject to the principle of proportionality, limitations are possible only if they are: necessary; genuinely meet objectives of general interest recognised by the EU; or aid in the protection of the rights and freedoms of others. In addition, Article 52 (3) of the Charter stipulates that, insofar as Charter rights are derived from the rights set out in the ECHR, a Charter right is to have the same scope and meaning as the ECHR right in question.36 Therefore, the FRA will draw especially on the case law of the ECtHR for the interpretation of the Charter.

[83] EU Fundamental Rights Agency, Opinion of the European Union Agency for Fundamental Rights on the Proposal for a Directive on the use of Passenger Name Record (PNR) data for the prevention, detection, investigation and prosecution of terrorist offences and serious crime (COM(2011) 32 final) (FRA Opinion 1/2011 – Passenger Name Record), Vienna, 14 June 2011, available at: http://www.statewatch.org/news/2011/jun/eu-pnr-fra-opinion.pdf

[84] FRA Opinion, section 2.2, first paragraph, with reference to the EU Commission's Impact Assessment of the EU PNR scheme.

[85] *Idem*, section 2.2, footnote omitted.

> In sum, limitations of certain fundamental rights are possible according to the Charter, but such limitations need to meet certain specified conditions, especially including "objectives of general interest recognised by the EU", "provided for by law", necessity and proportionality.

The opinion next notes the basic requirement of both the Charter and the Convention that any any limitation on the exercise of the rights and freedoms recognised by the Charter/the Convention be "provided for by law" or, in ECHR terminology, be "in accordance with the law"; and stresses the standard "quality" tests of foreseeability and accessibility, developed in the Strasbourg case-law but also accepted as "general principles of EC [now EU] law":[86]

> These requirements of accessibility and foreseeability as developed by the ECtHR[87] constitute an essential legal protection against arbitrariness when fundamental rights are being limited.[88] The ECtHR has held that protection against arbitrariness is even more important as regards surveillance measures, due to the heightened risks of arbitrariness in such circumstances.[89] This is relevant for the proposed EU PNR system, because it could be considered as a surveillance measure.[90] Individual passengers may be generally aware that their flight details are being recorded and exchanged but will typically know neither the assessment criteria applied nor whether or not they have been flagged by the system for further scrutiny.
>
> Therefore, any measure giving the authorities power to interfere with fundamental rights should contain explicit, detailed provisions which are sufficiently clear, sufficiently foreseeable and meet the required degree of certainty[91] with respect to their application.

Given the concurrence of the views of the EDPS and the Article 29 Working Party, the FRA is rather timid in its statement that the proposed EU PNR system "could be considered as a surveillance measure". In our view, this is beyond dispute.

[86] *Idem*, section 2.2.2.

[87] See also ECtHR, *Malone v. UK*, No. 8691/79, 2 August 1984; ECtHR, *Kruslin v. France*, No. 11801/85, 24 April 1990; ECtHR; *Khan v. UK*, No. 35394/97, 12 May 2000; ECtHR, *Vetter v. France*, No. 59842/00, 31 May 2005. [original footnote 43]

[88] ECtHR, *Liberty and Others v. United Kingdom*, No. 58243/00, 1 July 2008, paragraph 69. [original footnote 44]

[89] See ECtHR, *Klass and others v. Federal Republic of Germany*, No. 5029/71, 6 September 1978. [original footnote 45]

[90] Both the EDPS and the Article 29 Working Party concur: EDPS, *Opinion on the Proposal for a Directive of the European Parliament and of the Council on the use of Passenger Name Record data for the prevention, detection, investigation and prosecution of terrorist offences and serious crime*, 25 March 2011, p. 4; Article 29 Working Party, WP 181, 5 April 2011, p. 4. [original footnote 46]

[91] In the case *Malone v. UK*, a case concerning telephone tapping, the UK Government argued that the requirements of the Convention, notably in regard to foreseeability, cannot be exactly the same in the special context of interception of communications for the purposes of police investigations as they are where the object of the relevant law is to place restrictions on the conduct of individuals. In particular, the requirement of foreseeability cannot mean that an individual should be enabled to foresee when the authorities are likely to intercept his communications so that he can adapt his conduct accordingly. The ECtHR accepted this argument but held nevertheless that the law must be sufficiently clear in its terms to give citizens an adequate indication as to the circumstances in which and the conditions on which public authorities are empowered to resort to this secret and potentially dangerous interference with their right to respect for private life and correspondence. See ECtHR, *Malone v. UK*, No. 8691/79, 2 August 1984, paragraph 67. [original footnote 47]

The FRA opinion also discusses the “necessity” and “proportionality” principles that underpin much of the Strasbourg and Luxembourg case-law, with reference to the leading Strasbourg cases, such as *Handyside*,[92] adding a useful reference to a summary by the General Secretariat of the EU Council:[93]

> The General Secretariat of the Council sums up the principles of necessity and proportionality as follows: “It is settled case-law of the Court of Justice of the European Union that the principle of proportionality, which is one of the general principles of European Union law, requires that measures implemented by acts of the European Union are appropriate for attaining the objective pursued and do not go beyond what is necessary to achieve it. Furthermore, the necessary and proportionate limitations must respect the essence of the fundamental rights concerned.”[94]

We will return to this requirement of “appropriateness” or “suitability” of the measures in terms of achievement of the relevant “legitimate aim”, and to the question of the “essence” of the right in our summary of the core issues, in sub-section *V.vi*, below. There, we will also discuss the specific references to the (flimsy and unconvincing) evidence provided to support the claim of the measures’ efficacy.

IV. <u>Further relevant case-law</u>

European Court of Human Rights judgments:

There have been a number of cases in the last seven years or so that are relevant to this report. Below, we first note to cases with some relevance, before discussing at greater length another case that in our view has some major implications in relation to surveillance, “strategic surveillance” and data mining.

First of all, in the case of *Cemalettin Canli v. Turkey*,[95] the Court reaffirmed (with reference to *Amann* and *Rotaru*) that “public information” can fall within the scope of “private life” where it is systematically collected and stored in files held by the authorities (para. 33) – i.e., that the creation of such systematic records in itself constitutes an “interference” with the right to private life. The fact that the information was already in the public domain does not change this.

It furthermore held that if errors in such a record are not corrected, as they should be by law, the failure to do so means that the interference is not “in accordance with law”, and thus in violation of the Convention (paras. 42 – 44).

The “systematic records” referred to will always include any automated filing systems (since these are by their nature “systematic”), and will also cover what in the EC Data Protection Directive (Directive 95/46/EC) is called a “personal data filing system”, which is defined as:

> any structured set of personal data which are accessible according to specific criteria, whether centralized, decentralized or dispersed on a functional or geographical basis (Art. 2(c))

[92] Section 2.2.3, first para.

[93] *Idem*, third paragraph.

[94] Council of the European Union (2011), *Guidelines on methodological steps to be taken to check fundamental rights compatibility at the Council’s preparatory bodies*, Doc No. 10140/11, 18 May 2011. [original footnote 56]

[95] *Cemalettin Canli v. Turkey*, judgment of 18 November 2008.

In *Ciubotaru v. Moldova*,[96] the Court reaffirmed (with reference to *S and Marper v. the UK*) that "along with such aspects as name, gender, religion and sexual orientation, an individual's ethnic identity constitutes an essential aspect of his or her private life and identity"; and said that this "must be particularly true" in situations such as existed in the country of the applicant (the Republic of Moldova), "where the problem of ethnic identity has been the subject matter of social tension and heated debate for a long time" (para. 53).

This echoes the provisions in the Council of Europe Data Protection Convention (Convention No. 108) and in the EC Directive that require special safeguards in relation to processing of what is usually referred to as "sensitive data", defined respectively as:

> Personal data revealing racial origin, political opinions or religious or other beliefs, as well as personal data concerning health or sexual life [or] relating to criminal convictions. (Art. 6 DP Convention)
>
> Personal data revealing racial or ethnic origin, political opinions, religious or philosophical beliefs, trade-union membership [or] data concerning health or sex life (Art. 8(1) DP Directive)

It also found that where a law created "insurmountable barriers" for a person to prove his ethnic identity for an official record, when there was in fact "objectively verifiable" evidence in support of his claim, this failed to comply with the State's "positive obligations" to safeguard the person's right to respect for private and family life, and thus violated Article 8 of the Convention (paras. 57 – 59).

The case of *Segerstedt-Wiberg and Others v. Sweden*[97] is the most important. It concerned the extreme limitations that are in place in Sweden in relation to access to certain files held by the Security Police (this followed a previous period when there was a rule of absolute secrecy). The applicants all believed that quite extensive information might be held on them in such files, but were given access only to minimal numbers of documents. The Swedish courts rejected their demands for greater access, on the grounds that secrecy was essential for the work of this branch of the police. However, it did become apparent that the file on the first applicant had been opened in relation to threats that had been made against her life, in relation to her political activities; and that the files on the others had been created for "national security" purposes, largely related to the cold war and their (real or assumed) Communist sympathies.

The Court again affirmed that:

> the information about the applicants that was stored on the Security Police register and was released to them clearly constituted data pertaining to their "private life". Indeed, this embraces even those parts of the information that were public, since the information had been systematically collected and stored in files held by the authorities. (para. 72) -

and that the storage of the information at issue constituted an "interference" with the right to private life of the applicants (which was not contested by the respondent government) (para. 73).

[96] *Ciubotaru v. Moldova*, judgment of 27 April 2010.

[97] *Segerstedt-Wiberg and Others*, judgment of 6 June 2006.

The first main issue was whether the Swedish law in question, the Police Data Act, met the "quality" requirements of "law", developed by the Strasbourg Court, and in particular whether the law did not give too much discretion to the police in deciding on whom they would open a secret file. The central question in that respect concerned section 33 of the Act which is set out in para. 49 of the judgment in English translation as follows:

> The Security Police's register may contain personal information only if:
>
> 1. The person concerned by the information is suspected of having engaged in or of intending to engage in criminal activity that entails a threat to national security or a terrorist offence;
> 2. The person concerned has undergone a security check under the Security Protection Act; or
> 3. Considering the purpose for which the register is kept, there are other special reasons therefor.
>
> The register shall indicate the grounds for data entry. The government may lay down further regulations on the type of data that may be entered (Act 2003:157).

The most problematic is of course the inclusion of the open category of "special reasons" in sub-paragraph 3. In that respect, the preparatory documents to the Act explain the following:

> In order to enable the Security Police to perform the tasks assigned to them by the relevant legislation, it could in certain cases be deemed necessary to register persons also for reasons other than those laid down in sub-paragraphs 1 and 2 of section 33: for instance, persons who are connected with other persons registered under sub-paragraphs 1 and 2 of section 33; persons who could be the targets of threats; and persons who could be the object of recruitment attempts by foreign intelligence services. In order for the Security Police to be able to prevent and uncover crimes against national security, it was necessary to survey and identify potential threats and recruitment attempts. It should also be possible for the Security Police to identify links between persons who move to Sweden after participating in oppositional activities in their home countries. Moreover, it should be possible for the Security Police to register information about persons who have been smuggled into Sweden on assignment from foreign non-democratic regimes with the task of collecting information concerning fellow countrymen. There was a need to update information concerning such informers continuously. Also, information concerning contacts with foreign missions in Sweden was relevant in this context. (para. 49 of the ECtHR judgment)

The judgment goes on as follows:

> The Government stated that the fact that an individual's name had been included in the register did not necessarily mean that he or she was suspected of an offence or other incriminating activities. Other than the examples already mentioned above from the preparatory work, the Government gave the following illustrations:
>
> – he or she is in contact with someone suspected of a crime;
>
> – he or she is in contact with personnel from a foreign mission;

> – he or she has attracted the attention of a foreign intelligence service or is used by such a service;
>
> – he or she is active in a circle that has attracted the attention of a foreign intelligence service;
>
> – he or she is used by an organisation whose activities are the subject of an investigation regarding threats to security;
>
> – he or she is the referee of a foreign citizen seeking a visa;
>
> – he or she has contacted the Security Police and provided information;
>
> – he or she is contacted by the Security Police.
>
> The Government stated that information in respect of the person in question may be needed in order to determine the interests of an entity (State, organisational or individual) constituting a threat to Swedish security, and the extent and development of that threat.

The Court discussed the compatibility of the above with the Convention as follows:

> [A]s to the question regarding the quality of the law, the Court notes that, as is made clear by the terms of section 33 of the Police Data Act, "[t]he Security Police's register may contain personal information *only*" (emphasis added) on any of the grounds set out in sub-paragraphs 1, 2 or 3. The Court considers that an issue may arise, but only in relation to the apparent broadness of the ground in sub-paragraph 3 of section 33: "Considering the purpose for which the register is kept, there are other special reasons therefor". The Government stated that a person may be registered without his or her being incriminated in any way. Here the preparatory work gives some specific and clear examples: in particular, a person who is connected with another person who has been registered, a person who may be the target of a threat and a person who may be the object of recruitment by a foreign intelligence service. The Government have also given examples of wider categories, for instance "a person in contact with someone suspected of a crime". It is clear that the Security Police enjoys a certain discretion in assessing who and what information should be registered and also if there are "special reasons" other than those mentioned in sub-paragraphs 1 and 2 of section 33 (a person suspected of a crime threatening national security or a terrorist offence, or undergoing a security check).
>
> However, the discretion afforded to the Security Police in determining what constitutes "special reasons" under sub-paragraph 3 of section 33 is not unfettered. Under the Swedish Constitution, no entry regarding a citizen may be made in a public register exclusively on the basis of that person's political opinion without his or her consent. A general prohibition of registration on the basis of political opinion is further set out in section 5 of the Police Data Act. The purpose of the register must be borne in mind where registration is made for "special reasons" under sub-paragraph 3 of section 33. Under section 32 of the Police Data Act, the purpose of storing information on the Security Police register must be to facilitate investigations undertaken to prevent and uncover crimes against national security or to combat terrorism. Further limitations follow from section 34 governing the manner of recording data in the Security Police register.
>
> Against this background, the Court finds that the scope of the discretion conferred on the competent authorities and the manner of its exercise was indicated with sufficient clarity, having regard to the legitimate aim of the

> measure in question, to give the individual adequate protection against arbitrary interference.
>
> Accordingly, the interference with the respective applicants' private lives was "in accordance with the law", within the meaning of Article 8.
>
> (paras. 79 – 80, cross-references to para. 49 of the judgment omitted)

We do not agree with this reasoning. We note in particular the reference to "identification" of "potential threats", which is reminiscent of the terminology we discussed earlier, in section *IV*. It would also appear from the examples given that any contact between a person in Sweden and "persons who move to Sweden after participating in oppositional activities in their home countries" can be regarded as *ipso facto* reason for the opening of a record in the secret files, if the Security Police feels that this in any way relates to "national security". In our opinion, this is dangerously elastic.

The Court accepted, at least in principle:

> that the storage of the information in question pursued legitimate aims, namely the prevention of disorder or crime, in the case of the first applicant, and the protection of national security, in that of the remainder of the applicants. (para. 87)

However, it did then go on to look, to some extent, into whether these aims justified the retention of very old information, decades later. It is worth quoting the Court's considerations in these regards in full.

The Court started with a recapitulation of its general approach:

> While the Court recognises that intelligence services may legitimately exist in a democratic society, it reiterates that powers of secret surveillance of citizens are tolerable under the Convention only in so far as strictly necessary for safeguarding the democratic institutions (...). Such interference must be supported by relevant and sufficient reasons and must be proportionate to the legitimate aim or aims pursued. In this connection, the Court considers that the national authorities enjoy a margin of appreciation, the scope of which will depend not only on the nature of the legitimate aim pursued but also on the particular nature of the interference involved. In the instant case, the interest of the respondent State in protecting its national security and combating terrorism must be balanced against the seriousness of the interference with the respective applicants' right to respect for private life. (para. 88, references to *Klass* and *Rotaru* omitted)

It then went on to apply this to the different applicants:

> In so far as the first applicant is concerned, the Court finds no reason to doubt that the reasons for keeping on record the information relating to bomb threats in 1990 against her and certain other personalities were relevant and sufficient as regards the aim of preventing disorder or crime. The measure was at least in part motivated by the interest in protecting her security; there can be no question of any disproportionate interference with her right to respect for private life thus being entailed. The Court has received no particulars about the precise contents of the documents released to the applicant on 13 December 2002 and will not therefore examine that matter.

However, as to the information released to the second applicant (namely, his participation in a political meeting in Warsaw in 1967), the Court, bearing in mind the nature and age of the information, does not find that its continued storage is supported by reasons which are relevant and sufficient as regards the protection of national security.

Similarly, the storage of the information released to the fifth applicant could for the most part hardly be deemed to correspond to any actual relevant national security interests for the respondent State. The continued storage of the information to the effect that he, in 1969, had allegedly advocated violent resistance to police control during demonstrations was supported by reasons that, although relevant, could not be deemed sufficient thirty years later.

Therefore, the Court finds that the continued storage of the information released to the second and fifth applicants entailed a disproportionate interference with their right to respect for private life.

The information released to the third and fourth applicants raises more complex issues in that it related to their membership of the KPML(r), a political party which, the Government stressed, advocated the use of violence and breaches of the law in order to bring about a change in the existing social order. In support of their argument, the Government submitted a copy of the KPML(r) party programme, as adopted on 2-4 January 1993, and referred in particular to its Clauses 4, 22, 23 and 28 (...).

The Court observes that the relevant clauses of the KPML(r) party programme rather boldly advocate establishing the domination of one social class over another by disregarding existing laws and regulations. However, the programme contains no statements amounting to an immediate and unequivocal call for the use of violence as a means of achieving political ends. Clause 23, for instance, which contains the most explicit statements on the matter, is more nuanced in this respect and does not propose violence as either a primary or an inevitable means in all circumstances. Nonetheless, it affirms the principle of armed opposition.

However, the Court reiterates that “the constitution and programme of a political party cannot be taken into account as the sole criterion for determining its objectives and intentions; the contents of the programme must be compared with the actions of the party’s leaders and the positions they defend” (...). This approach, which the Court has adopted in assessing the necessity under Article 11 § 2 of the Convention of the dissolution of a political party, is also pertinent for assessing the necessity in the interests of national security under Article 8 § 2 of collecting and storing information on a secret police register about the leaders and members of a political party.

In this case, the KPML(r) party programme was the only evidence relied on by the Government. Beyond that, they did not point to any specific circumstance indicating that the impugned programme clauses were reflected in actions or statements by the party’s leaders or members and constituted an actual or even potential threat to national security when the information was released in 1999, almost thirty years after the party had come into existence. Therefore, the reasons for the continued storage of the information about the third and fourth applicants, although relevant, may not be considered sufficient for the purposes of the necessity test to be applied under Article 8 § 2 of the Convention. Thus, the continued storage of the information released to the respective applicants in

> 1999 amounted to a disproportionate interference with their right to respect for private life.
>
> In sum, the Court concludes that the continued storage of the information that had been released was necessary with respect to the first applicant, but not for any of the remaining applicants. Accordingly, the Court finds that there has been no violation of Article 8 of the Convention with regard to the first applicant, but that there has been a violation of this provision with regard to each of the other applicants.
>
> (paras. 89 – 92, references to other paragraphs and to other, mainly Turkish, cases omitted)

On essentially the same basis, the Court found that the rights to freedom of expression and association (Arts. 10 and 11 ECHR) had been violated:

> the Court considers that the storage of personal data related to political opinion, affiliations and activities that is deemed unjustified for the purposes of Article 8 § 2 *ipso facto* constitutes an unjustified interference with the rights protected by Articles 10 and 11. Having regard to its findings above under Article 8 of the Convention with regard to the storage of information, the Court finds that there has been no violation of these provisions with regard to the first applicant, but that there have been violations of Articles 10 and 11 of the Convention with regard to the other applicants. (para. 107)

We disagree in particular with respect to the first applicant. Specifically, the Court did not examine why "the prevention of disorder or crime" in her case required the opening of a secret Security Police file, rather than only the opening of ordinary criminal files.

