The Politics and Biases of the "Crime Anticipation System" of the Dutch Police

Serena Oosterloo[1,2] and Gerwin van Schie[1,2]

[1] Utrecht Data School, Drift 13, 3512 BR, Utrecht, The Netherlands
[2] Utrecht University, Drift 15, 3512 BR, Utrecht, The Netherlands

Abstract. In line with developments in many areas of business and governance, where bureaucracies of all sorts are increasingly datafied for budgetary reasons and the additional possibilities for automated analysis, the Dutch Police started with so-called Intelligence-Led Policing. This development led to the creation of the Crime Anticipation System (CAS). This data-driven system tries to predict crimes with statistics based on three data sources: BVI (Central Crime Database), GBA (Municipal Administration) and CBS (Demographics from Statistics Netherlands). By analyzing the used data categories with a critical data studies approach, we will show that the epistemological question concerning predictive policing systems turns into an ontological one: how are living environments and police work mutually shaped and determined by data? We will argue that intelligence-driven policing is not only a qualitative shift, but also has its continuities, since already existing ideas and biases concerning suspects and crimes are reproduced in the information and system of CAS.

Keywords: Predictive policing, intelligence-driven policing, critical data studies, data visualization, information bias.

1 Introduction

During an interview, Dick Willems, data scientist at the Amsterdam Police Department and the creative brain behind the Dutch predictive policing system, shared an anecdote about the negative correlation he found between the occurrence of flashers and the amount of burglaries in a neighborhood. As a result he once, jokingly, suggested police officers to sometimes take off their clothes in certain neighborhoods in order to make the number of burglaries go down. In this case it is easy to spot the confusion of causation and correlation. Because of the occurrence of flashers, the police decided to patrol more often in certain areas. This increase in surveillance was the reason for a drop in the number of burglaries, not the fact that exhibitionists gave in to their urges. As we will show in this paper, the Dutch Crime Anticipation System (from now on CAS) hosts a number of other assumptions about the relationship between crime, human characteristics, locations, and the interpretations police officers might have. Since CAS is developed in-house (rather than being a commercially available product, like PredPol), the

Dutch police itself is able to shape and tweak the system. In addition, because of the open attitude of the Dutch police and existing transparency laws, researchers are able to request information regarding the types of data that are used and the way in which information is presented.

In this paper we will critically analyze the data categories and data sources that are employed in CAS. With a critical data studies approach (see Dalton and Thatcher 2014; Dalton et al. 2016), which helps us discuss the underlying influences from people, theory and organizational structures that frame and color a data-driven system, we try to reconstruct this assemblage of actors related to CAS. As Andrew Illiadis and Frederica Russo argue, these data-driven systems are not neutral, but should be seen "as always-already constituted within wider data assemblages" (Iliadis and Russo 2016, p.1). Our task, in relation to CDS, is to uncover these ideas that are implemented in a system, which eventually also influence decisions and situations in the world.

We will perform an infrastructural inversion, a way of "recognizing the depths of interdependence of technical networks and standards, on the one hand, and the real work of politics and knowledge production on the other" (Bowker and Star 1999, p. 34). We will show that categorization systems are not free from the history and culture of the organizations in which they are embedded. CAS will be understood as a *data assemblage*, a term introduced by Rob Kitchin and Tracey Lauriault, which they explain as "a complex socio-technical system, composed of many apparatuses and elements that are thoroughly entwined, whose central concern is the production of data" (Kitchin and Lauriault 2014, p. 6). By analyzing these apparatuses one focusses on "the technological, political, social and economic apparatuses and elements that constitute and frame the generation, circulation and deployment of data" (Kitchin and Lauriault 2014, p. 1), in order "to track the ways in which data are generated, curated, and how they permeate and exert power on all manner of forms of life" (Iliadis and Russo 2016, p. 2). By combining the notion of CAS as a data assemblage with an infrastructural inversion we intent to reflect both on the socio-technical network of CAS within the police organization as on its dependence on categorization practices in Dutch society, science and policy as a whole. Through our analysis we will show that the epistemological question concerning predictive policing systems turns into an ontological one: how are living environments and police work mutually shaped and determined by data?

