Tilburg Law School LLM Law & Technology
2017 - 2018

Master Thesis

Predictive policing;
An investigation into the use of the Crime Anticipation System by the Amsterdam police department and the safeguard against discrimination

Author: E.L. van Kooten
SNR: 2009627

Supervisor: I. Skorvanek
Second reader: B. van der Sloot

Tilburg, May 2018

Table of Content

List of Abbreviations

CAS	Crime Anticipation System
CBS	Central Bureau of Statistic
ECHR	European Convention of Human Rights
GDPR	General Data Protection Regulation
iRN	Internet Recherche (& Onderzoek) Netwerk / Internet criminal investigation department & investigation network
USA	United States of America
VU	Vrije Universiteit Amsterdam / Free University of Amsterdam
Wbp	Wet bescherming persoonsgegevens / Law protecting personal data
Wpol/Wpg	Wet politiegegevens / Law police data
Wiv	Wet op de inlichtingen- en veiligheidsdiensten / Law information and security services
Wjsg	Wet justitiële en strafvordelijke gegevens / Law judicial and criminal data

1. Introduction

It is 2054 and people are constantly monitored by the authorities. Crime has almost disappeared in Washington D.C. thanks to an elite law enforcing squad called 'PreCrime'. The team consists of 'PreCogs', mutants with a certain gift. They can predict future murders by dreaming of the perpetrator and the victim. Many crimes are stopped this way before they are even committed. Based on their intentions, the alleged criminals are subsequently convicted. "PreCrime, it works!" Is the campaign slogan. Police Chief John Anderton firmly believes in the success of PreCrime, until the system predicts that in the next 36 hours he will kill someone. He is forced to flee the authorities. Anderton suspects that he is being misled by opponents of PreCrime. In his mission to prove his innocence, Anderton begins to see more and more the fundamental weaknesses of the system. Could it be that the prediction system is not waterproof?[1]

This plot of Steven Spielberg's film Minority Report and the book The Minority Report (1956) by the American science fiction writer Philip Dick, is usually called upon when discussing predictive policing. It seems like science fiction, but the film also contains methods and techniques that could be used in real-life in the future. Our society is characterized by the large amount of data collected and processed. Most of the time we generate personal data ourselves, for example when we use the internet. Digital developments such as increased storage capacity, growing computer power and new data analysis techniques have made it possible to store large amounts of data.[2]

These large amounts of data, called big data, can isolate patterns which are used for predictive policing. Predictive policing makes statistical predictions about potential criminal activity using analytical techniques.[3] The predictions can be divided into events and people, which means that the crime can be predicted when and where crimes might occur and might even go so far as to anticipate who is will likely to become a victim or perpetrator.[4]

1.1 Motivation

Big data can be very useful. Companies have a better understanding of customers' needs and the police can use it to predict crimes. But this development also has a downside. Question marks are raised in the cases of privacy, the transparency of personal data and the legitimacy of organizations that use big data.

The first idea of using big data in combatting crime came from a police chief of the Los Angeles Police Department. In 2008, they developed software called "PredPol" in collaboration with the University of California. This software applies an algorithm, which was used to predict earthquakes, to old crime statistics. All kinds of influencing factors, such as type of crime, place and time, were predicted more accurately than police analysts previously did.[5] More police departments followed, amongst them the Amsterdam police department in the Netherlands.

---

[1] http://www.imdb.com/title/tt0181689/ (last visited 23 November 2017)
[2] Rienks, R., *Predictive Policing; Kansen voor een veiligere toekomst* (2015), p.18-19
[3] Perry, W. L., McInnis, Brian, Price, C.C., Smith, S.C., and Hollywood, J.S., Predictive Policing: The Role of Crime Forecasting in Law Enforcement Operations. *RAND Safety and Justice Program* (2013), p.1-2
[4] Perry, W. L., McInnis, Brian, Price, C.C., Smith, S.C., and Hollywood, J.S., Predictive Policing: The Role of Crime Forecasting in Law Enforcement Operations. *RAND Safety and Justice Program* (2013), p.8
[5] Smit, S., de Vries, A., van der Kleij, R., van Vliet, H.; *Van predictive naar prescriptive policing; Verder dan vakjes voorspellen* (2016), p. 15

In 2005, the Dutch police published a report stating that they wanted, among other things, to no longer have a passive executive role, but also that of authoritative teacher and pacemaker. The police want to be able to identify and advise the Public Prosecution, the government and partners to make the Netherlands safer. A strategic vision, published by the police in 2006, shows that the strategy is aimed at increasing the police officers' ability to identify criminal behaviour through generic alerting principles and obtaining knowledge and signals from (police) data. This strategic vision mentions predictive analyses for the first time, which are helpful to find out who might be involved in an incident.[6]

Nowadays, several police departments in the Netherlands are using big data information for predictive policing. The Amsterdam police department uses the specially designed Crime Anticipation System (CAS), created by the police department in Amsterdam in association with the 'Vrije Universiteit Amsterdam'[7] (VU). The system is mostly used to predict the High Impact Crimes, like street robberies and domestic burglaries. The system works with algorithms that can generate statistical information on the probability that a burglary or robbery will take place in a particular residential area. For CAS to work, Amsterdam is divided into squares of 125 by 125 meters. Areas where the probability of an incident can be estimated to be at a low level, such as meadows and open water, are removed. For the remaining boxes, the Central Bureau of Statistics (CBS) provides a large amount of data: distance to the nearest highway, type and number of companies known to the police, as well as demographic and socio-economic data. Furthermore, the algorithm uses the data already available to the police, like crime history and distance to known suspects.[8]
After the development of the system in Amsterdam in 2014 and a pilot in four police departments, more than 90 basic teams will be working with CAS. It is unique that a national police force uses predictive policing nationwide.[9]

1.2 Hypothesis

The police focus on the discussion about entrusting the technology instead of the described juridical/ethical issues and believes that there is no dispute about the possibility of discrimination. The primary aim of this Master Thesis will be to investigate whether there is indeed no dispute about the possibility of discrimination when using systems for predictive policing. Because the information above could suggest otherwise. The hypothesis therefore will be:

Ethnicity plays a role when the Amsterdam police department is using the Crime Anticipation System for predictive policing because the used algorithms and analyses are biased, not transparent and generate a self-fulfilling prophecy.

This Master Thesis will elaborate on the juridical and ethical issues, like privacy, self-fulfilling prophecy, ethnic profiling, transparency and discrimination. The main research question will therefore be 'Is it possible for the Amsterdam police department to deploy the Crime Anticipation System used for predictive policing in a way that does not result in discrimination based on ethnicity?'

To answer this question, the following sub-questions need to be answered:
1. What is big data and how is it used in the Crime Anticipation System?

[6] Rienks, R. *Predictive Policing; Kansen voor een veiligere toekomst* (2015), p.11
[7] Free University of Amsterdam
[8] Willems, D., Doeleman, R.; *Predictive Policing – wens of werkelijkheid* (2014), p. 41
[9] https://www.politie.nl/nieuws/2017/mei/15/05-cas.html (last visited 23 November 2017)

In order to talk about big data, knowledge is needed about what big data is. Big Data expresses that there is an incomprehensible amount of data on the internet which expands every second.[10] Large amounts of data are not interesting by themselves, but can be by carrying out the correct analyses.[11] Therefore, the features of big data need attention. Furthermore, insight needs to be into how CAS works in order to judge the possibility of rights being infringed, like discrimination. Afterwards, the relation between big data and the use of CAS will be looked into.

2. What risks of ethnic discrimination arise with the use of the Crime Anticipation System?
Here will be discussed the notion of ethnicity, and especially the discrimination by CAS on ethnicity. To do this, the meanings of ethnicity and discrimination will be discussed. After that, the possible infringements that predictive policing, and therefore CAS, might make, will be discussed.

3. Is it possible to use the Crime Anticipation System without infringing rights?
A framework will be given to see if the use of CAS is possible without infringing rights. The infringements discussed in the previous sub-question will be used for this.

A possibility for police to prevent crime is to be present at locations designated as High Impact Crime areas. This presence will work preventively. When the police are present at those areas, it will probably not violate a citizen's right per se, but when they start undertaking action it might be. For example, when the police start with body searches, before a crime has occurred, a citizen's privacy is violated. Preventive searches can be done in the Netherlands in areas with a security risk and approval of the mayor.[12] Will predictive policing make it possible for the police to start using preventive searches to prevent crimes?[13]

It is difficult to select random people from a flow of people that is larger than the processing capacity. Despite the fact that guidelines are explicitly directed at the random designation of persons, it is still a person who selects someone. This selection is based on the knowledge in his head that automatically starts to look at characteristics that are viewed as an increased risk.[14]

A similar thing can occur in traffic controls. Without a suspicion, vehicles can be forced to stop, the driver's driving license may be asked for.[15] The so-called rifle judgment[16] makes it possible that, if an investigating officer encounters facts and circumstances that involve a reasonable suspicion of a criminal offense, the investigating officer may then apply investigative powers and, for example, can search a vehicle for the suspicion of drug possession. This is not based on an instrument that makes use of objectified knowledge. Furthermore, it is not the intention that biased choices are made, just as with preventive

[10] Lodder, A. R., van der Meulen, N. S., Wisman, T. H. A., Meij, L., & Zwinkels, C. M. M., *Big Data, Big Consequences? Een verkenning naar Privacy en Big Data gebruik binnen de opsporing, vervolging en rechtspraak* (2014), p.16
[11] Lodder, A. R., van der Meulen, N. S., Wisman, T. H. A., Meij, L., & Zwinkels, C. M. M., *Big Data, Big Consequences? Een verkenning naar Privacy en Big Data gebruik binnen de opsporing, vervolging en rechtspraak* (2014), p.20
[12] Artikelen 151b en 174b Gemeentewet / Articles 151b and 174b Municipal Law
[13] Mali, B., Bronkhorst-Giesen, C., den Hengst, M., *Predictive policing: lessen voor de toekomst: Een evaluatie van de landelijke pilot*. Politieacademie (2017), p.69
[14] Rienks, R. *Predictive Policing; Kansen voor een veiligere toekomst* (2015), p.140 - 142
[15] Artikel 160 Wegenverkeerswet 1994 / Article 160 Road Traffic Law 1994
[16] Supreme Court, ECLI:NL:HR:2011:BT6402, November 22th 2011

searches. But how random is random? Would it not be much better if the police could use objective criteria in a traffic control that have been tested and determined in advance. Selection based on gut feeling will be avoided.[17]

This brings a risk of ethnic profiling where skin colour and ethnicity are used as distinctive features to select from a population. It is frequently argued that this is in conflict with the non-discrimination principle.[18] Unequal treatment based on personal characteristics such as age, religion, race, gender or religion, skin colour or origin is not permitted. The fact that crime is mainly committed by young men does not mean that it is lawful to place the focus on young men in a preventive search. Ethnic minorities, such as non-Western immigrants, are also overrepresented in crime statistics. If you extend this line, you could also question the selective surveillance in deprived neighbourhoods based on information.[19]

Furthermore, with the use of big data a lot of data of innocent citizens are also collected. The privacy of those citizens are less protected because of this under Article 8 of the European Convention of Human Rights (ECHR). There also needs to be looked into the so-called 'chilling effect' that predictive policing can have. The chilling effect comes into play when people know that they are watched or think that they might be watched. People will act differently because they are afraid of the potential negative consequences it might have. This chilling effect is mostly connected to the right freedom of speech, but Article 8 ECHR is also applicable in relation to discrimination or stigmatization of certain groups in society.[20]

Another important issue is the self-fulfilling prophecy. Primarily, the algorithms used for predictions may not take into account the inaccuracies reflected by historical data. In the data that is being used, information can be overrepresented or underrepresented which leads to biased statistics, which can lead to self-fulfilling prophecies. Furthermore, when police is more active at certain areas, there will be more arrests.[21]

Because more and more data is collected about citizens, governments and companies can develop profiles of citizens, divide them into different categories and then treat them differently. This makes predictive policing mainly a privacy and discrimination issue.

"Ethnicity does not play a role in the prediction, because that is 'a politically incorrect variable'."

Inventor of Crime Anticipation System, Dick Willems

As the police states themselves, they cannot profile based on ethnicity and emphasizes, apart from the multiple offenders, not to focus on people, but on tackling specific sorts of crimes. In addition, police departments do not get to see why a certain area is flagged by the algorithm, because police agents need to interpret the data themselves.[22]

---

[17] Rienks, R. *Predictive Policing; Kansen voor een veiligere toekomst* (2015), p.140 - 142
[18] Article 2 Universal Declaration of Human Rights
[19] Rienks, R. *Predictive Policing; Kansen voor een veiligere toekomst* (2015), p.140 - 142
[20] van der Sloot B., *Is the Human Rights Framework Still Fit for the Big Data Era? A Discussion of the ECtHR's Case Law on Privacy Violations Arising from Surveillance Activities*, in: Gutwirth, S., Leenes, R., De Hert, P. (eds.), *Data Protection on the Move* (2016), p.12-13
[21] http://www.slate.com/articles/technology/future_tense/2016/11/predictive_policing_is_too_dependent_on_historical_data.html (last visited 06 December 2017)
[22] Willems, D., Doeleman, R.; *Predictive Policing – wens of werkelijkheid* (2014), p. 42

1.3 Relevance

Predictive policing is still a new phenomenon. Different studies have shown that predictive policing is useful. The police themselves never really elaborate on the juridical and ethical aspects. This Master Thesis will focus on a specific aspect of predictive policing, namely CAS, in combination with a specific form of discrimination, ethnicity. This Master Thesis can hopefully be seen as a contribution to the overall picture and aims to be used as a recommendation for policymakers and the police when using CAS.

