Machine justice

Peeters, Rik; Schuilenburg, Marc

published in
Information Polity
2018

DOI (link to publisher)
10.3233/IP-180074

document version
Publisher's PDF, also known as Version of record

Information Polity 23 (2018) 267–280
DOI 10.3233/IP-180074
IOS Press

Machine justice: Governing security through the bureaucracy of algorithms

Rik Peeters[a,*] and Marc Schuilenburg[b]
[a]*Centro de Investigación y Docencia Económicas, División de Administración Pública, Colonia Lomas de Santa Fe, C.P. 01210, Ciudad de México, México*
[b]*Department of Criminal Law and Criminology, VU Universiteit Amsterdam, 1081 HV Amsterdam, The Netherlands*

Abstract. The use of algorithms to predict behaviour is becoming the gold standard in criminal justice in various countries. This article critically analyses the algorithm-driven risk assessment tools used in predictive policing and predictive justice. First, we propose to see algorithms as essentially bureaucratic instruments. They are the digital offspring of the classic bureaucratic procedure, creating classification through standardised and impersonal decision-making. Second, we argue that the application of algorithms in criminal justice expands the bureaucratic field to areas previously understood as bulwarks of professional judgement. Third, we analyse the shift in purpose of algorithmic decision-making: instead of determining a citizen's status of beneficiary or obligate, we now see algorithmic anticipation of behaviour. This shifts the logic of decision-making over investigations, probations, and sentencing from individual judgement to bureaucratic classification based on the algorithms that are designed into risk assessments tools. This article is both a bureaucratic critique of algorithm-driven risk assessment tools in criminal justice and a call to rethink bureaucracy and bureaucratisation beyond the boundaries of public administration.

Keywords: Predictive policing, predictive justice, algorithms, information and communication technology

1. Introduction

In the key scene in the episode *Nosedive* of the British series *Black Mirror*, Bryce Dallas Howard's character, a young woman named Lacie Pound finds out that she can rent her fiendishly expensive dream apartment in the 'Pelican Cove Lifestyle Community' with a personal score higher than 4.5. According to her leasing agent, a score above 4.5 would qualify her for a 20% discount on the rent. Such a score is created by reviews on a scale of one to five by other people. Socially desirable behaviour, saying hello to the woman walking past you on your morning commute for instance, creates a higher rating. Also, dealing with persons with a higher rating increases your personal score. Through special lenses, a sort of Google Glass which contextual information projected on your glasses, you immediately see the ranking of the people around you and you can give them points for their behaviour.

Black Mirror is a British sci-fi series about the murky relationship between humans and technology. In the episode *Nosedive*, we are confronted with Lacie's doomed attempt to improve her personal rating. Having a high score comes with advantages and having low status can bar you from restaurants, work,

*Corresponding author: Rik Peeters, Centro de Investigación y Docencia Económicas (CIDE), División de Administración Pública, Carretera México-Toluca 3655, Colonia Lomas de Santa Fe, C.P. 01210, Ciudad de México, México. Tel.: +52 55 5727 9800 (ext. 2324); E-mail: rik.peeters@cide.edu.

and even health care. When you end up below 2.5, you are part of the social underclass. Ultimately, Lacie, a moulded 4.2 star rated citizen, fails to be a 4.5. She slides down to the gritty and grim bottom of society and ends up in prison where she is stripped of technology. Broken and tired, but relieved to be in a place where she can finally tell how she feels and what she thinks, without fear of being ranked down.

Watching the episode *Nosedive*, we are inevitable reminded of the neoliberal logic of competition and individualism, which is expressed by the fact that on social media one accumulates 'likes' (Facebook), 'views' (YouTube) and 'followers' (Twitter). But we are also reminded of the way our life is influenced by all kinds of consumer scores, created by score modellers who feed raw data into an algorithm designed to trawl through reams of data to detect patterns. Through consumer scores, people are 'ranked, separated, shifted, and otherwise categorized and also predicted their potential future actions' (Dixon & Gellman, 2014: 6). Although many of these consumer scores are in the field of social media, they are also used in health, the labour market, and justice. Our interest in this article is the criminal justice system, in which authorities increasingly use computer-generated scores for purposes of crime prevention and to decide whether a criminal defendant is likely to reoffend or not. This phenomenon has far-reaching social consequences and brings up all sorts of questions, from 'which kind of scores are used and for what purpose?' to 'how can we guarantee that the algorithms underlying the scores make fair decisions?'

In the following, we discuss several examples of 'ranking citizens' in the criminal justice system. From predictive policing to predictive justice, the criminal justice system in various countries relies increasingly on algorithm-driven risk assessments, which produce all kind of scores (e.g. Fitzgibbon, 2008; Hannah-Moffat, 2013; Van Eijk, 2017). Within the academic debate on the use of these tools and the role of algorithms, the focus is often on their negative or positive impacts on citizens, such as potential discriminatory effects. In this article, however, we aim to broaden the discussion by focusing on their effect on the bureaucratic field – broadly defined as the rules of the game governing the actions of public institutions entrusted with the maintenance of the legal order (Bourdieu et al., 1994; Wacquant, 2004). Our argument is threefold. First, we propose to see algorithms as essentially bureaucratic instruments. They are the digital offspring of the classic bureaucratic procedure, creating classification through standardised and impersonal processing of data. Second, we argue that the application of algorithms in criminal justice expands the bureaucratic field to areas previously understood as bulwarks of professional judgement. And third, new technological possibilities to analyse big data lead to a shift in the objectives of governments: instead of merely determining a citizen's status as beneficiary or obligate, information technology is now used for the algorithmic anticipation of behaviour of citizens. As a consequence, the 'datafication of criminal justice' (Smith et al., 2017: 260) shifts the logic of decision-making over investigations, probations, and sentencing from individual judgement to bureaucratic classification based on the algorithms that are designed into risk assessments tools.

We begin our argument by analysing the existing literature on the relation between bureaucracy and information and communication technology. In public service, information and communication technology is often used to increase efficiency by reducing red tape. However, critical studies show that information and communication technology relies on mechanisms that increase control over decision-making. Next, we apply this perspective to the field of criminal justice and describe two algorithm-driven risk systems, which have become a common element in decisions over surveillance and sentencing: predictive policing and predictive justice. We analyse the way both systems stress bureaucratic classification over individual judgement. In the discussion, we reflect on the bureaucratic nature of algorithms and propose to rethink the state-centred conception of contemporary bureaucracy and bureaucratisation.

