Predictive policing: The risks associated with risk assessment

Litska Strikwerda
Faculty of Law, Open University, Utrecht, the Netherlands

Abstract
Focusing on the Dutch tools SyRI and CAS, this paper describes predictive policing against the background of the broader development toward a pre-crime society, the accompanying culture of control and the new penal logic it gives rise to. It will explain the risks associated with the risk assessments predictive policing tools provide and end with the recommendation to use predictive policing not only for police deployment, but also to target problem-oriented responses to crime to the right persons and places.

Keywords
Predictive policing, pre-crime society, transparency

Introduction
This paper describes the phenomenon of predictive policing, focusing on two Dutch predictive policing tools that recently generated a lot of public attention, namely: SyRI and CAS. The Dutch government stopped using SyRI, a predictive policing tool for predicting fraudsters, because the District Court of The Hague ruled it violates the right to privacy as contained in article 8 of the European Convention on Human Rights (Rb. Den Haag, 05-02-2020, ECLI: NL: RBDHA:2020:865, available at http://www.rechtspraak.nl). The use of CAS, a predictive policing tool for predicting crime locations, just increased as it has been implemented nationwide since 2019 (Halfjaarbericht politie 2019, bijlage 5: 3), making the Netherlands the first country in the world to deploy predictive policing on a national scale.

It will be argued predictive policing is part of the broader development toward a pre-crime society (Zedner, 2007), the accompanying culture of control (Garland, 2001) and
the new penal logic (Feeley and Simon, 1992) it gives rise to and that the ethical issues surrounding this phenomenon can be better understood against this background. Predictive policing aims to prevent crime by providing risk assessments, but these risk assessments have risks of their own. Since the risk analyses often lack transparency and explainability, it is not possible to weigh the crime risks to be prevented and the risks of crime prevention properly, which may lead to disproportionate intrusions with the right to privacy and violation of the related rights to equal treatment in equal cases and to protection against discrimination, stereotyping and stigmatization.

Proponents claim that predictive policing tools for predicting offenders and crime locations actually restore instead of erode the rights of persons thought to be likely to commit a crime or living in a crime-prone neighborhood, because they provide neutral, quantitative evidence that a person is indeed likely to commit a crime or that an area is indeed a high-crime area. But the evidence predictive policing tools for predicting offenders or crime locations provide, is not entirely neutral, because it is based on assumptions that do not need to be true. Moreover, it is important to realize that numbers do not explain reasons. In the longrun, the causes of crime are an important key to crime prevention. Therefore, this paper will come to the conclusion that predictive policing should also be used to direct interventions to locations and persons where they are most needed if we really want to reduce crime.

Predictive policing: Definition, scope and effects

According to Perry et al., who have written a baseline document on this topic, ‘predictive policing is the application of analytical techniques—particularly quantitative techniques—to identify likely targets for police intervention and prevent crime (…) by making statistical predictions’ (Perry et al., 2013: 1–2). This definition can be clarified as follows. From a data science perspective, predictive policing is a predictive model, which can be described as ‘a formula for estimating the unknown value of interest’ (Provost and Fawcett, 2013: 45). The formula can be mathematical, but it can also be a logical statement, like a rule, or a hybrid of the two. The unknown value to be predicted can be something in the future, but it can also be something in the present or in the past (Provost and Fawcett, 2013: 45). In case of predictive policing, the unknown value to be predicted is a crime location or offender. For the sake of completeness, it should be added that the techniques used in predictive policing can also have other objectives, namely: the prediction of perpetrator’s identities (when likely offenders are matched with specific past crimes) or the prediction of victims (Perry et al., 2013: 8–9). However, these are beyond the scope of this paper.

Recently, the Dutch SyRI (short for: Systeem Risico Indicatie, which translates as System Risk Indication) generated a lot of attention, because the district court of The Hague ruled it violates the right to privacy as contained in article 8 of the European Convention on Human Rights (Rb. Den Haag, 05-02-2020, ECLI: NL: RBDHA:2020:865, available at http://www.rechtspraak.nl). SyRI can be considered a predictive policing tool for predicting offenders under the above-mentioned definition. It was used by the government to prevent and combat fraud in the areas of social security and income-related schemes, tax and social security contributions and labor laws.
It inputs, among other things, employment data, civic integration data, debt data, health insurance data (whether or not someone has a health insurance) and personal data (name, address, date of birth etc.). These data are encrypted and tested against a risk model with several indicators, which generates potential hits: natural or legal persons who are at risk of committing fraud. The data on potential hits are decrypted and handed over to an analysis unit, which can forward them to the police or prosecutor’s office if they warrant criminal investigation (Rb. Den Haag, 05-02-2020, ECLI: NL: RBDHA:2020:865, available at http://www.rechtspraak.nl).

Nor the indicators, neither the risk model used by SyRI are known. Therefore, it cannot be described how it works. To indicate how a predictive policing tool for predicting offenders can work, another tool, namely: the Custom Notification program, is briefly discussed here. The Custom Notification program is used on a citywide basis by the Chicago Police Department since July 2013 and has been developed by the Chicago Police Department itself in collaboration with the Illinois Institute of Technology. It identifies potential victims and perpetrators associated with the continuum of gun violence in the city of Chicago. Based on empirical data it creates a Strategic Subjects List (SSL): a rank-order list of individuals that have an increased likelihood of engagement in violent criminal activity. These data include data on demographics, arrest history and social network variables. They are analyzed by means of a prediction model that uses co-arrests to previous homicide victims to predict the likelihood that a person becomes a homicide perpetrator or victim (Saunders et al., 2016: 354–357). For each individual on the list a letter is created which incorporates factors known about him or her. This letter is delivered and explained by police officers. The Custom Notifications program serves as notice that law enforcement action will be targeted specifically to the individual and continuation to participate in gun violence will have cognizable penalties (Chicago Police Department 2015).