1.4 Methodology

To give an answer to the first sub-question, clarifying what big data is, literature concerning these subjects will be investigated. Sources about big data will be searched for in scientific journals and literature using (specialised) search engines, like Mendeley and Legal Intellegence. For the information about the working of CAS will be taken into account the evaluation of the national pilot of CAS, as well as two documents, one by two persons close involved with the development of CAS and one by an intelligence professional of the police. Also, the results of a research program carried out on the instructions of the Ministry of Security and Justice will be used.

For the other two sub-questions, a literature study shall be held as well. The sources used in the two documents mentioned above, will be taken into account. Furthermore, information about the correlation between discrimination and ethnicity will be searched for in scientific articles by using (specialised) search engines, like Mendeley and Legal Intellegence.

Literature from the United States of America (USA) will also be taken into account. This is because predictive policing is being used for a while by a few police departments in the USA. No comparison will be made between the USA and the Netherlands. It is only for informative purposes. This literature will also be searched by using (specialised) search engines, like Mendeley and Legal Intellegence.

Each chapter will be devoted to answer one of the sub-questions so that in the end the research question, whether the Amsterdam police department can deploy CAS in a way that does not result in discrimination based on ethnicity, can be answered.

The first chapter will give an explanation about big data and CAS. It will clarify the systems and any problems that arise. The second chapter will explain ethnicity and discrimination. Furthermore, possible infringements that CAS might make will be discussed. A framework to see if CAS may be used without possibly infringing rights will be given in the last chapter. This Master Thesis will end with a conclusion, the answer to the research question.

To clarify, when using the wording ‘predictive policing’, it is regarding predictive policing in general. When using ‘Crime Anticipation System’ or ‘CAS’, it is regarding the system used by the police in Amsterdam.

2. Big data and predictive policing

As mentioned in the previous chapter, this thesis is focused on predictive policing and the possible infringement the system makes with discrimination based on ethnicity. This chapter will focus on explaining what big data is by describing 'the four Vs' and the connection and what influence big data has on predictive policing and CAS. There will also be a paragraph dedicated to the legal grounds as they are in the Netherlands.

2.1 Big data analysis

People are relying on technologies and the internet almost every moment of the day. They use a (smart)phone, (smart)watch, (smart)tv, laptop, etcetera. However, this does not go without consequences. Cybercrimes emerged and people are traceable through using the internet. Big data is becoming more and more important. It collects a large amount of data, analyses this data and uses algorithms.[23] These technological developments make it possible to discover and predict criminal behaviour.

Big Data refers to practices of creating and analysing vast datasets, which can also include personal information.[24] When trying to define the term big data, literature often refers to the four "Vs": Volume, Variety, Velocity and Veracity.[25] The four V's are covered by the technical elements of big data.[26] Some literature writes to a fifth "V", which is Value. Yet this factor seems rather speculative and is thus best omitted.[27]

2.1.1 The four Vs

Volume stands for all the data that is collected. The data comes from all of the devices people use in their homes, like televisions, thermostats and even smoke detectors. The 'Internet of Things' ties them together and produces data. According to Moore's law, data density doubles approximately every 18 months while storage capacity is only doubling every 14 months. This means that data is becoming exponentially uncountable.[28]

Variety refers to the type of data. Big data (tries to) give structure to all the collected data from the different types of data such as social media interactions between people. This analysis can give an understanding of the behaviour of people, groups, systems, diseases, etcetera.[29]

Velocity is the speed at which the analysis of the data can develop. It refers to the collection and the demand for real-time insights and responses. Decisions must be made on the spot regarding which data is kept (to process) and which data is not. Furthermore, data analysis can be taken out of context or just be wrong. Real-time decisions can have serious consequences, especially when a wrong decision is made.[30]

Veracity is the accuracy of the data which could be achieved through the analytical process. It is important that datasets are reliable. Reliability can be achieved by processing increasing amounts of data, keeping in mind the quantity and quality. This can be done by increasing the

[23] van der Sloot, B., Broeders, D., Schrijvers, E. (eds.), *Exploring the Boundaries of Big Data* (2016), p.27-29
[24] Zarsky, T., *Incompatible: The GDPR in the Age of Big Data* (2017), p.996
[25] Zarsky, T., *Incompatible: The GDPR in the Age of Big Data* (2017), p.998-999
[26] van der Sloot, B., Broeders, D., Schrijvers, E. (eds.), *Exploring the Boundaries of Big Data* (2016), p.29
[27] Zarsky, T., *Incompatible: The GDPR in the Age of Big Data*, p.999 (footnote)
[28] Moore's Law is an observation made by Gordon Moore, co-founder of Intel
[29] van der Sloot, B., Broeders, D., Schrijvers, E. (eds.), *Exploring the Boundaries of Big Data* (2016), p.28
[30] van der Sloot, B., Broeders, D., Schrijvers, E. (eds.), *Exploring the Boundaries of Big Data* (2016), p.28-29

sample size to reduce statistical errors. Another option is adding independent sources to reduce systemic errors.[31]

2.1.2 Defining big data

Article 29 Working Party[32] gives a more specific definition of big data. According to the Working Party:

"Big data refers to the exponential growth both in the availability and in the automated use of information: it refers to gigantic digital datasets held by corporations, governments and other large organizations, which are then extensively analysed (hence the name: analytics) using computer algorithms. Big data can be used to identify more general trends and correlations but it can also be processed in order to directly affect individuals."[33]

2.2 From big data to predictive policing

A prerequisite for the successful use of predictive policing is that data of past crimes is available in a 'data warehouse', and that other data can be linked to this data warehouse. All sorts of data can be relevant in pointing out crime patterns in an area, like data on demographics or company types. The amount of data that can be discovered, is potentially so gigantic that a human probably will not be able to extract all the useful information. So computers are needed to process the information faster and more accurately. The automated discovery of relevant patterns in large amounts of data is called data mining.[34] Elaborating on data mining, this is the process where patterns can be found between different people or outcomes to determine what aspects make them similar or different.[35] With data mining, profiles are created. Generally, profiling involves categorising individuals according to their characteristics. These characteristics can be gender, age, ethnicity, habits and preferences. To create a profile, the technique 'behavioural analysis' is used. This technique makes connections between patterns of behaviour and characteristics. To get this profile, anonymous data and information are collected and stored in a data warehouse. An algorithm connects relevant information and creates new information, thus data mining. This new information is then used to 'predict' behaviour.[36]

Data mining has already been in use as a method of detection before CAS by the police in the Netherlands. Although data mining can possibly be a valuable and effective means of investigation, it also has an undesirable side. An important aspect is that in the process of data mining, a lot of data from innocent citizens are collected and this can have the consequence that the privacy of those citizens can be affected. And the right to privacy is a fundamental right protected by Article 8 of the ECHR and Articles 7 and 8 of the Charter of Fundamental Rights of the European Union (Charter).[37]

---

[31] van der Sloot, B., Broeders, D., Schrijvers, E. (eds.), *Exploring the Boundaries of Big Data* (2016), p.29

[32] Article 29 Working Party is an abbreviation for 'Data Protection Working Party established by Article 29 of Directive 95/46/EC'. The Working Party provides the European Commission with independent advice on data protection matters and helps the development of harmonised policies for data protection in the EU Member States

[33] Article 29 Working Party, Opinion 03/2013 on purpose limitation, WP203, 2 April 2013

[34] Mali, B., Bronkhorst-Giesen, C., den Hengst, M., *Predictive policing: lessen voor de toekomst: Een evaluatie van de landelijke pilot* (2017) p.30-31

[35] Selbst, A.D.; *Disparate Impact in Big Data Policing* (2017), p.13-14

[36] European Union Agency for Fundamental Rights (FRA); *Towards More Effective Policing. Understanding and Preventing Discriminatory Ethinic Profiling: A Guide* (2010), p. 8-9

[37] Brinkhoff, S.; *Big data datamining door de politie. IJkpunten voor een toekomstige opsporingsmethode* (2016), p.1400-1401

An aspect that needs attention from the perspective of privacy is the collection of data. Consideration should be given to what data can be used to address the problem, whereby in principle the more data that is collected, the better the outcomes. This means that data is collected relatively unrestricted in the first instance and is only sorted during the analysis. This can create tension with, among other things, the purpose limitation principle which stipulates that it must be clearly defined for what purpose data is collected, certainly in combination with the data minimization principle. This principle stipulates that no more data must be collected than is necessary for the intended purpose.[38]

With big data analysis it can be stated that data will be used for a specific purpose, but the question is to what extent such a purpose is sufficiently concrete. The formulation of a specific purpose is difficult when using big data, especially since a target specification requirement is that, first of all, the purpose must be defined and, secondly, must be explicitly defined. But the purposes are often not defined in advance and/or explicitly described. It is often difficult to ask for permission of the person whose data is going to be processed, if the responsible person cannot yet clearly define the purpose of processing. Nevertheless, the purpose of the necessary data processing must be specific and as accurately as possible. The basis for processing personal data is often a statutory duty.[39] Additionally, it will almost always be the intention in the performance of tasks of safety or investigation authorities that the person concerned does not know that he is subject to an investigation.[40]

2.2.1 Legal grounds

Security is high on the political agenda due to terrorist attacks of recent years and the government is taking significant measures. A lot of these measures stand opposite of privacy. Two developments, which are important for the safety measures, are the use of new technologies (DNA research, cameras, data mining, file coupling) and the use of digital information and technology by the police, the justice department and security services. The security services are getting more and more power and almost have unrestricted access to all personal data of citizens. National security seems more important than privacy of the citizens.[41]

When police is using big data to analyse personal data, this use falls under the scope of the Dutch laws 'Wet bescherming persoonsgegevens' (Wbp)[42], the 'Wet politiegegevens' (Wpg)[43], the 'Wet op de inlichtingen- en veiligheidsdiensten' (Wiv)[44] and the 'Wet justitiële

[38] Lodder, A. R., van der Meulen, N. S., Wisman, T. H. A., Meij, L., & Zwinkels, C. M. M., *Big Data, Big Consequences? Een verkenning naar Privacy en Big Data gebruik binnen de opsporing, vervolging en rechtspraak* (2014) p.35

[39] Article 8 sub c Personal Data Act: "the processing of data is necessary to comply with a legal obligation to which the controller is subject;"

[40] Lodder, A. R., van der Meulen, N. S., Wisman, T. H. A., Meij, L., & Zwinkels, C. M. M., *Big Data, Big Consequences? Een verkenning naar Privacy en Big Data gebruik binnen de opsporing, vervolging en rechtspraak* (2014) p.34

[41] Vedder, A., van der Wees, A., Koops, B.J., de Hert, P., *Van privacyparadijs tot controlestaat? Misdaad- en terreurbestrijding in Nederland aan het begin van de 21ste eeuw* (2007), p. 66-68

[42] Dutch law about protecting personal data

[43] Dutch law about police data

[44] Dutch law about information and security services

en strafvordelijke gegevens' (Wjsg)[45]. Which law is applicable depends on which institution uses the data.[46] Most important for police are the Personal Data Act and the Police Data Act.

A privacy principle that contradicts big data analysis is data minimization. Data minimization states that data may be processed only as they are sufficient, relevant and not excessive. No more data may be processed than necessary for the purpose to be achieved. This requirement applies to the Personal Data Act and the Police Data Act. With this principle, it is important to note that it is applicable to personal data. Big data does not only use personal data. Furthermore, the Personal Data Act claims that data may no longer be "stored in a form that allows the persons concerned to be identified any more than necessary for the purposes of which they are collected or subsequently processed."[47] This principle should, in theory, lead to the removal of personal data processed under the Personal Data Act as soon as it is no longer relevant to the original purpose for which they were processed. However, practice shows a completely different picture, in which the rule seems to be that personal data will be stored longer than necessary.[48]

The use of data mining is becoming increasingly favourable for the police and public prosecutions. Information of citizens is more often available because of covenants and partnerships with other government bodies, such as the Tax Authorities. Of course, the freely accessible data on the internet also contributes. In this way, the situation that is increasingly more apparent, is that in which police have access to real big data. The data can then be subjected to automated data analysis or can be linked to police data files. Articles 9, 10 and 11 of the Police Data Act offer the possibility to submit data files to automated data analysis.[49]

The legal basis for police data mining is, in addition to the aforementioned provisions, often found in Article 3 of the Police Data Act.[50] From jurisprudence[51] on this point it can be deduced that the current Article 3 of the Police Data Act can provide sufficient legal basis for non-specific legally regulated means of detection, such as data mining, as long as it only involves a limited infringement of fundamental rights (including the right to privacy) of citizens. If a more than limited infringement is made, a specific or adequate legal basis must exist for this.[52]

In addition, the Internet also offers more and more freely available information about citizens. With this, a situation unfolds in which big data is becoming more accessible to the police and it is therefore easier to link big data to police databases. Which increases the possibilities of predictive policing.