We will first reflect on the (not always voluntary) choices of data categories, which are not only shaped by availability, but also through the existing culture of policing which already has ideas about what characteristics are prevalent in potential suspects. Second, we will show how CAS is in no way exhaustive or all-inclusive. It is tied to types of crimes that are location and time based. Third, by highlighting the data-visualization in the form of a map and the possibilities and impossibilities for the users of CAS (information specialists), we will show that the data cannot speak for itself; it needs an interpreter. Finally we will argue that intelligence-driven policing is not only a qualitative shift, but also has its continuities, since already existing ideas and biases concerning suspects and crimes are reproduced in the information and system of CAS.

2 The Crime Anticipation System

In line with developments in many areas of business and governance, where bureaucracies of all sorts are increasingly datafied for budgetary reasons and the additional possibilities for automated analysis (see Kitchin 2014; Schäfer and van Es 2017), the Dutch Police started with so-called Intelligence-Led Policing (Kop and Klerks 2009). The groundwork for this development was laid out by the formation of a national police force from the formerly locally organized "base teams", which also resulted in a standardized central national crime database (Strategische Beleidsgroep Intelligence 2008). This development led to the creation of the Crime Anticipation System (CAS). This data-driven system tries to predict crimes through statistics based on three data sources: BVI (Central Crime Database), GBA (Municipal Administration) and CBS (Demographics from Statistics Netherlands). CAS uses different kinds of data, which date back up till three years, which can be classified in three types (Willems and Doeleman 2014, p. 41). The first type of information is socio-economic data from the Central Bureau of Statistics (CBS) (Willems 2017c; Willems 2017b). This data focusses on people's age, incomes and the amount of social benefits in an area. The second type is data from the *Basisvoorziening Informatie* (basic information provision, BVI), this is the data that is gathered by the police force itself and focusses on previous crimes, locations and known criminals (Willems 2017b). The third type of data comes from the Municipal Administration (BAG). The data from this source, consisting of streets and addresses, is not used as a predictor but as structure for the map whereon predictions are made.

Fig. 1. Picture of the interface of CAS. Color code for risk factors: red = burglary; green = street robbery; blue = disturbance by youth; pink = bicycle/scooter theft

CAS can be categorized as a spatiotemporal prediction system, meaning it focuses on hot spots and hot times in a city, and not on high-risk individuals. Through this data, CAS constructs heat maps (see Fig. 1), that illustrate which places have a higher risk for high-impact crimes. These heat maps show blocks which are left blank when the risk is either low or nonexistent, and the color increases in intensity, when the amount of risk increases (Mali et al 2017, 91). Different colors correspond to different crimes.

Only the top three percent of high-risk areas will be colored in one of the colors. This is to relate the amount of colored fields to the capacities of a police force.

The goal of CAS specifically, is to predict more at-risk areas in a city, and improve efficient distribution of manpower (Mali et al. 2017, p. 91; Willems and Doeleman 2014). Dick Willems initially constructed CAS for the Police of Amsterdam in 2013 (Mali et al 2017, 18). After testing this system in Amsterdam, a pilot project was started in the Dutch cities of Enschede, Hoorn, Hoefkade, and Groningen-Noord. Despite non-conclusive results with respect to user-experience of police officers in the four before mentioned cities (see Mali et al. 2017), CAS was made available for all police teams in The Netherlands in May 2017. Currently CAS is in use at 110 base teams, out of a total of 167, in 6 out of 10 districts (Willems 2017c). It could therefore be seen as a major factor in determining where and how the Dutch police do their surveillance and patrols.

3 Location and Time

The CAS maps are created with GBA-data (municipal database with addresses, zip codes etc.). The map is divided in blocks of 125 x 125 meters, a size which was determined through a process of trial and error by each time halving squares, starting from 1000 x 1000 meters. It proved to be a good compromise between precision and practicality. Too big of a square would be unworkable for patrolling officers, too small of a square would have a low hit rate and render it useless in the prediction (Willems 2017c). After plotting the squares on a map, all "empty" squares, such as squares with no houses or companies, and squares consisting of only water, forest or farmland are removed. Then of each square, it is determined to which zip code it belongs, and which addresses would fall within its borders.

Next, historic crime data from the BVI-database (addresses, locations and time/date of recent incidents of a certain type) is added. The number of incidents within each square and within its surrounding squares for several time periods before the prediction is determined (see Fig. 2).