2. Bureaucracy, information and communication technology, and algorithmic anticipation

2.1. Bureaucracy: Knowledge and control

One of the defining characteristics of the modern state is the invention of the bureaucracy. Ever since Weber (1921/1922) noted the emergence of the bureaucracy as part of a broader process of rationalisation of government and economy, it has been viewed upon as a Janus-faced organisation, looking two ways at once. According to Gouldner, 'on the one side, it was administration based on expertise; while on the other, it was administration based on discipline' (1954: 22). Without notable exceptions, the bureaucracy is known as the prime organisational 'vehicle' for governing through knowledge of citizens and the population as a whole. Consequently, Weber speaks of the bureaucracy in terms of 'authority through knowledge' (2006a: 226). Bureaucracy's other face – its disciplining nature – is symbolised in Weber's notion of the 'iron cage' (Weber, 2006b: 201):[1] a system in which humans are reduced to the instrumental cogs of a mechanistic machine and where reflection on values and objectives has been pushed out in favour of self-perpetuating procedures and control mechanisms.

In both expertise and discipline, the bureaucracy embodies instrumental rationality: decisions are made according to coherent rules, such as procedural rules that both define the ends and regulate the means of action, clear expectations, specific time frames, and goals and measures of success (Clegg & Lounsbury, 2009). Importantly, this should not be equated with modern-day 'efficiency', but above all seen as a way to realise predictable outcomes and reduce uncertainty (Gajduschek, 2003). While this stands in sharp contrast with value rationality, in which decisions are made according to an absolute value of belief (such as ideology or religion), it also makes rational-legal authority possible: rule exercised by officials and institutions by virtue of formal rules and laws (Weber, 2006a: 219-222).

2.2. Information and communication technology: Infocracy and screen level bureaucrats

Technological advancements are a well-known means to increase a state's capacity to gather knowledge and exert power (Popitz, 1992: 181). A crucial development in the history of the state bureaucracy is the digitalisation of its activities. The modern-day influence of information and communication technology on public bureaucracies has been well studied. For instance, Zuurmond (1994) speaks of the emergence of the 'infocracy', a bureaucracy that revolves around the collection of data and automated decision-making – including automatic number plate recognition for speeding tickets and automatic decisions on the height of students grants based on the income of parents (Zouridis, 2000). In the current age of big data, the possibilities are even bigger: larger amounts of data can be analysed than ever before, data can be more easily shared among organisations, and algorithms can automate the analysis process. This accelerates and expands the potential of public organisations to predict and profile people's behaviour (Sadin, 2009; Reigeluth, 2014).

Advocates of the use of information and communication technology in public administration argue that it can be a valuable tool to combat several bureaucratic pathologies. Classic critiques already showed that organisations with high levels of bureaucratisation tend to become rigid in their structure, closed off from external influences, and have an alienating effect on their employees because of the abstract and impersonal nature of bureaucratic work (Merton, 1940; Crozier, 1964; Tullock, 1965; Downs, 1967). 'Neutral' bureaucratic mechanisms transform into bureaucratic 'pathologies' (Acar & Aupperle, 1984).

[1] Though this is hardly a precise translation of the original German 'stahlhartes Gehäuse' (Klagge, 1997; Baehr, 2001).

For instance, formalisation becomes red tape, impersonality becomes indifference, and specialisation becomes fragmentation. Advocates of e-government suggest that information and communication technology can be used to cut red tape, speed up procedures, reduce fraud, and facilitate data sharing across organisations (e.g. La Porte et al., 2002; OECD, 2005). Interestingly, these arguments for the use of information and communication technology are, by and large, not anti-bureaucratic, but proposals to make bureaucracies work more efficient.

Information and communication technology is not just a tool that bureaucracies can use, but is, in itself, fundamentally bureaucratic. This point has been made in various studies, pointing out that information and communication technology is a mechanism for simplification and closure (Kallinikos, 2005). Simplification is the process through which information technology breaks down a task or problem into sets of operations that need to be performed sequentially in order to solve it. Closure is the necessary complement to simplification. It entails, according to Cordella and Tempini, 'the isolation and black boxing of the sequential operations, ensuring their execution is protected from external interference' (2015: 281). The use of information and communication technology has, therefore, triggered a further centralisation and rationalisation of decision-making processes (e.g. Margetts, 1999). Information technology-driven bureaucracies reduce the autonomy and discretionary space of professionals and bureaucrats. Street-level bureaucrats transform into 'screen level bureaucrats' (Landsbergen, 2004) that classify citizens rather than make factual assessments and decisions (Bovens & Zouridis, 2002). Web sites replace window clerks and information system replace case managers. As a consequence, human interference in decision-making procedures is eliminated.

2.3. Algorithmic anticipation: surveillance capitalism and governance of security

So far, we have limited our discussion to the administrative domain – to public organisations charged with deciding over eligibility to state benefits and or with enforcement of administrative law (e.g. Zouridis, 2000). However, recent rapid developments in the widespread social use of information and communication technology are opening up new fields for the use of algorithms. Companies like Apple, Facebook, and YouTube are using customer data to offer personalised services or sell information to companies for personalised marketing. Credit card companies use consumer scores to decide whether to offer a client an extended credit line. Car insurance companies are tracking their clients' driving style to offer discounts for safe driving (Schuilenburg & Peeters, 2017). New technological possibilities allow the quick analysis of vast amounts of data for the purpose of finding behavioural patterns (Mayer-Schoenberger & Cukier, 2013). The algorithmic anticipation of behaviour is a powerful business model (Sadin, 2009; Reigeluth, 2014) that Zuboff (2015) has called 'surveillance capitalism'.

