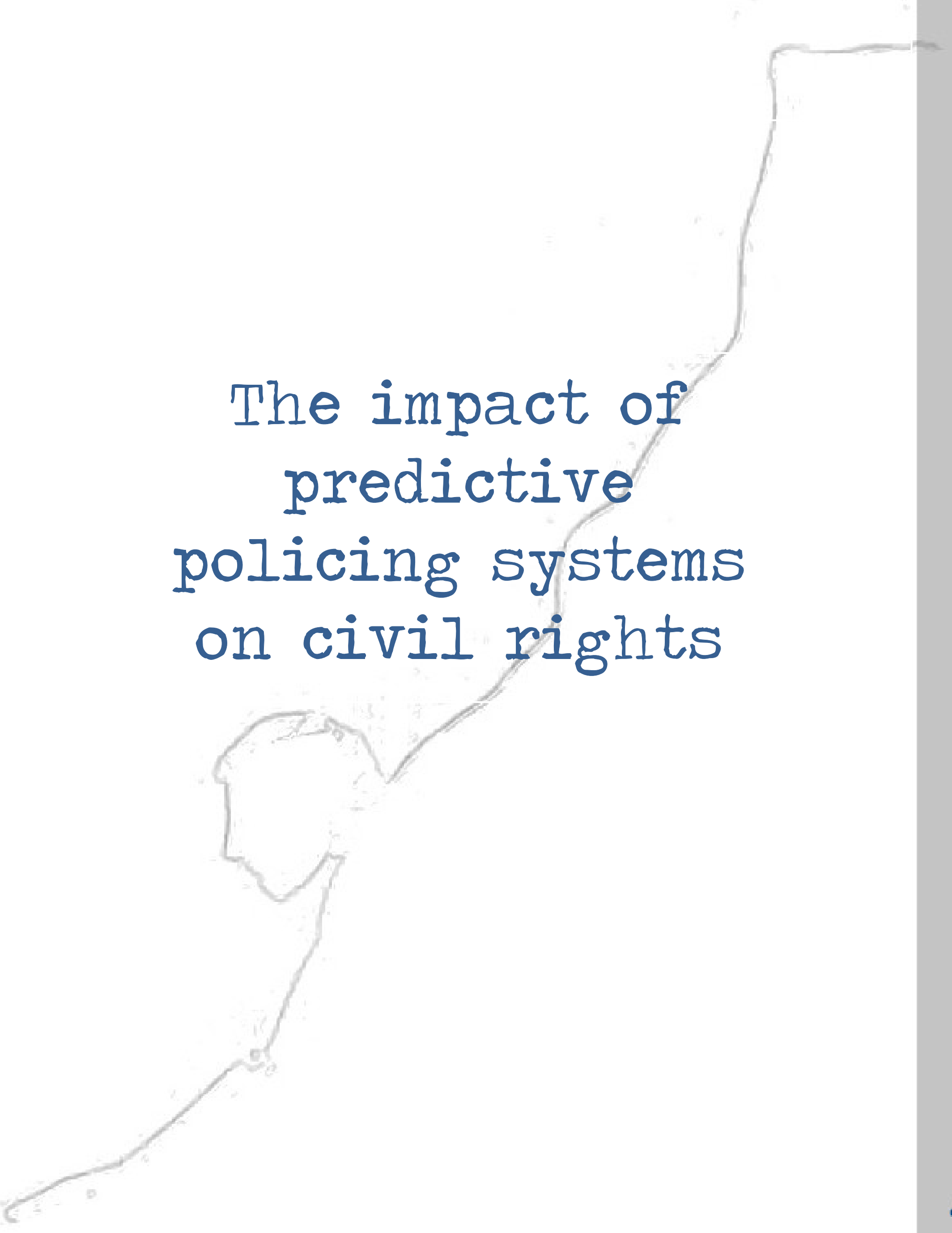
PAPER N. 33

a.a. 2018/2019



The impact of predictive policing systems on civil rights

BEATRICE PEREGO

Trento BioLaw Selected Student Papers

Jean Monnet
EUROPEAN CENTRE
UNIVERSITY OF TRENTO

Co-funded by the
Erasmus+ Programme
of the European Union

UNIVERSITÀ
DI TRENTO
Facoltà di
Giurisprudenza

I paper sono stati selezionati a conclusione del corso *BioLaw: Teaching European Law and Life Sciences (BioTell)* a.a. 2018-2019, organizzato all'interno del Modulo Jean Monnet “BioLaw: Teaching European Law and Life Sciences (BioTell)”, coordinato presso l'Università di Trento dai docenti Carlo Casonato e Simone Penasa.