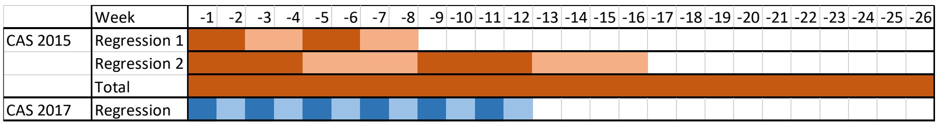

Fig. 2. Conceptualization of time periods before prediction in different versions of CAS

In CAS 2015 the resolution of BVI was two weeks, resulting in crime numbers of four consecutive two-week-periods, four consecutive four-week-periods, and one total number of incidents for a period of 26 weeks. Two directional coefficients were calculated. The first directional coefficient is derived from the number of incidents in the previous four two-week-periods (see Fig. 2, CAS 2015, Regression 1); the second is derived from the number of incidents in the previous four four-week-periods (see Fig. 2, CAS 2015, Regression 2).

In CAS 2017, the resolution of BVI has increased to one week. A choice was made for the calculation with twelve consecutive one-week-periods. In addition, the way in which crime trends are calculated has changed as well. Only one directional coefficient based on the twelve one-week-periods was used (see Fig. 2, CAS 2017). The change from two-week numbers to one-week numbers has increased the resolution of the system by a factor of two.

3.1 Limitations and biases of location and time variables.

Because CAS is based on locations, it is only suitable for the prediction of crimes that can be tied to a very specific place, such as burglary, pickpocketing, mugging, etc. It is not made to predict or identify incidents of, for example, fraud. It therefore also only targets a specific demographic amongst criminals, something which will be discussed in more detail in the next paragraph.

For accurate predictions it is necessary to have a relatively precise time of previous incidents. For pickpocketing, and mugging the time is usually quite precise (within an hour or so), for burglary however it is often much harder to determine. People are often not home or not awake when a burglary happens. Since BVI only works with exact times, rather than with a time window, the time exactly in between leaving and returning home (or falling asleep and waking up), is registered in the database (Willems 2017c). It is unclear whether or not this affects outcomes of predictions.

Another limitation of CAS is the number of incidents of a given type that happen within a delimited timeframe. Incidents that do not happen very often, such as murder, or incidents that are not always reported, such as rape or the sale of fake drugs, cannot be predicted accurately within a spatiotemporal system.

4 Demographic Data

In order to make predictions about locations, CAS makes use of open demographic data provided by CBS (see Table 1). After determining the zip code of which a square is part, aggregated information regarding that zip code area is incorporated in the risk score of a square. CBS indicators 1 to 11 have to do with the people that live in an area; indicators 12 to 16 have a more economic nature.

Indicators 1, 2, 3, and 4 refer to the number of inhabitants, men, women and households in an area. Indicators 5, 7, 8, 9, 10 and 11 refer to the composition of the households, incorporating the average number of people in a household, whether or not there are children in the house and if there are one or two parents in the household. Indicators 10 and 11 refer to the property values in an area and how many properties are currently empty. The next three indicators, 14, 15, and 16, are about the number of people that receive an income, whether or not this income consists of social benefits and what the total income is in an area.

Given the size of the squares (125x125 m) it becomes clear that the system is more suited for an urban area, since in rural areas the number of people and households might not be statistically valid. In addition a predisposition can be seen towards the shape of a family (one or two parents) and economic factors. It gives the idea that both criminals and/or victims are more often part of “broken families”. Because of the indicator regarding the recipients of social benefits it seems as if poorer people have higher likelihood of criminal behavior or victimhood, as the other end of the spectrum is not part of the system (i.e. there is no indicator mentioning the number of millionaires).

Table 1. CBS indicators used for each CAS square.