We also find it somewhat ironic that the Court on the one hand quotes the prohibitions on the keeping of records on the political opinions of citizens in the Swedish Constitution and Police Data Act as important limitations on the secret Security Police files (Para. 79) – but then finds in the above paragraph 107 that, "personal data related to political opinion, affiliations and activities" of the applicants were held in the relevant files, in violation of the Convention. It would appear that the (in any case, in our opinion, rather theoretical) limitations in the Constitution and in the Police Act were not very real in these cases.

However, it is important that the Court has expressly clarified that its approach to the dissolution of political parties also informs its approach to the question of whether it is legitimate to open secret "national security" files on leaders or members of political parties (and, we assume, other kinds of political or social groups or movements): those parties (and groups and movements) should be judged by their actions, not by (perhaps rather strongly – perhaps even objectionably-worded) phrases in their formal constitutions or declarations.

We believe that this principle should be carried through also to the "strategic surveillance" issues at the heart of this report: people should not be subjected to "filtering" or datamining based on tenuous links with organisations which do not pose any real, active threats to national security. This has obvious implications in relation to allegedly "extreme" – but not actively violent – Islamist groups too. We will return to that point in our later analyses and conclusions.

The judgment is also important in relation to the effectiveness of any available remedies (taken alone or in combination). Again, as usual, the Court first summed up its approach in earlier cases, in particular *Rotaru* and *Leander*:

> Article 13 guarantees the availability at national level of a remedy to enforce the substance of the Convention rights and freedoms in whatever form they might happen to be secured in the domestic legal order. It therefore requires the provision of a domestic remedy allowing the "competent national authority" both to deal with the substance of the relevant Convention complaint and to grant appropriate relief, although Contracting States are afforded some discretion as to the manner in which they conform to their obligation under this provision. The remedy must be "effective" in practice as well as in law ([*Rotaru*], § 67).
>
> The "authority" referred to in Article 13 may not necessarily in all instances be a judicial authority in the strict sense. Nevertheless, the powers and procedural guarantees an authority possesses are relevant in determining whether the remedy is effective. Furthermore, where secret surveillance is concerned, objective supervisory machinery may be sufficient as long as the measures remain secret. It is only once the measures have been divulged that legal remedies must become available to the individual (ibid., § 69).
>
> Turning to the present case, the Court observes that the Parliamentary Ombudsperson and the Chancellor of Justice have competence to receive individual complaints and have a duty to investigate them in order to ensure that the relevant laws have been properly applied. By tradition, their opinions command great respect in Swedish society and are usually followed. However, in the above-cited *Leander* judgment (§ 82), the Court found that the main weakness in the control afforded by these officials is that, apart from their competence to institute criminal proceedings and disciplinary proceedings, they lack the power to render a legally binding decision. In addition, they exercise general supervision and do not have specific responsibility for inquiries into secret surveillance or into the entry and storage of information on the Security Police register. As it transpires from the aforementioned judgment, the Court found neither remedy, when considered on its own, to be effective within the meaning of Article 13 of the Convention (ibid., § 84).
>
> In the meantime, a number of steps have been taken to improve the remedies, notably enabling the Chancellor of Justice to award compensation, with the possibility of judicial appeal against the dismissal of a compensation claim, and the establishment of the Records Board, replacing the former National Police Board. The Government further referred to the Data Inspection Board.
>
> Moreover, it should be noted that, with the abolition of the absolute secrecy rule under former Chapter 5, section 1(2), of the Secrecy Act (when it is deemed evident that information could be revealed without harming the purposes of the register), a decision by the Security Police whether to advise a person of information kept about him or her on their register may form the subject of an appeal to the county administrative court and the Supreme Administrative Court. In practice, the former will go and consult the Security Police register and appraise for itself the contents of files before determining an appeal against a refusal by the Security Police to provide such information. ...
>
> However, the Court notes that the Records Board, the body specifically empowered to monitor on a day-to-day basis the Security Police's entry and storage of information and compliance with the Police Data Act, has no

> competence to order the destruction of files or the erasure or rectification of information kept in the files.
>
> It appears that wider powers in this respect are vested in the Data Inspection Board, which may examine complaints by individuals. Where it finds that data is being processed unlawfully, it can order the processor, on pain of a fine, to stop processing the information other than for storage. The Board is not itself empowered to order the erasure of unlawfully stored information, but can make an application for such a measure to the county administrative court. However, no information has been furnished to shed light on the effectiveness of the Data Inspection Board in practice. It has therefore not been shown that this remedy is effective.
>
> What is more, in so far as the applicants complained about the compatibility with Articles 8, 10 and 11 of the storage on the register of the information that had been released to them, they had no direct access to any legal remedy as regards the erasure of the information in question. In the view of the Court, these shortcomings are not consistent with the requirements of effectiveness in Article 13 (see *Rotaru*, cited above, § 71, and *Klass and Others*, cited above, § 71) and are not offset by any possibilities for the applicants to seek compensation (...).
>
> In the light of the above, the Court does not find that the applicable remedies, whether considered on their own or in the aggregate, can be said to satisfy the requirements of Article 13 of the Convention.
>
> Accordingly, the Court concludes that there has been a violation of this provision.
>
> (Paras. 117 – 122)

The above gives some indication of elements of redress systems that are important in judging whether a remedy provided by a particular body is "effective" in terms of the Convention:

- The "competent national authority" need "not necessarily in all instances" be a judicial authority in the strict sense – although that is clearly the preferred option in the Court's view. But if it is not a judicial body, "the powers and procedural guarantees an authority possesses are relevant in determining whether the remedy is effective";
- The body must be able to both deal with the substance of the relevant Convention complaint and to grant appropriate relief. Since several of the bodies examined "lack[ed] the power to render a legally binding decision" on a complaint, and only "exercised general supervision", or "ha[d] no competence to order the destruction of files or the erasure or rectification of information kept in the files", they could not be regarded as offering an effective remedy;
- In respect of one body, the Data Inspection Board (the Swedish Data Protection Authority), the Court found that while it seemed to have relevant powers on paper, "no information ha[d] been furnished to shed light on the effectiveness of the Data Inspection Board in practice", and that "[i]t has therefore not been shown that this remedy is effective."

In sum, any effective remedial body must have full powers to fully investigate a complaint about secret files or secret surveillance; and full powers to order the destruction or correction of the file, and/or its release to the individual concerned – and the State must provide evidence that those powers are also actually and effectively exercised in practice.

Interestingly, the Court mentioned that "a decision by the Security Police whether to advise a person of information kept about him or her on their register may form the subject of an appeal to the county administrative court and the Supreme Administrative Court", and that "in practice, the former [i.e., a lower administrative judge] will go and consult the Security Police register and appraise for itself [him- or herself] the contents of files before determining an appeal against a refusal by the Security Police to provide such information" – but still did not find that this constituted an "effective remedy", perhaps again because there was no convincing evidence that this amounted to real scrutiny and effective remedial action on the part of the courts in practice.

Finally, there is the remark of the Court, with reference to its earlier case-law, that:

> where secret surveillance is concerned, objective supervisory machinery may be sufficient as long as the measures remain secret. It is only once the measures have been divulged that legal remedies must become available to the individual (para. 117, referring to *Rotaru*)

This is somewhat obscure. The paragraph in *Rotaru* to which the Court here refers (para. 69) in fact itself refers further back to the seminal case of *Klass*, paras. 70 – 71. Those paragraphs in *Klass* deal with the fact that under the German "G10" law, "there can be no recourse to the **courts** in respect of the ordering and implementation of restrictive measures" (in that case: secret, targeted interception of telephone communications) – although other remedies are available to any individual believing himself to be under surveillance; and that, under general civil law, **judicial redress** (in the form of a declaration of unlawfulness or the awarding of damages) can only be obtained after the surveillance has ended (although in practice this will only be possible in cases in which the individual is informed that surveillance on him has taken place).

In particular, the reference to "the measures [having been] divulged" here refers to the formal informing of an individual that he has been under surveillance.

The above remark should therefore in our opinion not be read as suggesting that non-independent internal supervisory mechanisms suffice while secret surveillance is carried out, or that redress need only be available after such surveillance has ended. That is in particular not how the case-law should be applied to long-term, untargeted "strategic surveillance" using bulk datasets, such as compulsorily obtained bulk PNR data. Rather, such surveillance should be subject to the above-mentioned full and independent and impartial remedies, with full remedial powers, indicated in the Court's consistent case-law.

The judgment of the Court of Justice of the European Union on compulsory suspicionless retention of electronic communications data (and similar national-constitutional court rulings)

In its by now famous judgment in the *Digital Rights Ireland* case,[98] the Grand Chamber of the CJEU ruled that the EC Data Retention Directive (Directive 2006/24/EC), which

[98] Joined Cases C-293/12 and C-594/12, *Digital Rights Ireland* and *Seitlinger and others*, judgment of 8 April 2014.

required the compulsory retention of electronic communications data by communication service providers, for the benefit of law enforcement investigations, to be invalid *in toto* and *ab initio*, because it was incompatible with the rights to private life and data protection, enshrined in the EU Charter of Fundamental Rights.

The case was briefly discussed, within its broader context, in the recent *Issue Paper* of the Council of Europe Commissioner for Human Rights on The Rule of Law on the Internet and in the wider digital world, as follows:[99]

> **Data protection and suspicionless data retention**
>
> Basic data-protection principles are also undermined by compulsory suspicionless untargeted retention of communications data "just in case" those data might be helpful later in a criminal investigation. This practice was imposed in the EU by the Data Retention Directive.[100] As noted in a Council of Europe publication:[101]
>
>> [Compulsory suspicionless, untargeted retention of communication records] "just in case" the data might be useful in some future police or secret service enquiry … ought to be viewed as mass surveillance of citizens without due cause: a fundamental departure from a basic principle of the rule of law.
>
> It is also fundamentally contrary to the most basic data-protection principles of purpose limitation, data minimisation and data-retention limitation.
>
> This issue is seriously aggravated by the fact that even metadata (recording when what links and communications were made in the digital environment, by whom and from what location, etc.) can be highly sensitive and revealing, often exposing, for instance, a person's race, gender, religious beliefs, sexual orientation or political and social affiliations.[102]
>
> What is more, extensive research has failed to show any significant positive effect on clear-up rates for crime, and especially not for terrorism-related crime, as a result of compulsory data retention.[103]

[99] Council of Europe Commissioner for Human Rights, The Rule of Law on the Internet and in the wider digital world (footnote 9, above), pp. 114 – 117. See also the more detailed Information Note on the case prepared for the EU Council (Council document 9009/14, 5 May 2014, leaked on the *Statewatch* website), at:
[ADD]
We quote the Legal Opinion of the European Parliament Legal Service in the text, below.

[100]. Directive 2006/24/EC of the European Parliament and of the Council of 15 March 2006 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, OJ L.105, p. 54ff. As the title shows, technically this amends the e-Privacy Directive (Directive 2002/58/EC).

[101]. Korff and Brown, "Social media and human rights", Chapter 6 in *Human rights and a changing media landscape* (Council of Europe 2011), p. 184. [original footnote 261]

[102]. See the expert witness statement of Prof. Edward Felten in the case of *ACLU vs. the NSA et al.*, at https://www.documentcloud.org/documents/781486-declaration-felten.html. The "Article 29 Working Party" opinion on surveillance, noted below, also refers to the Felten statement and usefully adds further references to judgments of the European courts stressing that metadata are equally protected under European human rights law as is content: Article 29 WP Opinion 04/2014 (note 284), pp. 4 – 5. [original footnote 262]

[103]. *Schutzlücken durch Wegfall der Vorratsdatenspeicherung? Eine Untersuchung zu Problemen der Gefahrenabwehr und Strafverfolgung bei Fehlen gespeicherter Telekommunikationsverkehrsdaten*, Max

Civil society has strongly and convincingly argued for the replacement of suspicionless data retention by data preservation (also referred to as quick-freeze of data) – the possibility for law-enforcement agencies to obtain an order requiring e-communications companies and the like to retain the communications data of people when there are factual indications that that may be helpful to the prevention, investigation or prosecution of crimes, with urgent procedures allowing for the imposition of such a measure without delay in appropriate cases, subject to *ex post facto* authorisation.[104]

Not surprisingly, laws introducing compulsory suspicionless data retention have been held to be unconstitutional in several EU member states, including Germany, with the Constitutional Court of Romania holding the very principle to be incompatible with fundamental rights.[105]

In April 2014, the Court of Justice of the EU similarly held that the Data Retention Directive violated basic principles of the EU Charter of Fundamental Rights and was invalid *ab initio*.[106] The Court criticised in particular the untargeted nature of the retention measures:

> Directive 2006/24 affects, in a comprehensive manner, all persons using electronic communications services, but without the persons whose data are retained being, even indirectly, in a situation which is liable to give rise to criminal prosecutions. It therefore applies even to persons for whom there is no evidence capable of suggesting that their conduct might have a link, even an indirect or remote one, with serious crime. ...
>
> Moreover, whilst seeking to contribute to the fight against serious crime, Directive 2006/24 does not require any relationship between the data whose retention is provided for and a threat to public security and, in particular, it is not restricted to a retention in relation (i) to data pertaining to a particular time period and/or a particular geographical zone and/or to a circle of particular persons likely to be involved, in one way or another, in a serious crime, or (ii) to persons who could, for other reasons,

Planck Institute for Comparative and International Criminal Law, 2nd enlarged report, prepared for the German Federal Ministry of Justice, July 2011, at: www.bmj.de/SharedDocs/Downloads/DE/pdfs/20120127_MPI_Gutachten_VDS_Langfassung.pdf?__blob=publicationFile. [original footnote 263]

[104]. See the Shadow evaluation report on the Data Retention Directive (2006/24/EC), produced by EDRi in April 2011, available at www.edri.org/files/shadow_drd_report_110417.pdf. [original footnote 264]

[105]. Eleni Kosta, "*The way to Luxemburg: national court decisions on the compatibility of the Data Retention Directive with the rights to privacy and data protection*", *Scripted*, Vol. 10 No. 3 (October 2013), p. 339ff, at http://script-ed.org/wp-content/uploads/2013/10/kosta.pdf. The Romanian Constitutional Court decision can be found at: www.legi-internet.ro/fileadmin/editor_folder/pdf/Decizie_curtea_constitutionala_pastrarea_datelor_de_trafic.pdf and an unofficial translation at http://www.legi-internet.ro/fileadmin/editor_folder/pdf/decision-constitutional-court-romania-data-retention.pdf (sources taken from Kosta). [original footnote 265]

[106]. Judgment of the Court of Justice of the European Union in Joined Cases C-293/12 and C-594/12, Digital Rights Ireland and Seitlinger and Others, 8 April 2014, available at: http://curia.europa.eu/juris/documents.jsf?num=C-293/12.
This follows the opinion of the Advocate-General, who had also concluded that the Directive "as a whole" was invalid and in violation of the Charter. See: http://curia.europa.eu/juris/document/document_print.jsf?doclang=EN&text=&pageIndex=0&part=1&mode=lst&docid=145562&occ=first&dir=&cid=218559. [original footnote 266]

> contribute, by the retention of their data, to the prevention, detection or prosecution of serious offences.[107]

Such untargeted compulsory data retention may therefore no longer be applied under EU law, or under national laws implementing EU law. Since most national data retention laws explicitly do exactly that, they will all have to be fundamentally reviewed and replaced with targeted surveillance measures.

Two points are worth noting after this important ruling. The Court described the legislation as a "particularly serious interference with those fundamental rights in the legal order of the EU". Despite this and despite the indication from the Court in 2007[108] that the legality of the legislation was questionable, it took eight years for the Directive to be overturned. It is also important to consider that the case only reached the Court as a result of a legal action taken by small NGOs whose very existence was threatened by the possibility of costs being awarded against them.

Second, since the ruling Member States have seemed to prefer to seek justifications to retain this serious interference with fundamental rights rather than repeal their national legal instruments transposing the Directive.

Two days after the CJEU judgement, the EU Article 29 Working Party that advises on the interpretation and application of EU data protection law issued its own opinion on state surveillance over electronic communications data, in which it cross-refered to the CJEU judgment:[109]

> From its analysis, the Working Party concludes that secret, massive and indiscriminate surveillance programs are incompatible with our fundamental laws and cannot be justified by the fight against terrorism or other important threats to national security. Restrictions to the fundamental rights of all citizens could only be accepted if the measure is strictly necessary and proportionate in a democratic society.

The CJEU judgment and the Art29WP Opinion came less than two weeks after the Human Rights Committee issued its Concluding Observations on the latest periodic

[107.] Judgment in Joined Cases C-293/12 and C-594/12 (previous note), paras. 58 – 59, emphasis added. The Court also criticised the lack of clarity over what constitutes "serious crime". [original footnote 267]

[108]. Opinion on the Promusicae/Telefónica de España case from an Advocate General, who pointed out that "there is reason to doubt, whether storing of personal data of all users – quasi on stock – is compatible with fundamental rights, in particular as this is done without any concrete suspicion" (Productores de Música de España (Promusicae) v. Telefónica de España SAU, case C-275/06, 29 January 2008). See Juliane Kokott, "Data retention – a critical side note by the Advocate General" at: http://www.libertysecurity.org/article1602.htm. [original footnote 268]

[109.] EU Article 29 Working Party, Opinion 04/2014 on surveillance of electronic communications for intelligence and national security purposes (WP215 of 10 April 2014), available at: http://ec.europa.eu/justice/data-protection/article-29/documentation/opinion-recommendation/files/2014/wp215_en.pdf. [original footnote 269]
Note that the opinion did not deal with "cable bound interception of personal data", i.e., with the alleged diversion of "full stream" data from the major high-capacity fibre-optic cables that are a major part of the backbones of the Internet. Rather, it focused in particular on access to precisely the kind of data – metadata – that are the main object of European data retention laws, and the CJEU judgment. The cross-reference to (and brief summary of) the CJEU judgment is on p. 5.

report under the ICCPR by the USA, in which it took the same view, and called upon the country to "refrain from imposing mandatory retention of data by third parties".[110]

In sum, compulsory retention of communications data is fundamentally contrary to the rule of law, incompatible with core data-protection principles, and ineffective. The EC Data Retention Directive and all national data-retention laws should be repealed and replaced by data-preservation laws.

We can only reiterate and endorse the above.