The surveillance capacity of modern information and communication technology has also been picked up in criminal justice. The 'proliferation of scoring and ranking citizens' (Harcourt, 2015: 205) extends to how likely a person is to commit a crime or whether a defendant is likely to commit new crimes. Here, the objective is less financial and more deliberately control-oriented. While the use of statistical techniques in criminal justice is not new, the number of algorithm-driven risk systems has vastly increased over the past decades. According to Harcourt (2007), profiling and predicting have become a second nature in the quest for security and prevention. A core element of these risk systems is their preoccupation with behaviour. For the purposes of prevention and punishment, the behaviour of individuals and specific groups is monitored and predicted, and, where necessary, subjected to interventions (e.g. Peeters, 2013). It seems rather cliché and old hat to accentuate that surveillance has been deeply implicated in all of this. However, it is important to realise that new forms of data analysis not only expand the opportunities for surveillance, but also add a new layer: the use of algorithms quantifies the singling out

of places, individuals or specific groups with a higher chance of showing criminal activity, according to the perceived risks of the used risk assessment tools. Instead of determining an individual's status, data is used to determine the risk individuals or groups pose. Just as a credit card company sees a client who suddenly starts buying cheaper brands as a financial risk factor, the police can target individuals that check the boxes of risk factors for criminal behaviour, such as criminal history, gender, age, low level of education, no stable housing and unemployment.

The use of algorithms in criminal justice is in part a consequence of increased technological opportunities, and in part of governmental efforts to predict and pre-empt risks as well as to broaden the governmental concern beyond crime control to a wider security context (Schuilenburg, 2015). The convergence of these developments creates a powerful incentive for the application of information technology and algorithmic anticipation in previously uncharted domains. In the following, we analyse the existing literature on the emerging use of algorithms in criminal justice – more specifically: in predictive policing and in predictive justice. We focus on how the bureaucratic nature of algorithms changes decision-making in policing and criminal justice. As we move beyond the administrative domain – where information technology and algorithms were used to determine a citizen's status as beneficiary or obligatee – we encounter new areas of bureaucratic 'colonisation' (Habermas, 1987). Judgements and decisions in policing and sentencing are being subjected to the 'violent abstraction' (ibid.: 363) we, until recently, only knew from classification mechanisms in public service organisations.

3. Predictive policing

The use of data to detect crimes has long been a central feature in security policies decision-making processes. Smith and others (2017) mention examples of automated license plate recognition systems and the emergence of the 'techno-cop' who is communicatively wired into command and control via radio device, camera, and GPS tracker. Furthermore, data is used to detect and punish traffic violations, such as speeding. Infractions are digitally captured and fines are automatically processed, thereby mediating 'the transaction between prosecutor and offender [...] by the transfer of data' (Smith & O'Malley, 2017: 278; O'Malley, 2010). More interesting for the purposes of this paper are the efforts to produce representations of reality that 'convert social acts into flows of data' (Smith & O'Malley, 2017: 275). A prominent application is traffic management, in which movements of vehicles are digitally converted to be able to regulate the flow and circulation of cars through variable speed limits. Security assemblages consisting of measures such as tracking devices, motion sensors, drones, radar guns, and smart software are used to govern safe driving and road use.

The use of 'predictive algorithms' and 'crime sensing' (Smith & O'Malley, 2017; Williams et al., 2017) takes this logic one step further. Instead of merely representing reality in flows of data, representations of reality are projected into the future. By bringing together 'big data' from various sources – ranging from criminal statistics to mining of individual's social media texts and images – and automatically analysing them, diagnostic conclusions are drawn to identify crime patterns and suspicious individuals, as well as their geographical placement and their networked nature. The ultimate objective of these algorithm-driven risk assessment systems is to change the outcomes as they are predicted through preventative action (Perry et al., 2013). Conceptually speaking, predictive policing is closely related to other data-infused policing practices, such as intelligence-led policing, data-driven policing, risk-based policing, hotspot policing, and evidence-based policing. What sets it apart is the explicit projection into the future, including the use of models that can predict changes in crime patterns instead of mere extrapolation of existing data (Bennett Moses & Chan, 2016).

Predictive policing can be used to make inferences about future crime on four different levels: 1) the location where crimes will likely be committed (often called 'hotspots'); 2) the identification of persons more likely to commit crimes; 3) the time when crimes are likely to be committed; and 4) the combination of 'where', 'who', and 'when' (Asquer, 2014: 20). The basic assumption is that crimes, such as theft and robberies, are to a large extent predictable, because criminals with a distinguishable profile tend to commit the same type of crime, at roughly the same location and time of the day. Statistical methods use existing police records on number and type of arrests and their geographical location to recognise patterns in criminal behaviour, which in turn can facilitate a more preventative approach as well as a more focused and quicker police response (Schuilenburg, 2016).

An example from the US is the system 'PredPol' (Predictive Policing), which uses an earthquake prediction model to predict crimes such as gang violence and residential burglary. Similar to earthquake aftershocks, these crimes tend to be temporarily high in a certain location after an initial report of a 'background event' (Bennett Moses & Chan, 2016). A similar Dutch example is the 'Crime Anticipation System' (CAS), which is used by the police in Amsterdam. In this system, the city of Amsterdam is divided in blocks of 125 by 125 meters. For each of these blocks, information about past crimes, distance to known suspects, the type and number of crimes, and the demographic and socio-economic structure is gathered. With regular two-week intervals, this information is updated and analysed for incorporation in daily police operations (Schuilenburg, 2016; De Vries & Smit, 2016).

An interesting new data source for predictive policing is 'big social data' (Williams et al., 2017), i.e. data voluntarily exposed by social media users (Harcourt, 2015). Governments use these data to track down expressions of allegiance to extremist groups or disorder-related posts that might indicate future criminal activity. A similar tactic is the analysis of Twitter posts to map crime patterns. For instance, big data analysis of tweets on minor public incivilities can serve as confirmations of a 'broken window theory' in a certain neighbourhood (Williams et al., 2017). Others have analysed tweets to map the spread of cyberhate (Williams & Burnap, 2016), used the geolocation of tweets to pinpoint crime hot spots (Malleson & Andresen, 2015) or showed that an absence of tweets is predictive of assaults, theft, and disturbances of the social order (Bendler et al., 2014).