Nr.	Explanation
1	Number of inhabitants in zip code area
2	Number of men in zip code area
3	Number of women in zip code area
4	Number of households in zip code area
5	Average size of households in zip code area
6	Number of non-Western allochthones in zip code area (removed in 2017)
7	Number of one person households in zip code area
8	Number of one parent households in zip code area
9	Number of multiple person households without children in zip code area
10	Number of two parent households in zip code area
11	Average age in zip code area
12	Housing stock in zip code area
13	Average property value in zip code area
14	Number of income recipients in zip code area
15	Number of social benefits recipients in zip code area
16	Fiscal monthly income in zip code area

The final indicator that should be mentioned is number 6, which consists of the number of so called *non-Western allochthones* and was removed in CAS 2017. The term *allochthone*, together with the term *autochthone* are used to describe the national origins of Dutch citizens. The term autochthone is used for people with two parents born in the Netherlands. Allochthone is used for people with at least one foreign-born parent. Allochthones are then split up in “Western” and “non-Western” people based on the “different socio-economic and cultural position of Western and non-Western countries” in comparison to the Netherlands. This results in the categorization of European countries (except Turkey), North American countries, Japan and Indonesia as “Western” , and South American, African and Asian countries (except Japan and Indonesia) as “non-Western” (CBS 2000). The inconsistency of classifying some former colonies of the Netherlands, such as Suriname and the Antilles, as non-Western, and others, such as Indonesia, as Western, combined with the ways in which the word allochthone is used in everyday language make the categorization system can be understood as race-ethnically determined (Yanow and Van der Haar 2013). This part of the Dutch governmental knowledge infrastructure was taken over in the CAS system unchanged.

What is striking in CAS 2015 is not only the presence of an indicator about the number of non-Western allochthones, but also the absence of an indicator for Western allochthones or autochthones. In CAS 2017 the indicator regarding non-Western Allochthones was removed, because "it didn't add predictive value" (Willems 2017a). This remark could be understood in at least two ways: 1) race-ethnic origins are not suitable as a predictor for crime, or 2) the race-ethnic origins of possible suspects are still implicitly woven in the system through the other indicators, making it unnecessary to have it in there explicitly. An argument could be made for both statements. A 2014 study on criminal behavior of youth aged 12-18 in the Dutch city of Rotterdam could not find a relationship between race-ethnic origins and criminal behavior (see Driessen et al. 2014). In addition, in their research on NEET-youth (UK youth aged 16-24, who are Not in Employment, Education or Training), Helen Thornham and Edgar Gómez Cruz (2018) found that, although this category is officially not gendered, the lived reality proved that mostly women fell into this classification. Similarly, CAS 2017 could, while no longer explicitly take into account the race-ethnic origins of people living in a neighborhood, implicitly, through the other demographic indicators, and location based approach, be race-ethnically biased. In addition to the question of effectivity, Oscar H. Gandy (2016, p. 62) suggests that some indicators, such as race or ethnicity, could best not be used in predictive systems, not because they are bad predictors, but because it would be unethical to use non-causal factors.

4.1 Known Offenders

The final three indicators come from BVI and refer to data about previously convicted criminals or so called "known offenders". For each square in CAS, first the distance to the closest known offender is calculated (see Table 2, indicator 1). The remaining two indicators refer to the number of recently active (within the past six months), known offenders that live within 500 and a 1000 meters respectively of a square.

Table 2. BVI indicators used for each CAS square.

Nr.	Explanation
1	Distance in km of the address of the closest known offender (suspect) of an incident who has been active in the past 6 months
2	Number of known offenders (suspects) of an incident who have been active in the past 6 months that live within 500 meters
3	Number of known offenders (suspects) of an incident who have been active in the past six months that live within 1000 meters

Interestingly enough in CAS 2017 the term "known offenders" has been replaced with the term "suspects". Although the label still refers to exactly the same category, one could argue that this is indicative of the mentality of predictive policing in which someone is a suspect before a (new) crime has happened. One could even question whether or not the use of these indicators is in conflict with the principle of letting someone start over with a clean slate after doing time. In addition, legally speaking, the word "suspect" is reserved for people of who "a reasonable suspicion of guilt to a criminal offense is assumed through facts and circumstances" (CBS 2018, translation by the authors). This classification is therefore not valid before a crime has happened.

4.2 Continuities, police culture and feedback loops

The indicators used to predict crime are not new discoveries, but rather datafied continuities of the indicators that police officers use in their work on the street. In his extensive ethnographical fieldwork, cultural anthropologist Sinan Çankaya (2012; 2015) joined police officers on patrol and observed their decision making process for several months. He found out that "street level bureaucrats" use three physical categorizations to determine whether or not to stop and interrogate someone (see Fig. 3). The characteristics of what Çankaya calls the "biological body" and the "disobedient body" can be seen as real life parallels of the demographic information of CBS and the "known offenders" data from the BVI. Police officers match this information with the time and location to determine if there is a mismatch and a need for action (Çankaya 2012 p. 74-76). In addition, Çankaya found that race-ethnicity plays an important role both in decision making during patrols (Çankaya 2012, p. 46-51; Çankaya 2015) and in the self-identification of the Police force as (white) Dutch (Gowricharn and Çankaya 2017).