Moreover, this has clear implications for demands for PNR data too. As the Legal Service of the European Parliament put it in its answer to one of the questions on the judgment put to it:[111]

> **III.B.2 On the second question: What are the consequences on legislative proposals requiring mass collection of personal data other than traffic data, storage of the data of a very large number of unsuspected persons and access to and use of such data by law enforcement authorities?**
>
> All new and pending legislative proposals which concern the special context of general programmes of surveillance must clearly now take account of the reasoning of the Court of Justice in the DRI judgment.
>
> Indeed, the Court has declared that the EU legislature's discretion is "*reduced*" in such cases, with the result that review of that discretion should be strict. Great care must therefore be taken in such cases to ensure full respect, at all stages of the legislative procedure, for the Charter. The European Parliament, Council and Commission .must all therefore act in a spirit of mutual cooperation to this end.
>
> **The proposed ED PNR[112] and Entry/Exit System[113] (both mentioned in the request for a legal opinion) can both clearly raise such issues.[114] The data in question here are also to be processed for use by the competent national**

[110.] Human Rights Committee, Concluding observations on the fourth report of the United States of America (note 107), para. 22(d). [original footnote 270]

[111] European Parliament, Legal Opinion *re* LIBE – Questions relating to the judgment of the Court of Justice of 8 April 2014 in Jolned Cases C-293/12 and C-594/12, Digital Rights Ireland and Seitlinger and others - Directive 2006/24/EC on data retention - Consequences of the judgment, 22 December 2014, paras. 61 – 64, original italics; emphasis in bold added. The document was leaked on the *Statewatch* website at:
http://www.statewatch.org/news/2015/apr/ep-ls-opinion-digital-rights-judgment.pdf

[112] Commission proposal for a directive of the European Parliament and of the Council on the use of Passenger Name Record data for the prevention, detection, investigation and prosecution of terrorist offences and serious crime, COM(20 11) 32 final. [original footnote 47]

[113] Commission proposal for a regulation of the European Parliament and of the Council establishing an EntrylExit System (EES) to register entry and exit data of third country nationals crossing the external borders of the Member States of the European Union, COM(20 [3) 95 final. [original footnote 48]

[114] As regards the Commission proposal for the Entry/Exit System (EES) it is to be underlined that the main objective of this proposal is to improve the management of the external borders and to combat irregular immigration (Article 4 of the draft regulation). However, recital 23 leaves open the possibility of a subsequent processing of the data collected for law enforcement purposes, if such a decision is taken 2 years after the start of operation of this system. The reasoning presented in this legal opinion is mostly relevant for this potential extension of the purpose of the Entry/Exit System, in case such an extension would be considered, given that the DR]judgment itself concerned the case of personal data retained and . processed for law enforcement purposes. Nevertheless, it goes without saying that even within their primary purpose, i.e. the management of external borders, the draft regulation must respect the Charter, and in particular Articles 7 and 8 thereof. [original footnote 49]

authorities in respect of large numbers of individuals, for whom there is no evidence capable of suggesting that their conduct might have a link, even an indirect or remote one, with crime. Accordingly, these cases also fall into the category of "*general programmes of surveillance*" covered by the case-law of the European Court of Human Rights to which the Court of Justice referred in the DRI judgment.

Great care must therefore be taken to ensure that the EU legislature does not exceed its "*reduced*" discretion in these cases and that adequate safeguards and objective limits are provided for, to avoid any risk that such legislation could later be declared "*invalid*" by the Court, as in the DRI judgment. The "*strict*" method of judicial review - outlined above - which was followed by the Court of Justice in the DRI judgment will also apply in these cases also and so every effort must be made to ensure full compliance with all the various factors identified by the Court in its reasoning, where applicable due to the nature and content of each particular legislative proposal.

The Legal Service adds, in its answer to the next question:[115]

III.B.3. On the third question: What are the consequences on Union's international agreements under negotiation regarding requiring mass personal data collection other than traffic data, storage of the data of a very large number of unsuspected persons and access to and lise of such data by law .enforcement authorities?

The same considerations as just set out above will apply also in the case of international agreements under negotiation, given that the EU legislature's discretion, in external relations, to conclude international agreements, under the Treaty and in accordance with the Charter, cannot be wider than the discretion, in internal matters, to adopt EU legislation applying within the ED legal order.

As a matter of principle, equal respect for the fundamental rights of individuals which are protected by Article 7 and 8 of the Charter must be ensured in all cases, whether there is an internal or external dimension of the application of EU law.

As concerns international agreements, we may just add that Article 218(11) TFEU foresees a special procedure by which the Parliament - as well as the Council, the Commission and the Member States - "*may obtain the opinion of the Court of Justice as to whether an agreement envisaged is compatible with the Treaties.*" Where the opinion of the Court is adverse, the agreement envisaged may not enter into force unlessit is .amended or the Treaties are revised.

In cases of doubt, the Parliament may thus consider this procedure for obtaining the opinion of the Court of Justice as to whether an agreement envisaged is compatible with the Charter. This would ensure that any doubts as to the compatibility of an envisaged international agreement with the Charter may be resolved, one way or another, by the Court before the agreement is concluded and thus binds the Union under international law. This obviates any future problems and difficulties that may later arise.[116]50

[115] European Parliament, Legal Opinion (footnote 110, above), paras. 65 – 68, original italics and underlining.

[116] See Opinion 1/09 of the Court of Justice dated 8 March 201 1, paragraph 48.

See also the Resolution of the European Parliament of 25 November 2014 on seeking an opinion from the Court of Justice on the compatibility with the Treaties of the Agreement between Canada and the

The CJEU *Data Retention Judgment* thus clearly has fundamental implications also for both the EU-Third Country PNR agreements and for the proposed EU PNR scheme.

We will discuss both these matters in the light of the legal standards adduced above in the next part of this report.

- o – O – o -

European Union on the transfer and processing of Passenger Name Record data (2014/2966(RSP)). [original footnote 50]

PART IV. THE LAW APPLIED

IV.i EU-third country PNR agreements

Several agreements for the transfer of PNR data have been concluded between the EU and the USA, Canada and Australia – but not without serious and extended controversy, especially in relation to the EU-US agreements, of which there have been three. For an insight into the problems, it suffices to focus on these EU-US agreements.[117]

As explained by Hornung and Boehm:[118]

> The transfer of Passenger Name Record (PNR) data has been heavily discussed in recent years and appears to be a prototypic example of the conflicts between security interests and privacy fundamental rights which has evolved since the attacks of 11 September 2001.
>
> As PNR data [on passengers on flights to or from the EU] is usually collected by a controller which is based in an EU Member States, the respective national data protection laws apply in accordance with Article 4 (1) Directive 95/46. Companies are thus bound by both US law and the law of the respective EU Member State. As the US do not, as such, ensure an adequate level of protection as defined by Article 25 Directive 95/46, it is in principle, illegal for air carriers to transfer the data to the US. However, US law precisely obliges the air carriers to do so. There is thus a conflict of law to which there was no solution prior to the respective PNR agreements. The first PNR agreement tried to solve this problem in 2004,[119] but it was squashed by the European Court of Justice due to the lack of a legal basis for the decision of the Council.[120] In July 2007, a follow-up agreement was signed.[121] In the absence of ratification, it has since only been applied provisionally. After the entry into force of the Treaty of Lisbon, the European Parliament was requested to give its consent. The Parliament did not do so, but instead called on

[117] For an extensive overview of the problems, with comprehensive links and references to all official documents and opinions and academic and civil society criticisms, see the *Statewatch* "PNR Observatory" webpage, at:
http://www.statewatch.org/pnrobservatory.htm

[118] Gerrit Hornung & Franziska Boehm, Comparative Study on the 2011 draft Agreement between the Unites States of America and the European Union on the use and transfer of Passenger Name Records (PNR) to the United States Department of Homeland Security (study funded by the Greens/EFA group in the European Parliament), Passau/Luxembourg, March 2012. The Executive Summary of the study is set out in the text. The full text can be found at:
http://www.uni-muenster.de/Jura.itm/hoeren/itm/wp-content/uploads/PNR-Study-FINAL-120313.pdf

[119] Agreement between the European Community and the United States of America on the processing and transfer of PNR data by air carriers to the United States Department of Homeland Security, Bureau of Customs and Border Protection, OJ 2004, L 183/84 (in the following: the 2004 Agreement). [original footnote 1]

[120] Both Article 95 and Article 300 TEC were not considered to be the appropriate basis, cf. ECJ, Joined Cases C-317/04 and C-318/04, European Parliament v Council and Commission; cf. Ulrich Ehricke, Thomas Becker and Daisy Walzel, "Übermittlung von Fluggastdaten in die USA", Recht der Datenverarbeitung 2006: 149- 156; see also the case notes of Westphal, Europäische Zeitschrift für Wirtschaftsrecht 2006: 406-407 and Peter Szczekalla, Deutsches Verwaltungsblatt 2006: 896-899. [original footnote 2]

[121] Agreement between the European Union and the United States of America on the processing and transfer of Passenger Name Record (PNR) data by air carriers to the United States Department of Homeland Security (DHS), OJ 2007, L 204/18 (in the following: the 2007 Agreement 2007). [original footnote 3]

> the Commission to renegotiate and substantially improve the agreement with regards to data protection standards in its resolution of 5 May 2010.[122] After negotiations with the US, the Commission initialised the agreement and recommended to the Council so sign it.[123] The Council adopted the agreement on 13 December 2011.
>
> There are thus three succeeding PNR agreements: those of 2004 and 2007, as well as the current 2011 draft. As the Parliament had argued against the 2004 agreement, not only with regards to the lack of competence, but also in relation to the violations of fundamental rights, and requested in its resolution of 5 May 2010[124] that certain "minimum requirements" must be respected when exchanging PNR, it is of particular interest whether the current document improves the privacy and data protection rights of travellers.

The European Parliament in fact, in the end, in April 2012, "consented" to what the authors refer to as the "current 2011 draft [EU-US Agreement]";[125] and that agreement was therefore in fact adopted, and is still in force. However, since the Snowden revelations, there have been renewed calls for the suspension of the agreement. It is therefore not wrong to recall the serious criticisms, reflected not just in the Hornung/Boehm study but also in the opinions of the EU Article 29 Working Party on data protection and the European Data Protection Supervisor.

Below, we set out the summary of the conclusions reached by Hornung and Boehm, as contained in their Executive Summary (in *italics*), with brief comments (indicated by "**Comment**: ..."):[126]

1. Purpose and use of the data have been extended

When comparing the 2004, 2007 and the [2012] agreements, the purposes for which the PNR data can be used have been considerably extended. According to Article 4 of [the 2012 agreement], PNR data can be used for other purposes not related to terrorist or related crimes (i.e. border control, use if ordered by a court, other violations of law). This extension is not in line with the demands of the European Parliament formulated in its resolution of 5 May 2010.

[122] European Parliament resolution of 5 May 2010 on the launch of negotiations for Passenger Name Record (PNR) agreements with the United States, Australia and Canada, final edition B7-0244/2010. [original footnote 4]

[123] COM(2011) 807 final. [original footnote 5]

[124] See above [footnote 67]. [original footnote 6]

[125] European Parliament legislative resolution of 19 April 2012 on the draft Council decision on the conclusion of the Agreement between the United States of America and the European Union on the use and transfer of Passenger Name Records to the United States Department of Homeland Security, available at: http://www.europarl.europa.eu/sides/getDoc.do?pubRef=-//EP//TEXT+TA+P7-TA-2012-0134+0+DOC+XML+V0//EN

[126] Gerrit Hornung & Franziska Boehm, o.c. (footnote 63, above), Executive Summary. In the quoted text, we have replaced references to "the draft 2011 agreement" with "[the 2012 agreement]" since, as explained, the European Parliament "consented" to the draft text, without being able to amend it.

Comment: According to Article 4(3) of the 2012 Agreement:

> PNR may be used and processed by DHS to **identify** persons who would be subject to closer questioning or examination upon arrival to or departure from the United States or who may require further examination.

Furthermore, according to Article 16, PNR data, and "analytical data obtained from PNR", may also be shared for this purpose with other US authorities, such the NSA, provided only that the sharing is in "cases" (read: circumstances) in which this is permitted by US law, and provided that any conditions in US law in this respect are met (Art. 16(2)).

The crux lies in the term "identify" in Article 4(3). As will hopefully be clear from our discussions in section I.iii, above, this now includes "rule-based" "identification" of individuals, in the sense of computerised labelling them on a risk scale, including "high risk" when deemed appropriate.

Furthermore, Article 7 of the 2012 Agreement states that:

> The United States shall not make [*sic*] decisions that produce significant adverse actions affecting the legal interests of individuals based solely on automated processing and use of PNR.

This does not forbid the use of PNR data in the taking (or "making") of fully-automated decisions with "significant adverse" consequences, as long as the PNR data are used to this end *in combination with other data* (such as electronic communications data and/or financial transaction data, held in wider anti-terrorist databases). It also does not prohibit the use of PNR data for the purpose of "improving" the criteria used in "rule-based" risk-designations, i.e., for the purpose of "improving" the underlying (dynamic) algorithms.

In other words, the EU-US PNR Agreement can be read as allowing for the use of EU PNR data for anti-terrorist datamining and profiling by the US authorities, and is in our opinion certain to be read in this way by the US authorities (whatever the EU officials may have believed when they agreed to it).

2. Retention period has been extended

The comparison of the data retention periods show that they were constantly extended until [the 2012 agreement] eventually abolished the time limit at all, bearing the risk of repersonalization after the "anonymization" envisaged after 15 years. The indefinite retention period (in particular for data of unsuspected individuals which have never been accessed) is, however, not in line with European data protection standards. The use of undefined terms such as "anonymization", "masking out" and "repersonalization" leads to uncertainty as regards the content of those terms.

Comment: As noted in section IV, the measures envisaged as achieving "anonymisation" of data in the PNR "Big Data" dataset are meaningless, and serve as little more than figleaves to hide the actually easy reidentifiability of the data.

3. Transfer to third parties has been broadened

Although some safeguards, including the information duty and express understandings incorporating data privacy protections, are contained in [the 2012 agreement], the purpose of onward transfers is not particularly specified and not directly linked even to the very broad purposes mentioned in Article 4 [of the 2012 agreement] (as it was in the former agreements by identifying the respective paragraph). Even if the purpose of transfer is linked to the overall purpose of [the 2012 agreement], the justifications for transfers would nonetheless be wider than those of the former agreements as the provisions on purpose limitation in Article 4 have been extended.

Comment: See our comment under point 1, above.

4. Independence of supervision is still not guaranteed

The provisions regarding review and oversight have been clearly improved in [the 2012 agreement]. However, they are considerably weakened by the fact that there is no truly independent authority and indeed no mandatory oversight from outside the DHS at all. This is however, again not in line with European data protection standards.

5. Amount of data sets has not been reduced; less protection for sensitive data

There is no change or reduction of the data categories transferred to the U.S. since 2004. The already weakened protection for sensitive data from the 2007 agreement is further weakened in [the 2012 agreement].

Comment: As noted in Part III, the limitations on the use of "sensitive data" in the Agreement are hardly limiting. More specifically, they do absolutely nothing to prevent discriminatory outcomes of the "Big Data" (including "Big PNR Data") analyses by the TSC. In fact, the "non-discrimination" clause in the Agreement (Article 9) is disingenuous in this respect. It stipulates that:

> The United States shall ensure that the safeguards applicable to processing and use of PNR under this Agreement apply to all passengers on an equal basis without unlawful discrimination.

However, since these "safeguards applicable to processing and use of PNR under this Agreement" do not include any prohibition on the discriminatory labelling of certain categories of people as "high-risk", or even any requirement on the part of the US authorities to check for such discrimination-by-computer, **this article is meaningless** in that respect. **It gives an appearance of an anti-discrimination clause without actually being one.**

6. Data subject's rights and judicial review still not enforceable

Although the provisions on data subject's rights and on judicial review are more detailed than in the former agreements, it is doubtful whether the provisions of the agreement grant any new rights to EU citizens, in particular with regard to Article 21, stating that the agreement does not confer any new right to individuals. In the other provisions, the proposal mostly refers to U.S. laws which would apply to the data subjects in any case. As according to the pre-vailing opinion, U.S. laws as such do not ensure an adequate level of data protection, the reference to U.S. law in force can hardly be deemed to ensure an adequate level of data protection (as stated in Article 19).

7. Comparison between the provisions of the draft agreement and the draft Police and Justice Data Protection Directive

[The 2012 agreement] clearly does not comply with the standards of the proposed directive in many respects. Many of these shortcomings relate to the points mentioned before. Basic data protection standards are not respected. Provisions relating to the wide-ranging purposes, the very long retention period, the independency of supervision and the rights of individuals (access, correction, rectification, compensation) are far from being comparable to those of the draft police and criminal justice data protection directive. With regard to the adequacy standards in Article 34 of this proposal, it is barely understandable that Article 19 of [the 2012 agreement] states that DHS provides an adequate level of protection for PNR processing and use, "within the meaning of relevant EU data protection law".

8. Conclusion

[The 2012 PNR agreement] provides only very few improvements when compared to the 2004 and 2007 agreements and in some regards even lowers the data protection standards of the former agreements. Data transferred under the agreement can be used for purposes not related to terrorist and serious transnational crimes, retention periods have been extended, and data subject rights are still not enforceable. [The 2012 agreement] also clearly does not meet the data protection standards envisaged in the proposed directive on data protection in the field of police and criminal justice.

The **European Data Protection Supervisor**'s critical opinion focussed on the same issues, and is neatly summarised as follows in the press release issued with the EDPS opinion:[127]

- the 15-year retention period is excessive: data should be deleted immediately after its analysis or after a maximum of 6 months;
- the purpose limitation is too broad: PNR data should only be used to combat terrorism or a well-defined list of transnational serious crimes;
- the list of data to be transferred to the DHS is disproportionate and contains too many open fields: it should be narrowed and exclude sensitive data;
- there are exceptions to the "push" method: these should be removed, the US authorities should not be able to access the data directly ("pull" method);
- there are limits to the exercise of data subjects' rights: every citizen should have a right to effective judicial redress;
- the DHS should not transfer the data to other US authorities or third countries unless they guarantee an equivalent level of protection.

We have already touched on some of these issues in Part III, and will return to some of them in section IV.v, below. Suffice it to note here that we feel that, for all the serious criticism, the matter of the passing on of PNR data transferred to the TSA and DHS in the USA, by the TSA and the DHS to the more general Terrorist Screening Center for inclusion of the data in the data mining/profiling operations involving the Terroris Screening Database, has not been sufficiently addressed.

[127] The full opinion can be found here: https://secure.edps.europa.eu/EDPSWEB/webdav/site/mySite/shared/Documents/Consultation/Opinions/2011/11-12-09_US_PNR_EN.pdf

Moreover, we believe that the supposed safeguards against such further – dangerous – uses of the data are weak and effectively meaningless, both in their own terms and because, as Edward Hasbrouck has shown, the USA can in any case obtain access to essentially all (full) PNRs, through the Computerized Reservation Systems used by all the main airlines, as described next.

IV.ii How the USA by-passes the EU-US PNR Agreements

In April 2010, Hasbrouck produced a series of slides for his presentation on the global PNR systems to the European Parliament hearing on the proposed EU-US Agreement in that month.[128] It must suffice here to only set out the summary information on the last slide (but anyone seriously concerned about what happens to PNR data globally should look at the presentation and the slides in full):

PNR bypass and “leakage”

- Standard airline business processes completely bypass the DHS-EU “agreement”.
- Most PNRs follows paths that are not controlled by the DHS-EU agreement.
- Most PNRs are not stored or controlled by airlines. They are hosted by CRSs [Computerized Reservation Systems].
- In most cases, data in PNRs is transferred to a CRS in the USA, and a PNR is created in the USA, before the data reaches an airline or CRS in the EU. Once the data is in the USA, it can “leak” or bypass the agreement, without legal controls.
- CRSs are not mere messengers. The CRS in the USA retains a copy of the PNR.
- There is no US data protection law for CRSs or other travel companies. CRSs can legally share PNR data with other companies and government agencies worldwide.
- Government agencies or other third or fourth parties in the USA or other countries can obtain PNR data, in secret, from CRSs or other travel companies.
- CRSs do not keep access logs. Nobody knows who has retrieved your PNR.
- None of these activities are regulated or controlled by the DHS-EU

In our opinion, plugging this massive “hole” in the EU-US PNR arrangements should be a top priority for anyone concerned about the uses of these big datasets by the US “intelligence” agencies. Until this is done, and strong safeguards are in place that actually in practice prevent the transfer of PNR data by European airlines to the USA – and to other countries: see below – the existing EU-US PNR Agreement is simply window-dressing: it is meaningless in practice.