---

[45] Dutch law about judicial and criminal data

[46] Lodder, A. R., van der Meulen, N. S., Wisman, T. H. A., Meij, L., & Zwinkels, C. M. M., *Big Data, Big Consequences? Een verkenning naar Privacy en Big Data gebruik binnen de opsporing, vervolging en rechtspraak* (2014) p.33

[47] Article 10 paragraph 1 Personal Data Act

[48] Lodder, A. R., van der Meulen, N. S., Wisman, T. H. A., Meij, L., & Zwinkels, C. M. M., *Big Data, Big Consequences? Een verkenning naar Privacy en Big Data gebruik binnen de opsporing, vervolging en rechtspraak* (2014) Pages 36-37

[49] Brinkhoff, S.; *Big data datamining door de politie. IJkpunten voor een toekomstige opsporingsmethode* (2016), p.1401

[50] Paragraph 1: "Police data are only processed to the extent that this is necessary for the purposes formulated by or pursuant to this Act."

[51] For example: Supreme Court, ECLI:NL:HR:2014:1563, July 1st 2014

[52] Brinkhoff, S.; *Big data datamining door de politie. IJkpunten voor een toekomstige opsporingsmethode* (2016), p.1401

2.2.2 From post-crime to pre-crime

A shift from post-crime to pre-crime may be detected within society. With this shift, the purpose and scope of the investigation of criminal offenses changes significantly. Predictive policing stands on the border of thought and behaviour. A bit exaggerated, but a criminal offense must, with this shift, contain revealed thoughts and those revealed thoughts must play an ever-increasing role. This is only in a situation where it would be about having thoughts and preparations do not have to be made. This creates the danger that having certain thoughts about whether or not to commit a criminal offense will become more and more the object of police control. Assessing the risk of someone's behaviour as a criminal actor increases the police's tendency to suspect the mental attitude or inner conviction of persons without any reasonable suspicion of guilt to a criminal offense.[53]

In Dutch criminal law acts are liable criminally. There is no punishment without an act that is forbidden by law. This is guilt as an element: guilt in the sense of culpability.[54] The act must be seen as a specific decision by a person. This person then commits an offense. This usually means to say that a person can be held criminally accountable for an act. The person has created the actual event through behaviour, like preparation or an act. That behaviour makes him the perpetrator of the fact. For the legislator of the Penal Code of 1886, the behaviour of the perpetrator in the various penal provisions was a physical movement. Consequentially, an action is only relevant to criminal law if the outcome of the physical movement triggers is also wanted. A punishable act must reveal itself by means of a specific decision.[55]

Taking into account the shift from 'post-crime' to 'pre-crime', it might happen that it is not the physical movement that is being prosecuted but the intention of doing something as interpreted by the police.[56] Although, the incentive of the police with CAS is not to arrest citizens but to anticipate and remove the opportunity to commit crimes before they can occur.

Coming back to Minority Report. In the film, someone has been arrested for the future murder of two people, which would have taken place at a moment in the near future. A critical question raised is if we can speak of actual murder if the act itself was not yet committed? Well, the fact that you prevent something does not mean that it would not happen if you had not intervened, according to John Anderton played by Tom Cruise.[57]
This illustrates that the mere thought of committing a criminal offense does not have to lead to actually committing that fact. Someone can come to repentance in time

2.3 Data mining by the police

Data mining by the police does not always happen within the current legal frameworks in the Netherlands. This means that the police can do what they want because there is no legal basis. But before going any further into this, two systems that the police use in combination with data mining are described. The first system that is described is a sort of precursor of CAS and the second system is CAS.

2.3.1 iRN / iColumbo

The iRN / iColumbo project is a trajectory supported by the six eastern police regions to provide controlled access to the Internet for Dutch government agencies with an investigative

[53] Schuilenburg, M.; *Predictive policing: De opkomst van een gedachtepolitie?* (2016), p. 935
[54] Supreme Court, (Milk and water) ECLI:NL:PHR:1916:BG9431, February 14th 1916
[55] Schuilenburg, M.; *Predictive policing: De opkomst van een gedachtepolitie?* (2016), p. 936
[56] Schuilenburg, M.; *Predictive policing: De opkomst van een gedachtepolitie?* (2016), p. 936
[57] Minority Report (2002)

and supervisory task. The so-called 'Internet Recherche (& Onderzoek) Netwerk' or iRN, enables detection and open source research on the Internet in a forensically secured manner.

iRN contains the infrastructure and access to composers including the process to use the processed data in a forensic way in the detection chain. iColumbo, an intelligent, automated and near real-time Internet monitoring service, contains all development activity to provide iRN end users with intelligent tools to be able to carry out more effective Internet research and investigation.[58]

The use of data available on the Internet is not new when looking at CAS. But CAS takes it a step further.

2.3.2 CAS

CAS is the most recent development in the field of predictive policing in the Netherlands. As mentioned earlier, the system is inspired by PredPol. CAS uses historical crime data and other input variables, like data about the nearest highway access, known criminal businesses in the area and other socio-demographic issues about the inhabitants. By jointly processing this, the system makes a prediction where certain crime will take place.[59]

CAS gets a lot of information from the Central Bureau of Statistic (CBS). But what other information does CAS use, besides the information described above? The CBS has access to all government data for the production of official statistics. In addition, they also use new data sources which sometimes have a very large volume. They first observe, measure and describe developments in society. This happens at a high frequency, so eventually they will be able to process real time statistics.[60] Additionally, CAS uses information already known to their system, like recorded crimes and calls for service.[61]

Furthermore, real time data can be purchased by manufacturers or developers of mobile phones or other devices. Open source intelligence can also be included. Examples are Twitter, Facebook, Instagram and other social media.[62] This means that CAS has the possibility to know very accurately what each individual is doing, where that person is, with whom that person communicates, and so on. If CAS also uses cameras placed in Amsterdam, then the system can keep an eye on its citizens as if in a panopticon.[63] For the time being, however, the data the police have cannot be used to detect individual behaviour. It does not look like that this will happen in the future, but the possibility exists with systems as CAS. Nevertheless, questions can be raised about the right to privacy regulated in Article 8 ECHR. This will be discussed further in Chapter 4.

It would be ideal for the police if CAS was available real time, which is possible with the current available data. If this would happen, other data that can have an influence can also be included, like the weather and incidental events such as soccer matches.[64]

---

[58] Ministerie van Veiligheid en Justitie, *Projectplan Verduurzaming iRN / iColumbo* (2011), p.7
[59] Willems, D., Doeleman, R., *Predictive Policing – wens of werkelijkheid* (2014), p.41-42
[60] https://www.cbs.nl/nl-nl/onze-diensten/innovatie/big-data (last visited on 21-02-18)
[61] Willems, D., Doeleman, R., *Predictive Policing – wens of werkelijkheid* (2014), p.41-42
[62] Sanders, C.B., Sheptycki, J., *Policing, crime and 'big data'; towards a critique of the moral economy of stochastic governance* (2017), p.8
[63] A panopticon is a prison where the prisoners sit in a dome around the central tower of guards. The guards are behind blinded glass, so that the prisoners cannot see them. However, the prisoners are watched full time and can be penalized at any time.
[64] Willems, D., Doeleman, R., *Predictive Policing – wens of werkelijkheid* (2014), p.42

In order to successfully use all the analysed information, Amsterdam is divided into squares of 125 by 125 meters. Uninteresting areas are omitted, like areas of which the probability of an incident can be estimated in advance, such as meadows and open water.[65] This means that what CAS predicts is the probability of at least one incident per square. Amsterdam can be divided into 22,734 squares and after deduction of uninhabited areas, 9918 squares are left. The system focusses on areas with a high risk of incidents so only the top 3% of the squares will be used on the map. This top 3% includes 298 squares. To determine the correctly predicted percentage, a distinction is made between the percentage 'direct hit' and the percentage 'near hit'. The direct hit is the percentage of home burglaries that took place in a high-risk compartment. Each high-risk box is surrounded by eight other boxes. If in such a box, if it is not itself a high risk compartment, a burglary has taken place, then that counts as a near hit. The percentage near hit is the percentage of home burglary in addition to a high-risk compartment that itself is not a high-risk compartment.[66]

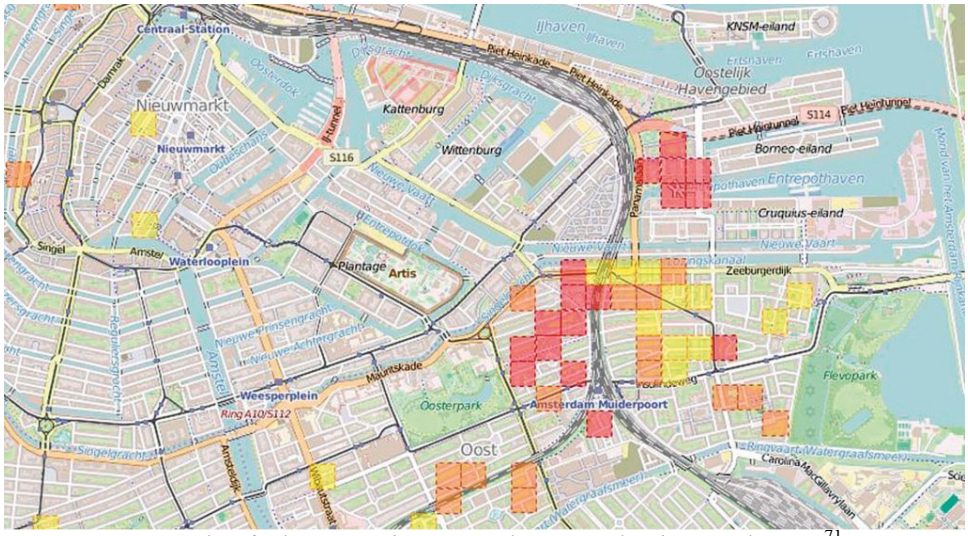
The known data of each square is recorded at various moments. Then the system records what happened to incidents in the two weeks after the measuring moment to check itself and learn from the mistakes. To see which combinations of characteristics are indicative of crime in the near future artificial neutral networks need to be used. So algorithms come into play. These algorithms are capable of learning to recognize patterns. This is done analogous to how it is assumed in science that people do this. If a person learns a new task, the brain will connect existing neurons. New connections are made and older connections are changed or broken up. The same happens with the artificial neural network. The network checks itself with the biweekly measuring moments by looking at where the biggest deviations were.[67]

The network can thus predict the probability of incidents in the future. It assigns a risk score to each square in Amsterdam. These scores can be used to create heat maps. The colours used for the heat maps are red, orange and yellow. The higher scores get a darker colour and the lower scores a brighter. The system aims to identify small surveillance areas with a high risk of incidents so it has been decided to only colour the top 3% of the squares.[68] The map can even show exactly how big the chance is of an incident for each period (often 2 to 4 hours) on a specific day.[69]

The area with the highest risk is coloured red, the next orange and the area with the least high risk is yellow. Below is an example of how such a heat map would look like. If these high risk locations are selected, in a similar way it is determined when the risk is most likely to be expected. Based on this information, a schedule can be drawn up and the police can be at the right place at the right time. With this heat map approximately 40% of the home burglaries and 60% of the street robberies in Amsterdam can be predicted.[70]

---

[65] Willems, D., Doeleman, R., *Predictive Policing – wens of werkelijkheid* (2014), p.41
[66] Mail, B., Bronkhorst-Giesen, C., den Hengst, M., *Predictive policing: lessen voor de toekomst: Een evaluatie van de landelijke pilot* (2017), p.96
[67] Willems, D., Doeleman, R., *Predictive Policing – wens of werkelijkheid* (2014), p.41
[68] Willems, D., Doeleman, R., *Predictive Policing – wens of werkelijkheid* (2014), p.41-42
[69] Smit, S., de Vries, A., van der Kleij, R., van Vliet, H., *Van predictive naar prescriptive policing; Verder dan vakjes voorspellen* (2016) p.16
[70] Willems, D., Doeleman, R., *Predictive Policing – wens of werkelijkheid* (2014), p.41-42



Example of a heat map in Amsterdam on a day in March 2013[71]

For the percentage of 'direct hit', CAS predicted an average of 13,5% of the home burglaries in Amsterdam within the top 3%. This is 4,5 times better than what may be expected on a random basis. In an analysed period of 44 months, a total of 1470 (15%) different squares were assigned to the top 3% high-risk boxes. The same boxes are often in the top 3%. This means that 8,448 squares (85%) have not been a high-risk in almost four years. Within the 15% of boxes that have been a high-risk box at least once, to almost a quarter (24.4%) applies that this has only been the case once or twice (3.6% of all squares) and for a large 5% of these squares (0.8% of all squares), they have been a high-risk box more than 73 times.[72]

An investigation by the Police Academy showed that the interpretation of the heat map is problematic. The persons working with the map do not always have enough knowledge of how to read such a map. The research shows how complicated data interpretation can be. The most common action was deploying more police to a specific area, but that is not always the best option. But there is a risk with predictive policing. Crimes are more often committed in deprived neighbourhoods, so the algorithm focuses mostly on those neighbourhoods. As stated above, in total 1470 squares were assigned to the top 3% high-risk boxes. Then deploying more police at these areas, will give the result that more crimes are found.[73] Which may result in self-fulfilling prophecy.