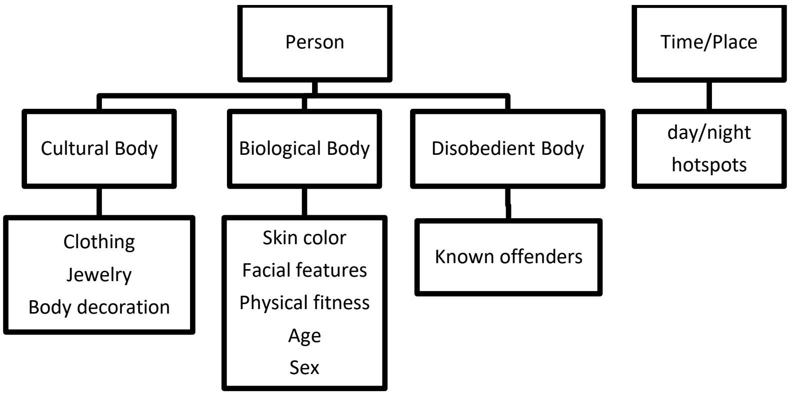

Fig. 3. Part of decision making process of Dutch police officers (Çankaya 2012, p. 78)

Because "street level bureaucrats" and CAS are largely aligned in their decision making processes, the risk for a feedback loop emerges. Subconscious biases regarding less economically advantaged people and intended and unintended race-ethnic profiling can result in biased data concerning certain groups in society. This biased data is then used to determine where to patrol etc. This process could be seen as a form of datafied "cumulative disadvantage" (Gandy 2016). Cumulative disadvantage creates and reinforces differences is the quality of life of different groups. Discrimination in this sense automatically creates inequality which builds up on itself, hence *cumulative* disadvantage (see Gandy 2016, p. 55).

5 Visualization and Framing

CAS visualizes results in a standardized way for all police departments throughout The Netherlands that use this system. This is done in two ways. Firstly, CAS provides a simple line-graph where the x-ax shows time (of a single day, or week) and the y-ax shows the likelihood of a high-impact crime happening (Willems 2017a). Secondly, the system presents its results in a grid map. Both ways of visualizing results can be questioned critically because of the data that they use and the way results are simplified.

The graph that shows time and risk is very easy to read – which is a good thing. However, the problem with this type of visualization is not the visualization itself, but the data behind it. Because of the aforementioned problem to determine the time and date of a burglary, the police use an average time in their report. Say, one leaves their home at 8 A.M. and comes home at 6 P.M., the report will state that the crime happened at 1 P.M. (Willems 2017c). Eventually, the graph in CAS will maybe not show a reliable prediction for when the risk is high for a crime as burglary, but more when the risk is higher of 'people being halfway through their work day'. Because CAS is a closed system information officers using CAS do not have the option to look up the data behind this graph.

The second visualization within CAS is a grid map, which is often used in other predictive policing systems. As stated before, the squares that are colored red represent the areas with the highest risk for a certain crime to take place, orange squares have a medium risk, and yellow squares have a low, but mentionable risk.

The only specifics or filter a user of CAS can add within the grip-map visualization, is selecting a type of crime. This map will then only show the predicted amount of risk for that selected type of crime. A user cannot, for example, select a square or a crime and receive more in-depth information on why the risk is predicted as high or low. In other words, the user cannot see if a risk is predicted on the basis of a stable variable or because of a more changeable or unstable variable as the presence of a known-offender. Especially in the early stages of CAS, often the same three percent of squares would light up, which gave the impression that some undeterminable factor was causing the risk to always be high for certain areas.

5.1 The role of the data officer

The role of the data officer is to use the visualizations in CAS as a starting point, and from there to try to explore and retrieve the background of the risks in certain areas. Data officers are able to use other databases or sources from within police departments. At the end of their search, they come up with a possible explanation for the certain amount of risks. In a way, they enrich the results presented by CAS, because they do not only use the results from CAS, or information from the BVI, but also use their contextual knowledge on the area and their expertise on crime to look for possible causational effects. In addition, they can add unpredictable events that are not part of CAS, such as football matches of the cities home team or music festivals, and determine by themselves how much this influences the prediction.