IV.iii PNR demands are spreading

As already noted, demands for bulk access to PNR data are spreading, and are now made by not just the USA, Australia and Canada, and the EU – but also by Russia,

[128] The slides are available here:
http://hasbrouck.org/IDP/IDP-PNR-BRU-8APR2010.pdf
Hasbrouck’s full presentation can be seen on video via the links provided here:
http://hasbrouck.org/blog/archives/001855.html
See also Hasbrouck’s general page on PNR, with numerous further references:
http://hasbrouck.org/articles/PNR.html

Mexico, the United Arab Emirates, South Korea, Brazil, Japan, Argentina and Saudi Arabia.[129]

This is of course completely unsurprising: why would such countries not follow suit and grab these data for themselves too, in bulk, either to "match" them against their own "watchlists" – or indeed for use in their own versions of the data mining and profiling operations we have discussed in relation to the USA? There have as yet been no Russian or Chinese "Edward Snowdens", but it would be surprising if these two "superpowers", at least, would not be building similar "rule-based" surveillance and analysis systems.

Yet, as Olga Enerstvedt points out:[130]

> [Having accepted PNR agreements with several countries, including the USA], and having a pending proposal on the EU PNR system, the EU has weakened its position in negotiations with [further] third countries. How will the EU deal with the Russian as well as with all the future requests for PNR?

The state of play with regard to the above-mentioned countries is not very clear. Several of them have adopted the relevant legislation, requiring the compulsory bulk provision of PNR data, also from EU airlines, but then postponed the application of the requirement in respect of those airlines – but some of these postponements have apparently now run out.

Thus, Russia adopted a PNR requirement in its laws in 2013, but postponed imposing the requirement on EU carriers, first to December of that year, and then apparently for longer. On 27 January 2014, the European Commissioner then responsible for the issue, Ms Malmström, provided the following answer to questions from MEPs:[131]

> The Commission understands that EU carriers do not transfer PNR data to the Russian authorities.
>
> The Commission is not negotiating a framework agreement with Russia. The Commission services have informed the Russian authorities about the EU legal framework for transferring personal data in the course of two meetings which took place at technical level.
>
> Parliament's Civil Liberties Committee has been informed of developments through letters sent by Commissioner Malmström to its Chairman dated 28 June, 17 September and 9 December 2013.

The issue is apparently still pending. In January 2015, the EU's foreign policy chief, Federica Mogherini, wrote in an Issue Paper on EU relations with Russia, presented at the 19 January 2015 EU Foreign Affairs Council, *inter alia* that:[132]

[129] These are the countries mentioned in the Spanish Note to the EU Council, already referred to (footnote 61, above). According to one researcher, Olga Enerstvedt, "At least six countries have PNR systems; [and] over thirty are planning to introduce them." (see next footnote for the reference). However, this thirty presumably includes the EU countries that are in the process of introducing them.

[130] Olga Mironenko Enerstvedt, Russian PNR system: Data protection issues and global prospects, available at:
http://www.sciencedirect.com/science/article/pii/S0267364913001994

[131] See:
http://www.europarl.europa.eu/sides/getAllAnswers.do?reference=P-2013-014030&language=EN

[132] EU foreign policy chief's paper on EU-Russia relations – text, Reuters quoted in the *Daily Mail Online*, available at:

> Movement on finalising negotiations on an upgraded Visa Facilitation Agreement could be coupled with demands for Russia to move on the Passenger Name Records (PNR) requirements introduced by 1.12.14, which remain unacceptable from a data protection angle.

Otherwise, according to the Spanish Note to the EU Council of last March:[133]

> **Mexico** adopted PNR legislation in 2012 requesting the transfer of passenger data from the air carriers that operate in the country. To this day, the legislation has not entered into force as Mexico has postponed its application on three occasions. **The present moratorium will expire on the 1st of April** and carriers will have to face financial sanctions of up to 30.000 dollars per flight if they do not comply and transfer the required passenger data.
>
> Until today, Mexico has given proof of flexibility to the EU, postponing the entry into force of the legislation and reducing the amount of the sanctions. However it has clearly stated that it will not extend the actual moratorium unless the EU commits to negotiating a PNR agreement setting the legal framework for the transfer of PNR data.
>
> The **Republic of Argentina** has also adopted new PNR legislation on 24 September 2014 which will enter into force on 24 March 2015.

Spain urged the EU Commission and the EU's European External Action Service to "engage urgently in a constructive dialogue with the Mexican and Argentinian authorities", in order to resolve the issues and reach PNR agreements with them.

On 27 March 2015, the EU Commissioner for Migration, Home Affairs and Citizenship, Dimitris Avramopoulos, issued a statement indicating that he did indeed want to start negotiations for an EU-Mexico PNR agreement, but also saying that he was:[134]

> considering a horizontal approach for cooperation with third countries on the use of PNR data, in the context of the preparation of the European Agenda on Security, which will be presented at the end of April.

The Agenda was duly adopted on 28 April 2015, and contains the following on future PNR agreements:[135]

> The EU has concluded **PNR agreements** with the United States, Canada and Australia. Such cooperation has real added value in identifying and apprehending foreign terrorist fighters, drug traffickers or travelling sex offenders. The Union's future approach to the exchange of PNR data with non-EU countries will take into account the need to apply consistent standards and specific fundamental rights protections. Once the European Court of Justice has issued its opinion on the draft PNR Agreement with Canada, and based on the Court's conclusions, the

http://www.dailymail.co.uk/wires/reuters/article-2911726/EU-foreign-policy-chiefs-paper-EU-Russia-relations--text.html

[133] Spanish Note to the EU Council (footnote 61, above), original emphases.

[134] Statement of Commissioner Avramopoulos on EU-Mexico PNR, 27 March 2015, available at: https://ec.europa.eu/commission/2014-2019/avramopoulos/announcements/statement-commissioner-avramopoulos-eu-mexico-pnr_en

[135] Communication from the Commission to the European Parliament, the Council, the European Economic and Social Committee and the Committee of the Regions: The European Agenda on Security, Strasbourg, 28 April 2015 (COM(2015)185final), p. 7, original emphasis, available at: http://ec.europa.eu/dgs/home-affairs/e-library/documents/basic-documents/docs/eu_agenda_on_security_en.pdf

> Commission will finalise its work on legally sound and sustainable solutions to exchange PNR data with other third countries, including by considering a model agreement on PNR setting out the requirements third countries have to meet to receive PNR data from the EU.

This will not be easy. As Tony Bunyan, the Director of *Statewatch*, pointed out:[136]

> It should come as no surprise to the EU that having put three PNR agreements in place that other countries now want the same. What is surprising is that with just two to three weeks to go until Mexico and Argentina implement their national laws the Commission is being ask to take "urgent" action. They have known about the Mexican law since 2102 and that of Argentina in September last year.
>
> Reaching agreement on new PNR deals, which meet EU data protection standards, is on past evidence going to take years especially for countries whose democratic standards and privacy laws may be questionable.

IV.iv EU PNR

Proposals for surveillance over all EU passengers go back many years.[137] The current proposals for an EU PNR Directive date from 2011.[138] They have been critically assessed by the EU Fundamental Rights Agency,[139] the EU Standing Committee of experts on international immigration, refugee and criminal law (The Meijers Committee),[140] the EU "Article 29 Working Party" on data protection,[141] the European Data Protection Supervisor,[142] academics[143] and civil society groups.[144]

[136] European Commission in a pickle over PNR, *Statewatch*, 8 March 2015, available at: http://www.statewatch.org/news/2015/mar/eu-mexico-agentina-pnr.htm

[137] For general detailed information, see the *Statewatch* "EU PNR Observatory", at: http://www.statewatch.org/Targeted-issues/eu-pnr/eu-pnr-observatory.htm
For the period 2003 – 2008, see:
http://www.statewatch.org/eu-pnrobservatory.htm

[138] Proposal for a Directive of the European Parliament and of the Council on the use of Passenger Name Record data for the prevention, detection, investigation and prosecution of terrorist offences and serious crime, 2 February 2011 (COM(2011)32final), available at:
http://www.statewatch.org/news/2011/feb/eu-com-eu-pnr-com-32-11.pdf
Impact assessment:
http://www.statewatch.org/news/2011/feb/eu-com-eu-pnr-ia-sec-132-11.pdf
Staff working paper (summary of the impact assessment):
http://www.statewatch.org/news/2011/feb/eu-com-eu-pnr-staff-working-paper-sec-133-11.pdf

[139] EU Fundamental Rights Agency, Opinion of the European Union Agency for Fundamental Rights on the Proposal for a Directive on the use of Passenger Name Record (PNR) data for the prevention, detection, investigation and prosecution of terrorist offences and serious crime (COM(2011) 32 final) (FRA Opinion 1/2011 – Passenger Name Record), Vienna, 14 June 2011, available at:
http://www.statewatch.org/news/2011/jun/eu-pnr-fra-opinion.pdf

[140] Standing Committee of experts on international immigration, refugee and criminal law (Meijers Committee), Note on the PNR Directive, available at:
http://www.statewatch.org/news/2011/jul/maijers-cttee-eu-pnr-opinion.pdf
The assessment of the Committee was prepared by Evelien Brouwer, who published this analysis with minor adjustments in a paper for the Centre for European Policy Studies (CEPS), under the title Ignoring Dissent and Legality: The EU's proposal to share the personal information of all passengers, available at:
http://www.ceps.eu/publications/ignoring-dissent-and-legality-eu%E2%80%99s-proposal-share-personal-information-all-passengers

[141] EU Article 29 Working Party, Opinion 10/2011 on the proposal for a Directive of the European Parliament and of the Council on the use of passenger name record data for the prevention, detection,

The EU Commission has given high priority to the adoption of the proposed directive. In the recently adopted Security Agenda, it said that:[145]

> Tracking the movements of offenders is key to disrupting terrorist and criminal networks. It is now urgent that the co-legislators finalise their work on the establishment of an **EU Passenger Name Record (PNR)** system for airline passengers that is fully compatible with the Charter of Fundamental Rights while providing a strong and effective tool at EU level. Analysis of PNR information provided at the time of booking and check-in helps to identify high risk travellers previously unknown to law enforcement authorities. PNR data has proven necessary to identify high risk travellers in the context of combatting terrorism, drugs trafficking, trafficking in human beings, child sexual exploitation and other serious crimes. Once adopted, the PNR Directive will ensure better cooperation between national systems and reduce security gaps between Member States. Common risk indicators for the processing of PNR data will help to prevent criminals escaping detection by travelling through another Member State. Europol and Frontex can again play a key role in developing and distributing such risk indicators on the basis of information received from Member States.

We will assess the claims of "proven necessity" and effectiveness separately below, in sub-section *VI.iii*, under the heading "*Does It Work?*". Here, we will focus on the claim that the proposed system is "fully compatible with the Charter of Fundamental Rights".

In this, we again cannot provide a full analysis of our own in this short paper. Rather, we refer to the excellent analysis of the EU PNR Directive by Evelien Brouwer for the Meyers Committee.[146] Below, we focus on a number of her conclusions (in quotes), again adding our own comments (in the main text before and after the quotes). In section IV.v, we will further discuss what we believe to be the core issues posed by PNR when looked at it the wider contexts from a European perspective, in terms of data protection and more broadly.

Purpose-specification and -limitation

Brouwer writes:[147]

> [Compared to the earlier, 2007, proposal,] the new proposal does not really narrow the scope of its application, nor does it provide extra safeguards. On the contrary, instead of limiting the goals for which member states may use PNR data, the current proposal extends the purpose of this instrument further. Whereas the

investigation and prosecution of terrorist offences and serious crime (WP181, adopted on 5 April 2011), available at:
http://ec.europa.eu/justice/policies/privacy/docs/wpdocs/2011/wp181_en.pdf

[142] Opinion of the European Data Protection Supervisor on the Proposal for a Directive of the European Parliament and of the Council on the use of Passenger Name Record data for the prevention, detection, investigation and prosecution of terrorist offences and serious crime, 25 March 2011, available at:
https://secure.edps.europa.eu/EDPSWEB/webdav/shared/Documents/Consultation/Opinions/2011/11-03-25_PNR_EN.pdf

[143] E.g., Ms Brouwer (footnote 140, above).

[144] See, e.g., the comments by European Digital Rights, at:
https://edri.org/files/101212-EU-PNR-EDRicomments.pdf

[145] European Agenda on Security (footnote 135, above), p. 7, original emphasis.

[146] See footnote 140, above.

[147] Brouwer, o.c. (footnote 140, above), section 2.

> earlier draft Framework Decision on the use of PNR data was limited to the purpose of "preventing and combating terrorist offences and organised crime", this has been changed in the new PNR proposal to "the prevention, detection, investigation and prosecution of terrorist offences and serious crime".
>
> Especially in the definitions of "prevention, detection and investigation" and "serious crimes" the national authorities are left with a wide margin of discretion, which will result in large differences among the member states implementing this Directive.
>
> ...
>
> [Moreover], consideration 28 of the preamble provides that the possibility remains for member states to oblige air carriers to transfer PNR data for purposes other than those specified in the Directive.

The FRA and the EDPS also note the latter, unacceptable, dilution of the purpose-specification and –limitation principle; and we can only agree. However, we will also go further in our criticisms of the proposals in this respect, in section IV.v, below, where we will note that even leaving preamble consideration 28 aside, the proposals lump together a variety of quite different purposes (plural), that should be separately assessed.

Using PNR data for the purpose of datamining and profiling

However, the most worrying issue in this regard is that the proposal is explicitly aimed at allowing the use of PNR data for the kind of "rule-based" "identification" of people as posing certain "risks" (e.g., as "high-risk") we described earlier, in subsection *III.ii*, above. As Brouwer puts it:[148]

> According to the explanatory memorandum, **the draft PNR Directive is aimed at achieving information on "unknown criminals or terrorists"**. Unlike other databases, such as the Schengen Information System (SIS) or Visa Information System (VIS), which provide information solely on identified persons regardless of whether they are being reported for specific goals (arrest warrants or refusal of entry), the transfer and especially analysis of PNR data should assist national authorities of the member states in identifying criminal offenders or associates or persons suspected of terrorist or serious crimes.
>
> The Commission distinguishes among three possible ways PNR data can be used: "re-active", "real-time" and "pro-active" use.
>
> **"Re-active" use** refers to use of the data in investigations, prosecutions and the unravelling of networks after a crime has been committed.
>
> With **"real-time" use**, the Commission refers to national authorities using data prior to the arrival or departure of passengers in order to prevent a crime, watch or arrest persons before a crime has been committed or because a crime has been or is being committed. In such cases PNR data may be used for running such data against predetermined assessment criteria to identify persons who were previously "unknown" to law enforcement authorities, or for running the data against various databases.

[148] Brouwer, o.c. (footnote 140, above), section 5, introduction, some hard returns and emphases added.

> Finally, **"pro-active" use** concerns the use of the data for analysis and the creation of (new) assessment criteria, which could then be used for a [future] pre-arrival and pre-departure assessment of passengers.

We note the reference by the Commission to attempts to "identify" "unknown criminals or terrorists", on the basis of "rule-based" analyses and computer-generated "criteria": this is of course exactly the same as is being done in the USA, as described in section *III.ii*, above – and subject to the same dangers as we discussed in section *IV*.

Brouwer, too, notes the risks inherent in such profiling, and relates them to "privacy and data protection, non-discrimination rights, and the right to free movement". She expands on the issue of non-discrimination, with reference to the case-law of the European Court of Human Rights, the 2010 FRA report on ethnic profiling and the Council of Europe Recommendation on profiling of the same year (CM/Rec(2010)13) and further important sources.[149] She is rightly sceptical about the Commission claims that computerised profiling will reduce discriminatory treatment:

> According to the Commission, the proposed use of PNR data has the advantage that it enables national authorities to perform "a closer screening only of persons who are most likely, based on objective assessment criteria and previous experience, to pose a threat to security". This would [the Commission claims] facilitate the travel of all other passengers and reduce the risk of passengers being subjected "to screening on the basis of unlawful criteria such as nationality or skin colour which may wrongly be associated with security risks by law enforcement authorities, including customs and border guards".
>
> The Commission addresses an important problem of current border controls and the risk that these controls are led by discriminatory considerations. It nonetheless seems questionable whether the aforementioned use of "pre-determined criteria" will actually result in less discrimination at the borders or whether it just changes the moment of screening by the PIUs. Both methods will have the same result, namely that a person may be refused entry or subjected to further investigation measures on the basis of "pre-determined criteria", or in other words, the use of profiling.

We would put it stronger: the EU Commission's claim that computerised profiling on the basis of "pre-determined criteria" – i.e., algorithms entered into a computer, and then dynamically "improved" – is likely to *reduce* discrimination is not just fanciful; it is seriously misleading, contrary to the evidence, and dangerous.

Brouwer also notes the weaknesses in the supposed safeguards in this respect:[150]

> Art. 5(6) of the current proposal provides that competent authorities may not take any decision that produces an adverse legal effect on a person or significantly affects a person "on the basis of a person's race or ethnic origin, religious or philosophical belief, political opinion, trade union membership, health or sexual life". **Although this general prohibition of discriminatory decision-making is to be welcomed, it does not exclude that the analysis or assessment of PNR data by the PIU may be based on one or more of the aforementioned criteria. This means that indirectly, on the grounds of this Directive, decision-**

[149] See section 5.1 and the footnotes to that section. We will note some even wider implications – including the undermining of "respect for the human identity" – in our later analyses, at IV.v, below.

[150] Last paragraph of section 5.1, emphasis added.

> **making by competent authorities based on one of these discrimination grounds is still possible.**
>
> **Furthermore, the reference to "decisions" does not make clear that this prohibition also applies to the measures of national authorities, including physical measures such as searches or preventing persons from entering the territory.**

These are important conclusions, which we fully endorse. We would only add that, as explained in section I.iii (with reference to the seminal work by Oscar Gandy) **there can be seriously detrimental discriminatory outcomes of automated data mining and profiling, even if no "sensitive data" are used: discrimination can be embedded at deeper levels (even deeper than just through "proxies" for sensitive elements, such as meal preferences), in ways that are extremely difficult to even discover, let alone counter**.

Specifically, the supposed safeguard requiring an "individual review by non-automated means" before any "action" is taken against an individual (Art. 4(2)(a) and (b)) is becoming increasingly meaningless in relation to complex, dynamically-created, algorithm-based analyses. As we have noted with regard to the fourth "high-risk" list maintained by the US TSA, there are no real, effective remedies against such computerised labelling of people. The stipulated "reviews" will be – cannot be other than be – meaningless.

The EDPS suggests that statistical reviews can be helpful in this regard, and we will return to this in section IV.v, in relation to remedies. Here, we may already note that by the time such a review is carried out, already **many people will be labelled on the basis of dubious analyses which are *inevitably* subject to the "base rate fallacy" – i.e., which will *inevitably* produce hundreds if not thousands of "false positives".**

We find such a prospect unacceptable.

Transfers to other domestic authorities

PNR data are in principle to be processed in the Member States primarily by so-called Passenger Information Units or PIUs. However, Brouwer notes that:[151]

> Art. 5(1) of the proposal obliges member states to adopt a list of competent national authorities entitled to request or receive PNR data or the results of processed PNR data by the PIUs. The Directive does not give any further specifications, however, other than that these authorities should be "competent for the prevention, detection, investigation or prosecution of terrorist offences and serious crime". A comparable mechanism has been identified in the Data Retention Directive (2006/24/EC).[152]

[151] Brouwer, o.c. (footnote 140, above), section 4.4.

[152] See Art. 4 of Directive 2006/24/EC of the European Parliament and of the Council of 15 March 2006 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, OJ L 105/54, 13.4.2006. [original footnote 11]

> The list of authorities having access to telecommunications data, published in the European Commission's recent evaluation report on the implementation of the Data Retention Directive, reveals many differences among the member states.[153]
>
> These differences concern in the first place the scope of 'competent national authorities'. According to this evaluation, 14 member states include security and intelligence services, 6 member states list tax or customs authorities (or both) and 3 list border authorities. Second, the list makes clear many differences with regard to the procedure for gaining access to the telecommunication data.
>
> Among the member states, 11 require judicial authorisation for each request for access to retained data and 3 require judicial authorisation in "most cases". In 4 member states the authorisation of a senior officer is required but that not of a judge, and in 2 member states the only condition is that the request is made in writing. In the evaluation report, the Commission states that it is necessary to assess the need for a greater degree of harmonisation with respect to the authorities having access and the procedure for obtaining access to retained data. In our opinion, the adoption of comparable mechanisms with regard to PNR data or other proposals granting national law enforcement authorities access to personal data (for example Eurodac) should wait for the outcome of such an evaluation.