Besides the above-mentioned predictive policing tool for predicting offenders SyRI, there is also a predictive policing tool for predicting crime locations in use in the Netherlands, namely: CAS (short for: Criminaliteits Anticipatie Systeem, which translates as Crime Anticipation System). CAS has generated a lot of public interest as well since it has been implemented nationwide in 2019 (Halfjaarbericht politie 2019, bijlage 5: 3). Although predictive policing tools are used by several police forces, mainly in the US, but also in Europe, the Netherlands were the first country in the world to deploy predictive policing on a national scale. CAS predicts future crime hotspots on a so-called ‘heat map’ which is divided in squares of 125 by 125 meter. Each square is assigned a risk score for the next 2 weeks. Squares with a high risk score will color red. Moreover, CAS does not only indicate where the crime risk is high, but also when. These ‘hot times’, as I will call them, are presented in a line chart. To come to these predictions CAS makes use of a large amount of data in a data warehouse, including crime rates, but for example also the distance from the location to the nearest highway (Mali et al., 2017: 93–94). That is probably because research reveals that features of the urban environment affecting the accessibility of places shape patterns of offending (Johnson et al., 2009).

For the rest it is not clear how CAS works exactly, but it is known it makes use of a near-repeat concept (Drenth and Van Steden, 2017: 6). The near-repeat concept is based on the empirical phenomenon that there is an increased risk of crime occurring within a
certain geographical and time window once a crime has taken place (Rummens et al., 2017: 264–266). The best predictor of victimization is prior victimization. ‘Repeat victimization’, as it is called, tends to occur quickly after the initial crime (Johnson and Bowers, 2004b). Victims are at an elevated risk of crime in the months directly following an event (Johnson et al., 2007). In case of burglary, not only the burgled home is at risk of being burgled soon again, but also other, similar properties in the neighborhood (Johnson and Bowers, 2004b). There is evidence that specifically burglary is geographically concentrated, as it clusters in space and time more than one would expect if the perceived ‘patterns’ of crime would simply result from the attractiveness of places to offenders (Johnson et al., 2007, 2009). This can be explained as follows. Once a perpetrator has committed a crime, it is easier to repeat that crime than to identify a new location and/or criminal act. In case of burglary, especially in wealthy housing areas where people have the means to do so, counter-measures are often taken against revictimization. That makes neighboring properties easier to invade than the original burglarized house (Johnson et al., 2007: 204). Properties within 400 meters of a burgled house are at a significantly elevated risk of burglary for up to two months after the initial event (Bowers et al., 2004). The location of clusters of burglary is not predictable over periods of 3 or more months, however (Johnson and Bowers, 2004a). There are a number of plausible explanations for this: the characteristics of an area can change over time (e.g. because of interventions from those responsible for crime reduction) and the perpetrator’s memory for the features of a particular house and those nearby may decay (Johnson et al., 2007: 215).

An example of another predictive policing tool for predicting crime locations that makes use of the near-repeat concept is PredPol. Predpol was launched in 2013 and, like CAS, it predicts crime hotspots and hot times which are presented on a map. Predpol makes use of a prediction model called ‘epidemic-type aftershock sequence (ETAS)’, which is analogous to models used to predict seismic activity (Mohler et al., 2016: 1400). Not only models from seismology are used to predict crime locations and times, but also statistical techniques that were originally developed to study the transmission of disease (Johnson and Bowers, 2004b).

Prior victimization has a greater predictive power than any other variable (Johnson and Bowers, 2004b). A study of burglary patterns in two different areas in five separate countries reveals that ‘near repeat’ has also good predictive power (Johnson et al., 2007: 215). But the effects of Predpol and CAS on crime reduction (i.e. the outcome of the police response employed) are disappointing. Mohler et al performed two randomized controlled trials of Predpol; one in the Los Angeles Police Department and one in the Kent Police Department. They found an average crime reduction of 7.4% (Mohler et al., 2016: 1409). The results of an independent evaluation by Kent Police (UK) are less positive. After a 4-month trial period a 4% reduction in all crime was observed. This was not sustained and after a 15-month trial period no further overall reductions were observed. However, police officers reported that during the trial they often did not have time to use Predpol (Kent Police 2013). A recent evaluation of CAS by the Dutch Police Academy shows that the number of burglaries in Amsterdam indeed reduced, but the researchers did not find a correlation with the predictions made by CAS (Mali et al., 2017: 98).

Little is known about the effectiveness of predictive policing tools for predicting offenders. There are, to my knowledge, no data available about SyRI. In a trial Saunders,
Hunt & Hollywood found that individuals on the SSL produced by the Custom Notification program were not more (or less) likely to become a victim of a shooting than the comparison group, but that they were more likely to be arrested for one. A possible explanation is that the list was used as an intelligence-gathering source: when a shooting happened, the police looked at the SSL for possible suspects (Saunders et al., 2016: 365–367).