2.3.3 From predictive policing to prescriptive policing

CAS is not an oracle, it only says where something is likely to happen, not what you have to do against it. Prescriptive policing is the step that follows predictive policing. With predictive policing the places with the greatest chance of crime are predicted on the basis of incident data, while in prescriptive policing the effectiveness of a specific deployment of police resources is predicted based on the knowledge of the effects of certain interventions.

[71] Smit, S., de Vries, A., van der Kleij, R., van Vliet, H., *Van predictive naar prescriptive policing; Verder dan vakjes voorspellen* (2016) p.16
[72] Mail, B., Bronkhorst-Giesen, C., den Hengst, M., *Predictive policing: lessen voor de toekomst: Een evaluatie van de landelijke pilot* (2017), p.97-98
[73] Mail, B., Bronkhorst-Giesen, C., den Hengst, M., *Predictive policing: lessen voor de toekomst: Een evaluatie van de landelijke pilot* (2017), p.35-37

Both systems use each other, but are not by definition interwoven. The systems itself can operate apart from each other, but without predictions you cannot determine what you have to do, without intervention a prediction has no effect.[74]

Prescriptive policing predicts how effective police resources can be from a number of interventions. If there is a dark street where many robberies take place, a street lamp or a security camera may be the best solution. When many burglaries are taking place, then another patrol route can make a difference. The system calculates the effect of each intervention and gives that information to agents. They then choose what the best option is, based on the system and their experience on the street.[75]

2.4 Conclusion

The most important part of this chapter was to give an answer to the question 'what is big data and how is it used in CAS?'.

Big data entails creating and analysing vast datasets. Algorithms are released on these datasets and create profiles. The automated discovery of relevant patterns in large amounts of data is called data mining. Data mining is the process where patterns can be found between different people or outcomes to determine what aspects make them similar or different.[76] In the Netherlands, the police also uses data mining for the detection of crime. The most recent system is CAS. CAS uses historical crime data and other input variables, like data about the nearest highway access, known criminal businesses in the area and other socio-demographic issues about the citizens. By jointly processing this, the system makes a prediction about where certain crimes will take place.[77]

The use of CAS can be problematic because of the possibility of infringements of rights. First of all, the right to privacy regulated in Article 8 ECHR. With data mining, a lot of data is collected. No distinction is made (yet) between guilty and innocent citizens. This can have consequences for the privacy of, most of all the innocent, citizens. In the Netherlands, the data collecting must meet the standards of the purpose limitation principle and the data minimization principle. It must be clearly defined for what purpose data is collected and no more data must be collected than necessary for the intended purpose.[78] CAS uses police data like crime history and data from the CBS. When using CAS, there is not a specific case or person about which or whom information is sought. Information from everyone, guilty or not, is collected to be able to make the best possible predictions. This goes against the purpose limitation principle. With the consequence that more data is collected than necessary. So there is also an infringement of the data minimization principle.

As mentioned in paragraph 2.3.2., an outcome of CAS is that the chance of being caught when committing burglaries and robbery has increased. But is this development because of CAS or because more police have been deployed to deal with such crimes? Mostly deprived neighbourhoods come out as a 3% high-risk boxes. This makes that the police deploy more

[74] Smit, S., de Vries, A., van der Kleij, R., van Vliet, H., *Van predictive naar prescriptive policing; Verder dan vakjes voorspellen* (2016), p. 51
[75] Smit, S., de Vries, A., van der Kleij, R., van Vliet, H., *Van predictive naar prescriptive policing; Verder dan vakjes voorspellen* (2016), p. 51-52
[76] Selbst, A.D., *Disparate Impact in Big Data Policing* (2017), p.13-14
[77] Willems, D., Doeleman, R., *Predictive Policing – wens of werkelijkheid* (2014), p.41-42
[78] Lodder, A. R., van der Meulen, N. S., Wisman, T. H. A., Meij, L., & Zwinkels, C. M. M., *Big Data, Big Consequences? Een verkenning naar Privacy en Big Data gebruik binnen de opsporing, vervolging en rechtspraak* (2014) p.35

people to those areas and eventually more crime is found and the algorithm uses this with the next biweekly prediction. This will keep the self-fulfilling prophecy alive.

In conclusion, it is likely that infringements of multiple rights are made on the purpose limitation principle, data minimization principle and privacy, and self-fulfilling prophecy can become real, when big data is used for predictive policing, and thus CAS.

3. Ethnicity and discrimination

When data is used for predictive policing, the results can affect the fate of groups of people in critical ways. Data miners need to be careful, otherwise the outcomes can result in a disproportional focus on historically disadvantaged groups. This may look a lot like discrimination.[79] To investigate if ethnic groups are disadvantaged by CAS, it needs to be clear what ethnicity and discrimination mean. When this is clear, the effects on predictive policing and CAS will be discussed in combination with the disadvantages of CAS.

Furthermore, predictive policing can lead to early deployment of methods and techniques for detection and more inspections without a judge supervising this. Investigations are extended to persons who are not a suspect, and without even knowing it. Consequently, it can lead to an arbitrary and difficult to control deployment of criminal investigation means.[80]

The use of data to create a heat map to identify individuals is a privacy concern. Citizens with no criminal record can be monitored, even though they have not done something wrong (yet).[81] In this chapter, problems related to CAS and minorities are discussed. This is to see if CAS van be used without infringing rights per se. Topics that will be discussed are, among others, bias data, presumption of innocence, social sorting, self-fulfilling prophecy, transparency and privacy.

3.1 Ethnicity

"Ethnicity" is defined in the Oxford Dictionary as: "the fact or state of belonging to a social group that has a common national or cultural tradition".[82] Another important term for the purpose of this Master Thesis is the term "ethnic profiling". This refers to the use of race, ethnicity, religion, or national origin instead of individual behaviour when it comes to law enforcement and/or investigative decisions about who may be involved in criminal activities.[83]

Law enforcement should be based on individual conduct, not on a membership of a certain group. With ethnic profiling stereotypes about criminal offenders and minority groups are used. Those groups, allegedly, are more prone to commit a crime. The individuals of these groups are reviewed based on group attributes and not on potentially suspicious behaviour.[84]

The police use ethnic profiling across a range of police operations and tactics, for example stop-and-search tactics. Ethnic profiling is the result of targeting certain forms of crime and certain areas. Individuals within the police department may contribute to ethnic profiling by arresting more people with an ethnic background.[85]

Three factors for the justification of the use of ethnic profiles are important, namely effectiveness, proportionality and necessity. Looking at effectiveness, ethnic profiles must be

---

[79] Barocas, S., and Selbst, A.D., *Big Data's Disparate Impact* (2016), p. 673
[80] Lodder, A.R., Schuilenburg, M.B., *Politie-webcrawlers en Predictive policing* (2016), p.154
[81] Available at https://www.floridatechonline.com/blog/criminal-justice/4-problems-with-predictive-policing/ (last visited on 30 October 2017)
[82] https://en.oxforddictionaries.com/definition/ethnicity (last visited on 22 Februari 2018)
[83] Open Society Justice Initiative, *Addressing Ethnic Profiling by Police; A Report on the Strategies for Effective Police Stop and Search Project* (2009), p.17
[84] Open Society Justice Initiative, *Ethnic Profiling in the European Union: Pervasive, Ineffective, and Discriminatory* (2009), p.19
[85] Open Society Justice Initiative, *Ethnic Profiling in the European Union: Pervasive, Ineffective, and Discriminatory* (2009), p.21-22

evidently effective as a means of preventing crimes to be used. This means that the created profiles must have a high probability of identifying criminals. Furthermore, to test the proportionality, there must be a consideration between using ethnic profiles and whether it makes law enforcement more efficient versus the probability of harm done by using ethnic profiles and a possible perceived discrimination of individuals or groups. Necessity can be justified by looking at other laws, which could achieve the same, but do not carry the risk to discriminate. If so, then the use of ethnic profiles is unnecessary and no ground for justification is achieved.[86]

3.1.1 Group profiles

It can be argued that three groups of people may be affected by the use of profiles. First, the people whose data are used to create the profile. Second, the people to whom the profiles apply, and third, the people who are exposed to the decisions based on the profile. There are risks following these groups.[87]

With surveillance, and the use of profiles, there is a risk of intrusive interferences to privacy. With big data and all the technological possibilities more and more data becomes available, which may reveal behavioural data. The impact this may have is not per se due to the data collection itself, but more due to the transformation of said data into knowledge about people. Applying a profile to someone can be threatening to his autonomy. This is a risk all citizens have whose data is used to make the profile as well as citizens to whose data the profile is applied.[88]

When profiles are applied to groups by companies, this is usually to find, for example, shared features between members of a certain community. When the police use this technique, it may be to find common characteristics in the group of convicted murderers. This is not without consequences. Group profiling can lead to discrimination.[89]

Putting people into boxes or using stereotypes, also called social sorting, may create risks. People aren't judged anymore on their behaviour, but on their characteristics. This may be an infringement of the principle of equality and the presumption of innocence, which may lead to discrimination. This discrimination may result in more control exercised on groups who match the profile. It can also be discrimination by algorithms because of the lack of transparency.[90]

With profiling based on sensitive characteristics as race, the problem is not per se the creation of a group profile, but what is done with the information. This problem is not only present for the people included in the profile, but might also have consequences for the people who are excluded from the profile. Bearing in mind the groups of people like the homeless or illegals which are excluded from the databases.[91]

---

[86] Open Society Justice Initiative, *Addressing Ethnic Profiling by Police; A Report on the Strategies for Effective Police Stop and Search Project* (2009), p. 22-23
[87] van der Sloot, B., Broeders, D., Schrijvers, E. (eds.), *Exploring the Boundaries of Big Data* (2016) p.16-17
[88] van der Sloot, B., Broeders, D., Schrijvers, E. (eds.), *Exploring the Boundaries of Big Data* (2016) p.152
[89] Hildebrandt, M., *Defining Profiling: A New Type of Knowledge?,* in Hildebrandt, M. and Gutwirth, S. (eds.), *Profiling the European Citizen* (2008), p.35
[90] van der Sloot, B., Broeders, D., Schrijvers, E. (eds.), *Exploring the Boundaries of Big Data* (2016) p.152
[91] van der Sloot, B., Broeders, D., Schrijvers, E. (eds.), *Exploring the Boundaries of Big Data* (2016) p.274

3.2 Discrimination

The data used for predictive policing must be approached with care. When using certain data, the prejudice of prior decision makers or the widespread biases of society can result in patterns of discrimination in the new data. With the use of algorithms these tendencies can be shown, even when they are not being programmed manually. So discrimination can be an object of the data mining process. But the fact that this may happen, is not recognized by most scholars and policy makers. It seems like they tend to fear hidden intentions or the overlooked effects of human bias or error in hand coding algorithms. This injustice might be hard to identify and address because the disadvantages for groups are less obvious.[92]

It might be useful to define "discrimination". The Oxford Dictionary gives two definitions which apply to discrimination in the legal context: (1) "the unjust or prejudicial treatment of different categories of people, especially on the grounds of race, age, or sex" and (2) "recognition and understanding of the difference between one thing and another".[93] The goal of a data miner is to build a system that can discriminate in the sense that the system can recognize and understand the difference between one thing and another. But the other definition, the first one, is the one that might give a problem. In this sense, discrimination can happen by the system or persons. If the system treats someone unfairly, hence discriminate, the discrimination might be unintentional. But when a person treats someone unfairly, most of the time it is seen as intentional and is protected by Article 1 of the Dutch Constitution and Article 14 ECHR.[94]

Article 14 ECHR[95] prohibits discrimination in the enjoyment of rights and freedoms protected by the Convention. Discrimination is a wide understanding in this article. It entails discrimination on any ground such as sex, race, colour, language, religion, political or other opinion, national or social origin, association with a national minority, property, birth or other status.[96] Protocol No. 12 of the ECHR adds in Article 1[97] to Article 14 ECHR that discrimination on any ground in respect of any right set forth in national law by any public authority is prohibited. This protocol is signed by the Netherlands.[98] In addition, the Explanatory Report to Protocol No.12 states that this prohibition is applicable to discrimination by a public authority in the exercise of discretionary power.[99] This means that by law it is prohibited in the Netherlands to discriminate. This does not only apply to the citizens and companies, but also to the public authorities like the police. Which means that the police legally may not discriminate with CAS.