We could not agree more. In particular, **<u>the possibility that PNR data are passed on to "security and intelligence agencies", possibly without a judicial warrant and with minimal procedural guarantees, is shocking in the light of the Snowden revelations – the more so since, as noticed under the previous heading, the proposed directive appears to be specifically drafted in such a way as to allow the use of PNR data in anti-terrorist datamining and profiling systems of the kind described earlier with regard to the USA – but which are operated in close partnership with at least the UK's GCHQ</u>.**

<u>Transfers to foreign authorities</u>

Indeed, the proposed directive would appear to allow for – or at least not prevent – the tying in of the domestic disclosures with the global surveillance and database systems exposed by Snowden.

Brouwer writes:[154]

> [In principle,] Art. 8 of the 2011 proposal allows member states to transfer PNR data and the results of the processing of PNR data, **only on a case-by-case basis <u>and</u> if**
>
> - it is in accordance with the conditions Art. 13 of the Framework Decision 2008/977/JHA;
> - the transfer is necessary for the purposes of this Directive specified in Art. 1(2); and
> - the third country agrees to transfer to third states only when necessary for the purpose of this Directive, and only with the express authorisation of the member state.

[153] European Commission, *Evaluation report on the Data Retention Directive*, COM(2011) 225, Brussels, 18 April 2011 – see pp. 9-12. [original footnote 12]

[154] Brouwer, <u>o.c.</u> (footnote 140, above), section 4.5.

> The inclusion of the condition of "case-by-case basis" prohibits the systematic transfer to third countries; however, to ensure its effective application, this provision will need close supervision. Whereas the 2007 proposal only provided for the further transfer of PNR data, this draft also allows for the transfer of the results of the PNR analysis by the PIUs or national authorities. The reference to Art. 1(2) of the proposal excludes the transfer of PNR data for "other purposes" as mentioned in the preamble, but it does include the very wide definition of purposes as provided in Art. 4(2) of the Directive.
>
> Whereas the 2007 proposal explicitly stated that transmission to third countries may only take place in accordance with the national laws of the member state concerned and any applicable international standard, the 2011 proposal only refers to the Framework Decision 2008/977/JHA.[155]

We again fully agree that this will have to be very closely supervised. The crux lies in the word "and": under the terms of this provision, transfer of PNR data covered by the directive (i.e., received by a PIU but then possibly transferred in bulk to a national security agency), from by a Member State to a third country must meet the specifications contained in the three bullet-points *and* still be on a case-by-case basis. This is important because, as Brouwer notes, the Framework Decision referred to contains a sweeping exception clause, Article 13(3), that would otherwise allow for effectively unlimited transfers to foreign national security agencies (provided these are involved in the fight against terrorism and other serious crime, as these days they increasingly are), as long as the Member State in question thinks this serves "important public interests" (as it will always claim), and as long as the sending state "deems" the safeguards provided by the recipient third country to be "adequate".

We have three reasons to be cautious in this respect. First of all, it is nowhere spelled out what "case-by-case" means. In the traditional border control and law enforcement contexts, this would mean in relation to the case of an "identified" wanted person or suspect – i.e., a person "identified" in the traditional sense, usually by name and other details. But we assume that it can now also relate to people "identified" in the completely different sense discussed in section I.iii, above: people labelled as "high-risk" (or some other level of risk) on the basis of algorithmic analyses. And the text of the proposal makes clear that this "identification", this labelling, can have been done by some other, indeed possibly foreign, agency. Article 4(2)(b) stipulates that:

> In carrying out such an assessment [i.e., in order to "identify" any persons who *may* be involved in a terrorist offence or serious crime and who require further examination by the competent authorities referred to in Article 5 – with "identify" here having the wider, "labelling" meaning] the Passenger Information Unit may compare PNR data against relevant databases, **including** international or national databases or national mirrors of Union databases, where they are established on the basis of Union law, on persons or objects sought or under alert, in accordance with Union, international and national rules applicable to such files. (emphasis added)

The word "including" makes clear that the PIUs may match the PNR data they hold against any "relevant" database, as long as that is done "in accordance with Union,

[155] Council of the European Union, Framework Decision 2008/977/JHA of 27 November 2008 on the protection of personal data processed in the framework of police and judicial cooperation in criminal matters, OJ L 350, 30.12.2008. [original footnote 13] Emphasis added.

international and national rules applicable to such files." Note also that some countries, notably the UK, argue that EU law ("Union law") quite simply does not apply to any actions by its national security agencies. In that case, the only rules that need to be complied with are therefore international human rights- and data protection rules to which the country is subject, and the country's national laws.

In other words, we believe the text expressly allows – or at least can be read as allowing, and thus very probably will be read by at least the UK as allowing – the passing on, by EU Member States, of full PNR data, and of any relevant analysis of PNR data, relating to people on the "fourth" "high-risk" list used by the US authorities, discussed in section I.ii, above. **This means that PNR data on any person labelled "high-risk" on the seriously defective US "rule-based" watchlists can be passed on by any EU Member State – and in particular by the UK – to the US authorities, for further processing and profiling.**

Secondly, we have noticed that the UK tends to argue that all processing that in some way *relates* to the activities of its national security agencies is outside the scope of EU law – and it seems to imply that this applies not only to processing of personal data by those agencies, but also to any processing of personal data by any other entity in support of the activities of the agencies. More in particular, the UK appears to take the view that any disclosures of personal data that would otherwise be subject to EU law, is not subject to Union law if it relates to the activities of the UK national security agencies. In other words, in this view, none of the EU rules on the processing of personal data – neither the current data protection directives, nor the old "Third Pillar" instruments, including the Framework Decision, nor indeed the EU PNR Directive if it is adopted – would apply to disclosures of personal data by public- or private-sector bodies to the UK's national security agencies. The UK has perhaps not stated the above in such blunt terms – but its refusal to even discuss the activities of its national security agencies with the European Parliament, or to give any other clarification in these respects, makes us suspect that in practice that is how it applies the rules. **At the very least, before the EU PNR Directive is adopted, this should be clarified.**

And then, of course, there is the massive "hole" in the system we noted in section IV.ii, above: the fact that the US's – and quite possibly also the UK's – national security agencies in any case already have full access to almost all PNR records. In the USA, this would be, as Hasbrouck has shown, by either direct (tapped-into) access by the NSA and its sister agencies into the airlines CRSs, or by means of legal orders demanding access (while also imposing secrecy about this access). Given that Snowden has revealed that the UK's GCHQ (working closely with the US's NSA) has also tapped directly into the Internet cables running under the Atlantic Ocean (and possibly into other such cables), it would also appear highly likely that the UK, too, already has almost full access to PNRs sent through such cables to the CRSs (and even to the PIUs).

This would mean that the entire EU PNR scheme would be as easily and as completely by-passed as the EU-US PNR Agreement apparently already is. Like that agreement, the EU PNR scheme would be essentially nothing more than an empty legal facade.

Even that is not the end of it; as Brouwer notes:

> Finally, the draft Directive allows the further transfer of personal data from the third state to other third countries. Even if this requires the explicit consent of the member state concerned, it does not give other member states, national

> supervisory authorities, the EDPS or the Commission any power to control this further dissemination of passenger data.

We find this particularly worrying in the light of the fact that terrorism-related information, including data resulting from the "mining" of databases and "profiling", is apparently routinely shared between at least the countries included in the "5EYES" intelligence partnership: the USA, the UK, Australia, Canada and New Zealand. But in fact, as noted in sub-section , above, the USA apparently shares the data on its "high-risk" terrorism-related lists (including the "fourth", "rule-based" list) with as many as 22 countries.[156]

Without a very serious tightening-up of the rules on transfers of PNR data to other domestic authorities (in particular national security authorities, more in particular the UK's GCHQ) and to third-country authorities, and unless the "hole" in the system exposed by Edward Hasbrouck is closed, the EU PNR scheme will simply feed more bulk data into the massive surveillance- and data mining/profiling schemes revealed by Edward Snowden, and replicate and perpetuate them in the EU.

III.iv The core issues

In this section, we focus more specifically on the four issues that we believe are central to the concerns over "PNR":

- the problems in relation to purpose-specification and –limitation;
- the difficulty of providing effective remedies against algorithm-based labels;
- the fact that algorithmic data mining and profiling touch on the most fundamental foundation for data protection: respect for human identity –

and an issue that sometimes gets somewhat lost in the more technical discussions of human rights- and data protection law:

- whether the data mining and profiling operations that we have highlighted actually work.

Purpose-specification and -limitation

Purpose specification and –limitation is always the first and most crucial issue to address in data protection: the adequacy, relevance and updating of personal data; the necessity and proportionality of personal data and of the various forms of processing of personal data; and indeed the limitations of the use of personal data – all can be assessed only in relation to a clearly specified purpose.

We feel that amidst the numerous discussions of PNR this point, while not ignored, has not been put sufficiently centrally to the debate. In particular, we believe that full account should be taken of the recent trend toward the use of PNR data in data mining and "profiling", as a means of "identifying" people as posing "risks" in terms of terrorism. This sub-section seeks to remedy this.

The overall purpose of the 2012 EU-US PNR Agreement is extremely broadly phrased:

> The purpose of this Agreement is to ensure security and to protect the life and safety of the public. (Article 1(1))

[156] see the quote from the ACLU Report (footnote 17, above) at the bottom of p. 20, above.

However, the specific articles in the Agreement spell out a range of further, more specific purposes for which the PNR data covered by the Agreement may be used or shared. Thus, the 7th preamble consideration states that:

> DHS processes and uses PNR for the purpose [*sic*] of preventing, detecting, investigating and prosecuting terrorist offenses and transnational crime.

Other references relate to data sharing in relation to "international police and judicial cooperation" (8th preamble) and to "border security" (14th preamble). It is unclear whether this includes immigration control. There is no reference to the protection of national security as a purpose of the permitted transfers.[157]

Interestingly, when the main list of purposes is repeated in the body of the agreement, in Article 4(1), the plural is used:

> The United States collects, uses and processes PNR for the purpose**s** of preventing, detecting, investigating, and prosecuting:
>
> (a) Terrorist offenses and related crimes ...

The agreement expressly stipulates that PNR data may, for these purposes:

> be used and processed by DHS to identify persons who would be subject to closer questioning or examination upon arrival to or departure from the United States or who may require further examination.
>
> (Article 4(3))

This is clearly intended to cover both meanings of "identify", discussed in section *IV*, above, i.e., verifying that a person is a particular person on a list, and "identifying" – meaning labelling – a person as posing a certain risk (e.g., "high-risk") on the basis of an algorithm.

Coupled with the fact that the TSA and DHS may now link the passenger data they receive to other databases, and make the passenger data themselves available for "improving the criteria" (read: algorithms) used in those other databases – all for the purpose of "identifying" targets, i.e., apparently in accordance with Article 4(3) of the Agreement – this means that PNR data transferred to the USA under the 2012 EU-US PNR agreement can now be used in the wider data mining/profiling operations of the US's NSA.

The proposed EU PNR Directive in its very title claims that it seeks to enable "the use of Passenger Name Record data *for the [purposes of] prevention, detection, investigation and prosecution of terrorist offences and serious crime*". The issue is also listed on the Consultative Committee's March 2015 agenda under the agenda item "data protection and police" (agenda item 9).

The first main point to be made is that the lists in the EU-US PNR agreement and in the proposed EU PNR Directive both cover **a range of purpose*s***, plural (as we have emphasised by adding the words in square brackets above). It is crucial to break these down.

Traditional Police purposes

[157] The term only appears once, in Article 11(2), where it is stipulated that access to a data subject's data can be limited *inter alia* on grounds of national security.

As already noted in section I.i, the core police role is, or used to be, the *solving* of crimes after they have been committed, by identifying the likely perpetrators and charging them, followed by *prosecution* and *trial* (by prosecution and judicial authorities). The powers granted to the police in this respect, and the procedures to be followed, are spelled out in great detail in police- and criminal procedure laws or -codes – precisely because, although they are aimed at finding criminals and bringing them to justice, those powers, if used inappropriately, threaten the rights and fundamental freedoms of all citizens.

Thus, in German law, there are different levels of "suspicion", corresponding to different categories of "suspects", against whom different (increasingly severe) levels of measures can be taken: *Anfangsverdacht*/*Verdächtige* (initial suspicion/suspect); *Hinreichender Tatverdacht*/*Angeschuldigte* (sufficient suspicion [to charge]/person charged with a crime but not yet committed to trial/accused); *Dringender Tatverdacht*/*Beschuldigte* (strong suspicion/defendant). To be considered a "suspect" in the formal sense, there must be sufficient "coherent indications" (*verdichtende Momente*) to indicate (i) that a crime has actually been committed and (ii) that the person in question is responsible for (or criminally involved in) the crime – mere "vague suspicions" or assumptions do not suffice. Seriously intrusive measures, such as house searches, may normally only be used against people under formal investigation; and the most serious (such as pre-trial detention) only against people formally charged – and even then usually subject to further requirements, such as the relevant crimes being of a certain seriousness, or threats to the investigation of witnesses, etc.. As the English translations in brackets already indicate, similar distinctions are made in other legal systems, such as English law. Those distinctions are not always the same, but distinctions of this kind – different minimum levels of suspicion and evidence being required for the use of different types of normal or special investigative measures, subject to different safeguards – there always are.

Preventing crimes by contrast has always been a much more contentious police role. There is little controversy over the legitimacy of the police countering "imminent threats to public order or public safety", and to the granting of appropriate (proportionate) powers to the police for those purposes. Thus, of course the police must have – and in all developed states does have – the power to move people away from a suspected explosive device; and to apprehend, on the spot, without warrant, anyone posing an immediate threat to the public.

Special police powers for wider purposes

Much more problematic and contentious has always been the idea that the police (and more shadowy state agencies) should have powers – especially intrusive powers – to try and identify *possible future risks* of crimes being committed, and to try and identify *possible future perpetrators* of such *possible future crimes*, not just before they have been committed, but even before sufficient "concrete acts" have been carried out by those thus targeted to even qualify them as suspects in the formal sense.

It was not that long ago that in Western Europe at least this idea of a "secret police" aimed at such "general prevention" was regarded as not, or barely, compatible with the rule of law – which meant that, if such special (sub-) agencies were to be allowed at all, their activities should be very tightly controlled and regulated. Even the mere creation of general files on people "of interest", but not necessarily suspected (even vaguely) of any crimes, without the use of special intrusive powers (such as the *Renseignements Généraux* in France) was regarded as incompatible with the rule of law – and data protection principles.

Terrorism and other "serious organised crime"

In a number of countries, terrorism, or rather the state response to terrorism, has fundamentally changed this. From the 1970s on, "the protection of the state" has been "brought forward", to counter not just real and immediate threats and actual or at least imminent crimes, but also to "nip threats in the bud".[158] To this end, the police, or such special forces, were given increased powers of intrusion and surveillance: undercover operations, infiltration and the use of informers; "special investigative measures" such as wire-tapping against people who are not formally suspects; the planting of listening devices and the hacking of computers. Predictably, such special powers, originally claimed to be necessary in the fight against (ill-defined) terrorism, soon spread to other areas, in particular the fight against "serious organised crime" (but which are also often vaguely defined).

There have been many serious problems with the use of such special powers, especially by special anti-terrorist or forces or units, or special anti-organised crime forces or units, in many countries. To mention just a few: the illegal actions of the Royal Mounted Police in Canada against Quebec separatists in the late-1970s;[159] the 1994 scandal in the Netherlands about the methods and tactics of the "Interregional Research Team";[160] and the recent scandals about undercover police officers acting as *agents provocateur* and being involved in sexual misbehaviour in the UK.[161] The full list is long, even in Western Europe.

One main problem is that such special investigations under special legislation (brought in either as special laws or as special amendments to police laws or criminal procedure codes) specifically target people (and groups of people) against whom there is "not yet" any or sufficient evidence to even categorise them as suspects in the formal sense.

This has predictably and inevitably led to the creation of "suspect communities": groups selected for "special attention" and intrusive measures, including intensive surveillance by various means, on the basis of stereotypes (or profiles reflecting stereotypes: see below).[162]

[158] See: Sebastian Cobler, Die Gefahr geht von den Menschen aus: der vorverlegte Staatsschutz, Berlin, 1976 (and especially the second edition *["Zweite, auf der Höhe des Rechtsstaates gebrachte Auflage"]*, 1978.

[159] See Part III of the Second Report of the Commission of Inquiry concerning Certain Activities of the Royal Canadian Mounted Police (the McDonald Report), August 1981. This dealt with "activities engaged in by members of the RCMP which might be described as institutionalized wrongdoings."

[160] See the entry on "*Holland – Major police scandal*" in the Statewatch Bulletin, Vol 4 no 3 (May – June 1994), p. 7, at:
http://www.statewatch.org/docbin/bulletin/bul-4-3.pdf

[161] See: "Undercover policing: Inquiry established by Theresa May", BBC News, 12 March 2015, at:
http://www.bbc.co.uk/news/uk-31852220

[162] See: Paddy Hillyard, Suspect Community: People's Experience of the Prevention of Terrorism Acts in Britain, London, 1993 (about the targeting of Irish republicans under anti-terrorism laws in the UK);

But that aside, it means that increasingly large sections of the population in Council of Europe Member States are made the subject of attention by police forces, or special, secret or semi-secret units within national police forces, in the absence of any specific, individualised suspicion.

National security and "foreign intelligence"

The surveillance systems exposed by Snowden take this "preventive" and "predictive" action to a yet higher and more dangerous level. The even more secretive agencies involved – the USA's NSA; the UK's GCHQ; their sister agencies in the other "5EYES" states, Australia, Canada and New Zealand; and their counterparts in other countries, including Germany and Sweden – are no longer content with analysing and predicting the behaviour of groups identified as sources (or even incubators) of "potential criminals" or "potential terrorists" – which has caused plenty of problems in any case for the "suspect communities" in question, and for community relations' with the state. No, they want more. **As already noted, the whole point of the "hoovering up" of almost unimaginably large amounts of personal data, *on nearly everyone*, without distinction (except perhaps in respect of the countries' own nationals) is to "mine" the massive databases thus created, to make "risk assessments" of everyone, and to label those deemed, on the basis of these assessments – which are hidden in complex, secret algorithms – as "low", "medium" or "high risk" of "being implication in terrorism" (or other major crime).**

This is a relatively new and contentious activity. Crucially, it is important to distinguish this new purpose – the computer-assisted assessment of whole populations in terms of "risk" – clearly from the previous law enforcement-related purposes. In our view, this is central to the issue of what is, and what is not, acceptable in relation to PNR data, also in data protection-legal terms.

How the above relates to the core issue

As already noted in section I.iii above, we believe we can identify the following distinct purposes for which it is proposed that PNR data should be obtained in bulk:

1. checking the identity and credentials (e.g., visa) of an airline passenger for the purpose of verifying whether that person is entitled to enter the country [identity check and immigration control];

   NB: some countries may also have exit requirements, but these are usually related to the purposes listed at (2) and (3), below.

2. identifying "known" wanted criminals (for which one should read persons convicted of criminal offences) and persons properly categorised as suspects within the meaning of the relevant national criminal law and criminal procedure law ("known suspects");

   NB: this of course includes such identification of people convicted or formally suspected of terrorism.