**The broader development: The shift from harm to risk**

Crime forecasting is not a new phenomenon; it has been around for decades. Scientists have used statistical and geospatial analyses to determine crime risk levels ever since the sociologist Shaw and the criminologist McKay studied the persistence of juvenile crime in specific neighborhoods of Chicago, and later also in 20 other American cities, and found juvenile delinquency to be highly correlated with, among other things, changes in population, inadequate housing, poverty, tuberculosis and mental disorders (Shaw and McKay, 1942). From the mid-1980s onward risk factor prevention became a dominant paradigm in crime control (Mehozay and Fisher, 2019: 524). In the 1990s algorithms for crime forecasting were computerized. In recent years, these algorithms have become more sophisticated and, due to an increase in computer power and storage, bigger data sets can be analyzed (Mehozay and Fisher, 2019: 524; Perry et al., 2013: 3–4). Meanwhile, the amount of data available has grown exponentially and still does: it doubles in volume every 2 years. Moreover, the investigatory utility of the data available improves because they are networked: law enforcement agencies and private companies connect their databases and, thereby, aggregate their data (Ferguson, 2015: 354, 360). These developments have led to predictive policing as defined and described in the last section.

Predictive policing does not stand on its own but is part of a broader development and the ethical issues surrounding this phenomenon can be better understood against this background. According to the criminologist Zedner we are on the cusp of a shift from a post-crime society, within which the dominant ordering practices arise post hoc, to a pre-crime society, within which ordering practices are pre-emptive. Where post hoc ordering practices respond to wrongs done, pre-emptive ordering practices shift ‘the temporal perspective to anticipate and forestall that which has not yet occurred and may never do so’ (Zedner, 2007: 262). By definition, crime prevention is pre-emptive action. Especially the impact of 9/11 has pressed governments to think and act pre-emptively (Zedner, 2007: 264). It is, therefore, not surprising that most examples of pre-emptive ordering practices are found in the area of post 9/11 anti-terrorism legislation. Antiterrorism laws generally broaden criminal liability because they criminalize the preliminary stage of terrorist acts, before they become harmful (Borgers and Van Sliedrecht, 2009: 175). Consider, for instance, laws that prohibit the collection or possession of information that could be useful to commit or prepare an act of terrorism. Predictive policing can also be seen as a pre-emptive ordering practice, for it aims to prevent crime by predicting where and when or by whom it will be committed.

The shift from post hoc ordering practices to pre-emptive ordering practices goes hand in hand with a shift from a perception of crime as harm to a perception of crime as risk (Zedner, 2007: 262). It is a more or less established fact that contemporary society
can be characterized as a ‘risk society’ (Borgers and Van Sliedregt, 2009: 172). The concept of the risk society originally derives from the sociologist Beck (Beck, 1986). In broad terms, it entails that society is focused on the control and management of risks. According to the criminologist Garland this has led to a ‘culture of control’, within which criminal law is seen as an instrument to control the risk of crime (Garland, 2001).

The view of crime as a risk to be calculated has also led to a ‘new penal logic’ (Feeley and Simon, 1992), inspired by the theory of selective incapacitation, which seeks not to punish criminals for what they have done in the past, but to prevent them from doing it again in the future (Mehozay and Fisher, 2019: 531). A Dutch example of this new penal logic is the Reoffender Institutionalization Measure (Inrichting voor Stelselmatige Daders)², which entered into force in 2004. It incorporated the Penal Detention of Addicts Measure (Strafrechtelijke Opvang Verslaafden), which had entered into force a few years earlier and was aimed at drug-dependent offenders. The scope of the former is broader than the latter: it aims to prevent ‘persistent offenders’ (people who have committed over three crimes in a period of 5 years, drug-dependent or not) from reoffending by creating the possibility to imprison them for a period of up to 2 years. In this way the Reoffender Institutionalization Measure converts incarceration from the ultimate repressive instrument into an instrument of prevention, which is also pre-emptive action by definition (Moerings, 2016: 65).

From a criminological point of view, the Reoffender Institutionalization Measure can be seen as part of the managerial movement, within which risk assessments are conducted on an individual basis and based on clinical judgment by professionals. Predictive policing is part of a new phase in this movement, the actuarial phase, which can be described as an evolution toward evidence-based practices and mathematical tools to assess risk. The actuarial phase, again, represents a new penology as it focuses less on things like responsibility, guilt, intervention and rehabilitation, but instead focuses on setting out techniques for the identification, classification and management of groups according to risk levels. Proponents of these techniques argue that they introduce a new level of accuracy and may even eliminate forms of bias that were inherent in previous methods of risk assessment. The motivation for accuracy and bias-free analysis has led to more and more sophisticated analytical tools that incorporate big data and machine learning algorithms. The downside of this development is that these algorithms are often ‘black box’ and it is impossible to explain how the risk score was established (Mehozay and Fisher, 2019: 524, 531–533), as is the case for the two Dutch predictive policing tools discussed in the last section: SyRI and CAS.