4. Predictive justice

Algorithms are not only transforming police practices, but the administration of the criminal justice system as well. Against the background of a penology that is increasingly utilitarian instead of merely retributive (Hamilton, 2015: 5), data is used by courts to predict 'future dangerousness' of individuals (Kleiman et al., 2007; Berk, 2012; Berk & Bleich, 2013). The application of algorithms in the criminal justice system, also referred to as 'actuarial justice' (Hamilton, 2015: 3), is used to estimate a defendant's flight risk and to assess his or her threat to public safety. 'Actuarial prediction instruments' (Sjöstedt & Grann, 2002), 'recidivism risk algorithms'[2] and 'risk assessment tools'[3] are instruments that use variables about a defendant's criminal history and socio-demographic characteristics to estimate whether someone poses a high, moderate or low risk. Key variables are identified based on records of past criminal cases. The objective is to improve criminal sentencing, bail decisions, and rehabilitation programmes

[2]http://www.businessinsider.com/harvard-mathematician-reveals-algorithms-make-justice-system-baised-worse-black-people-crime-police-2017-6 (accessed 12 January 2018).

[3]https://epic.org/algorithmic-transparency/crim-justice/ (accessed 12 January 2018).

through evidence-based risk assessment. More specifically, algorithms aim to increase equity (treating all risks the same), transparency (in the form of an objectified score), and efficiency (reducing lead times) (Goel et al., 2016).

There are currently more than 200 risk assessment tools available in criminal justice and forensic psychiatry, which are widely used to inform sentencing, parole decisions, and post-release monitoring (Douglas et al., 2017: 134). In the US specifically, there are three main systems in use (COMPAS, PSA, and LSI-R) with their own specific sets of variables. Interestingly, statisticians, mathematics, computer scientists and software engineers working for private companies develop these systems and the exact workings of the algorithms are, therefore, proprietary and not made public.[4] Christin et al. (2015) identify four criminal justice areas in which predictive algorithms are used:

- Pre-trial and bail decisions: assessment of the risk that a defendant will engage in violent crime or fail to show up in court. According to 2016 data, less than 10% of U.S. jurisdictions use pre-trial risk assessments. However, support for their introduction is growing.[5] A study of a pre-trial risk score in Baltimore[6] shows how the assessment of who will be a public safety or flight risk is based on an algorithm that includes mitigating (age, employment, etc.) and aggravating (substance abuse, previous arrests, etc.) risk factors. In a different tool, a set of scores is derived from 137 questions about criminal past, criminal networks, drug use, a defendant's violent character, and a defendant's view on criminal behaviour.[7] These tools are supposed to help pre-trial services recommend bail decisions and offer support programmes (such as drug rehabilitation) based on standardised and objective criteria.
- Criminal sentencing: assessment of potential recidivism can play a role in determining the height of a sentence. Its history goes back to at least 1984, when Sentencing Tables were created, a mandatory federal instrument to determine ranges of incarceration for serious federal crimes based on offense level and criminal history of a defendant. Since 2005, these Tables are advisory, but still have significant impact. Moreover, new risk assessment tools go beyond them and include factors not directly connected to the crime, such as socio-demographic background. A well-documented example is the case of Wisconsin v. Loomis, in which the defendant was given a long sentence for his role in a drive-by shooting partially because of his 'high risk' score.[8]
- Parole and probation: assessment of prison inmates' security classification, their eligibility for parole, and levels of probation supervision. Instruments may measure both a defendant's current living situation and behaviour, and variables outside his or her direct control such as having parents with a criminal background. Typically, this is where a broader notion of rehabilitation is introduced in risk assessments. Besides risk of recidivism, a defendant's need can be included as well to determine treatment programmes, services, and supervision.[9]
- Juvenile justice: assess the eligibility of minors for specific types of detention or for release home. In 2014, more than 300 jurisdictions in 39 US states are working with a risk assessment instrument specifically designed to determine the most suitable form of detention and support programme for juvenile delinquents.

[4] https://epic.org/algorithmic-transparency/crim-justice/ (accessed 12 January 2018).

[5] http://www.citylab.com/crime/2016/12/justice-by-algorithm/505514/ (accessed 12 January 2018).

[6] http://www.citylab.com/crime/2016/12/justice-by-algorithm/505514/ (accessed 12 January 2018).

[7] https://www.propublica.org/article/machine-bias-risk-assessments-in-criminal-sentencing (accessed 12 January 2018).

[8] The case is well known because the defendant challenged his sentence on the basis that he was not allowed to assess the algorithm. In 2016, the state supreme court ruled that knowledge of the algorithm's output was a sufficient level of transparency, see: https://www.wired.com/2017/04/courts-using-ai-sentence-criminals-must-stop-now/ (accessed 12 January 2018).

[9] https://fas.org/sgp/crs/misc/R44087.pdf (accessed 12 January 2018).

5. Advocates and critics

In literature, advocates in favour of the use of algorithm-driven risk assessments to automate decisions in police practices and the criminal justice system see the use of actuarial prediction tools as ways of 'scientization, economitization and democratization' (Smith et al., 2017: 261). First, advances and progress in science and technology open the door to new approaches for the assessment of risks and dangerousness. As such, there is a strong belief that these new tools are capable of finding blind spots and solving complex social problems. Second, actuarial prediction tools can improve government efficiency. Law enforcement agencies, according to this economical argument, 'can detect more crime with the same resources if they investigate citizens who are at greater risk of criminal offending; and sentencing bodies can reduce crime if they incapacitate citizens who are more likely to recidivate in the future' (Harcourt, 2005: 3). Third, there is a popular belief that these new tools are value-neutral ways to generate knowledge and expertise within law enforcement. They not only save time and money, but also make sentencing more accountable by protecting it against human bias and curb discriminatory and racist sentencing practices (Christin et al., 2015).

In sharp contrast, a more critical perspective on algorithm-driven risk assessments points out the serious concerns regarding reliability of the algorithms in predicting crime or recidivism. Full prediction is unattainable, since 'universal and perfect surveillance' (Smith et al., 2017: 267) would be needed to gather a truly complete set of data. Moreover, the sheer volume of data necessary to make a reliable prediction 'can overwhelm the critical signals in a fog of possible correlations' (Crawford, 2014). A model is always needed to select and sample relevant data, which necessarily increases chances of distortion and ending up with 'only speculative projections of historic patterns' (Smith et al., 2017: 268). But reliability in predictive algorithms is also far from self-evident on a more modest level. A practical constraint is that existing biases in police data (such as unreported crime and previous arrests) are designed into algorithms, thereby leading to an emphasis on already over-policed high-poverty and non-white areas.[10] The available evaluations show mixed results in predictive reliability (Sjöstedt & Grann, 2002; Douglas et al., 2017), and, moreover, reveal that non-white defendants are more often subjected to false positives (regarded as high risk, but not re-offending) and white defendants more to false negatives (regarded as low risk, but re-offending) (Rehavi & Starr, 2014).[11]

Others have raised concerns that actuarial prediction tools inscribe and alter rehabilitative approaches to punishment (Brown, 2009: 182–189; Robinson, 2008): risks for the general public weigh heavier than the needs of defendants, and, as a consequence, risk assessment tools trigger punishment rather than weeding out low-risk defendants. This might be an unintended consequence of risk assessment tools, but critics claim that these tools are politically and ideology-driven (Simon, 1988). Instead of the 'get tough' ideology and discourse of the punitive turn in crime policies during the last decades (Garland, 2001), the actuarial turn presents a 'much cooler, more nuanced foundation' (Clear & Frost, 2014: 16) for a punitive approach under the guise of objectivity and empiricism.