Liberty, A new suspect community, London, 2003 (about the more recent targeting of Muslim communities), about which, see:
https://www.liberty-human-rights.org.uk/sites/default/files/a-new-suspect-community-october-2003.pdf

3. identifying other “known” persons on the basis of specific laws permitting action against the individuals, e.g., preventing a person from leaving a country because he has failed to pay child maintenance;

4. using PNR data to facilitate the *ex post facto* investigation of criminal offences and the *ex post facto* “identification” and prosecution of the perpetrators;

5. pro-active “identification” of “possible suspects”, i.e., the marking of people as a “probable criminal” or “possible criminal”, without those people being yet formally categorised as suspects in the criminal law/criminal procedure law sense (i.e., in the absence of any evidence against them that would suffice to properly designate them as formal suspects, in accordance with criminal procedure law); and

6. pro-active “identification” of people for “preventive targeting” on national security grounds, in cases in which no action can (yet) be taken against them under the criminal law.

In that section I.iii, we noted the misleading different meanings of the term “**identify**” in relation to these different purposes: In (1), (2) and (3), above, the term means “*confirming that a certain person (e.g., a person stopped at a border check) is a specific person named [or otherwise identified] on a list or official record or document*”; in (4), the aim is to see if there is clear evidence against a person that could suffice to formally make him a “suspect” in terms of the law; while (5) and (6) are not really about “identification” in the traditional sense at all, but relate to the labelling of a person on a risk scale (e.g., as “high-risk”), on the basis of a computer assessment.

With this clarification, we believe it is now possible to assess the compatibility of the collection and further processing of PNR data for each of the above purposes.

1. checking the identity and credentials (e.g., visa) of an airline passenger for the purpose of verifying whether that person is entitled to enter the country [identity check and immigration control]

*We can be quite straight-forward about this: **there is no need whatsoever to use full PNR data for this purpose**. Rather, the traditional API records are sufficient for this purpose.*

Any rules on the use of PNR data should therefore not apply to this purpose (except insofar as they clarify that the PNR rules are not intended to stop the use of API data for identify checks and immigration control).

On the other hand, we believe it is “necessary” and “proportionate” to ask airlines to provide the API data to immigration control in the destination country in advance, so that people already identified as not entitled to enter the destination country can be prevented from even boarding the flight; and so that checks on arrival can be speeded up.

Data retention rules should reflect the normal data protection rules: API data should not be retained for longer than necessary to carry out the above checks, except that a record can be kept of any identification of people not allowed to enter, and of any actual denial of entry (in accordance with the relevant national law).

2. **identifying "known" wanted criminals (for which one should read persons convicted of criminal offences) and persons properly categorised as suspects within the meaning of the relevant national criminal law and criminal procedure law ("known suspects")**

*Our conclusions in this respect are largely the same as in relation to the first purpose mentioned: **by and large, we can see no need to use full PNR data for this purpose. Again, the traditional API records are sufficient for this purpose.***

Any rules on the use of PNR data should therefore also not apply to this purpose (except insofar as they should again clarify that the PNR rules are not intended to stop the use of API data for such identifications of wanted convicted criminals or formal suspects).

It will also be helpful to border control and law enforcement officials to have the API data in advance, e.g., to check for aliases; and we feel that this is again "necessary" and "proportionate" for this purpose.

Beyond this, again, data retention rules should reflect the normal data protection rules: API data should not be retained for longer than necessary to carry out the above checks, except that a record can be kept of any identification of wanted "known" convicted criminals or formal suspects, and of any action taken (such as their arrest at the border).

3. **identifying other "known" persons on the basis of specific laws permitting action against the individuals, e.g., preventing a person from leaving a country because he has failed to pay child maintenance**

*The same applies as for the first two purposes: **traditional API data clearly suffice for this**; the rules on the use of PNR data should not apply to this purpose; and the normal data retention rules should apply.*

<u>In sum</u>: There is no need for the provision of PNR data to states for any of the above three purposes. For all these three purposes, the provision of API data suffices (but it can be said to be "necessary" and "proportionate" to require the API data to be sent some reasonable time in advance).

4. **using PNR data to facilitate the *ex post facto* investigation of criminal offences and the *ex post facto* "identification" and prosecution of the perpetrators**

PNR data may be relevant, at times even crucial, to criminal investigations, e.g., to determine whether a particular suspect flew to a specific place, and perhaps stayed in the same hotel, as another suspect; whether one person examined in the investigation perhaps paid for flights by another person, and what that might say about their relationship; etc..

However, these are measures that are (i) targeted, and (ii) *ex post facto*. They can be carried out subject to the normal rules for access to information in the criminal procedure code (e.g., subject to a warrant).

If it were to be shown that airlines destroy their PNRs very quickly after each flight, there might be a case for thinking about how to ensure that important evidence is not lost. However, first of all, as far as we know, airlines keep their PNRs for quite some time (possibly longer than needed for their legitimate business purposes, in contravention of European data protection law). Secondly, in this regard the same applies as has been proposed in respect of electronic communications data: ***an effective***

data preservation regime including data "freezing" orders, would suffice to overcome this possible (but in our view, in practice not arising) problem.

In sum: There is no need for the advance provision of PNR data in bulk to states in order to enable normal ex post facto, targeted criminal investigations.

5. **pro-active "identification" of "possible suspects", i.e., the marking of people as a "probable criminal" or "possible criminal", without those people being yet formally categorised as suspects in the criminal law/criminal procedure law sense (i.e., in the absence of any evidence against them that would suffice to properly designate them as formal suspects, in accordance with criminal procedure law); and**

6. **pro-active "identification" of people for "preventive targeting" on national security grounds, in cases in which no action can (yet) be taken against them under the criminal law.**

We believe that the questions of whether the use of PNR data for either of the above purposes should be allowed, or is "necessary" or "proportionate" to the above purposes, are actually the wrong questions.

This is because, in our opinion, **the very concept of "predictive policing" or "predictive protection of national security" of the above types – the *Vorverlegen* or "bringing forward" of state intrusion, to "deal" with people who are not (yet) breaking the law, but who are either labelled as "probably" or "possibly" being a terrorist or other criminal, or "predicted" to "probably" (or even "possibly") become one in future – is inherently incompatible with the rule of law.**

The whole edifice of the criminal law and police- and criminal procedure law has been built precisely to ensure that those who have not broken the law, and who are not obviously posing an immediate serious threat to live and limb or law and order, will be left alone. In a state under the rule of law, the law does not "target" such people.

The mantra: "*If you have done nothing wrong, you have nothing to fear*", apart from being abused to justify unwarranted intrusions into people's lives, should at least also have a flip-side: "*As long as you do nothing wrong, the long arm of the State will not touch you*". Otherwise, we all really do have to fear the state.

Yet such "predictive", "preventive" action is increasingly promoted as not just legitimate but somehow necessary. As the newly elected prime minister of the UK, David Cameron, said just a few days after the general election:[163]

> **For too long, we have been a passively tolerant society, saying to our citizens: as long as you obey the law, we will leave you alone. This government will conclusively turn the page on this failed approach.**

Regrettably, this was not some flippant, unconsidered remark – but spelled out in an official "No. 10" press release on an important, formal government meeting. It shows the extent to which the very concept of the rule of law is increasingly held in disdain by those fearful of "extremism".

[163] Press release: Counter-Extremism Bill - National Security Council meeting, 13 May 2015, available at: https://www.gov.uk/government/news/counter-extremism-bill-national-security-council-meeting

If the above are the only purposes for which PNR data can conceivably be useful, then any plans to allow them for those purposes should be scuppered on the basis of the above considerations alone:

<u>It cannot be acceptable in a society under the rule of law that intrusive measures are used to "target" people who have done no wrong – not even on the basis that "the computer says" that they are at some dubiously-calculated "risk" of doing some wrong in the future, or similarly dubiously calculated to have "possibly" or indeed "probably" been involved in any wrong, without the kind of evidence (even preliminary evidence) that states under the rule of law require for the imposition of repressive measures</u>.

This ought to suffice to reject any plans to allow PNR data, or any bulk data on general populations, for large-scale data mining and profiling.

However, we will still also consider the other three fundamental objections mentioned.

The problem with remedies

It should by now be clear that we believe that the central problem with the demands for the provision of PNR in bulk to the authorities, is that this is – that this can only be – aimed at facilitating data mining and profiling by means of these records, linked to other major datasets – as is clearly done in the USA and as is clearly also the main aim of the proposed EU PNR scheme: as noted above, full PNR data are simply not needed for any other, normal, legitimate law enforcement or border control purpose.

We are particularly concerned about the dangers in labelling people on a risk scale (e.g., as "high risk") on anti-terrorist lists, on the basis of such data mining and profiling, when such lists are by their very nature of highly dubious reliability, with inevitably many "false positives", i.e., people being wrongly labelled as "high risk" (cf. the discussion of the "*base rate fallacy*" in Part I, section I.iii). We are also deeply concerned about the high risk of such data mining and profiling resulting in "discrimination by computer" (as discussed under that heading in the same section).

In Part I, section I.ii, we noted that the US Government Accountability Office was disingenuous about the "mitigation processes" that are supposed to be operated by the TSA to remedy any "misidentification to" such algorithm-based lists. Not only does the GAO report show that even these processes are classified as "sensitive security information", we also noted that it is inherently near-impossible to provide serious remedies against such mis-labelling. As we put it in that section:

> **By the very nature of a list created by algorithms applied to inherently ambiguous and subjective "intelligence", such determinations are extremely difficult to challenge – and they become effectively unchallengeable if the underlying "intelligence" and the evaluations of the "intelligence" and the precise algorithm used to weigh the various elements of the "intelligence" cannot be challenged. As of course no victim of such a determination will ever be able to do.**

In Part I, section I.iii, we expanded on this, with reference to attempts by Daniel Keats Citron and Nicholas Diakopolous at trying to figure out how effective remedies could be provided against algorithm-based decisions.[164] They suggest "a new concept of technological due process" (Citron) or a new system of "algorithmic accountability" (Diakopolous). The former requires "a carefully constructed inquisitorial model of quality control", i.e., serious, deep examination of any algorithmic profiling system; the latter has tried "reverse engineering" of commercial algorithms, with some success. But as Rachel O'Connor pointed out, it is difficult to see how their suggestions could be applied to the use of algorithms in typically highly secret law enforcement and national security data mining operations.

Here, we can only repeat what we already said in that section, i.e. that:

> **We believe that trying to provide answers to that question [of how to apply such possible solutions to algorithm-based decision-making] must be one of the Consultative Committee's top priorities, in relation to commercial-, administrative-, law enforcement- and national security agencies' use of profiles – including in relation to the use of PNR data in such profiles.**

Until such answers have been found, the situation is as it is now: ***there simply are no currently available, let alone operational, remedies against the dangers of people being mis-labelled as "high risk" on an anti-terrorist list as a result of deficiencies in the algorithms used, or against discrimination-by-computer caused by the algorithms. Crucially, you simply cannot remedy such wrongs by "improving" the algorithm, or by adding more data: the dangers are inherent in the processes and can only be countered, if at all, by deep analyses and auditing of the results of the data mining. There is no indication whatsoever that such deep analyses and audits are actually carried out with the aim of protecting innocent people from being wrongly labelled. Until such analysis- and audit systems are in place, and are made transparent – with involvement of critical scientists and human rights and data protection advocates – "dynamic" algorithm-based profiling should not be permitted in a state under the rule of law.***

In simple human rights and data protection terms: there are no effective remedies available against anti-terrorist/national security "dynamic" algorithm-based data mining and profiling – and without such remedies such operations are simply not compatible with the European Convention on Human Rights, the EU Charter of Fundamental Rights, or the Council of Europe Data Protection Convention.

"Respect for human identity"

Die Würde des Menschen ist unantastbar. Sie zu achten und zu schützen ist Verpflichtung aller staatlichen Gewalt.

[164] The US National Research Council of the US National Academies has suggested trying to introduce restrictions on the use of "selectors" in querying the anti-terrorist databases ("Isolating bulk data"), and/or the auditing of the usage of bulk data. See: Bulk Collection of Signals Intelligence – Technical Options, report of the Committee on Responding to Section 5(d) of Presidential Policy Directive 28: The Feasibility of Software to Provide Alternatives to Bulk Signals Intelligence Collection, 2015, available at: http://www.nap.edu/catalog/19414/bulk-collection-of-signals-intelligence-technical-options However, these safeguards only relate to fairly straightforward querying of a bulk database, as in "find all telephone numbers called from [a specified phone number] in [a specified period]". As far as we can see, the report does not address the, in our opinion much more serious problem of providing accountability and remedies in relation to dynamically-"improved" algorithms, or decisions based on such algorithms (e.g., the classifying a person as "high risk" on an anti-terrorist list).

Jeder hat das Recht auf die freie Entfaltung seiner Persönlichkeit, soweit er nicht die Rechte anderer verletzt und nicht gegen die verfassungsmäßige Ordnung oder das Sittengesetz verstößt.

(German Constitution, Arts. 1 & 2)

L'informatique doit être au service de chaque citoyen. Son développement doit s'opérer dans le cadre de la coopération internationale. Elle ne doit porter atteinte ni à l'identité humaine, ni aux droits de l'homme, ni à la vie privée, ni aux libertés individuelles ou publiques.

(French Data Protection Law of 1978, Art. 1)

Data protection is a relatively recent, modern human right. It relates to the fear that modern computing technologies can pose a fundamental threat to basic human values. This is reflected in the German constitutional basis for data protection: the proto-right in Article 1 of the German Constitution to "[respect for] the human personality" (*das allgemeine Persönlichkeitsrecht*), and in the very first article of the French 1978 data protection law, that stipulates that information technology "may impinge neither on human identity, nor on human rights, nor on private life, nor on individual or public freedoms".

These highest values reflect the essence, the "untouchable core" of the right to data protection: whatever limitations may be permitted on the use of personal data for important public or private purposes – they should never go so far as to touch this untouchable core.

The very notion of the "untouchable core" of human rights is extensively developed in German and other constitutional case-law, and also reflected in the case-law of the European Court of Human Rights and the Court of Justice of the EU. It is expressly reflected in Article 52(1) of the Charter of Fundamental Rights of the EU, which stipulates that:

> Any limitation on the exercise of the rights and freedoms recognised by this Charter must be provided for by law and respect the essence of those rights and freedoms.

In its judgment on the Data Retention Directive (discussed earlier), the CJEU held that the essence of the rights to privacy and data protection had not been adversely affected in that case, because the directive did not permit access to the contents of the electronic communications it related to; and because it incorporated certain data protection principles, notably in relation to data security (§§ 39 and 40).

It could be argued that the PNR transfer agreements and the proposed EU PNR Directive provide for similar constraints and safeguards – that is certainly what the Commission would argue.

However, in the Data Retention case the Court did not examine the issue of "rule-based" profiling or "analytical" uses of the data. Rather, the Court noted more generally that:

> ... not only is there a general absence of limits in Directive 2006/24 but Directive 2006/24 also fails to lay down any objective criterion by which to determine the limits of the access of the competent national authorities to the data and their subsequent use for the purposes of prevention, detection or criminal prosecutions (§60)

We believe that "preventive" or "predictive" profiling of individuals on the basis of essentially unverifiable and unchallengeable "dynamic"-algorithm-based mining of bulk data, unrelated to any specific indications of wrongdoing, and without any targeting on the basis of such suspicions *does* touch on the "essence", the untouchable core of the right to privacy – and indeed violates the even more fundamental principle underpinning the right to privacy (and other rights), that states must respect "human identity". In our opinion, the PNR instruments allowing for such datamining and profiling are thus, on this basis too, incompatible with European legal principles of the most fundamental kind.

Does It Work?

It follows from the above that data mining and profiling of bulk data, without any targeted suspicion, should never be allowed in Europe (or in other regions of the world subscribing to the same principles), irrespective of their effectiveness.

However, for those who might not go along with that view, it is important to note that quite apart from the acceptability or otherwise of these actions as matters of principle, there is also **no serious, credible evidence that untargeted suspicionless data mining and profiling in general, or the use of PNR data in such activities, are effective in detecting ("identifying") terrorists or other serious criminals.**

This follows first of all from the very problems with trying to single out rare incidents from very large datasets, i.e., from **the "base rate fallacy"**, already explained at I.iii. But it simply cannot be repeated often enough, so repeat it here we will:

> **If you are looking for very rare instances in a very large data set, then no matter how well you design your algorithm, you will always end up with either excessive numbers of "false positives" (cases or individuals that are wrongly identified as belonging to the rare class), or "false negatives" (cases or individuals that do fall within in the rare, looked-for category, but that are not identified as such), or both. Such techniques should never be used to trying to "identify" (real, let alone potential) terrorists from a large dataset. Even the "identification" (in the traditional sense) of actual "known" suspects will be problematic – but any attempt to rate individuals on this basis is inherently doomed to serious failure, with many innocent people wrongly classified as "high risk", and still too many actual terrorists being left unidentified.** [165]

The Consultative Committee need not take our word for this – rather, we strongly recommend that it should seek advice from serious, disinterested statisticians on this important issue (who we are certain will confirm the above).

But two further matters are of interest. First of all, as again already noted in section *IV*, even some closely involved in the US operations, or asked to review them, are raising doubts over the efficacy of the bulk data collection and –mining exercises carried out in the fight against terrorism; while the EU Commission, still fighting a rear-guard action to facilitate such practices, is clearly incapable of providing any credible evidence of such efficacy.

[165] See again in particular the "security blog" on the issue by Bruce Schneier, Why Data Mining Won't Stop Terror (footnote 42, above).

In relation to the first, we should recall the conclusion of the US National Research Council that:

> Automated identification of terrorists through data mining (or any other known methodology) is neither feasible as an objective nor desirable as a goal of technology development efforts.[166]

And we also already noted the doubts expressed by several NSA analysts about the sense in building ever greater "SIGINT trucks".[167]

These doubts about the efficacy of the use of bulk data for preventive purposes are also reflected in an extensive study carried out by the Max Planck Institute for criminal law for the EU Commission (again also already referred to)[168] into whether compulsory suspicionless bulk retention of e-communication data had been effective in preventing crime.

Perhaps most astonishing is the finding of this 300-page study that those who support bulk data collection and –mining have singularly failed to provide any serious evidence to show that it is effective in preventing crime and terrorism – and were neither trying nor even planning to produce such evidence:[169]

> The results of the present study present a picture of a particular moment in time [*eine Momentaufnahme*, i.e., as opposed to a proper longitudinal analysis]. The current situation is hallmarked by a still very uncertain statistical basis, an absence of systematic empirical research and highly differing assessments on the part of the practitioners involved, as apparent from the qualitative interviews.
>
> ...
>
> ... appropriate [statistical] data, that would allow a quantitative analysis of the effect of compulsory suspicionless [e-communications] data retention [*Vorratsdatenspeicherung*, literally "just-in-case" collection of data] on crime clear-up rates, are up to now not recorded, and there are no plans to do this systematically either [in Germany] ... because [such statistical data collecting] is deemed to be too expensive.
>
> For the European Commission, this poses a special problem in this regard [i.e., in relation to compulsory bulk communications data collection]. No [scientific/statistical] data have yet been produced, and no such data can be produced, that would allow for a [proper, scientifically sound] evaluation of [the effectiveness in practice] of Directive 2006/24/EC [the Data Retention Directive], because no appropriate [i.e., scientifically sound] framework for the collection of such data has been planned.
>
> ...

[166] NRC, Protecting Individual Privacy in the Struggle Against Terrorists: A Framework for Program Assessment (footnote 45, above).

[167] See the quote on pp. 24-25, above, and footnote 44.

[168] *Schutzlücken durch Wegfall der Vorratsdatenspeicherung? Eine Untersuchung zu Problemen der Gefahrenabwehr und Strafverfolgung bei Fehlen gespeicherter Telekommunikationsverkehrsdaten*, Max Planck Institute for Comparative and International Criminal Law, 2012 (footnote 102, above).