This is problematic because it goes against an important moral value, namely: the value of transparency, which is among the key requirements for Trustworthy Artificial Intelligence (Ethics Guidelines for Trustworthy AI, Chapter II). The value of transparency is usually not seen as an end in itself, but as an important prerequisite for the realization of other values, such as the value of privacy. As AI and other computer-related technology is often ‘morally opaque’, people do not recognize that the practice raises ethical questions and because they do not know what moral values are at stake, they cannot act to protect them (Brey, 2010: 51). Therefore, designers should be encouraged to make their technology ‘morally transparent’; they should make understandable what moral values are at stake in relation to it and avoid advises or decisions that cannot
be explained to end users (Brey, 2010: 51). As a matter of fact, the Ethics Guidelines for Trustworthy AI claim that transparency merely concerns ‘explainability’ (§ 1.4).

The lack of transparency and hence explainability of the risk calculations provided by predictive policing tools is even more problematic because of the nature of the decisions that can be based on them, such as the deployment of investigatory powers. According to the 19th century philosopher John Stuart Mill ‘the only purpose for which power can be rightfully exercised over any member of a civilized community, against his will, is to prevent harm to others’ (Mill, 1865: 6). Post hoc ordering practices indeed find legitimation in this harm principle, but not (pre-emptive ordering practices based on the outcome of) predictive policing. They can instead be legitimated on the basis of the precautionary principle, which originates, and that is in light of Beck’s *Risk Society* not surprising, in environmental and public health law. There is no generally acknowledged definition of the precautionary principle. But a common definition is provided by (a forerunner of) the European Commission, according to which the precautionary principle entails that authorities may take action when scientific and objective research indicates that a phenomenon may have a dangerous effect but cannot determine the risk with sufficient certainty (Commission of the European Communities Communication from the Commission of 2 February 2000 on the precautionary principle).

The desirability of the above-mentioned shift from the harm principle to the precautionary principle is under discussion. The influential American legal scholar Cass Sunstein, who focuses specifically on the precautionary principle as a moral ground for the criminalization of certain behaviors, claims that it ‘becomes operational if and only if those who apply it wear blinders-only, that is, if they focus on some aspects of the regulatory situation but downplay or disregard others’ (Sunstein, 2003: 26). Risks that tend to be downplayed or disregarded are risks that come with regulation itself. Sunstein provides the following example. It is easy to see that arsenic is potentially dangerous; it is well-known as a poison. But there is also a risk associated with arsenic regulation; for it might lead people to use even less safe alternatives (Sunstein, 2003: 32). This argument actually applies to criminal legislation based on the harm principle as well. But the problem with the precautionary principle is that it does not give us the opportunity to weigh the risks of the regulatory situation and the risks of regulation properly (Pieterman, 2008: 184). As will be explained in the next section, the same goes for police action taken in response to a prediction by a predictive policing tool, especially when the prediction is based on a risk analysis that is not transparent nor explainable because the algorithm employed is a ‘black box’.

**The risk associated with predictive policing**

In the SyRI case, the District Court of The Hague compared the objectives of the legislation regulating the use of SyRI, i.e. to prevent and combat fraud in the interest of economic welfare, with the intrusion into private life that it makes. According to the court, the legislation does not meet the fair balance required by the ECHR in order to be able to speak of a sufficiently justified intrusion into private life, because it is insufficiently transparent and verifiable with regard to the use of SyRI. The legislation was therefore declared unlawful and non-binding (Rb. Den Haag, 05-02-2020, ECLI: NL:
RBDHA: 2020:865, available at http://www.rechtspraak.nl). This verdict can be explained as follows.

In European countries the right to privacy is protected by article 8 of the European Convention on Human Rights. It reads as follows: ‘everyone has the right to respect for his private and family life, his home and his correspondence. There shall be no interference by a public authority with the exercise of this right except such as is in accordance with the law and is necessary in a democratic society in the interests of national security, public safety or the economic well-being of the country, for the prevention of disorder or crime (…)’. The court considers that new technologies, including digital file linking and algorithmic analysis capabilities, increase the possibilities for public authorities to exchange data among themselves as part of their legal duty to prevent and combat fraud. The court shares the State’s view that these possibilities should be exploited. It is of the opinion that the SyRI legislation is in the interest of economic welfare and therefore serves a legitimate purpose (§ 6.4). However, as the risk indicators and risk model SyRI deploys are unknown, it is not possible to establish whether or not the interference with the right to privacy is necessary in a democratic society (proportional) in relation to the legitimate aim, the economic wellbeing of the country, served (§ 6.94). As was established in the last section, it is important that this consideration can be made, because of the interests that are at stake. The right to privacy in the case of data protection touches upon the right to equal treatment in equal cases and the right to protection against discrimination, stereotyping and stigmatization (§ 6.24). Given the large amounts of data eligible for processing in SyRI, including personal data, and the fact that risk profiles are used, there is a risk that inadvertent links are established with the use of SyRI on the basis of bias, such as a lower socio-economic status or an immigration background (§ 6.93).

The latter also goes for other predictive policing tools for predicting offenders. An evaluation of the Chicago Custom Notification program discussed before shows that there was ‘a statistically significant increase in police contacts’ with persons on the rank-order list it provides of individuals that have an increased likelihood of engagement in violent criminal activity (Saunders et al., 2016: 367). As it is only used in Chicago, the ECHR is not applicable, but in the US, the fourth amendment of the U.S. Constitution protects people’s right to privacy from arbitrary governmental intrusions. An important common legal restriction to interferences with the right to privacy by police officers, think for example of investigative detentions, searches and seizures, is that they always require reasonable suspicion of involvement in a crime. When a predictive policing tool for predicting offenders directs police officers to a particular person, that does not create a reasonable suspicion in itself; police officers need to relate the information it provides to actions they observe (Ferguson, 2015: 388). But in light of the fact that the predictive policing tool identifies the person concerned as likely to commit a crime, police officers may see almost every act he or she performs as suspicious. As a result, they will constantly think they have reason to interfere with his or her right to privacy by means of investigative detentions, searches, seizures etc.