6. Machine justice

While many scholars have pointed to the dangers and promises of algorithm-driven risk assessments, very little is known about how these tools influence the rules of the game governing the actions of public

[10] http://www.citylab.com/crime/2016/12/justice-by-algorithm/505514/ (accessed 12 January 2018).
[11] https://www.propublica.org/article/machine-bias-risk-assessments-in-criminal-sentencing (accessed 12 January 2018).

officials entrusted with the maintenance of the legal order. Nevertheless, we know from bureaucracy critiques that the use of information and communication technology tends to diminish the discretionary space of decision-makers at street-level (e.g. Bovens & Zouridis, 2002). Street-level bureaucrats become, as several studies showed, dependent on data and the algorithms through which it is processed. Additional to the above-mentioned arguments, we argue therefore that algorithm-driven risk assessments push decision-making in policing and the criminal justice system away from professional assessment and towards machine-like procedural classification. To put it paradoxically, although the machine learning ability of algorithm-driven risk tools suggests that the decision-making process is in constant flux, the use of these tools leads to more rigid and standardized behaviour.

Risk assessments were also used prior to the introduction of algorithms. For instance, judgements on bail, parole, and probation have always been infused with concerns about recidivism and failure to appear in court. However, these risk predictions were unstructured and unregulated. They were professional judgements by judges and mental health clinicians, largely free of bureaucratic procedures and sometimes infused by gut decisions (Hamilton, 2015: 5-9). In sharp contrast, as information and communication technology standardises the bureaucratic field, by reducing social complexities into predefined scores, algorithm-driven risk assessments will function according a machine-like logic: once an algorithm is designed, outcomes are produced automatically, without human interference and, thereby, jeopardising individualised justice. The diminishing of professional discretion by this kind of decision displacement, in which algorithms function as the 'digital recipe' through which data is processed, is strengthened by the fact that the currently available risk assessment tools are made by private companies and are therefore proprietary, which reduces them to black boxes closed for thorough evaluation (Bennett Moses & Chan, 2016). This makes accountability for decisions regarding police deployment, sentencing, probation, and bail highly problematic (Brucato, 2015; Hamilton, 2015).

One might object that the actors operating with systems of predictive policing and predictive justice have the freedom to ignore algorithms and make their own assessments of individual cases. While this is formally speaking the case, we argue that bureaucratic practice shows a tendency towards a self-imposed reduction of professional discretion. Algorithms combine the quest for reducing social complexities into predefined categories and risk factors with the quest for efficiency and objectivity. As a consequence, the outcomes of an algorithm are more determined by its input and design than by the individual assessment of a case. What is more, bureaucratic structures are highly disciplining for the behaviour of those subjected to it (e.g. Weber, 1921/1922; Merton, 1940; Crozier, 1964; Tullock, 1965; March & Olsen, 1989). Furthermore, the voluntary subjection to bureaucratic rules is a common behavioural mechanism. After all, 'rules protect the people who submit to them' (Crozier, 1964: 206). Rules provide a default for action, and thus legitimacy. To deviate from them can be dangerous: the very existence of 'objective' scores has a chilling effect on the use of professional discretion. According to Hamilton, '[i]nstead of caution, [...] policymakers and judges seem impressed by the guise of empiricism, and in lieu of critiquing the fitness, validity, and reliability of risk tools, officials are more likely to reify them' (2015: 49). Similarly, Nilsson and others (2009: 402) note a 'better safe than sorry' approach when it comes to following rather than deviating from algorithms. The consequence of this all is that algorithms have come to play the role of rules in Weber's theory of instrumental rationality.

7. Conclusion and discussion

As our society continues to become more complex, with the proliferation and extension of algorithm-driven risk assessments to all facets of life, cases like Lacie Pound's in the television series *Black Mirror*

will become increasingly common: automated scores based on the systematic collection of personal data with potentially negative effects on the lives of citizens. In the episode *Nosedive*, rating citizens takes the form of a softer, more invisible social control. Perceived anti-social behaviour is punished by fellow citizens through a lower social score, which ultimately limits a person's access to public and private goods, such as housing, transport and restaurants. This type of scoring mechanisms is becoming a reality in the criminal justice system, in which computer-generated scores produce new ways for the authorities to identify future criminal behaviour. According to experts, the use of these risk assessments may be the biggest shift in the governance of security since the criminal justice system began accepting social science and other expert evidence more than a century ago.[12]

In the above, we have shown how algorithm-driven risk assessments are used to (1) predict future criminal behaviour and to (2) administer a criminal justice outcome. Predictive policing and predictive justice adopt two of the most important recent developments in this respect. First, they express the desire to use knowledge of the population for assessing risk rather than determining status – in line with a more behavioural and preventative approach to late-modern security issues. Second, algorithm-driven systems are the main tools for the assessment of risks – whether to predict geographical locations of crime or levels of dangerousness in individuals. The implicit message in the algorithms used in predictive policing is: 'the higher the risk, the more surveillance'. For predictive justice, the credo is: 'the higher the risk, the higher the punishment'.

We have argued that advocates of predictive policing and predictive justice highlight the benefits associated with risk assessment tools. An important argument is that the automatic tools do a better job in ensuring public safety then ad-hoc and subjective assessments of police officers and judges. Critics respond that algorithms produce a cycle of self-fulfilling prophecies because they are based on historical crime data. They also share the concern that predictive algorithms produce discriminatory results because crime data is both incomplete (a large percentage of crime is not reported) as reflect longstanding institutional biases along income, race and gender lines.