Of course one should add that the function of visualizing a large amount of data is to make information more accessible and readable for people. This is also the function of the graph and the grid map within CAS. One could not process and calculate all the variables and data as quickly as CAS. However, as often is said about visualizing data, it does mean that information is simplified and some elements are left out. Trying to understand something as complex as crime, what is influenced by many factors, by 'reading' a simple visualization could be challenging and is prone to misinterpretations. Nonetheless, Dick Willems argues that by giving data officers so little extra information with the visualization and specifics of the predictions made by CAS, stimulates data officers to research the risk areas in a more open-minded way, because they have to rely on their own knowledge instead of being closely guided by the results of CAS (Willems 2017a).

The visualization of results within CAS is kept simple and the numbers of options given to users to explore these visualizations are small. One could conclude, through this brief analysis of the interface and visualizations used in CAS, that the Dutch police have consciously chosen to keep this system closed for users. This changes the role of the user from someone that merely looks at visualizations and reproduces these results, to someone that enriches these results with a kind of qualitative explanation. However, with this extra layer of interpretations by the data officers, it can be said that there is another moment where possible (personal) biases can come into the equation. As discussed before, the production of data, such as mainly arresting people from particular groups, is biased. In the case of CAS, after the data is processed through the algorithms, humans can add and contextualize these results through their own frame of reference. However, there is always the risk of skipping the time-consuming process of data enrichment, and the temptation of simply taking over the predictions of the system that keeps their users in the dark on purpose. Given the current discussions on issues of work pressure and the lack of personnel within the Dutch National Police, this scenario is not unlikely.

5.2 Epistemology or ontology?

As we have shown the CAS epistemology is deeply rooted in the existing (Dutch) culture of the Dutch police. As it starts to increasingly determine which places and people are under surveillance it could be approached as ontology as well. We understand ontology as " a social construction of reality, defined in the context of a specific epistemic culture as sets of norms, symbols, human interactions, and processes that collectively facilitate the transformation of data into knowledge" (Kuiler 2014, p. 312). Through this definition it becomes clear that CAS is not a neutral instrument, but a particular social construction of reality, shaped by what is seen by the Dutch police as deviant physical traits, economic situations and behavior. CAS does not merely represent these characteristics in a certain way, but it actively shapes the living environment of people that are considered not to be part of the norms set in Dutch society through knowledge infrastructures.

6 Conclusions

We have shown that a critical data studies approach can give us much needed insights into systems that are shaping our life world. With this perspective we have shown that the used data categories in CAS are not always voluntary choices of data categories, but are often determined by availability, not only in terms of what is measured, but also how it is measured. The CBS data is shaped by Dutch history, culture and politics regarding what to measure in terms of social benefits, household composition and race-ethnic origins of people. BVI data is a result of the existing Dutch culture of policing which already has ideas about what characteristics are prevalent in potential suspects and how to record crime (in terms of time and location). By using GBA data the maps of CAS are tied to existing formats such as streets and house numbers. Because of these predispositions, CAS is in no way exhaustive or all-inclusive. It is tied to types of crimes that are location and time based, and happen often enough to be able to make statistically valid predictions about.

By highlighting the data-visualization in the form of a map and the possibilities and impossibilities for the users of CAS (information specialists), we have shown that the data cannot speak for itself; it needs the interpretation of a data officer. However, the possibilities for the data officer to find out the origins for a prediction are severely limited. The aim to create uniform predictions creates the possibility for a less accurate but time saving approach at the end of the interpreter by only consulting CAS.

Another valid concern that needs much more research is the effects of the combination of possible system biases with human biases. Possible race-ethnic biases of CAS could be magnified when combined with police practice. For future research critical data studies approaches could be combined with ethnographic studies to get more insights in work processes and possible feedback loops.

Acknowledgements

We would like to thank Dick Willems, data scientist at the Amsterdam police department, for taking the time to answer our questions and providing us with a 2017 version of the CAS indicators list.