Similar concerns rise with regard to predictive policing tools for predicting crime locations. When for example CAS or PredPol directs police officers toward a place where crime is likely to occur at a time it is likely to occur, they cannot reasonably
assume that all persons there present are involved in a crime. Based on the circumstances, the police officers will have to determine which persons, if any, warrant further investigation. However, the fact that a predictive policing tool for predicting crime locations has identified the place as a place where crime is likely to occur, may influence their view on the situation. Imagine, for example, that a predictive policing tool for predicting crime locations has directed police officers to a certain area where property crime is likely to occur. There present they see a person carrying a duffel bag. That activity is by itself not obviously suspicious but in light of the fact that the predictive policing tool has identified the area as a place where property crime is likely to occur, the police officers find it suspicious and stop the person in order to search his bag for burglar’s tools or stolen property (Joh, 2014: 55–59). Then it turns out that the person, who has no criminal record, was on his way from his home to the laundromat and the duffel bag contains laundry. Had he lived in another area, police officers would probably not have thought they had reason to interfere with his right to privacy.

Some legal scholars belong to the proponents of predictive policing that claim that predictive policing tools for predicting offenders and crime locations actually restore instead of erode the right to privacy of persons thought to be likely to commit a crime or people living in high-crime neighborhoods (see e.g. Koss, 2015: 305). They argue as follows. Reasonable suspicion is, at its core, ‘a doctrine of predictive suspicion’ (Ferguson, 2015: 391). Police officers have always tried to identify crime-prone individuals and crime hotspots. Before predictive policing tools for predicting offenders and crime locations were there, they placed pushpins on paper maps to reveal clusters of criminal activity or just used their experience and intuition (Koss, 2015: 302). Unfortunately, these judgments include all kinds of biases. Race, class, choice of clothing, gender and age all factor into police officers’ discretionary decisions (Ferguson, 2015: 389). The police are often blamed for targeting individuals with particular types of characteristics and areas that have historically been defined as high-crime disproportionately represent low-income and minority neighborhoods (Koss, 2015: 304). Replacing those generalized intuitions with precise detail about actual people should result in a more accurate policing strategy (Ferguson, 2015: 389–390). A strong positive argument for the use of predictive policing tools for predicting offenders or crime locations is thus that they, in contrast with experience or intuition, provide neutral, quantitative evidence that an area is indeed a high-crime area or that a person is indeed likely to commit a crime (Koss, 2015: 305).

Other legal scholars contest this claim. The above-mentioned evidence predictive policing tools for predicting offenders or crime locations provide, is not entirely neutral. That is because the development of these tools necessarily involves human discretion (Joh, 2014: 58). Not all predictive policing tools for predicting offenders or crime locations are the same. They make use of different algorithms, requiring the input of different (amounts of) data and resulting in different statistical predictions of the likeliness that that a certain person will commit a crime or crime will occur at a certain place. Moreover, the developers of a particular predictive policing tool have selected the appropriate algorithm with a certain model or theory of crime prediction in mind (Bennett Moses and Chan, 2016: 6). As mentioned before, the models used in predictive policing have not been specifically developed for crime prediction, but originate from
seismology and epidemiology (Johnson and Bowers, 2004b; Mohler et al., 2016: 1400). But several theories of crime prediction exist. Well-known theories for predicting offenders are the strain theories, which assume people can be pressured into crime. Agnew recently made an attempt to develop a more general strain theory. According to Agnew the most important strains are if people are treated negatively by others, lose something that is valuable to them or cannot achieve their goals. These strains evoke negative emotions which in turn encourage delinquent behavior (Agnew, 2006: 19; Kolthoff, 2016: 85–86). Examples of theories for predicting crime locations are the opportunity theory and the routine activity theory (Johnson et al., 2007: 203). The opportunity theory reasons that crime rates will be the highest in locations that contain the best opportunities for crime. The routine activity theory assumes that crime will not take place ‘unless a motivated offender comes into contact with a suitable target (opportunity for crime) in the absence of a capable guardian’ (Ibid.). One of the most influential theories for predicting crime locations in contemporary criminology is the broken windows theory, which assumes that degradation and nuisance in a neighborhood invite deviant behavior and can ultimately lead to serious crime (Kolthoff, 2016: 162).

Assumptions underlying the algorithms at work in predictive policing

Not only is the selection of the appropriate algorithm, but also the design of the algorithms themselves based on assumptions. This section will discuss four common assumptions underlying the algorithms at work in predictive policing tools for predicting offenders or crime locations. It will also explain how they affect the accuracy of the predictions these tools offer.