---

[92] Barocas, S., and Selbst, A.D., *Big Data's Disparate Impact* (2016), p. 674

[93] https://en.oxforddictionaries.com/definition/discrimination (last visited 21 February 2018)

[94] Selbst, A.D., *Disparate Impact in Big Data Policing* (2017), p.13-14

[95] The enjoyment of the rights and freedoms set forth in this Convention shall be secured without discrimination on any ground such as sex, race, colour, language, religion, political or other opinion, national or social origin, association with a national minority, property, birth or other status

[96] Open Society Justice Initiative, *Ethnic Profiling in the European Union: Pervasive, Ineffective, and Discriminatory* (2009), p.21-22

[97] Article 1 – General prohibition of discrimination

1. The enjoyment of any right set forth by law shall be secured without discrimination on any ground such as sex, race, colour, language, religion, political or other opinion, national or social origin, association with a national minority, property, birth or other status.

2. No one shall be discriminated against by any public authority on any ground such as those mentioned in paragraph 1.

[98] Protocol No. 12 to the Convention for the Protection of Human Rights and Fundamental Freedoms (Treaty No.177)

[99] Explanatory Report to the Protocol No. 12 to the Convention for the Protection of Human Rights and Fundamental Freedoms (European Treaty Series - No. 177)

Not all distinctions or differences in treatment constitute a discrimination. The ECHR decided that discrimination is treating persons different in relevantly similar situations, without an objective and reasonable justification. We can speak of racial discrimination when discrimination is taking place on account of someone's actual or perceived ethnicity. Racial discrimination is invidious and requires special attention and strong reaction from the authorities.[100]

The Court has set forth a test to determine when a distinction or difference in treatment amounts to discrimination. Equality of treatment is violated if the distinction has no objective and reasonable justification. Article 14 ECHR is likely to be violated when a difference of treatment does not pursue a legitimate aim, but also when it is clearly established that there is no reasonable relationship of proportionality between the means employed and the aim sought to be realised.[101]

3.3 The influence of discrimination on ethnicity

Treating someone different in itself is not per se unacceptable. Everyone has a certain preference and uses these preferences to make choices. These choices are not interfered by law but may be affected by law when this different treatment takes place in a public context based on prohibited grounds. An example, a police officer treats an individual different from others who are in a similar situation and the reason for this is their ethnicity. This will be considered unlawful. In practice, it will be enormously difficult to determine on which grounds the decision is based.[102]

It can be argued that ethnic profiling is unlawful because it ignores that everyone is unique as an individual. It can contribute to deterioration of relations between different groups in society. Furthermore, it can create mistrust between these different groups. But more importantly, when action is taken based on unlawful profiling, like ethnic profiling, there can be an increase in racial tensions. This may lead to resentment towards police within minorities. Ethnicity can bring police on alert, but this is not the only ground they can act on. Other possible grounds will depend upon requirements of national law.[103]

Legitimate factors can be found to approach or act towards someone based on ethnic profiling. For example, if a system like CAS would suggests that a robbery will take place in a specific part of Amsterdam and that it would be carried out by a criminal organisation with Chinese origins, then the police could legitimately consider racial appearance as relevant to determining whether an individual becomes a potential suspect.[104] To be clear, this is not what CAS does. CAS only reproduces a heat map showing where a certain crime has high probability to take place.

---

[100] ECtHR Timishev v. Russia, no. 55762/00, 55974/00, December 13th 2005, par. 54-56.
[101] ECtHR Belgian Linguistics Case (No. 2), appl.nrs. 1474/62, 1677/62, 1691/62, 1769/63, 1994/63, 2126/64, July 23th 1968, par. 10
[102] European Union Agency for Fundamental Rights (FRA), *Towards More Effective Policing. Understanding and Preventing Discriminatory Ethinic Profiling: A Guide* (2010), p. 16
[103] European Union Agency for Fundamental Rights (FRA), *Towards More Effective Policing. Understanding and Preventing Discriminatory Ethinic Profiling: A Guide* (2010), p. 18-20
[104] European Union Agency for Fundamental Rights (FRA), *Towards More Effective Policing. Understanding and Preventing Discriminatory Ethinic Profiling: A Guide* (2010), p. 20-21

3.4 Bias data

Criminal activities are often more committed in deprived neighbourhoods, which results in focus on those neighbourhoods by the algorithm. Consequently, more crime will be found. This is an example of racially imbalanced outcome, because the potential harm is the result of having more police on the streets in specific neighbourhoods.[105] When this is the case, the algorithm creates, unintentionally, a racial bias with the consequence that self-fulfilling prophecy comes into play. After all, the police often go to those neighbourhoods and therefore arrest more members of certain groups. The use of information from the CBS in CAS increases the likelihood of unintentional racial profiling, instead of only using crime figures. It is difficult to prevent this, because these kinds of prejudices are not programmed. They arise from the statistics. That does not make them true, it is a fault in the statistics.[106]

3.4.1 Bias data place-based

Training data, the first data the algorithm gets accustomed with, used by the system can be a problem. The system learns by example and therefore uses its training data as fundamental truth. This is the only data the system knows. This means that decisions based on discoveries that rest on hit-or-miss labelled data or data labelled in a biased manner, will seem legitimate. Consequently, the fundamental truth and the subsequent decisions will be affected by some form of prejudice and the data mining will include rules with the same bias. To put this into context, consider the next example placed in a different context. An algorithm used to hire or reject job applications may learn to discriminate against certain groups of peoples, like people from Eastern Europe, if the algorithm is trained on previous hiring decisions in which an employer has consistently rejected jobseekers with degrees from Eastern Europe universities.[107]

An important source that can enable discrimination is past crime data. This is the most common source of data for algorithms used for predictive policing because this is often collected by the police themselves. Reliance on past crime data is problematic as accurate crime data hardly exists. The most important reason for this is the moment of contact police have with “criminals”. There is always contact when someone is arrested but the results after the arrest are often not included or updated in the report. This means that statistics and systems use arrest data as the best available proxy. Taking into account the possibility that an arrest might be racially biased. Even tough, if the statistics from after the arrests were collected, it would not be a good representation of the facts because a number of cases end in plea agreements that do not reflect the crime the arrestee committed or was arrested for. Resulting in that ultimately it will show a higher percentage of crime committed than reality by members of certain groups.[108]

Furthermore, what might facilitate discrimination is training data. Data mining’s main aim is to match patterns. Without representative samples it will end in sampling bias. So training data must be a representative sample of the whole population.[109]

Another source for discrimination is feature selection. Organizations and their data miners need to make choices about what elements they observe and use into their analyses. These

---

[105] Selbst, A.D., *Disparate Impact in Big Data Policing* (2017), p. 20
[106] Smit, S., de Vries, A., van der Kleij, R., van Vliet, H., *Van predictive naar prescriptive policing; Verder dan vakjes voorspellen* (2016), p. 38-39
[107] Barocas, S., and Selbst, A.D., *Big Data's Disparate Impact* (2016), p. 682
[108] Selbst, A.D., *Disparate Impact in Big Data Policing* (2017), p.23-26
[109] Selbst, A.D., *Disparate Impact in Big Data Policing* (2017), p.23-26

decisions can have a serious impact on the treatment of members in certain groups when factors that represent statistical differences between members of a certain group are not properly reflected in the set of selected features.[110] The police need to make decisions about where the geographic spots are, what features they should aim to contain and how big they should be. Furthermore, they must decide which features will be taken into account. Will it only be the crime type, location, date and time, like PredPol does.[111] Or will they include other variables like socio-economic indicators, weather, holidays, etcetera. These choices can greatly impact the outcome of the statistics.[112] As for CAS, this system also includes socio-economic data on top of the information PredPol uses.

A lack of insight into the veracity of the model can create problems. It is important to know which data played an important part in the flagging of an area. The system might also be considered as less reliable when there is data missing, bias data or wrong data used. This may result in an important problem, because when profiles are used to legitimize measures taken for predictive policing, it might jeopardize transparency.[113]

3.4.2 Bias data person-based

As said before, social media data can be included in predictive policing systems, just as other data like addresses of the people calling the police. This is not done with CAS (yet). But when this is the case, the system will look at the people most likely to be involved in a possible future crime. This type of data can come forward in the extra monitoring of subjects and when a crime occurs, police might be more likely to look at these subjects first. The racial imbalance is, in this case, different than in the situations mentioned above. When more data is used in the system, the system is more focussed on individuals. With the consequence of historically biased policing, the algorithm might think that a member of a minority is more likely to be a criminal.[114]

Feature selection is also more complicated than described in the previous paragraph. Police can use easy accessible data like race, gender or age. But choosing those features might not always result in a relevant distinction between people or locations and will make the outcome of the algorithm less accurate, with the consequence of discriminatory outcomes.[115]

3.5 Big data risks

It is almost impossible to reconcile privacy and big data. Big data often involves millions of minor individual privacy violations. Citizens often do not go to court, if they are even aware of these violations. All of these minor violations combined are a major violation of privacy. When citizens are aware of surveillance, there is the danger of a chilling effect.[116] As mentioned in the introduction chapter, the chilling effect results in people behaving differently because they are afraid of the potential negative consequences it might have acting ‘normally’. This chilling effect is mostly connected to the right freedom of speech, but Article 8 ECHR is also applicable in relation to surveillance measures, discrimination or stigmatization of certain groups in society. The ECHR accepted that someone may claim that

[110] Barocas, S., and Selbst, A.D., *Big Data's Disparate Impact* (2016), p.688
[111] http://www.predpol.com/about/ (last visited 22 February 2018)
[112] Selbst, A.D., *Disparate Impact in Big Data Policing* (2017), p.23-26
[113] van der Sloot, B., Broeders, D., Schrijvers, E. (eds.), *Exploring the Boundaries of Big Data* (2016) p.152
[114] Selbst, A.D., *Disparate Impact in Big Data Policing* (2017), p.27-29
[115] Selbst, A.D., *Disparate Impact in Big Data Policing* (2017), p.27-29
[116] Ministerie van Veiligheid en Justitie, *Kabinetsstandpunt over WRR-rapport Big Data in een vrije en veilige samenleving* (2016), p.5

he or she will suffer harm in the future based on an infringement of his or her right to privacy. This shows that the chilling effect is an issue that needs to be considered by the police when using predictive policing methods like CAS.[117] Another consequence of the use of big data is that citizens have no access to their data or the way their data is used for analysis purposes.[118]

Big data may be divided into roughly three phases: data, data analysis and the use of the analysis. The data involves large amounts of structured and unstructured data from various sources. The analysis is data-driven and automatically searches for correlations. And these analyses must lead to (useful) knowledge.[119] The potential benefits of big data lie in the area of efficiency, (crime) history, real-time analysis and the prediction of crime. But the analysis of these data done by algorithms are mostly not neutral. Frameworks may contain assumptions and errors might be found in big data analysis. When searching in large quantities of data, correlations will always be discovered. The question is, will these correlations be meaningful.[120]

With big data, patterns are being recognized in large amounts of data. This makes it suitable for crimes that have a regular character. If a problem does not occur much, like a terrorist attack, then there is insufficient material available to discover a meaningful pattern.[121]

Looking only at the technological part, predictive policing could become bigger in the sense that it could be used to predict more sorts of crimes. Predictive policing definitely has advantages, but how about the disadvantages that have not been mentioned before?[122]

3.5.1 Transparency

With the use of big data, information is collected in an invisible way for people. People should be aware of this and of any analyses that are made.[123] Big data promised to make everything more transparent, but the data collection is increasingly opaque and the algorithms that are used are often not transparent.

These risks or challenges can be seen as the transparency, identity and power paradox introduced by Neil M. Richards & Jonathan H. King. Within the framework of this Master Thesis, the transparency paradox is best applicable on predictive policing and will be discussed further.