In addition to these arguments, we have argued that police practices and the administration of justice are radically transformed by the technological opportunities and availability of data combined with the preventative ambitions inherent to many security policies. While operating under a cloak of rationality, algorithm-driven risk assessments present a specific threat because the outcomes do not 'argue' (they do not present an argument or a reasoning, which includes revealing sources and assumptions, but present a 'truth') and because they are increasingly generated automatically (not through human analysis, but through algorithms). In terms of the rules of the game governing the actions of public institutions entrusted with the maintenance of the legal order, the cumulative effect of the use of algorithm-driven risk tools makes bureaucracies operate more machine-like. Only this time around, algorithms instead of rules are the source of the rationalisation of decision-making, the formalisation of the procedures for reaching decisions, and the standardisation of the treatment of each case.

The bureaucratic nature of information and communication technology and its adoption in risk assessment tools not only provides new relevant insights for the criminal justice system, but it also urges us to broaden the scope of our thinking on bureaucracy. Weber thought of bureaucracy as a modern social phenomenon, which extended over both public and private organisations. Today, we hardly associate modern capitalistic organisations with the machine-bureaucracies of mass production on assembly lines anymore. However, modern public service organisations have proven to be remarkably resilient in their

[12] https://www.theyeshivaworld.com/news/general/1460874/ai-in-the-court-when-algorithms-rule-on-jail-time.html (accessed 12 January 2018).

bureaucratic nature (Meier & Hill, 2005). Most of the literature of the bureaucratic form is, therefore, focused on government and public administration (Peters, 2003; Olsen, 2006). It is, moreover, associated with a very specific interpretation of bureaucracy, namely as an organisational structure. This association of bureaucracy with the organisation of the state obscures the bureaucratic nature of non-state organisational forms and decision-making mechanisms.

A substantial body of literature challenges the relevance of 'bureaucracy' to understand the functioning of public organisations (Clegg & Baumeler, 2010). Some describe these contemporary organisations as 'post-bureaucratic' (Heckscher & Donnellon, 1994), others as 'soft bureaucracies' in which classic bureaucratic mechanisms are softened by cooperation, entrepreneurship, and self-management (Courpasson, 2000). A more fundamental critique even suggests that the social conditions for bureaucratic organisations have eroded: based on the sociological analyses of Giddens and Beck (Beck, 1986; Beck et al., 1994), some argue that reflection rather than instrumentality – which blindly follows the desire for progress and efficiency – will form the future basis for legitimate decision making. The authority of rational-legal institutions is no longer taken for granted, but needs to be proven in constant deliberation with social actors and individual citizens (Hoogenboom & Ossewaarde, 2005).

Regardless of the contested validity of these claims (cf. Peters, 2001; Hodgson, 2004; Alvesson & Thompson, 2006; Courpasson & Clegg, 2006; Adler, 2012), our analysis shows that, first, bureaucratic forms can be understood more broadly than organisational structure. Instead, we can also find standardisation, formalisation, and rationalisation in information and communication technology. Algorithms are bureaucratic in the sense that they structure human behaviour and decision-making in similar ways as bureaucratic organisations do – or perhaps even more so, given the reduction of judgement on individual cases to classification. Second, bureaucratic forms can be found beyond the beaten paths of public service organisations. The new logic of surveillance capitalism and its reliance on algorithms suggest that bureaucratic forms are alive and well in the private sector. Moreover, the use of algorithms for the anticipation of behaviour – instead of merely determining a person's status as beneficiary or obligate – opens up new paths for bureaucratisation of the public sector beyond the confines of public service provision.

It seems likely that algorithmic tools are here to stay. It is, therefore, worth discussing several paths to mitigate the dangers of machine justice. Some suggest government regulation or an independent watchdog[13] to guarantee 'algorithmic accountability'[14] and the possibility to evaluate algorithms on their design, input data, and possible biases.[15] Others argue for including more diversity (people with different backgrounds) in the phase of algorithm design to prevent bias.[16] While useful, these suggestions do not tackle the problem of decision displacement, which is less a matter of design and more a matter of use in decision-making. To escape the temptations of algorithm-centred decision making, mechanisms are needed that stress human assessment of individual cases. We suggest focusing on the empowerment of citizens, such as strengthening the information position of citizens regarding the variables in algorithms and the personal data that decision makers use, introducing a right to correction of unfair algorithmic decision making, and introducing a stronger duty to motivate decisions based on algorithms regarding individual circumstances. Confronted with a further bureaucratisation of state-citizen interactions, citizens should be given the means to restore the balance.

[13] https://www.theguardian.com/technology/2017/jan/27/ai-artificial-intelligence-watchdog-needed-to-prevent-discriminatory-automated-decisions (accessed 12 January 2018).

[14] http://harvardmagazine.com/2017/01/the-watchers (accessed 12 January 2018).

[15] http://www.techrepublic.com/article/biases-in-algorithms-the-case-for-and-against-government-regulation/ (accessed 12 January 2018).

[16] http://www.techrepublic.com/article/bias-in-machine-learning-and-how-to-stop-it/ (accessed 12 January 2018).

Conflict of interest

The authors declare that they have no conflict of interest.

References

Acar, W. & Aupperle, K.E. (1984). Bureaucracy as organizational pathology. *Systems Research*, *1*, 157-166.

Adler, P.S. (2012). The Sociological Ambivalence of Bureaucracy: From Weber via Gouldner to Marx. *Organization Science*, *23*(1), 244-266.

Alvesson, M. & Thompson, P. (2006). Post-bureaucracy?. In S. Ackroyd, R. Batt, P. Thompson & P.S. Tolbert (eds.). *The Oxford Handbook of Work and Organization* (pp. 485-507). New York: Oxford University Press.

Asquer, A. (2014). Big Data and Innovation in the Delivery of Public Services: The Case of Predictive Policing in Kent. Retrieved from https://ssrn.com/abstract=2429395.

Baehr, P. (2001). The 'Iron Cage' and the 'Shell as Hard as Steel': Parsons, Weber, and the Stahlhartes Gehäuse Metaphor in the Protestant Ethic and the Spirit of Capitalism. *History and Theory*, *40*(2), 153-169.

Beck, U. (1986). *Risikogesellschaft. Auf dem Weg in eine andere Moderne*. Frankfurt am Main: Suhrkamp.