The first general assumption underlying the algorithms at work in predictive policing tools of any type is that the data inputted accurately reflect reality. But this does not need to be the case, especially with regard to crime data. Whether or not something constitutes a crime and how that crime is classified or categorized is a matter of discretion and may differ for different police officers (Bennet Moses and Chan, 2016: 5). This problem is exacerbated because predictive policing tools do not only make use of data collected by police departments in their normal course of business, but also of data that come from other sources (Perry et al., 2013: 13). Certain classifications or categories may have different meanings in different organizations. Moreover, crime data are necessarily limited to reports by victims and police observations. On the one hand, there is a lot of unreported and unseen crime, especially in the area of domestic violence. Because there are no data on these hidden crimes, they will not be inputted into predictive policing tools and they will continue to ignore them (Bennet Moses and Chan 2016: 4–5). On the other hand, there are crimes that have a greater chance of being seen than others, because they are committed by an individual that is considered crime-prone or in a neighborhood that is considered high-crime. When data on these crimes are inputted into a predictive policing tool for predicting crimes or offenders, its predictions may simply reinforce stereotypes that certain neighborhoods or individuals need heavier police attention (Joh, 2014: 58). This way, prior police contacts become a kind of digital scarlet letters (Ferguson, 2015: 401). That is problematic because predictive policing
results in ‘social sorting’: it sorts people into categories assigning risk (Van Brakel and De Hert, 2011: 176). As was mentioned in the last section, people of color and poor people have had disproportionate contact with the criminal justice system in the past. If race and class become data points for risk factors, predictive policing tools can lead us to believe our own worst instincts (Ferguson, 2015: 402). In conclusion, the statistical predictions provided by predictive policing tools can be based on biased statistics and police officers should, therefore, be wary that predictive policing does not perpetuate ‘prejudice in a dangerous new way, by shrouding it in the legitimacy accorded by science’ (Hvistendahl, 2016).

The second general assumption underlying the algorithms at work in predictive policing tools of any type is that history repeats itself. The data analyzed in the context of predictive policing always consist of historical data. The algorithmic procedures used to analyze them look for patterns (Perry et al., 2013: 17). They assume that factors relevant in the past will continue to be relevant in the future (Chan and Bennett Moses, 2016: 32). As was established before, there is evidence that specifically burglary is geographically concentrated, as it clusters in space and time more than one would expect if the perceived patterns of crime would simply result from the attractiveness of places to offenders (Johnson et al., 2007, 2009). But that does not apply (to the same extent) to other crimes (Bennett Moses and Chan, 2016: 5). Moreover, the location of clusters of burglary is not predictable over periods of 3 or more months (Johnson and Bowers, 2004a), because the characteristics of an area change over time and the perpetrator’s memory for the features of a particular house and those nearby may decay (Johnson et al., 2007: 215). Predictive policing itself may spur the recording of crime in a neighborhood, but that may only mean that criminals change their work area and crime rates increase in adjacent areas (Bennett Moses and Chan, 2016: 5). The same goes for predictive policing tools for predicting offenders. If police would, for example, assume that tattoos correlate with crime and, therefore, focus surveillance on people with tattoos, people planning to commit crimes may decide against getting tattoos (Chan and Bennett Moses, 2016: 33). It should be added that advanced, machine-learning algorithms are able to predict such changes as well though (see e.g. Berk and Bleich, 2013: 541). But it is probably the reason why the risk indicators and risk model behind SyRI have never been made public.

The third general assumption underlying the algorithms at work in predictive policing tools of any type is that they focus on a relevant set of data and that the data omitted are irrelevant (Bennet Moses and Chan, 2016: 5). But the fact that certain data are omitted, does not mean that they are irrelevant. They might simply not be available, expensive or difficult to procure. Or it was not realized that they would be relevant when the predictive policing tool in question was being developed. Clearly, the omission of relevant data affects the accuracy of predictive policing tools (Bennet Moses and Chan, 2016: 6).

The three assumptions mentioned above show that we should be aware that the statistical predictions predictive policing tools for predicting crime locations or offenders provide ‘are only as good as the underlying data’ (Perry et al., 2013: 116). It is also good to realize that numbers do not explain reasons: predictive policing tools only predict that crime is likely to occur at a certain place at a certain time or that a certain person is likely to commit a crime, they do not tell why (Chan and Bennett Moses, 2016: 22; Levine, 2006: 46). Predictive policing allows us to let statistical algorithms find all
kinds of patterns, but these patterns do not always indicate causation, while an understanding of causation is necessary in order to predict the impact of police intervention (Chan and Bennett Moses, 2016: 22, 32). Accurate predictions about where and when the crime risk is high or who are at risk to commit a crime only enable police officers to achieve effective short-term crime prevention (Mohler et al., 2016: 1410). But in the long run, the causes of crime are an important key to crime prevention. Problem-oriented responses, such as mentor programs, youth sports programs and neighborhood meetings, can be a more effective way to prevent crime than police deployment (Bennett Moses and Chan, 2016: 8).