References

Bowker, G.C., Star, S.L.: Sorting things out: Classification and its consequences. MIT press, Cambridge, MA (1999)

Çankaya, S.: De controle van marsmannetjes en ander schorriemorrie. Het beslissingsprocess tijdens proactief politiewerk. Boom Lemma, Amsterdam (2012)

Çankaya, S.: De politiële surveillance van ras en etniciteit. In: Moor, L.G., Janssen, J., Easton, M., and Verhage, A. (eds.) Ethnic profiling en interne diversiteit bij de politie. pp. 13–33. Maklu-Uitgevers, Antwerpen (2015)

CBS: Allochtonen in Nederland. CBS, Voorburg (2000)

CBS: Begrippen, http://www.cbs.nl/nl-nl/onze-diensten/methoden/begrippen
Dalton, C.M., Taylor, L., Thatcher, J.: Critical Data Studies: A dialog on data and space. Big Data Soc. 3, 2053951716648346 (2016). doi:10.1177/2053951716648346
Dalton, C.M., Thatcher, J.: What does a critical data studies look like and why do we care?, http://societyandspace.org/2014/05/12/what-does-a-critical-data-studies-look-like-and-why-do-we-care-craig-dalton-and-jim-thatcher/, (2014)
Driessen, F.M.H.M., Duursma, F., Broekhuizen, J.: De ontwikkeling van de criminaliteit va n Rotterdamse autochtone en allochtone jongeren van 12 to 18 jaar: De rol van achterstanden, ouders, normen en vrienden. Politie & Wetenschap and Bureau Driessen, Apeldoorn and Utrecht (2014)
Gandy, O.H.: Coming to Terms with Chance: Engaging Rational Discrimination and Cumulative Disadvantage. Routledge, London (2016)
Gowricharn, R., Çankaya, S.: Policing the Nation: Acculturation and Street-Level Bureaucrats in Professional Life. Sociology. 51, 1101–1117 (2017). doi:10.1177/0038038515601781
Iliadis, A., Russo, F.: Critical Data Studies: An Introduction. Big Data Soc. 3, 205395171667423 (2016). doi:10.1177/2053951716674238
Kitchin, R.: The Data Revolution: Big Data, Open Data, Data Infrastructures and their Consequences. Sage, London (2014)
Kitchin, R., Lauriault, T.: Towards Critical Data Studies: Charting and Unpacking Data Assemblages and Their Work. Social Science Research Network, Rochester, NY (2014)
Kop, N., Klerks, P.: Doctrine Intelligencegestuurd Politiewerk. Politieacademie, Apeldoorn (2009)
Kuiler, E.W.: From Big Data to Knowledge: An Ontological Approach to Big Data Analytics. Rev. Policy Res. 31, 311–318 (2014). doi:10.1111/ropr.12077
Mali, B., Bronkhorst-Giesen, C., Den Hengst, M.: Predictive Policing, Lessen voor de toekomst. Politieacademie, Apeldoorn (2017)
NOS: Politie gaat misdaad voorspellen met nieuw systeem, https://nos.nl/artikel/2173288-politie-gaat-misdaad-voorspellen-met-nieuw-systeem.html
Schäfer, M.T., van Es, K.: New Brave World. In: Schäfer, M.T. and van Es, K. (eds.) The Datafied Society: Studying Culture Through Data. pp. 13–22. Amsterdam University Press, Amsterdam (2017)
Strategische Beleidsgroep Intelligence: Waakzaam tussen Wijk en Wereld: Nationaal Intelligence Model Sturen op en met Informatie. Politieacademie (2008)
Thornham, H., Gómez Cruz, E.: Not just a number? NEETs, data and datalogical systems. Inf. Commun. Soc. 21, 306–321 (2018). doi:10.1080/1369118X.2017.1279204
Willems, D.: Interview, (2017)(a)
Willems, D.: Email Variabelen CAS, (2017)(b)
Willems, D.: Predicting criminal events with Data Science. Data Science Utrecht Meet Up. , Utrecht (2017)(c)
Willems, D., Doeleman, R.: Predictive Policing: Wens of Werkelijkheid. Tijdschr. Voor Politie. 4, (2014)
Yanow, D., Van der Haar, M.: People out of place: allochthony and autochthony in the Netherlands' identity discourse—metaphors and categories in action. J. Int. Relat. Dev. 16, 227–261 (2013)