[169] *Idem*, introductory sentence to the Conclusions and main text from points 3 – 6 and 8, p. 218 (our translation), emphasis added.
A rare political acknowledgment from a conservative German politician is reported here:
http://bendrath.blogspot.co.uk/2011/06/conservative-hardliner-admits-lack-of.html

> The arguments based on anecdotal evidence [*Einzelfälle*] declare such individual cases to be "typical", without this being empirically proven, or indeed provable.
>
> **To this are added references to the special dangers emanating from Islamic terrorists. [Yet] precisely in this regard there are no indications that compulsory suspicionless [e-communications] data retention has in the last years led to the prevention of any terrorist attack.** Traffic- and location data might perhaps be useful in assisting [*ex post facto*, criminal] investigations into terror attacks; they have however merely raised the question why already existing and known digital [electronic] communication traces might not have sufficed to prevent attacks.

The situation has not changed: **there is still no serious effort on the part of those who clamour, not just for continuing communications data retention, but also for further bulk "just-in-case" collections, such as the compulsory provision of full PNR data, to actually provide any serious, meaningful, scientifically valid evidence to show the efficacy of the measures in fighting serious crime or terrorism**.[170]

Rather, they keep on providing anecdotal evidence and unsubstantiated, unscientific claims of cases in which they simply say the relevant data (communications data, or PNRs) were useful, or even crucial, in solving serious crimes or preventing terrorist attacks, without further explanation.

Several critical commentators have noted this in respect of the EU Commission's "Communication" in support of the EU PNR proposals. As Brouwer puts it:

> The reasons for the (extended) use of PNR data are not clarified. In the explanatory memorandum, the Commission refers to trafficking in human beings and drug-related crime, and illustrates the human and economic costs of these crimes using rather random data from various sources, including data of the UK Home Office on costs incurred "in anticipation of crime" of 2003. Moreover, **the Commission does not provide real evidence of the added value of using PNR data for the prevention or prosecution of these crimes.** The European Commission only refers to examples in three countries (Belgium, Sweden and the UK) in which a substantial number of drug seizures would have been "exclusively or predominantly" due to the processing of PNR data. These data are not further specified, and surprisingly not mentioned at all in the impact assessment of this proposal. It also seems odd that according to the Commission, Belgium reported that 95% of all drug seizures in 2009 exclusively or predominantly stemmed from the processing of PNR data, while according to the same impact assessment Belgium would not have implemented any PNR scheme by that time.

Moreover:

[170] In its EU PNR proposal, the Commission obscures the reason for the absence of properly recorded statistics on the effect of the compulsory suspicionless data retention on clear-up rates, by saying that:

> In the absence of harmonised provisions on the collection and processing of PNR data at EU level, detailed statistics on the extent to which such data help prevent, detect, investigate and prosecute serious crime and terrorism are not available. (COM(2011) 32 final, p. 6)

But that is again typically disingenuous. The problem is not differences in statistics between the Member States, but ***a collective refusal to collect meaningful, scientifically verifiable and challengeable statistics***. The suggestion made to the MPI that this is because it is "too expensive" is of course ludicrous in the face of the many millions of dollars and euros that are spent on the creation of the bulk databases; it is wisely not repeated by the Commission.

> The Commission does not provide information on the implementation of the Directive on the use of advanced passenger information (API), which was adopted in 2004 and for which the implementation date was exceeded in September 2006.[171]
>
> ...
>
> During negotiations on earlier drafts of the API Directive, the use of API was originally planned for immigration control purposes alone. Shortly before the final adoption of the Directive, however, a provision was added according to which member states may use the passenger data for law enforcement purposes (Art. 6). One would have expected an evaluation by the Commission of the current use of the API Directive, together with the existing large-scale databases in the EU, before proposing new measures of data collection. Although Directive 2004/82/EC does not include a sunset clause or obligation for the Commission to evaluate this instrument itself, it is in line with the general policy of the Commission to assess "the initiative's expected impact on individuals' right to privacy and personal data protection and set out why such an impact is necessary and why the proposed solution is proportionate to the legitimate aim of maintaining internal security within the European Union, preventing crime or managing migration".[172] This failure to first identify the security gaps of existing systems and methods of cooperation has similarly been pointed out by the Article 29 Working Party in its opinion of April 2011.[173] According to the Working Party, if any gaps exist, then the next step should be to analyse the best way to fill these gaps by exploiting and improving the present mechanisms, without necessarily introducing a whole new system.

The EU Fundamental Rights Agency noted the same deficiencies:

> In the Explanatory Memorandum attached to the proposal, the European Commission included some examples which provide evidence of the necessity of a PNR system leading to critical progress in combating serious crime, in particular in the fight against drugs and human trafficking.[174] It is important to further analyse these examples to assess the necessity of an EU PNR system.

The FRA is quite right to demand "further analysis" of these anecdotal examples; and the same must apply to the claim by the Commission that:

> The necessity of using PNR data is ... supported by information from third countries as well as Member States that already use such PNR data for law enforcement purposes. (p. 6)

From the (in any case somewhat confusing) summaries in the Commission proposal it is not at all clear that these cases could not have been resolved without any demand for bulk PNR data; rather, it would appear that in both examples on p. 5 the authorities were already aware of the groups of criminals (human traffickers) and, indeed, of the stolen credit cards they used, and targeted requests for PNR data, or a request to all airlines and travel agencies to look out for the stolen card details, might well have been more than sufficient. The statistics from Belgium, Sweden and the UK (provided on p. 6) are as useless: they fail to distinguish between cases in which PNR data was useful in a case, but acquired for the case by

[171] Council of the European Union, Directive 2004/82/EC of 29 August 2004 on the obligation of carriers to communicate passenger data, OJ L 261/24, 6.8.2004. In June 2010, the Commission started an infraction procedure against Poland for failure to adopt the necessary laws implementing the Directive, Case C-304/10, OJ C 246/22, 11.9.2010. [original footnote 6]

[172] European Commission, Communication on Overview of information management in the area of freedom, security and justice, COM(2010) 835, Brussels, 20 July 2010, p. 25 [original footnote 7]

[173] Article 29 Data Protection Working Party, Opinion 10/2011, [adopted on 5 April 2011 (http://ec.europa.eu/justice/policies/privacy/docs/wpdocs/2011/wp181_en.pdf] [original footnote 8]

[174] COM(2011) 32 final, pp. 5-6. [original footnote 57]

> Examples of the value of PNR data can also be found in other European Commission documents. In 2010, the Commission published a communication on information management in the area of freedom, security and justice. In this communication, the Commission provided further examples for the necessity of PNR data relating to child trafficking, trafficking in human beings, credit card fraud and drug trafficking, but it did not disclose the source of its information.[175]
>
> However, examples relating to terrorism or to many of the other types of crimes defined as serious crime in Article 2 (h) of the proposal cannot be found in the Explanatory Memorandum or in the accompanying documents.

Moreover, the FRA noted:[176]

> The examples provided by the European Commission relate only to cases in which PNR data were successfully used in the course of investigations. For a more complete picture, it would also be necessary to analyse those cases in which the use of data proved to be misleading and led to the investigation of innocent people. Such a case is included by the European Union Committee of the UK House of Lords in its 2007 report on the EU/US Passenger Name Record (PNR) Agreement: the case of Maher Arar.[177]

a targeted demand, and cases in which compulsory suspicionless [bulk] data was proven to be essential. The "third country" mentioned is presumably the USA – but in that respect no information on the effectiveness of its watchlists or datamining exercises is provided at all.

[175] European Commission (2010), *Overview of information management in the area of freedom, security and justice*, COM(2010) 385 final, Brussels, 2010. [original footnote 58]

[176] FRA Opinion on the EU PNR scheme (footnote xxx, above), section 2.2.3, p. 16.

[177] The EU/US Passenger Name Record (PNR) Agreement, 21st Report, Session 2006-07, HL Paper 108, paragraph 24-27; for more details on the *Maher Arar* case, see the website of the [Canadian] Commission of Inquiry at: http://epe.lac-bac.gc.ca/100/206/301/pco-bcp/commissions/maher_arar/07-09-13/www.ararcommission.ca/default.htm [original footnote 62]

The case of Mr Arar was in fact one of the most scandalous instances of an innocent person being classified as a terrorist on a US watchlist. Amnesty International summarises the case as follows:

> Maher Arar, a Canadian citizen, was travelling home to Canada from visiting relatives in Tunisia in 2002. While changing planes at New York City's JFK airport, he was detained by U.S. authorities and then transferred secretly to Syria, where he was held for a year and tortured.
>
> https://www.amnestyusa.org/our-work/cases/usa-maher-arar
>
> Furthermore:
>
> An inquiry conducted by a Canadian judge found, among other things, that he had indeed been tortured, and that "it is very likely that, in making the decisions to detain and remove Mr Arar, American authorities relied on information about Mr Arar provided by [Canadian authorities]." The inquiry emphasized that Canadian authorities, having pursued all the information available to them, had failed to find "any information that could implicate Mr. Arar in terrorist activities."
>
> [Apert from making various recommendation, that have however still not been implemented], **Canadian officials have also requested that the US government remove Maher Arar's name from the US watch list. That request has been refused. As such, it remains impossible for him to travel to the USA or over US airspace, and he faces constant uncertainty about other countries that may have adopted the USA watch list.**
>
> https://www.amnestyusa.org/our-work/cases/maher-arar/i-apologize-action
>
> (emphasis added)

The FRA was however willing to be persuaded that evidence "might" exist, and referred to a UK House of Lords Committee that:[178]

> was persuaded [in 2008] by confidential evidence received from the Home Office that PNR data, when used in conjunction with data from other sources, could significantly assist in the identification of terrorists [and that, in 2011] had no hesitation in accepting the Home Office's assessment of the value of PNR data for the prevention and detection of serious crime and terrorism.

However, the FRA rightly noted that such "confidential evidence" was no evidence at all – certainly not of the scientific, statistical kind attempted by the Max Planck Institute. Rather:[179]

> In any case, the necessity and proportionality of the PNR system would need to be demonstrated [read: by means of published, academically verifiable evidence]

These critics have rightly linked the question of proof of efficacy of the measures of general surveillance to the question of "necessity" and "proportionality" in terms of human rights- and data protection law. As already noted, the EU Fundamental Rights Agency referred in this to the useful summary by the General Secretariat of the EU Council, which said that "the principle of proportionality, which is one of the general principles of European Union law, requires that measures implemented by acts of the European Union are **appropriate for attaining the objective pursued and do not go beyond what is necessary to achieve it**."[180]

The concept of "appropriateness" or "suitability", included in the CJEU's concept of "necessity", reflects the German constitutional- and administrative-legal principle of *Geeignichtkeit*: a measure that impinges on fundamental rights must be shown to be *suited* to the aim being pursued: a measure that is totally incapable of achieving that aim, or that is grossly ineffective in doing so, can never be "necessary" or "proportionate" to that aim.

Moreover, if any state (or, in the present case, European) body is proposing to introduce a measure that impinges on (in ECHR terminology: "interferes with") a fundamental right (in particular, any of the rights protected by the ECHR or the EU Charter), then the *onus* rests on that state (or European) body to demonstrate the suitability of the proposed measure in achieving the aim being pursued.

If a state or European body seeking to introduce a measure that impinges on fundamental rights fails to provide such evidence, then that in itself should suffice to

[178] FRA Opinion on the EU PNR scheme (footnote xxx, above), section 2.2.3, p. 15, emphasis added, with reference to, respectively: The Passenger Name Record (PNR) Framework Decision, 15th Report, Session 2007-08, HL Paper 106, paragraph. 49; and The United Kingdom opt-in to the Passenger Name Record Directive, 11th Report, Session 2010-11, HL Paper 113, paragraph 6.

[179] *Idem*, p. 16

[180] Council of the European Union (2011), *Guidelines on methodological steps to be taken to check fundamental rights compatibility at the Council's preparatory bodies* (footnote 93, above). The CJEU case-law to which this refers is noted by Brouwer: see Attachment x, footnote xxx.
The Council summary adds immediately after the words quoted that "*Furthermore, the necessary and proportionate limitations must respect the essence of the fundamental rights concerned*." We have already concluded in the previous section that the current arrangements and proposals allowing for algorithmic datamining of bulk data fail to do so.

declare the measure to be incompatible with European human rights law (and data protection law where relevant).

In the present case, in which it is accepted that the various PNR measures (PNR transfer agreements and the proposed EU PNR scheme) constitute "general surveillance", and that those data, and more especially the mining of those data, can be highly intrusive and revealing of a person's intimate private life; and in which, consequently, the measure must be subjected to particularly "strict" [181] scrutiny – in such a case the evidence of the efficacy of the measures should particularly strong, if the measure is to be deemed compatible with the European human rights (and data protection) standards. Yet it should by now be clear that the absolute opposite is the case: there is *no* serious, verifiable evidence to show that data mining and profiling by means of bulk data in general, or the compulsory addition of bulk PNR data to the data mountains already created (of communications data and financial transaction data in particular) more specifically, is even suitable to the ends supposedly being pursued.

We therefore fully agree with the other critics that the PNR measures (transfers and the EU scheme) are not "appropriate", or "suitable", and thus not "necessary" or "proportionate" in relation to any legitimate law enforcement or anti-terrorist actions.

- o – O – o -

[181] European Parliament, Legal Opinion *re* LIBE – Questions relating to the judgment of the Court of Justice of 8 April 2014 in Jolned Cases C-293/12 and C-594/12, Digital Rights Ireland and Seitlinger and others - Directive 2006/24/EC on data retention - Consequences of the judgment, (footnote 110, above).

PART V. SUMMARY, CONCLUSIONS & RECOMMENDATIONS

Summary of findings

The facts

- The central problem with the demands for the provision of PNR in bulk to the authorities, is that this is – that this can only be – aimed at facilitating data mining and profiling by means of these records, linked to other major datasets (such as bulk communications data, or bulk financial transaction data) – as is clearly done in the USA and as is clearly also the main aim of the proposed EU PNR scheme: *full PNR data are simply not needed for any other, normal, legitimate law enforcement or border control purpose*.

- The demands for bulk data for such purposes are part of what used to be called by the USA "Total Information Awareness" – a programme that has not died but rather, has been resurrected in the USA's new "New Collection Posture" under which the USA effectively seeks access to all information available through the Internet and global IT networks, as exposed by Edward Snowden.

- No serious, verifiable evidence has been produced by the proponents of compulsory suspicionless [bulk] data collection to show that data mining and profiling by means of the bulk data in general, or the compulsory addition of bulk PNR data to the data mountains already created in particular, is even *suitable* to the ends supposedly being pursued – let alone that it is *effective*. Yet in law (as noted under the next heading), the onus to proof o such suitability and effectiveness rests on those who demand the introduction or continuation of such measures.

- Such data mining and profiling is used in the USA, and is clearly intended to be used in the EU, at rating people on a risk scale (e.g., as "high risk") on anti-terrorist lists, on the basis of such data mining and profiling (see in particular the discussion of the "Fourth List" noted by the US GAO, in Part I, section I.ii, of the report).

  [NB: As noted below, the proposed EU PNR scheme is aimed at facilitating the creation of similar "dynamic"-algorithm-based lists.]

- However, such lists are by their very nature of highly dubious reliability, with inevitably many "false positives", i.e., people being wrongly labelled as "high risk" on an anti-terrorist database (cf. the discussion of the "*base rate fallacy*" in Part I, section I.iii of the report).

- Yet these lists are widely shared by the USA, with reportedly at least 22 other countries – without any of the recipient countries being in any way able to understand, let alone challenge, the "high-risk" designation of individual passengers.

- There have already been cases of people being wrongly labelled on such lists and, consequently, handed over to repressive regimes and tortured (see, e.g., the Maher Arar case discussed in the final section of the report).

- There is also a high risk of such data mining and profiling resulting in "discrimination by computer" (as discussed under that heading in Part I, section I.iii of the report). Crucially, given the misplaced focus on the use of "sensitive data" in profiling, such discrimination can result from profiling that does not use any such data, or even any proxies for such data (such as meal preferences). Rather, algorithms can reinforce much more deeply and insidiously embedded social distinctions, linked to almost any kind of matter (e.g., postcode or length of residency). This has implication in terms of human rights- and data protection law, as noted under the next heading.
- Yet at the same time, by the very nature of a list created by algorithms applied to inherently ambiguous and subjective "intelligence", such determinations, and such discriminatory outcomes, are extremely difficult to challenge – and they become effectively unchallengeable if the underlying "intelligence" and the evaluations of the "intelligence" and the precise algorithm used to weigh the various elements of the "intelligence" cannot be challenged. As of course no victim of such a determination will ever be able to do.
- Proposals to provide some form of "algorithmic accountability" (Citron), or to use "reverse engineering" to counter such dangers (Diakopolous) are in practice impossible to use in relation to secretive law enforcement/border control/national security databases. As noted under the next heading, this means that there are, in reality, no effective remedies against such wrong labels or discriminatory outcomes of the profiling by the relevant agencies.
- The latest (2012) EU-US PNR Agreement does not stand in the way of the PNR data transferred to the USA under the agreement being fed into these kinds of wider anti-terrorist databases, in order to "identify" "high-risk" passengers: the use of the data for such "identification" is clearly allowed, but the word "identification" is here used, misleadingly, not to match PNR data on lists of "known" terrorists or other serious criminals, but to rate the passengers on a risk scale, on the basis of dynamic-algorithm-based profiling.
- Edward Hasbrouck has shown that in any case, the USA are completely by-passing the EU-US PNR Agreement, in that they can already obtain full access to the vast bulk of PNR data – including full PNRs on most intra-European flights – from the Computerised Reservation Systems of the airlines and travel agencies, that are housed (or mirrored) in the USA.
- The proposed EU PNR Directive, read closely, is clearly aimed at facilitating the creation of similar "dynamic"-algorithm-mined databases, resulting in similar "identifications" of people as "high risk" (or as "posing serious danger", to use another euphemism that crops up in the literature), i.e., as similarly labelling them in this way on the basis of inherently fallible analyses (see Part II, section II.ii).
- Unsurprisingly, many other countries are now also beginning to demand the handing over of PNR data in bulk. So far, this includes Russia, Mexico, the United Arab Emirates, South Korea, Brazil, Japan, Argentina and Saudi Arabia.

- The EU intends to provide for "horizontal" rules on the provision of PNR data, by European airlines, to these (and any other) countries. However, how could these regulate the labelling of people by such countries according to their own definitions of "high risk"? If Western countries already want to extend close surveillance and other repressive measures to "extremists-who-have-not-yet-broken-the-law" (as David Cameron is explicitly suggesting), how will these "horizontal" rules prevent the targeting of non-criminal dissidents by those other countries, on the basis of similar algorithm-based profiling? And if Western countries already themselves fail to counter the danger of algorithms creating "suspect communities" and leading to "discrimination-by-computer", how will these rules address those wrongs in those other states?

- There have as yet been no Russian or Chinese "Edward Snowdens", but it would be surprising if China and Russia, at least, would not already be building – or already have in operation – such "rule-based" surveillance and analysis systems. Will the "horizontal" EU rules allow the feeding of PNR data from EU airlines into those systems? How would they prevent that?

The law

- The general requirements of the European Convention on Human Rights in relation to targeted surveillance, as developed by the European Court of Human Rights, are summarised in a text box in the report, on p. 46.

- These general principles are important, e.g., by clarifying that even targeted but secret use of PNR data would have to be restricted to particularly serious crimes, and to strictly limited categories of people (with at least some link to serious criminal or terrorist activity); and that any such uses should be subject to strict substantive and procedural safeguards and "effective remedies".

- Moreover, any "general surveillance" based on bulk PNR data should be based on statute law; and all the main rules on how it is to be carried out should be clear and made public, so that they can be "foreseeable" in their application.