Beck, U., Giddens, A. & Lash, S. (1994). *Reflexive Modernization. Politics, Tradition and Aesthetics in the Modern Social Order*. Oxford: Polity Press.

Bendler, J., Brandt, T., Wagner, S. & Neumann, D. (2014). Investigating Crime-to-Twitter Relationships in Urban Environments – Facilitating a Virtual Neighbourhood Watch. Proceedings of the European Conference on Information Systems (ECIS), Tel Aviv, June 9–11.

Bennett Moses, L. & Chan, J. (2016). Algorithmic Prediction in Policing: Assumptions, Evaluation, and Accountability. *Policing and Society: An International Journal of Research and Policy*, DOI: 10.1080/10439463.2016.1253695.

Berk, R.A. (2012). *Criminal Justice Forecasts of Risk: A Machine Learning Approach*. New York: Springer.

Berk, R.A. & Bleich, J. (2013). Statistical Procedures for Forecasting Criminal Behavior. *Criminology & Public Policy*, *12*, 513-544.

Bourdieu, P., Wacquant, L. & Farage, S. (1994). Rethinking the State: Genesis and Structure of the Bureaucratic Field. *Sociological Theory*, *12*(1), 1-18.

Bovens, M. & Zouridis, S. (2002). From Street-Level to System-Level Bureaucracies: How Information and Communication Technology Is Transforming Administrative Discretion and Constitutional Control. *Public Administration Review*, *62*(2), 174-184.

Brucato, B. (2015). Policing Made Visible: Mobile Technologies and the Importance of Point of View. *Surveillance & Society*, *13*, 455-73.

Christin, A., Rosenblat, A. & Boyd, D. (2015). Courts and predictive algorithms. *Data & Society*. Retrieved from www.datacivilrights.org (consulted: 7-9-2017).

Clear, T.R. & Frost, N.A. (2014). *The Punishment Imperative: The Rise and Failure of Mass Incarceration in America*. New York: NYU Press.

Clegg, S. & Baumeler, C. (2010). Essai: From Iron Cages to Liquid Modernity in Organization Analysis. *Organization Studies*, *31*(12), 1713-1733.

Clegg, S. & Lounsbury, M. (2009). Weber: Sintering the Iron Cage. Translation, Domination, and Rationality. In P. Adler (ed.), *The Oxford Handbook of Sociology and Organization Studies: Classical Foundations*. Oxford: Oxford University Press.

Cordella, A. & Tempini, N. (2015). E-government and organizational change: Reappraising the role of ICT and bureaucracy in public service delivery. *Government Information Quarterly*, *32*(3), 279-286.

Courpasson, D. (2000). Managerial strategies of domination: Power in soft bureaucracies. *Organization Studies*, *21*(1), 141-161.

Courpasson, D. & Clegg, S. (2006). Dissolving the Iron Cages? Tocqueville, Michels, Bureaucracy and the Perpetuation of Elite Power. *Organization*, *13*(3), 319-343.

Crawford, K. (2014). The Anxieties of Big Data. *The New Inquiry*, 30 May 2014.

Crozier, M. (2010 [1964]). *The Bureaucratic Phenomenon*. New Brunswick & London: Transaction Publishers.

Dixon, P. & Gellman, R. (2014). The Scoring of America: How Secret Consumer Scores Threaten Your Privacy and Your Future. *World Privacy Forum*, April 2 2014.

Douglas, T., Pugh, J., Singh, I., Savulescu, J. & Fazel, S. (2017). Risk assessment tools in criminal justice and forensic psychiatry: The need for better data. *European Psychiatry*, *42*, 134-137.

Downs, A. (1967). *Inside bureaucracy*. New York: Harper Collins Publishers.

Fitzgibbon, D.W. (2008). Fit for purpose? OASys assessments and parole decisions. *Probation Journal*, *55*(1), 55-69.

Gajduschek, G. (2003). Bureaucracy: Is It Efficient? Is It Not? Is That the Question? Uncertainty Reduction: An Ignored Element of Bureaucratic Rationality. *Administration & Society*, *34*(6), 700-723.