Conclusion

Focusing specifically on the Dutch predictive policing tools SyRI (for predicting offenders) and CAS (for predicting crime locations), this paper has described predictive policing as part of the shift from harm to risk that is effectuated by the development toward a pre-crime society (Zedner, 2007), the accompanying culture of control (Garland, 2001) and the new penal logic (Feeley and Simon, 1992) they give rise to. However, predictive policing carries a risk in itself, for the algorithms employed are often a ‘black box’, meaning that the risk calculations provided cannot be explained. This goes against an important moral value, namely: the value of transparency. This lack of transparency is even more problematic because of the nature of the decisions that can be based on the risk calculations provided by predictive policing, such as the deployment of investigatory powers. Where post-hoc ordering practices responding to wrongs done find legitimation in the harm principle, pre-emptive ordering practices meant to prevent crime can only be legitimated based on the precautionary principle, which does not give us the opportunity to weigh the risks of the regulatory situation and the risks of regulation properly. This is illustrated by the recent verdict of the District Court of The Hague (Rb. Den Haag, 05-02-2020, ECLI: NL: RBDHA: 2020:865, available at http://www.rechtspraak.nl), which declares the legislation regulating the use of SyRI unlawful and non-binding for that reason. The same goes for other predictive policing tools for predicting offenders and the police actions based on them. Similar concerns rise with regard to predictive policing tools for predicting crime locations.

Proponents claim that predictive policing tools for predicting offenders and crime locations actually restore instead of erode the rights of persons thought to be likely to commit a crime or living in a crime-prone neighborhood, because they provide neutral, quantitative evidence that a person is indeed likely to commit a crime or that an area is indeed a high-crime area. But the evidence predictive policing tools for predicting offenders or crime locations provide, is not entirely neutral. They make use of different algorithms, requiring the input of different (amounts of) data and resulting in different statistical predictions of the likeliness that a certain person will commit a crime or crime will occur at a certain place. Also, the developers of a particular predictive policing tool have selected the appropriate algorithm with a certain theory or model of crime prediction in mind. Not only is the selection of the appropriate algorithm, but also the design of the algorithms themselves based on assumptions. Moreover, it is important to realize that numbers do not explain reasons. In the long run, the causes of crime are an
important key to crime prevention. Problem-oriented responses, such as mentor programs, youth sports programs and neighborhood meetings, can be a more effective way to prevent crime than police deployment. Therefore, predictive policing should also be used to direct these interventions to locations and persons where they are most needed if we really want to reduce crime.

This conclusion gives rise to other questions, such as whether crime can be seen as a risk in itself, which is the underlying idea of predictive policing. Or are the causes of crime the actual risks that need to be prevented? Answering these questions is beyond the scope of this paper, however. Predictive policing and its underlying ideas are in need of further discussion and analysis not only in the field of legal philosophy, but also in the respective fields of police studies and criminology.

Acknowledgements
I would like to thank my brother Rense Strikwerda, data analyst, and two anonymous reviewers for their useful comments.

Declaration of conflicting interests
The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding
The author(s) received no financial support for the research, authorship, and/or publication of this article.

ORCID iD
Litska Strikwerda https://orcid.org/0000-0002-9001-5067

Notes
1. See e.g. the UK’s Terrorism Act 2000 sec. 58. Available at <http://www.legislation.gov.uk/ukpga/2000/11> (accessed 28 July 2020).
2. Wet van 9 juli 2004 tot wijziging van het Wetboek van Strafrecht, het Wetboek van Strafvordering en de Penitentiaire beginseleinwet (plaatsing in een inrichting voor stelselmatige daders), Stb. 2004, 351, inw.tr., Stb. 2004, 471.

References
Agnew R (2006) *Pressured Into Crime: An Overview of General Strain Theory*. Los Angeles: Roxbury.
Beck U (1986) *Risikogesellschaft: Auf dem Weg in eine andere Moderne*. Frankfurt am Main: Subrkamp Verlag.
Bennett Moses L and Chan J (2016) Algorithmic prediction in policing: assumptions, evaluation and accountability. *Policing and Society* 28: 806–822.
Berk RA and Bleich J (2013) Statistical procedures for forecasting criminal behavior. a comparative assessment. *Criminology & Public Policy* 12(3): 513–544.
Borger M and Van Sliedregt E (2009) The meaning of the precautionary principle for the assessment of criminal measures in the fight against terrorism. *Erasmus Law Review* 2(2): 171–195.
Bowers KJ, Johnson SD and Pease K (2004) Prospective hot-spotting: the future of crime mapping? *The British Journal of Criminology* 44(5): 641–658.

Brey P (2010) Values in technology and disclosive computer ethics. In: Floridi L (ed.) *The Cambridge Handbook of Information and Computer Ethics*. Cambridge: Cambridge UP, pp. 41–58.

Chan J and Bennett Moses L (2016) Is big data challenging criminology? *Theoretical Criminology* 20(1): 21–39.

Drenth A and Van Steden R (2017) Ervaringen van straatagenten met het Criminaliteits Anticipatie Systeem. *Het Tijdschrift voor de Politie* 79(3): 6–10.

Ferguson AG (2015) Big data and predictive reasonable suspicion. *University of Pennsylvania Law Review* 163(2): 327–410.

Feeley M and Simon J (1992) The new penology: notes on the emerging strategy of corrections and its implications. *Criminology* 30(4): 449–474.

Garland D (2001) *The Culture of Control. Crime and Social Order in Contemporary Society*. Oxford: Oxford UP.

Hvistendahl M (2016) Can ‘predictive policing’ prevent crime before it happens? *Science*. Available at: <http://www.sciencemag.org/news/2016/09/can-predictive-policing-prevent-crime-it-happens> (accessed 28 July 2020).

Joh EE (2014) Policing by numbers: big data and the fourth amendment. *Washington Law Review* 89(35): 35–68.

Johnson SD and Bowers KJ (2004a) The stability of space-time clusters of burglary. *The British Journal of Criminology* 44(1): 55–65.