- We conclude from this that, for instance, the meaning of the word "identification" should be made clear in the rules (and any accompanying documentation, such as Explanatory Memoranda to draft laws), in particular when the term is used, not to indicate finding a "known" person (typically, a person on a list), but to indicate a "risk" rating, a labelling, rather than such direct "identification".

- Also, as the Venice Commission has said, one "implication of the ECtHR's approach is that there must be [published] legal authority for issuing selectors as regards the content of the data, and as regards metadata, for issuing instructions for contact-chaining and otherwise analyzing this data." Of course, the exact terms used as "selectors" need not be published, but the basic structure of the analyses should be transparent.

- However, the implication drawn by the Venice Commission can relate only to fairly straight-forward use of pre-specific "selectors". It is in practice impossible to pre-specify any algorithm that might be used to "dynamically" "improve" the data mining/profiling, e.g., by creating further (combinations of) selectors by means of "artificial intelligence" and the adding of different (and also dynamically changed) "weight" to the different selectors.

- In this respect, it is important to note that it follows from the European Court of Human Rights judgment in the case of *Segerstedt-Wiberg and Others v. Sweden*, discussed in Part III, at IV, that people should not be subjected to "filtering" or data mining based on tenuous links with organisations which do not pose any real, active threats to national security. This has obvious implications in relation to allegedly "extreme" – but not actively violent – Islamist groups too.

- Any "selectors" that put under surveillance organisations, or anyone with links to organisations, that may appear to be "extremist" but that have not actually engaged in violence or terrorism would in our opinion be in contravention of this judgment.

- It is an essential requirement of the ECHR and the EU Charter, and indeed of the rule of law, that there must be "effective remedies" against violations of individual rights. In the *Segerstedt-Wilburg* case, the Court reaffirmed what it had already held in *Klass* and other earlier cases: that in relation to secret surveillance this "need not necessarily in all cases" require a judicial remedy (although that is clearly the best option) – but it expanded on the relevant requirements to stress that any effective remedial body must have full powers to fully investigate a complaint about secret files or secret surveillance; and full powers to order the destruction or correction of the file, and/or its release to the individual concerned – and the State must provide evidence that those powers are also actually and effectively exercised in practice.

- In our opinion, a somewhat obscure remark in the judgment relating to the sufficiency of internal supervisory mechanisms while secret surveillance is carried out is clearly limited to brief, targeted telephone interception, and does not apply to long-term analyses of bulk data: the obtaining and further processing, including any data mining/profiling of such data must always be subject to the full powers of fully independent bodies, just mentioned.

ALL OF THE ABOVE IS IMPORTANT. HOWEVER, WE HAVE FOUND THAT THE KIND OF "DYNAMIC"-ALGORITHM-BASED DATA MINING AND PROFILING WE HAVE FOCUSED ON RAISES EVEN MORE FUNDAMENTAL ISSUES IN TERMS OF THE EUROPEAN CONVENTION OF HUMAN RIGHTS AND THE EU CHARTER, AND THUS ALSO IN TERMS OF THE COUNCIL OF EUROPE DATA PROTECTION CONVENTION. SPECIFICALLY:

- Such special, dangerous processing must be assessed especially strictly in regards to the question of whether it serves – can ever be said to serve – a "legitimate aim" in a democratic society; or in data protection terms: whether there is a clear and acceptable "specified" purpose and whether the processing is indeed limited to that purpose – *if it does not, that means that it is* ipso facto *in violation of the ECHR and the EU Charter, and of the Data Protection Convention*;

- The effectiveness of any supposed remedies against such processing must also be especially strictly scrutinised – *if there are no actually effective remedies in place, or available, that too would in itself violate those instruments*;

- Most especially, such processing of personal data should never touch on the “essence”, on the “untouchable core” of the rights in question, i.e., of the right to private life and the right to data protection – *if it did, it would again be incompatible with these instruments at the most fundamental level*;

  And at a more prosaic (but still crucial) level:

- Such special processing must at the very least be capable of achieving the purported purpose for which it be used; it must be “suited to” that aim – *if it is not, the processing can never be regarded as “necessary” or “proportionate” to that aim, and would therefore also on that basis be in violation of these instruments*

We have concluded that in all four of these fundamental respects, “dynamic”-algorithm-based profiling, aimed at rating individuals on a “risk scale” (e.g., “high risk”) on an anti-terrorist database, fails to meet these requirements, as further explained in our Conclusions, below.

Conclusions

As noted above, we have drawn important conclusions on the use of bulk PNR data in respect of four fundamental issues:

The compulsory suspicionless provision of PNR data in bulk does not serve a legitimate aim:

As already noted, we found that bulk PNR data are not needed for any normal, legitimate law enforcement or border control purpose (API suffices for those). Rather, we concluded that the only real purposes of the demand for bulk PNR data is to serve either of the two following purposes:

- pro-active “identification” of “possible suspects”, i.e., the marking of people as a “probable criminal” or “possible criminal”, without those people being yet formally categorised as suspects in the criminal law/criminal procedure law sense (i.e., in the absence of any evidence against them that would suffice to properly designate them as formal suspects, in accordance with criminal procedure law); and
- pro-active “identification” of people for “preventive targeting” on national security grounds, in cases in which no action can (yet) be taken against them under the criminal law –
- on the basis of “dynamic”-algorithm-based data mining and profiling.

In other words, the demands for PNR data are part of an attempt at “predictive policing” or “predictive protection of national security”: the *Vorverlegen* or “bringing forward” of state intrusion, to “deal” with people who are not (yet) breaking the law, but who are either labelled as “probably” or “possibly” being a terrorist or other criminal, or “predicted” to “probably” (or even “possibly”) become one in future.

In our opinion, it cannot be acceptable in a society under the rule of law that intrusive measures are used to “target” people who have done no wrong – not even on the basis that “the computer says” that they are at some dubiously-calculated “risk” of doing some wrong in the future, or similarly dubiously calculated to have “possibly” or indeed "probably" been involved in any wrong, without the kind of evidence (even preliminary evidence) that states under the rule of law require for the imposition of repressive measures.

As the case of Maher Arar shows, being thus labelled on a list is not without consequences – indeed possible extreme consequences.

In other words: "dynamic"-algorithm-based data mining and profiling with the aim of such "predictive" or "preventive" labelling of people on a "risk scale" is not a "legitimate aim" in a democratic society, and is therefore inherently fundamentally incompatible with the European Convention of Human Rights and the EU Charter of Fundamental Rights.

This ought to suffice to reject any plans to allow PNR data, or any bulk data on general populations, for large-scale data mining and profiling.

However, we will still also consider the other three fundamental objections mentioned.

There are no effective remedies against the outcomes of "dynamic"-algorithm-based data mining and profiling:

We have concluded that there simply are no currently available, let alone operational, remedies against the dangers of people being mis-labelled as "high risk" on an anti-terrorist list as a result of deficiencies in the algorithms used, or against discrimination-by-computer caused by the algorithms.

Crucially, you simply cannot remedy such wrongs by "improving" the algorithm, or by adding more data: the dangers are inherent in the processes and can only be countered, if at all, by deep analyses and auditing of the results of the data mining.

There is no indication whatsoever that such deep analyses and audits are actually carried out with the aim of protecting innocent people from being wrongly labelled.

Until such analysis- and audit systems are in place, and are made transparent – with involvement of critical scientists and human rights and data protection advocates – "dynamic" algorithm-based profiling should not be permitted in a state under the rule of law.

In simple human rights and data protection terms: there are no effective remedies available against anti-terrorist/national security "dynamic" algorithm-based data mining and profiling – and without such remedies such operations are simply not compatible with the European Convention on Human Rights, the EU Charter of Fundamental Rights, or the Council of Europe Data Protection Convention.

Or to put it at its absolute mildest:

The conclusion must be that either "dynamically-improved" algorithms should be regarded as intrinsically contrary to the ECHR, because they cannot be properly controlled; or that actually effective means of controlling them must be found, e.g., to check on how reliable the application of the algorithms is: how many "false positives" and how many "false negatives" did they generate? And were the results (unintentionally) discriminatory?

As noted in the report that is a much bigger challenge than is acknowledged by the proponents of those systems.

“Dynamic”-algorithm-based datamining and profiling, in particular if aimed at rating people on a “risk scale” on an anti-terrorist list, violates the most fundamental duty of the State and the EU to “respect human identity”:

We believe that “preventive” or “predictive” profiling of individuals on the basis of essentially unverifiable and unchallengeable “dynamic”-algorithm-based bulk data, unrelated to any specific indications of wrongdoing, and without any targeting on the basis of such suspicions touches on the “essence”, the untouchable core of the right to privacy – and indeed violates the even more fundamental principle underpinning the right to privacy (and other rights), that states must respect “human identity”.

In our opinion, the PNR instruments allowing for such data mining and profiling are thus, on this basis too, incompatible with European legal principles of the most fundamental kind.

Trying to “identify” “possible” or “probable” terrorists by means of “dynamic”-algorithm-based datamining and profiling does not work:

Profiling and mining large datasets with the aim of “identifying” rare phenomena, such as the small number of terrorists in the general population (or even in more specific populations) inevitably suffers from the “*base rate fallacy*”, leading to unacceptably high number of “false positives” (people wrongly labelled a “possible” or “probable” terrorist, or generally as “high risk”), or “false negatives” (actually terrorists not being identified), or both.

It has been acknowledged by the US National Research Council and others that the US data mining operations have not stopped any terrorist attack.

The EU Member States and the European Commission have failed to provide any serious, scientifically verifiable data in support of their claims that bulk PNR data does work in identifying terrorists, or indeed that other bulk datasets, specifically compulsorily retained communications data, have had any impact on law enforcement clear-up rates.

The largest and most serious study into possible efficacy of bulk data retention, by the Max Planck Institute at the request of the European Commission, discussed in Part xxx of the report, found that:

> there are no indications that compulsory suspicionless [e-communications] data retention has in the last years led to the prevention of any terrorist attack.

There is still no serious effort on the part of those who clamour, not just for continuing communications data retention, but also for further bulk “just-in-case” collections, such as the compulsory provision of full PNR data, to actually provide any serious, meaningful, scientifically valid evidence to show the efficacy of the measures in fighting serious crime or terrorism.

Yet under the ECHR and the EU Charter, the onus is on them to show convincing evidence of the effectiveness of bulk data collection and –analyses. This duty is the more onerous in view of the very serious interferences with human rights inherent in such collection and analyses (as noted above).

The fact that they have not provided any such evidence, in our opinion, simply underlines the scientific doubts about the efficacy of data mining in these regards: the proponents of bulk data collection, -mining and –profiling do not provide any real evidence of the efficacy of their “dynamic”-algorithm-based system, because they simply DO NOT WORK.

This ought to suffice in simple practical terms to abandon these highly-intrusive and dangerous efforts. But in more legal terms, it means “dynamic”-algorithm-based data mining and profiling are simply not “appropriate”, not “suited” to the proclaimed aim of “identifying” terrorists from large datasets – and thus also not “necessary” or “proportionate” in relation to any legitimate law enforcement or anti-terrorist actions.

In other words, our overall conclusions are that:

- **The compulsory suspicionless provision of PNR data in bulk does not serve a legitimate aim;**
- **There are no effective remedies against the outcomes of “dynamic”-algorithm-based datamining and profiling;**
- **“Dynamic”-algorithm-based datamining and profiling, in particular if aimed at rating people on a “risk scale” on an anti-terrorist list, violates the most fundamental duty of the State and the EU to “respect human identity”;**

and on top of that:

- **Trying to “identify” “possible” or “probable” terrorists by means of “dynamic”-algorithm-based datamining and profiling does not work.**

Recommendations

NB: We have been asked by the Consultative Committee to draft recommendations that the Committee itself might wish to adopt. We provide a number of those below. However, it is of course entirely up to the Committee to decide whether to make any of these draft, tentative recommendations its own.

The Consultative Committee recalls that European human rights- and data protection law requires, *inter alia*, that:

- All requirements that personal data should be provided to law enforcement-, border control- or national security agencies "in bulk" should be clearly set out in clear and precise statute law; and all subsidiary rules that are necessary to enable individuals to foresee the application of the statutory rules, should be equally clear, and made public. Only the lowest, operational guidance-type rules might be kept secret, and even then only as long as they do not contradict of obscure the application of the published rules. This also applies to any requirements that PNR data be handed over to state (or international) authorities in bulk;
- The application of all those rules in practice should be subject to serious, meaningful transparency and accountability;[182] and that
- There should be full and effective remedies against the use of bulk data, including bulk PNR data, in "general surveillance".

In that regard, the Consultative Committee notes that the Secretary-General of the Council of Europe has been urged, *inter alia*, by the Parliamentary Assembly of the Council of Europe, to use his power under Article 52 of the European Convention to demand that all CoE Member States provide full account of any "general surveillance" of the kind exposed by Edward Snowden that they may be involved in, with clarification on how this accords with their obligations under the ECHR.

The Consultative Committee supports this call, and recommends that when the Secretary-General does issue such a demand, he specifically also asks the Member States:

- whether they use any bulk data they acquire for any data mining and profiling in order to "identify" "possible" (or "probable") terrorists – with full clarifications of what exactly this "identification" entails (i.e., whether it merely involves matching PNR data against lists of "known" people, or whether it involves rating people on "risk scales" that are reflected in anti-terrorist databases);
- what safeguards are in place against straightforward mis-identifications on such lists,

  but also especially:

[182] We have not addressed this issue in the report, because it would have exceeded our brief. We note however the very useful *Issue Paper* of the Council of Europe Commissioner for Human Rights on Democratic and effective oversight of national security services (May 2015), and the Venice Commission "Update of the 2007 Report on The Democratic Oversight of the Security Services and Report on the Democratic Oversight of Signals Intelligence Agencies (April 2015), which provide important indicators in this area, of which the Consultative Committee should take account.

- how they guard against erroneous risk ratings of such kind; and why they believe any such redress and remedial action is effective.

Pending the provision of information that might lead to another conclusion, the Consultative Committee believes that the use of "dynamic"-algorithm-based data mining and profiling with the aim of "predictive" or "preventive" labelling of people on a "risk scale" is *not a "legitimate aim" in a democratic society*, *touches on the "essence", the untouchable core, of the right to private life and the right to data protection*, and would appear to be *unsuited to* the aim of actually identifying real terrorists – and thus *neither necessary nor proportionate* to that aim; and is therefore *fundamentally incompatible* with the European Convention of Human Rights, the EU Charter of Fundamental Rights – and with the Council of Europe Data Protection Convention of which the Committee is a guardian;

And therefore recommends:

- That "dynamic"-algorithm-based data mining and profiling for the purpose of "identifying" "possible" (or "probable") terrorists on the basis of a computer assessment by any State party to the Data Protection Convention be stopped immediately; and
- That the passing on of PNR data to any non-State Party for the purpose of such "dynamic"-algorithm-based profiling, or that may result in the use of the data in such processing by the non-State Party be also stopped; and
- That serious scientific studies are commissioned as a matter of urgency of appropriate independent scientist, with the involvement of human rights- and data protection advocates and civil society, to evaluate the effectiveness or ineffectiveness of such processes for such purposes, in particular also in terms of "false positives" and "false negatives", and in relation to the question of whether such data mining and profiling can or did lead to discriminatory outcomes; and to examine if effective, scientifically sound, means can be developed to counter such negative outcomes (or whether this is impossible).

- o – O – o -

ATTACHMENT: API, SFPD & PNR data compared

SFPD data Secure Flight Passenger Data	**APIS data [EU / US*]** Advanced Passenger Information	**PNR data/EU-US Agrmt** Passenger Name Records **[Data not in EU-US Agrmt in [grey] and in square brackets]**
Basic Information:	*Basic Information:*	*Basic Information:*
Full Name	Full Names	All APIS data [18]
Date of Birth	Date of Birth	
Gender	-	
Passport Number	Type of travel document used (e.g., passport) & Number	
Passport Country	Nationality	
None of these data are included in the SFPD list	Country of residence*	
	For non-US persons travelling to the USA: Address of first night spent in the USA*	
	Initial point of embarkation	
	Border crossing point of entry into the EU	
	Code of transport (airline and flight number)	
	Departure and arrival time of the transportation (of the flight)	
	Total number of passengers carried on that transport (on the flight)	
Special Information:		PNR Record locator code [1]
Redress Control Number (to correct "mislistings")		Date of reservation/Issue of ticket [2]
Known Traveler number (for TSA Prev^{TM})		Date of intended travel [3] [NB: duplicates API]
		Name(s) [4] [NB: duplicates API]
		Other names & number of travellers on the PNR [6]
		Frequent Flyer/Benefits Information [5]
		Contact Info. [7]: Contact address [also of originator], billing address, emergency contact, email address, mailing address, home address, intended address [in State requiring PNR data transfer]

PNR list continued from previous page:

		PNR data (*continued*)
		Telephone details (May include mobile number) [NB: For EU-US Agrmt preumably covered by Contact Info [7])
		Payment information [8] (including credit card details; details of person/agency paying the ticket, etc.)
		Travel itinerary for PNR [9] **[Full travel itinerary** **(to the extent provided/covered by the booking)]**
		Travel agent information [10]
		Code share information [11] [NB For IATA presumably included in full travel itinerary; and/or duplicates API]
		Split/divided information [12] [NB: unclear what this means]
		Travel status/Check-in information [13] [NB IATA lists "Go-show" and "No-show" separately]
		Ticketing information [14], including one-way tickets and Automated Ticket Fare Quote [NB: may duplicate some API]
		Baggage information [15]
		Name of person who made the booking [NB: Often covered by the payment information [8]]
		Seat information [16]
		All historical changes to the PNR listed in numbers [1] to [18]
Note:: According to the IATA Guidelines, the "open fields" listed to the right: - **may not include:** any information that an aircraft operator does not need to facilitate a passenger's travel, e.g. racial or ethnic origin, political opinions, religious or political beliefs, trade-union membership, marital status or data relating to a person's sexual orientation; - **but they may include:** e.g. meal preferences and health issues as well as free text and general remarks, legitimately entered to facilitate a passenger's travel. In the EU-US PNR Agreement this is covered by Article 6 – which actually permits more processing of sensitive data in the PNRs than the IATA Guidelines. The acronyms OSI, SSI and SSR refer, respectively, to :Other Service related Information, Special Services Information, and Special Service Requests.		*Open fields:*
		General remarks including OSI, SSI and SSR information [17] [NB: IATA lists both a "General remarks" open field and "Free text/code fields in OSI, SSR, SSI", with the latter apparently allowing "remarks/history"]. See note on the left.]
		Free text/code fields (in OSI, SSR, SSI, remarks/history)

Sources: *see overleaf*

Sources:

SFPD list:
http://www.tsa.gov/sites/default/files/publications/pdf/SecureFlight_PassengerDataDefinitions.pdf

APIS list(s):
Council Directive 2004/82/EC of 29 April 2004 on the obligation of carriers to communicate passenger data, Article 3 (NB: the data are listed above in a different order from the one in the article, to allow easier comparison between the tables).

The fields marked * are not included in the EC Directive but are required by the US authorities in relation to travellers to the USA, see:
https://en.wikipedia.org/wiki/Advance_Passenger_Information_System

PNR list:

International Civil Aviation Organization (IATA) Guidelines on Passenger Name Record (PNR) Data, first edition, 2010, available at:
https://www.iata.org/iata/passenger-data-toolkit/assets/doc_library/04-pnr/New%20Doc%209944%201st%20Edition%20PNR.pdf

NB: The numbers given to the "PNR Data Types" listed in the Annex to the 2012 EU-US PNR Agreement are added to the table in square brackets. "Data Types" or fields included in the IATA list but not in the Annex to the EU-US PNR Agreement have been entered in square brackets and in **[grey]**.

- o – O – o -