Garland, D. (2001). *The Culture of Control: Crime and Social Order in Contemporary Society*. Chicago: University of Chicago Press.
Goel, S., Rao, J.M. & Shroff, R. (2016). Personalized Risk Assessments in the Criminal Justice System. *American Economic Review: Papers & Proceedings*, *106*(5), 119-123.
Gouldner, A.W. (1954). *Patterns of Industrial Bureaucracy*. Glencoe, IL: Free Press.
Habermas, J. (1987). *The Theory of Communicative Action. Volume II: Lifeworld and system: A Critique of Functionalist Reason*. Cambridge: Polity Press.
Hamilton, M. (2015). Adventures in Risk: Predicting Violent and Sexual Recidivism in Sentencing Law. *Arizona State Law Journal*, *47*(1), 1-62.
Hannah-Moffat, K. (2013). Actuarial Sentencing: An "Unsettled" Proposition. *Justice Quarterly*, *30*(2), 270-296.
Harcourt, B.E. (2005). Against Prediction: Sentencing, Policing, and Punishing in an Actuarial Age. Paper presented at the Criminal Justice Roundtable, Harvard Law School, May 13, 2015.
Harcourt, B.E. (2007). *Against Prediction: Profiling, Policing, and Punishing in an Actuarial Age*. Chicago: Chicago University Press.
Harcourt, B.E. (2015). *Exposed. Desire and Disobedience in the Digital Age*. Cambridge: Harvard University Press.
Heckscher, C. & Donnellon, A. (eds.) (1994). *The Post Bureaucratic Organization: New Perspectives on Organizational Change*. London: Sage.
Hodgson, D.E. (2004). Project Work: The Legacy of Bureaucratic Control in the Post-Bureaucratic Organization. *Organization*, *11*(1), 81-100.
Hoogenboom, M. & Ossewaarde, R. (2005). From Iron Cage to Pigeon House: The Birth of Reflexive Authority. *Organization Studies*, *26*(4), 601-619.
Kallinikos, J. (2005). The order of technology: Complexity and control in a connected world. *Information and Organization*, *15*(3), 185-202.
Klagge, J. (1997). Approaches to the Iron Cage: Restructuring the Bars of Weber's Metaphor. *Administration & Society*, *29*(1), 63-77.
Kleiman, M., Ostrom, B.J. & Cheeman II, F.L. (2007). Using risk assessment to inform sentencing decisions for nonviolent offenders in Virginia. *Crime & Delinquency*, *53*(1), 1-27.
Landsbergen, D. (2004). Screen level bureaucracy: Databases as public records. *Government Information Quarterly*, *21*(1), 24-50.
La Porte, T.M., Demchak, C.C. & Jong, M. de (2002). Democracy and Bureaucracy in the Age of the Web. *Administration & Society*, *34*(4), 411-446.
Malleson, N. & Andresen, M. (2015). The Impact of Using Social Media Data in Crime Rate Calculations: Shifting Hot Spots and Changing Spatial Patterns. *Cartography and Geographic Information Science*, *42*(2), 112-21.
March, J.G. & Olsen, J.P. (1989). *Rediscovering Institutions. The organizational basis of politics*. New York: Free Press.
Margetts, H. (1999). *Information Technology in Government: Britain and America*. London: Routledge.
Mayer-Schoenberger, V. & Cukier, K. (2013). *Big Data: A Revolution that will transform how we live, work, and think*. London: John Murray Publishers.
Meier, K.J. & Hill, G.C. (2007). Bureaucracy in the Twenty-First Century. In: E. Ferlie, L-E. Lynn Jr. & C. Pollitt, *The Oxford Handbook of Public Management* (pp. 61-71). Oxford: Oxford University Press.
Merton, R.K. (1940). Bureaucratic Structure and Personality. *Social Forces*, *18*(4), 560-568.
Nilsson, T., Munthe, C., Gustavson, C., Forsman, A. & Anckarsäter, H. (2009). The precarious practice of forensic psychiatric risk assessments. *International Journal of Law and Psychiatry*, *32*(6), 400-407.
OECD (2005). *E-Government for better government*. Paris: OECD.
O'Malley, P. (2010). Simulated Justice: Risk, Money and Telemetric Policing. *The British Journal of Criminology*, *50*(5), 795-807.
Peeters, R. (2013). *The Preventive Gaze: How Prevention Transforms Our Understanding of the State*. The Hague: Eleven International Publishing.
Peters, B.G. (2001). The Politics of Bureaucracy: An Introduction to Comparative Public Administration. London & New York: Routledge.
Peters, B.G. (2003). Dismantling and Rebuilding the Weberian State. In J. Hayward & A. Menon, *Governing Europe* (pp. 113-127). Oxford: Oxford University Press.
Perry, W.L., McInnis, B., Price, C.C., Smith, S.C. & Hollywood, J.S. (2013). *Predictive Policing: The Role of Crime Forecasting In Law Enforcement Operations*. Santa Monica: RAND Corporation.
Popitz, H. (1992). *Phänomene der Macht: Autorität, Herrschaft, Gewalt, Technik*. Tübingen: Mohr.
Rehavi, M.M. & Starr, S.B. (2014). Racial Disparity in Federal Criminal Sentences. *Journal of Political Economy*, *122*(6), 1320-1354.
Reigeluth, T. (2014). Why data is not enough: Digital traces as control of self and self-control. *Surveillance & Society*, *12*(2), 243-254.

Sadin, E. (2009). *Surveillance global*. Paris: Flammarion-Climats.
Schuilenburg, M. (2015). *The Securitization of Society: Crime, Risk, and Social Order*. New York: New York University Press.
Schuilenburg, M. (2016). Predictive policing: de opkomst van een gedachtenpolitie? *Ars Aequi*, *65*(12), 931-936.
Schuilenburg, M. & Peeters, R. (2017). Gift Politics: Exposure and Surveillance in the Anthropocene. *Crime, Law, and Social Change*, *5*(68), 563-578.
Simon, J. (1988). The ideological effects of actuarial practices. *Law and Society Review*, *22*(4), 771-800.
Sjöstedt, G. & Grann, M. (2002). Risk Assessment: What is Being Predicted by Actuarial Prediction Instruments? *International Journal of Forensic Mental Health*, *1*(2), 179-183.
Smith, G.J.D., Bennett Moses, L. & Chan, J. (2017). The Challenges of Doing Criminology in the Big Data Era: Towards a Digital and Data-driven Approach. *The British Journal of Criminology*, *57*(2), 259-274.
Smith, G.J.D. & O'Malley, P. (2017). Driving Politics: Data-driven Governance and Resistance. *The British Journal of Criminology*, *57*(2), 275-298.
Tullock, G. (1965). *The Politics of Bureaucracy*. Washington, DC: Public Affairs Press.
Van Eijk, G. (2017). Socioeconomic marginality in sentencing: The built-in bias in risk assessment tools and the reproduction of social inequality. *Punishment & Society*, *19*(4), 463-481.
Vries, A. de & Smit, S. (2016). Predictive policing: politiewerk aan de hand van voorspellingen. *Justitiële verkenningen*, *3*(42), 9-22.
Wacquant, L. (2004). Pointers on Pierre Bourdieu and Democratic Politics. *Constellations*, *11*(1), 1-15.
Weber, M. (2006a [1921/1922]). *Wirtschaft und Gesellschaft*. Paderborn: Voltmedia GmbH.
Weber, M. (2006b [1904/1905]). *Die protestantische Ethik und der Geist des Kapitalismus*. München: Verlag C.H. Beck.
Williams, M.L. & Burnap, P. (2016). Cyberhate on Social Media in the aftermath of Woolwich: A Case Study in Computational Criminology and Big Data. *The British Journal of Criminology*, *56*(2), 211-238.
Williams, M.L., Burnap, P. & Sloan, L. (2017). Crime Sensing With Big Data: The Affordances and Limitations of Using Open-source Communications to Estimate Crime Patterns. *British Journal of Criminology*, *57*(2), 320-340.
Zouridis, S. (2000). *Digitale Disciplinering: Over ICT, organisatie, wetgeving en het automatiseren van beschikkingen*. Delft: Eburon.
Zuboff, S. (2015). Big other: surveillance capitalism and the prospects of an information civilization. *Journal of Information Technology*, *30*, 75-89.
Zuurmond, A. (1994). *De infocratie: een theoretische en empirische heroriëntatie op Weber's ideaaltype in het informatietijdperk*. Erasmus Universiteit Rotterdam (diss.).