Johnson SD and Bowers KJ (2004b) The burglary as clue to the future: the beginnings of prospective hot-spotting. *European Journal of Criminology* 1(2): 237–255.

Johnson SD, Bernasco W, Bowers KJ, et al. (2007) Space–time patterns of risk: a cross national assessment of residential burglary victimization. *Journal of Quantitative Criminology* 23: 201–219.

Johnson SD, Bowers KJ, Birks DJ, et al. (2009) Predictive mapping of crime by promap: accuracy, units of analysis, and the environmental backcloth. In: Weisburd D, Bernasco W and Bruinsma GJ (eds) *Putting Crime in its Place*. New York, NY: Springer.

Kolthoff E (2016) *Basisboek criminologie*. Den Haag: Boom Criminologie.

Koss KK (2015) Leveraging predictive policing algorithms to restore fourth amendment protections in high-crime areas in a post-Wardlow world. *Chicago-Kent Law Review* 90(1): 301–334.

Levine N (2006) Crime Mapping and the *CrimeStar* Program. *Geographical Analysis* 38(1): 41–56.

Mali B, Bronkhorst-Giesen C and Den Hengst M (2017) Predictive policing: lessen voor de toekomst. Een evaluatie van de landelijke pilot. Politieacademie. Available at: <http://www.politieacademie.nl/kennisonderzoek/kennis/mediatheek/PDF/93263.PDF> (accessed 28 July 2020).

Mohozay Y and Fisher E (2019) The epistemology of algorithmic risk assessment and the path towards a non-penology penology. *Punishment & Society* 21(5): 523–541.

Mill JS (1865) *On Liberty*. London: Longmans, Green and Co.

Moerings M (2016) The Risk of the Persistent Offender and the Special Security Measure. In: Malsch M and Duker M (eds) *Incapacitation: Trends and New Perspectives*, Farnham: Ashgate, pp. 63–75.

Mohler GO, Short MB, Malinowski S, et al. (2016) Randomized controlled field trials of predictive policing. *Journal of the American Statistical Association* 110(512): 1399–1411.
Perry WR, McInnis B, Price CC, et al. (2013) Predictive Policing. The Role of Crime Forecasting in Law Enforcement Operations. RAND. Available at: <http://www.rand.org/pubs/research_reports/RR233.html> (accessed 28 July 2020).

Pieterman R (2008) De Voorzorgcultuur. Streven naar veiligheid in een wereld vol risico en onzekerheid. Den Haag: Boom Juridische Uitgevers.

Provost F and Fawcett T (2013) Data Science for Business: What you need to know about data mining and data-analytic thinking. Sebastopol: O’Reilly Media.

Rummens A, Hardyns W and Pauwels L (2017) A scoping review of predictive analysis techniques for predicting criminal events. In: Vermeulen G and Lievens E (eds) Data Protection and Privacy under Pressure. Transatlantic tensions, EU surveillance, and big data. Antwerp: Maklu, pp. 253–288.

Saunders J, Hunt P and Hollywood JS (2016) Predictions put into practice: a quasi-experimental evaluation of Chicago’s predictive policing pilot. J Exp Criminol 12: 347–351.

Shaw CR and McKay HD (1942) Juvenile Delinquency and Urban Areas. Chicago, IL: University of Chicago Press.

Sunstein CR (2003) Beyond the Precautionary Principle. John M. Olin Law & Economics Working Paper No. 149, Public Law and Legal Theory Working Paper No. 38 (2003). Available at SSRN: <http://ssrn.com/abstract_id=307098> (accessed 28 July 2020).

Van Brakel R and De Hert P (2011) Policing, surveillance and law in a pre-crime society: Understanding the consequences of technology based strategies. Cahiers Politiestudies 3(20): 163–192.

Zedner L (2007) Pre-crime and post-criminology. Theoretical Criminology 11(2): 261–281.

Table of Legal Documents

Chicago Police Department, Custom Notifications in Chicago, Special Order S10-05, Issue Date: 6 October 2015. Available at: http://directives.chicagopolice.org/directives/data/a7a57bf0-1456fa9-bfa14-570a-a2deebf33c56ae59.html (accessed 28 July 2020).

Commission of the European Communities Communication from the Commission of 2 February 2000 on the precautionary principle (COM (2000) 1). Available at: <http://europa.eu/legislation_summaries/consumers/consumer_safety/l32042_en.htm> (accessed 28 July 2020).

Ethics Guidelines for Trustworthy AI, Independent High-Level on Artificial Intelligence set up by the European Commission, available at <https://ec.europa.eu/digital-single-market/en/news/ethics-guidelines-trustworthy-ai> (accessed 28 July 2020).

European Convention on Human Rights, Rome, 4 November 1950. Available at: <http://www.echr.coe.int> (accessed 28 July 2020).

Halfjaarbericht politie 2019, bijlage 5, ICT-vernieuwing bij de politie, Ministerie van Justitie en Veiligheid, Directoraat-Generaal Politie en Veiligheidsregio’s.

Kent Police (2013). PredPol operational review-initial findings. Available at: <http://www.state watch.org/docbin/uk-2013-11-kent-police-pp-report.pdf> (accessed 28 July 2020).

U.S. Constitution. Available at: <http://constitution.findlaw.com> (accessed 28 July 2020).