Private information is unescapably collected by big data, but the further processing of this data is almost done entirely hidden in legal and commercial secrecy. This can be called the

---

[117] van der Sloot B., *Is the Human Rights Framework Still Fit for the Big Data Era? A Discussion of the ECtHR's Case Law on Privacy Violations Arising from Surveillance Activities*, in: Gutwirth, S., Leenes, R., De Hert, P. (eds.), *Data Protection on the Move* (2016), p.12-13 and van der Sloot, B., Broeders, D., Schrijvers, E. (eds.), *Exploring the Boundaries of Big Data* (2016) p.191

[118] Ministerie van Veiligheid en Justitie; *Kabinetsstandpunt over WRR-rapport Big Data in een vrije en veilige samenleving* (2016), p.5

[119] van der Sloot, B., Broeders, D., Schrijvers, E. (eds.), *Exploring the Boundaries of Big Data* (2016) p.11

[120] van der Sloot, B., Broeders, D., Schrijvers, E. (eds.), *Exploring the Boundaries of Big Data* (2016) p.20-23

[121] van der Sloot, B., Broeders, D., Schrijvers, E. (eds.), *Exploring the Boundaries of Big Data* (2016) p.20-23

[122] Smit, S., de Vries, A., van der Kleij, R., van Vliet, H., *Van predictive naar prescriptive policing; Verder dan vakjes voorspellen* (2016), p.32

[123] Article 29 Working Party, Guidelines on Automated individual decision-making and Profiling for the purposes of Regulation 2016/679, WP251, 3 October 2017, p.17

transparency paradox.[124] An ordinary example of this is 'Home automation'.[125] Paragraph 2.1.1 mentions that, when explaining what volume is, household devices like televisions, thermostats and even smoke detectors collect data. This collected data is interesting for the consumer, but also for the manufacturer. However, it is not always clear to the consumer whether and how this data is used by the manufacturer.
Some big data collected is done in secrecy because of highly sensitive intellectual property and national security assets.[126] This way of thinking can be extended to CAS. CAS uses the information to make decisions about locations where a possible crime might take place. It is debatable if people, affected by the decisions or living in the concerned neighbourhood, not have a right to know on what basis the decisions are made.

Another side of transparency is the possibility of a collaboration between the police and private company with several predictive policing systems. PredPol for instance cannot correct or change their software and depend on a private company. One may question the transparency because it is unclear who decides on the features of the algorithm and who provides the information.
A good thing about CAS is that they don't collaborate with a private company. The Amsterdam police department developed their own software. This means that the system is more transparent and provides control.[127]

3.5.2 Presumption of innocence

A risk of predictive policing is that the police might arrest citizens only based on the algorithm. This might be a bit exaggerated, but it almost goes that far as that you will be stopped by the police when you walk on the street with a toolbox in an area where a lot of robberies are taking place. Something that prevents this from happening is the presumption of innocence. According to the presumption of innocence, a suspect is innocent until proven otherwise. Article 6 paragraph 2 of the ECHR prescribes the presumption of innocence as follows: "Any person against whom a prosecution has been brought shall be presumed innocent until his guilt has been established in law". The idea behind the presumption of innocence is that the suspect has the right to remain silent and the burden of proof lies with the prosecutor. The result of large-scale surveillance and the use of predictive policing however, is that a lot of personal data is in the hands of the authorities. The suspect has no control over his own data. Moreover, part of the evidence can be collected during the time when the suspect was not yet a suspect. This may be seen as a lack of transparency. In addition, this may cause a shift of the burden of proof in large-scale surveillance, which puts the presumption of innocence under pressure.[128]

A judge should not settle for an arrest without having seen any proof. This means that police in the field will have to assess how much value can be attributed to a prediction of the algorithm.[129]

---

[124] Richards, N.M. & King, J.H., *Three Paradoxes of Big Data* (2013), p.42
[125] Home automation is the application of electronics and home networks for the automation of processes in and around the home. When connected with the Internet, it is an important element of the Internet of Things. <https://en.wikipedia.org/wiki/Home_automation> (last visited 11 April 18)
[126] Richards, N.M. & King, J.H., *Three Paradoxes of Big Data* (2013), p.43
[127] van der Sloot, B., Broeders, D., Schrijvers, E. (eds.), *Exploring the Boundaries of Big Data* (2016), p.124
[128] Bonicci, J.P.F. & Milaj, J., *Unwitting subjects of surveillance and the presumption of innocence* (2014), p.420
[129] Smit, S., de Vries, A., van der Kleij, R., van Vliet, H., *Van predictive naar prescriptive policing; Verder dan vakjes voorspellen* (2016), p.32-33

3.5.3 Human interpretation

This brings the risk that humans need to interpret the outcomes of the algorithm. This makes humans an important link in predictive policing. The algorithm gives a basis, but police officers determine what happens with the outcome. With CAS, the outcome of the algorithm is not binding. So humans will always play an important role in CAS. Nevertheless, it is questionable whether you want to put almost all of the responsibility in a machine, because then the 'thought police' from Minority Report will be very close to becoming reality.[130] As for human interference, the data can be interpreted in a coloured or wrong way. This results in coloured predictions and may lead to discrimination. This is mostly of risk when predictive policing is used for predications about who is going to commit a crime and not per se when it is about locations. Transparency is very important here.[131]

3.5.4 Social sorting

Social sorting, mentioned in paragraph 3.1.1, is the process were individuals are divided into different social categories on the basis of similar characteristics, for example gender, race or profession.[132] The analysis is fully automated, nevertheless, there is a change of discrimination. The algorithm filters and classifies data with the aim to assess people. It therefore also affects the lifestyle of people and the choices they make. For example, if you do not accept the cookies of an Internet page, you are denied access to or certain information on this page. The consequence with predictive policing is that crime statistics about the neighbourhood where you live may influence matters such as insurance premiums and the extent to which police are present in this district. In other words, social sorting of people is decisive for the way you are treated. Classifying into categories based on the risk that people form is related to the assumption that there is a certain standard of behaviour. Not meeting this standard may then be seen as suspicious.[133] With predictive policing, social classification translates into filtering the identity of people in categories of inclusion or exclusion of (extra) surveillance. The question of how these categories come into play becomes a political and ethical issue. There is a danger that these categories will be based on social stereotypes, which will become entrenched and institutionalized in the long term. The profiling of passengers at airports is an example of this. Certain marginalized and already discriminated groups in society are subjected to more security investigations before they are allowed to board the aircraft than other groups.[134]

3.6 Self-fulfilling prophecy

According to the Cambridge Dictionary, self-fulfilling prophecy means ‘something that you cause to happen by saying and expecting that it will happen’.[135] This may be the case with the use of predictive policing when big data is involved.

---

[130] Smit, S., de Vries, A., van der Kleij, R., van Vliet, H., *Van predictive naar prescriptive policing; Verder dan vakjes voorspellen* (2016), p.52

[131] Smit, S., de Vries, A., van der Kleij, R., van Vliet, H., *Van predictive naar prescriptive policing; Verder dan vakjes voorspellen* (2016), p.39

[132] Bodenhausen, G., Kang, S., Peerey, D., *Social Categorization and the Perception of Social Groups* (2011), p.318

[133] Lyon, D., *Surveillance as Social Sorting: privacy, risk, and digital discrimination* (2003), p.20-22

[134] Van Brakel, R. & De Hert, P., *Policing, surveillance and law in a pre-crime society: Understanding the consequences of technology based strategies* (2011), p.176

[135] https://dictionary.cambridge.org/dictionary/english/self-fulfilling-prophecy (last visited 15 april 2018)

Data on which predictions are based are not always of good quality. Error is therefore a crucial problem with big data. The police data that are used can be faulty, like hastily made notes or half references. The poorer the data, the more unreliable the prediction. Reliability of the data also depends on how the technology is implemented by the police. Since CAS uses analysis conducted at aggregate level it is less of an issue. But that does not mean that data can be coloured which may lead to ethnic profiling. The historical crime figures used in CAS might contain certain prejudices, for example because certain groups are more often arrested for the same offense than others. If an algorithm starts looking for patterns in it, those same prejudices may roll out of the system again. This leads to self-fulfilling prophecy, a prediction that makes itself true.[136]

When a lot of surveillance is being carried out in a neighbourhood, those neighbourhoods will appear more prominently in the crime figures because more crime will be found, as mentioned earlier. It must be mentioned that (extra) police attention might be positive, because people feel more safe. But the extra police attention might also increase the existing problems further, which in turn can be the basis for new policy, which in turn further strengthens the (negative) image. This is the self-fulfilling prophecy. It is a threat for CAS. Although CAS does not use data to identify specific individuals, the way CAS works may, indirectly, have the same effect. When police are deployed more in neighbourhoods with 'bad reputations', this might be accompanied by ethnic profiling. With the consequence that police may arrest someone walking around in a hoody with a coloured skin or stop an expensive car with someone in it with a coloured skin, because the appearance does not add up. This means that a high rate of false information increases the risk that certain citizens, individuals and groups, are more often seen as potential criminals. This might stigmatize individuals to start showing criminal behaviour.[137]

3.7 Right to privacy

What does the use of algorithms mean for the privacy of individuals? What will change in that regard when real-time data, like social media, is used? Well, exact movements of persons or groups can be monitored then. When personal data is at stake, data protection laws are applicable. The right to privacy is protected under Article 8 ECHR and Articles 7 and 8 of the Charter. Under Article 8 paragraph 2 ECHR a violation of the right to privacy is permitted, but in those cases there must be a legal basis, and the conditions of proportionality and subsidiarity need to be met.[138]

As stated in paragraph 2.2, data mining is the process where patterns can be found between different people or outcomes to determine what aspects make them similar or different.[139] With data mining, profiles are created. But this investigative method may intervene with the right to privacy of citizens.

Bearing in mind, the algorithm used with predictive policing requires information like the time, date, location and type of offense. This is then compared with historical crime data and environmental factors to come to a prediction. The conclusions drawn on the basis of

[136] van der Sloot, B., Broeders, D., Schrijvers, E. (eds.), *Exploring the Boundaries of Big Data* (2016), p.126
[137] van der Sloot, B., Broeders, D., Schrijvers, E. (eds.), *Exploring the Boundaries of Big Data* (2016), p.126
[138] Smit, S., de Vries, A., van der Kleij, R., van Vliet, H., *Van predictive naar prescriptive policing; Verder dan vakjes voorspellen* (2016), p.35
[139] Selbst, A.D., *Disparate Impact in Big Data Policing* (2017), p.13-14

predictive policing can have an influence on the privacy of the person who suffers from these conclusions by, for example, being watched or arrested.[140]

3.7.1 Foundations for the use of big data

Justification for the use of big data and big data analysis can be found in Articles 3, 9, 10 and 11 of the Police Data Act. According to Articles 9, 10 and 11 of the Police Data Act, data and data files can be subjected to automated analysis. Data mining is regulated in Article 3 of the Police Data Act. This law states that the Dutch Police is burdened with the investigation of criminal offences, among other things. When investigative methods are not specifically regulated by the law, case law of the Dutch Supreme Court can be used as a legal basis.[141] But, this may only lead to a limited interference with fundamental rights like the right to privacy.[142] If this infringement is more than limited, there must be a specific or appropriate legal basis for this.[143]

For the use of CAS, there is no explicit legal provision. Without a specific legal provision, there is no clarity about the limitations of data mining of personal information. This means that sensitive information such as ethnicity can easily be obtained and used. Furthermore, there is no supervision over the system. These 'shortcomings' can lead to interference with the right to privacy.[144]

Considering the above, the legal framework of data mining in the Netherlands can be seen as too limited. This constitutes that there is no accordance with the exceptions of paragraph 2 of Article 8 ECHR. It is unclear to what extent data mining is permitted under the legal framework and who has supervisory authority. This means that interference with the right to privacy is unavoidable.[145]

3.7.2 Conditions for the use of big data

In the Netherlands, one of the judgements that gives conditions concerning the use of technology in relation to privacy is the 'Zwolsman' judgement.[146] This judgement states that the exercise of the powers should be reasonable and moderate in relation to the intended purpose. The progressive development of the fundamental right to the protection of privacy and the increasing technical refinement and intensification of research methods and techniques require a more precise legitimation in the law.[147] According to this judgement, the existing powers seem sufficient for predictive policing. With regard to big data analysis, Article 11 of the Police Data Act is important. This article gives conditions, for when police data processed for a specific investigation can be automatically compared with other police

---

[140] Lodder, A. R., van der Meulen, N. S., Wisman, T. H. A., Meij, L., & Zwinkels, C. M. M., *Big Data, Big Consequences? Een verkenning naar Privacy en Big Data gebruik binnen de opsporing, vervolging en rechtspraak* (2014) p.69

[141] Supreme Court, ECLI:NL:HR:1995:ZD0328, December 19th 1995 and Supreme Court, ECLI:NL:HR:2014:1563, July 1st 2014

[142] Brinkhoff, S., *Big Data Data Mining by the Dutch Police: Criteria for a Future Method of Investigation* (2017), p.61

[143] Brinkhoff, S., *Big data datamining door de politie. IJkpunten voor een toekomstige opsporingsmethode* (2016), p.1401

[144] Brinkhoff, S., *Big Data Data Mining by the Dutch Police: Criteria for a Future Method of Investigation* (2017), p.62

[145] Brinkhoff, S., *Big Data Data Mining by the Dutch Police: Criteria for a Future Method of Investigation* (2017), p.62

[146] Supreme Court, ECLI:NL:HR:1995:ZD0328, December 19th 1995

[147] Zwolsman Judgement paragraph 6.4.4

data that are processed pursuant to Article 8 or 9 of the Police Data Act, in order to establish whether there are any links between the relevant data. First of all, the processing needs to be necessary for the investigation. Another condition is that it must be about police data. Looking at Article 1 of the Police Data Act, any personal data processed in the performance of the police task can be qualified as police data. This means that big data can only be used as police data if there is an identifiable reason why they are being sought.[148]

The random collection of data without limitations and without a clear goal might be interesting for detection because the unfocused collection of data can produce in an earlier stage unrecognized patterns or results. For purposes of investigation, completely untargeted retrieval of information cannot be reconciled with work based on suspicions and is also at odds with the proportionality test that must always be taken into account in case of privacy breaches.[149] Something important to keep in mind is that the law prescribes that data may not be used for any purpose other than for which it was collected, and that it may not be used more than is strictly necessary. The advantage of big data research lies in the non-focused collection and combining of endless amounts of data, so that unexpected patterns can emerge.[150]

3.9 Conclusion

The goal of this chapter was to give an answer to the question 'How are minorities disadvantaged by CAS?'.

First summarizing the chapter. The police use ethnic profiles when targeting certain forms of crime. Ethnic profiling refers to the use of race, ethnicity, religion, or national origin instead of individual behaviour when it comes to law enforcement and/or investigative decisions about who may be involved in criminal activities. This is connected to discrimination, because the use of ethnic profiles can enable discrimination. Discrimination is criminalized in Article 1 of the Dutch Constitution and Article 14 ECHR. Article 14 ECHR is likely to be violated when a difference of treatment does not pursue a legitimate aim, but also when it is clearly established that there is no reasonable relationship of proportionality between the means employed and the aim sought to be realised.

When police are deployed to certain locations where, according to the algorithm, criminal activities are taking place, more crime will be found. With this happening, the algorithm creates unintended a racial bias. But the question is, is it discrimination when there are more police present at places where really is more crime? Police presence can aggravate the already existing crime problem, but on the other hand the police needs to do their job. This issue can create a self-fulfilling prophecy which can be harmful.

Regarding place-based bias, training data, past crime data and feature selection, these can give a problem with discriminating minorities. Regarding person-based bias, when more data of citizens are used in the system, the system is more aimed at individuals. With the

[148] Lodder, A. R., van der Meulen, N. S., Wisman, T. H. A., Meij, L., & Zwinkels, C. M. M., *Big Data, Big Consequences? Een verkenning naar Privacy en Big Data gebruik binnen de opsporing, vervolging en rechtspraak* (2014), p.62-63

[149] Lodder, A. R., van der Meulen, N. S., Wisman, T. H. A., Meij, L., & Zwinkels, C. M. M., *Big Data, Big Consequences? Een verkenning naar Privacy en Big Data gebruik binnen de opsporing, vervolging en rechtspraak* (2014), p.79-80

[150] Lodder, A. R., van der Meulen, N. S., Wisman, T. H. A., Meij, L., & Zwinkels, C. M. M., *Big Data, Big Consequences? Een verkenning naar Privacy en Big Data gebruik binnen de opsporing, vervolging en rechtspraak* (2014), p.35

consequence being that the algorithm might think that being a member of a minority is more likely to be a criminal. Feature selection is also an issue.

We might assume that big data commits different violations, which might be small, but more and more minor violations result in a major violation. The most important violation for this Master Thesis is discrimination. An important risk, which may lead to discrimination, is social sorting. People are placed in boxes based on a characteristic by the algorithm. Furthermore, with the use of big data, information is gathered of people who are not a suspect (yet). This collides with the principles of data minimization and purpose limitation. It might even be argued that there is no explicit legal provision thus the legal framework is too limited. This results in an infringement of Article 8 paragraph 2 ECHR. But, on the other hand, in Article 1 Police Data Act legal restrictions of data collection are defined. The collected data must be regarded as police data. And this brings us back to the principles of data minimization and purpose limitation.

The consequence of predictive policing on the presumption of innocence is very interesting. Information collected by the authorities about a suspect, which might be useful, can be used against the suspect. The suspect has no control over this data and it might even be collected before the suspect was a suspect. This may cause a shift of the burden of proof from the prosecutor to the suspect when big data is used in predictive policing.

The human interpretation of the big data analysis is an important link in predictive policing. Because of human interference, an outcome of an analysis may be interpreted in a coloured or wrong way. Consequently, predictions might be coloured, which may lead to discrimination. With CAS, outcomes of analyses are not binding. This means that people continue to play a major role in interpreting the data and that there will be a risk of discrimination. The risk human interference brings with it, is a type of error difficult to completely rule out.

Another important risk of predictive policing is self-fulfilling prophecy. Self-fulfilling prophecy means that something will happen because people say and expect it to happen. With CAS, it might be less of an issue because the system conducts analysis at an aggregated level. Even though, the historical crime figures used by CAS may contain certain prejudices. If the algorithm uses this data to look for patterns, the same prejudices may roll out of the system again.

The transparency paradox means that the collection of private information is done by big data, but the processing of this data is almost done entirely hidden. It should not be desirable that surveillance is done in secret and decisions are made by a system of opaque decision makers. Important to notice, CAS does not collaborate with private companies and does not exchange data or cooperate with other institutions, besides the CBS.

It might be concluded that the police might do minorities wrong on several points. But transparency is a recurring term that plays a very important role with the use of big data. Will transparency be the solution to infringements when using CAS? An answer to this question will be given in the next chapter.

4. Framework for CAS

In the previous chapter different possibilities of infringements by using predictive policing systems were discussed. This chapter will give a framework to see if CAS can be used without infringing rights, and mostly focus on the rights in which minorities are disadvantaged.

4.1 Ethnicity and discrimination

4.1.1 Ethnic profiling

With ethnic profiling stereotypes about criminal offenders and minority groups are used. This causes police to look at group attributes and not at individual behaviour. This kind of group profiling may lead to discrimination and social sorting. Furthermore, it might infringe the principle of equality and presumption of innocence. Another possible consequence is that when action is taken based on unlawful profiling, there might be an increase in racial tensions. These tensions may lead to resentment towards police within minorities.

CAS does not give a profile of who might commit a crime, so it could be said that ethnic profiles are not an issue here. But that does not mean that a police unit does not make use of ethnic profiles or that police on the street take racial appearance into consideration. This means that ethnic profiles might be a factor in the use of CAS, but it is not something facilitated by the system per se. The police need to look at the human factor with this problem, because when ethnic profiles are used, it will come from humans.
There is a possibility to use racial profiles legitimately. This happens when ethnicity is an important factor to identify a suspect, however, it may not be the only ground.

Furthermore, even though CAS does not give a profile, it may create a stigma of people in areas where high-crimes are located. For example, when the police go to an Indonesian neighbourhood, the stigma may be that it must be an Indonesian looking person.

4.1.2 System and human factor

Discrimination is the unjust or prejudicial treatment of different categories of people, especially on the grounds of race, age, or sex. Discrimination can be done by a system or by a person. Due to discrimination, people are no longer seen as unique individuals. It may also contribute to deterioration of relations between different groups in society.

There is a possibility that discrimination might happen when using predictive policing systems. As said above, it might happen that police on the street take racial appearance in consideration. So looking at the human factor in discrimination, it might be an issue in the use of CAS. I am not sure how this should be prevented and if it would be possible at all to do so. As said in paragraph 1.2, it is difficult to randomly select someone out of a group of people. A police officer needs to make a selection of who might be a risk based on his knowledge. He will look at characteristics of people and then decide if that person might be a risk. So how random will this selection be? Not every police officer will, of course, discriminate. However, the possibility remains that a police officer does discriminate.

Looking at the system factor, CAS, using certain data may create a discriminating outcome. Data from CBS might increases the likelihood of unintentional racial profiling. Past crime data needs to be accurate and correct, otherwise it might corrupt the system. Furthermore, the information used as training data needs to be neutral, otherwise it will keep discriminating.

With feature selection, the choices that are made may have a serious impact on minorities when statistical differences are not properly reflected in the set of selected features.

The possible unintentional racial profiling, because of using data from CBS, might be prevented by using only crime figures. The consequence is that it does not give an accurate heat map. The system of predictive policing depends a lot on the data that goes into it. So the data needs to be accurate and correct. But reported crimes are not always that accurate. Some neighbourhoods are more likely to report crimes and some crimes are more likely to be reported. When a car is stolen, people often report this to the police. When a drugs deal went bad, most of the time it will not be reported. Is this also taken into consideration by CAS?

The pitfall of training data is that it is the fundamental truth for the algorithm. If there is wrongful data or discriminating data used, decisions will be made based on that information. It will not be a pitfall if the creators of CAS used the right data. This is, of course, a big if. But I do not have enough information to give it a constructive judgement. Nevertheless, data miners need to be extra careful when they setup a predictive policing system and start to work with it.

The feature selection is important for the system; it determines which data will be used. I am not sure on which grounds the selection is done with CAS because I could not find that in the literature. But I think it would be best if an objective criterion is used. There should be guidelines deciding which features may and may not be considered. And these guidelines need to state to which extend CAS may grow. There should be national, or even European, rules about this and perhaps an independent commission whose sole purpose it to check compliance.

4.2 Predictive policing risks

Another risk is the chilling effect. With this chilling effect, infringements of the right to freedom of speech and the right to privacy in relation to discrimination of minorities might be made. To prevent this from happening, transparency is important. The police should be more open about CAS. They should make information about the system available to people and inform them about the data that is used. Of course, in case of an active investigation, not everything has to be transparent at first. But then a special commission might come into play. They can monitor the cases when the police cannot be transparent.

The presumption of innocence means that a suspect is considered innocent until proven otherwise. A lot of information is collected while a person is not a suspect (yet) nor does the person has any control over this data. This shows lack of transparency. Again, it might be helpful if a commission comes into play to exercise control.
Furthermore, the possibility exists that the presumption of innocence might be passed. When the presumption of innocence would be passed, then the police may have more freedom to act without judicial control. For example, the police need permission from the examining magistrate to tap a telephone. This check is becoming vaguer in predictive policing. Of course, permission for a telephone tap is still needed, but it could be possible that the police perform a stop-and-search in a high risk area because they think it is needed without the needed permission from a magistrate. This is not legal, but a judge might be more accepting if they explain why they did it. This might be far-fetched of course, but this should in no way be grounds for the police to pursue this tactic. However, it might be a good control if a judge or special commission examines certain actions taken by the police before it happens. Not only for aforementioned example, but also if a certain data should be collected or a feature should

be used for analysis. Otherwise, the police might be able to keep an eye on people without a purely legal ground for that. This creates an almost unrestrained freedom when they intervene earlier in the process. A suggestion to prevent this from happening might be that the police need to be educated regularly about what kind of powers they do and do not have, and the consequences of misuse.

Human interpretation might also be a risk. The outcome of data analyses always needs to be interpreted by humans. The people working with heat maps do not always have enough knowledge of how to read such a map. A logical action is to deploy more police to a high risk area, but that is not always the best option. A solution is to train the people who interpret heat maps and analyse certain situations. Furthermore, working with a team would also be helpful. Then possible actions can be discussed. This might have the disadvantage that it takes a lot of time, but it might increase quality.
With CAS, the outcome of the algorithm is not binding. So humans will always play an important role in CAS and that is why the people interpreting the system need to work as a team and be on top of their game.

With social sorting, data is categorized to assess people. This may lead to discrimination when people are treated unjust or different, especially on the ground of race, age or gender. People are judged on their characteristics, instead of on their behaviour. This may be an infringement of the principle of equality and the presumption of innocence, which may lead to discrimination. Also, this discrimination may result in more control exercised on groups who match the profile. It may also be discrimination by algorithms because of the lack of transparency.
It must be said that social sorting will probably not play a big role in CAS. CAS only reproduces a heat map showing where a certain crime has high probability to take place. It does not focus on individuals per se. When CAS would predict who is likely to commit a crime, the issues that social sorting may bring might happen.

Self-fulfilling prophecy means that something happens because you expect it to happen. Eventually it means that the same prejudices may roll out of the system repeatedly and disadvantage minorities.
Furthermore, when the police carry out more surveillance in one neighbourhood than in another one, more crime will be found in the first neighbourhood. This might expand existing problems.
This self-fulfilling prophecy is a threat for CAS. Although CAS does not use data to identify specific individuals, the way CAS works may, indirectly, have the same effect. When police are more often deployed in neighbourhoods with ‘bad reputations’, this might be accompanied by ethnic profiling. This, of course, cannot happen. The police need to do everything they can to avoid this from happening.

According to the Police Data Act, there must be an identifiable reason why personal data is sought. Bearing in mind that CAS does not make predictions about people who are likely to commit a crime, it does not mean that the police do not use personal data. For example, they use crime history and information about distance to known suspects. When a person has done his time, does he still need to be treated as a risk? And does a suspect of a crime not have the right to the presumption of innocence? With CAS, there should not be a random collection of data without limitations nor a clear goal. It is not about collecting data from as many people as possible, limitations, like feature selection, must be made.

4.3 Transparency

Big data promised to make everything more transparent, but in most of the time this is not the case. When information is unescapably collected by big data, but the further processing is almost done entirely hidden in legal and commercial secrecy, we may speak of a transparency paradox.

To counter this paradox, the Article 29 Working Party proposes that consumers should have access to their data profile. Then there will be more transparency because they will know what data is collected about them. Furthermore, people also get the opportunity to correct data. This opportunity will provide more reliable and complete data. However, the data that can be corrected will not be data available in police systems. Citizens will not have access to this. But other data that is used by the police or have obtained might be available to correction, like the information from the CBS. An advantage of this is that when wrongful information is being processed, wrongful information will come out of the data processing. Nevertheless, incorrect results may still be obtained on the basis of correct data. To improve this, better algorithms need to be used and there always need to be looked critically at the outcomes.

Some predictive policing systems are a collaboration of police with private companies. CAS does not exchange data or cooperate with other institution, besides the CBS. This is a good thing for the transparency. The Amsterdam police department developed their own software which means that they have complete control over the system.

4.4 Conclusion

The goal of this chapter was to give an answer to the question 'Is it possible to use the Crime Anticipation System without infringing rights?'

I think CAS can be used without infringing rights. There are, although not specific for CAS, laws which CAS must meet. Nevertheless, when using CAS, risks as described above still exists. Of course, some more than others. In my opinion, the most vulnerable part of CAS are the people working with the system. The interpretations they make and the work they do on the streets. Discrimination will always lurk, even though it might not be intentional. Human error is something the system has to deal with, as it will always be present, but it needs to be minimized.

I think transparency is very important for CAS. This applies to the data that is used and the interpretation of the data analyses. Transparency may also be used to prevent several possible risks described in this chapter.

It might also be a good idea to increase more awareness among citizens and inform them about how CAS works. Police should focus more on possible infringements like discrimination and self-fulfilling prophecy, so these possible infringements may be prevented of at least minimized. Furthermore, I think it would be very useful to install a completely independent commission that will deal with monitoring transparency, used features, etcetera.

CAS is, for now, only used to produce a heat map. With predictive policing it could be possible to predict more, like who is likely to commit a crime. To do this, more (personal) information is needed. It even might result in real-time tracking of people. Looking at the possible risks this would bring, I do not think this should be desirable as the reward of decreasing crime does not outweigh the increase of infringements of rights.

5. Conclusion

The aim of this Master Thesis was to test my hypothesis. To do this the main research question was 'Is it possible for the Amsterdam police department to deploy the Crime Anticipation System used for predictive policing in a way that does not result in discrimination based on ethnicity?'

I think ethnicity does not have to play a role when the Amsterdam police department use CAS. Although, algorithms and analyses may be bias, lacking transparency or generating self-fulfilling prophecies, with the right safeguards in place, this does not have to be an issue.

What has emerged in this Master Thesis is that there are different stages were problems with big data and predictive policing may occur. First of all, the data itself. It might be collected illegally or might be wrong or incorrect. Secondly, the algorithm. It could be bias or the wrong features have been selected. Thirdly, the data may be interpreted in the wrong way by data miners.

Predictive policing in practice is giving a probability of a certain crime that might happen as accurate as possible, designed on a map. The starting point is a large amount of data from a neighbourhood, like historical crime figures, socio-geographical information, addresses of known offenders, etcetera. Implementing more types of data will increase the accuracy of the analysis. Self-learning algorithms will then search for patterns and make a map showing high risk areas were a robbery or burglary could take place. Predictive policing is effective. Simply because criminals like to avoid risk. If they succeed in stealing a car in one area, they might succeed again in the same area. The system tries to avoid this by giving risk scores to those areas.

Referring back to Minority Report and the 'thought police', there is no sign that the police are going to arrest people who might commit a crime. Minority Report will probably not become reality. Nevertheless, the film showed that that system is not 'fool-proof' and can be manipulated by humans. Which can be extended, to a certain extent, to predictive policing systems and the human interpretation. In the film and in reality, the human factor probably is the weak link.

Besides the human factor, predictive policing comes with more risks. The most risks may be prevented, but there must be an active awareness and sometimes action to actually prevent these risks from becoming reality. Not only by the police, but it might also be a good idea to invoke a third party to exercise control over certain actions taken by the police.

Theoretically, CAS can work without discriminating ethnicities, thereby disproving my hypothesis. However, humans will always remain a limiting factor. This means that in different stages people working with CAS may negatively influence the system.

Bibliography

Articles

Barocas, Solon and Selbst, Andrew D., Big Data's Disparate Impact, *California Law Review* (104) 2016, pp. 671-732

Bonicci, J.P.F. & Milaj, J., Unwitting subjects of surveillance and the presumption of innocence, *Computer Law & Security Review* (30) 2014, pp. 419-428

van Brakel, R. & De Hert, P., Policing, surveillance and law in a pre-crime society: Understanding the consequences of technology based strategies, *Technology-Led Policing: Journal of Police Studies* (3 20) 2011, pp. 163-192.

Brayne, S., Rosenblat, A., Boyd, D., Predictive Policing, *Nature* (541) 2017, pp. 548-460

Brinkhoff, S., Big Data Data Mining by the Dutch Police: Criteria for a Future Method of Investigation, *European Journal for Security Research* (2) 2017, pp. 57-69

Brinkhoff, S., Big data datamining door de politie. IJkpunten voor een toekomstige opsporingsmethode, *Nederlands Juristenblad* (994) 2016, pp. 1400-1407

Joh, E.E., Policing by Numbers: Big Data and the Fourth Amendment, *Washington Law Review* (89) 2014, pp. 35-68

Kennedy, L.W., Caplan, J.M. & Piza, E., Risk Clusters, Hotspots, and Spatial Intelligence: Risk Terrain Modeling as an Algorithm for Police Resource Allocation Strategies, *Journal of Quantitative Criminology* (27) 2011, pp. 339-362

Lodder, A.R., Schuilenburg, M.B., Politie-webcrawlers en Predictive policing, *Computerrecht* (3) 2016, pp. 150-154

Richards, N.M. & King, J.H., Three Paradoxes of Big Data, *Stanford Law Review Online* (66) 2013, pp. 41-46

Sanders, C.B., Sheptycki, J., Policing, crime and 'big data'; towards a critique of the moral economy of stochastic governance, *Crime, Law and Social Change* (68) 2017, pp. 1-15

Schuilenburg, M., Predictive policing: De opkomst van een gedachtepolitie?, Ars Aequi (AA20160931) 2016, 931-936

Selbst, Andrew D., Disparate Impact in Big Data Policing, *Georgia Law Review* (52) 2017, pp. 109-195

Willems, D., Doeleman, R., Predictive Policing – wens of werkelijkheid, *Tijdschrift voor de politie* (76) 2014, pp. 39-42

Zarsky, T., Incompatible: The GDPR in the Age of Big Data, *Seton Hall Law Review* (47) 2017, pp. 995-1020

Books

Bodenhausen, G. V., Kang, S. K., & Peery, D., Social Categorization and the Perception of Social Groups, *in* Fiske, S., and Macrae, C. N. (eds.), *The SAGE Handbook of Social Cognition*, Los Angeles, CA: SAGE Publications Ltd., 2012, pp. 311-329

Hildebrandt, M., Defining Profiling: A New Type of Knowledge?, *in* Hildebrandt, M. and Gutwirth, S. (eds.), *Profiling the European Citizen*, Dordrecht: Springer Science + Business Media, 2008, pp. 17-45

Lodder, A. R., van der Meulen, N. S., Wisman, T. H. A., Meij, L., & Zwinkels, C. M. M., *Big Data, Big consequences? Een verkenning naar privacy en big data gebruik binnen de opsporing, vervolging en rechtspraak*. Amsterdam: WODC - Vrije Universiteit, 2014

Lyon, D., *Surveillance as Social Sorting: privacy, risk, and digital discrimination*, Routledge, London and New York, 2003

Open Society Justice Initiative, *Addressing Ethnic Profiling by Police; A Report on the Strategies for Effective Police Stop and Search Project*, Open Society Institute, New York, 2009

Open Society Justice Initiative, *Ethnic Profiling in the European Union: Pervasive, Ineffective, and Discriminatory,* Open Society Institute, New York, 2009

Rienks, R., *Predictive Policing: Kansen voor een veiligere toekomst*, Brave New Books, 2015

van der Sloot, B., Broeders, D., Schrijvers, E. (eds.), *Exploring the Boundaries of Big Data*, wrr/Amsterdam University Press, The Hague/Amsterdam, 2016

van der Sloot B., Is the Human Rights Framework Still Fit for the Big Data Era? A Discussion of the ECtHR's Case Law on Privacy Violations Arising from Surveillance Activities, *in*: Gutwirth, S., Leenes, R., De Hert, P. (eds.), *Data Protection on the Move, Law, Governance and Technology Series*, *vol 24,* Dordrecht: Springer, 2016, pp. 411-436

Smit, S., de Vries, A., van der Kleij, R., van Vliet, H.; *Van predictive naar prescriptive policing; Verder dan vakjes voorspellen*, TNO, Den Haag, 2016

Vedder, A., van der Wees, A., Koops, B.J., de Hert, P., *Van privacyparadijs tot controlestaat? Misdaad- en terreurbestrijding in Nederland aan het begin van de 21ste eeuw*, Rathenau Instituut, Den Haag, 2007

Judgements

Supreme Court, (Milk and water) ECLI:NL:PHR:1916:BG9431, February 14th 1916

ECtHR Belgian Linguistics Case (No. 2), no. 1474/62, 1677/62, 1691/62, 1769/63, 1994/63, 2126/64, July 23th 1968

ECtHR Niemietz v. Germany, no. 13710/88, December 16th 1992

Supreme Court, (Zwolsman judgement) ECLI:NL:HR:1995:ZD0328, December 19th 1995

ECtHR Cissé v France (Admissibility) no. 51346/99, January 16th 2001

ECtHR Peck v. the United Kingdom, no. 44647/98, April 28th 2003

ECtHR Timishev v. Russia, no. 55762/00, 55974/00, December 13th 2005

Supreme Court, ECLI:NL:HR:2011:BT6402, November 22th 2011

Supreme Court, ECLI:NL:HR:2014:1563, July 1st 2014

Official documents

Amnesty International; *Proactief politieoptreden vormt risico voor mensenrechten: Ethnisch profileren onderkennen en aanpakken* (2013)

Article 29 Working Party, Opinion 03/2013 on purpose limitation, WP203, 2 April 2013

Article 29 Working Party, Guidelines on Automated individual decision-making and Profiling for the purposes of Regulation 2016/679, WP251, 3 October 2017

Directive (EU) 2016/680 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data by competent authorities for the purposes of the prevention, investigation, detection or prosecution of criminal offences or the execution of criminal penalties, and on the free movement of such data, and repealing Council Framework Decision 2008/977/JHA

European Union Agency for Fundamental Rights (FRA), *Towards More Effective Policing. Understanding and Preventing Discriminatory Ethinic Profiling: A Guide* (2010)

Explanatory Report to the Protocol No. 12 to the Convention for the Protection of Human Rights and Fundamental Freedoms (European Treaty Series - No. 177)

Mali, B., Bronkhorst-Giesen, C., den Hengst, M., *Predictive policing: lessen voor de toekomst: Een evaluatie van de landelijke pilot*, Politieacademie (2017)

Ministerie van Veiligheid en Justitie, *Kabinetsstandpunt over WRR-rapport Big Data in een vrije en veilige samenleving* (2016)

Ministerie van Veiligheid en Justitie, *Projectplan Verduurzaming iRN / iColumbo*

Protocol No. 12 to the Convention for the Protection of Human Rights and Fundamental Freedoms (Treaty No.177).

Perry, W. L., McInnis, B., Price, C.C., Smith, S., and Hollywood, J.S., *Predictive Policing: The Role of Crime Forecasting in Law Enforcement Operations*, RAND Safety and Justice Program (2013)

Websites

https://dictionary.cambridge.org/dictionary/english/self-fulfilling-prophecy (Last visited 15 April 2018)

https://www.floridatechonline.com/blog/criminal-justice/4-problems-with-predictive-policing/ (Last visited 02 November 2017)

http://www.predpol.com/thank/ (Last visited 23 November 2017)

http://www.predpol.com/about/ (Last visited 22 February 2018)

http://www.imdb.com/title/tt0181689/ (Last visited 23 November 2017)

https://www.politie.nl/nieuws/2017/mei/15/05-cas.html (Last visited 23 November 2017)

http://www.slate.com/articles/technology/future_tense/2016/11/predictive_policing_is_too_dependent_on_historical_data.html (Last visited 06 December 2017)

https://www.cbs.nl/nl-nl/onze-diensten/innovatie/big-data (Last visited 21 February 2018)

https://en.oxforddictionaries.com/definition/discrimination (Last visited 21 February 2018)

https://en.oxforddictionaries.com/definition/ethnicity (Last visited 22 February 2018)

https://www.eugdpr.org (Last visited 13 April 2018)