The impact of predictive policing system on civil rights

*Beatrice Perego**

ABSTRACT: Police departments around the world are embracing new predictive policing technology that will help them spot criminals before a crime ever takes place. However, communities often have little or no idea of why or how this technology is being used, and that raises some important privacy and human rights concerns. This paper wants to provide an overview of predictive policing – what it is, how it is supposed to work, some of the analytic tools used, and three key issues, linked to each other, that must be considered. Since the rise of Big Data and their widespread use in policing, problems are being encountered. That because the check-crossing techniques are able to trace the information back to individuals, rigging the anonymity of the data, and leading to potential stigmatization and discrimination practices. But there's more, the data that are collected could already contain bias, driving the software to mistakes: some reports could be incomplete, some of the demographic information not updated, victims and criminal profiling incomplete or even misleading. This policing practices shape the methodology by which data are collected and processed, stressing inaccuracies and forms of systematic bias integrated in predictive analytics. This leads to consider the accountability of police decisions and point out the need for some combination of transparency and evaluation of the algorithm work, in order to explain police deployment decisions. Finally, the link between unlawful and biased police practices and the data used to implement predictive systems is evident in the study case of the New Orleans Police Department, which is an example where the extensive evidence of dirty policing practices suggests an extremely high risk that dirty data was or will be used in any predictive policing application.

KEYWORDS: Predictive Policing; Civil Rights; Bias; Data; Machine Learning

1. Introduction

Since ancient times, the human being has nurtured the dream of predicting the future. Indeed, the primary goal of Science has always been to understand and explain natural events, in order to control and, possibly, foretell them. In this regard, it is questionable whether prediction techniques can be applied to human behaviour, which has always been kept separated from chemical or physical phenomenon, due to the man's ability to self-determination. The traditional assumption of the impossibility to predict human behaviour, with the same accuracy of the laws of nature, seems to crumble against the advent of new technology and the now widespread use of predictive analytics[1].

** Student at the University of Trento, Faculty of Law.*

[1] Professor Simoncini, by mentioning the Sophocles' tragedy Oedipus Rex -as interpreted by Sabino Cassese- highlights that the dream of knowing the future is inherent to the specific man's nature since antiquity. Both Laius and Oedipus have been questioning the Oracle of Delphi and, right following the response, started to act in a way that fulfill the prophecy. However, the prediction of the oracle doesn't provide any justification, but man's curiosity is such that Oedipus and his father accept it anyway. And by doing so, the man transcends his nature of rational human being. This situation is highly topical when we consider the increasing use of predictive algorithms: such algorithms, like the oracle's prophecies, are obscure, they fail to give an explanation that could be understandable to man (see paragraph 2.2). See: A. SIMONCINI, *L'algoritmo incostituzionale: intelligenza artificiale e il futuro delle libertà,* in *BioLaw Journal – Rivista di BioDiritto,* 2019; S. CASSESE, *Il diritto nello specchio di Sofocle*, in *Il Corriere della Sera*, 18

The term predictive analytics indicates several Machine Learning techniques, including Data Mining[2], as predictive models typically include a machine learning algorithm. The algorithms perform the data mining and statistical analysis, determining trends and patterns in data, so that it could be used to make predictive models. The algorithms are defined as 'classifiers', identifying which set of categories data belongs to. Those algorithms are usually divided into two groups: supervised learning models (also known as predictive modelling) and unsupervised methods (called descriptive modeling)[3]. Both predictive systems work from the inputs that are given to them. The difference between the two methods lies in the fact that in supervised models it is also provided the output, unlike in the case of unsupervised learning. While following a different learning process, both types of predictive analytics, starting from a basis of historical data, shall gross up patterns that facilitate predictions on future behaviours in the context from which the data come.

For organisations overflowing with data but struggling to turn it into useful insights, predictive analytics and machine learning can provide the solution.

At some level, most decision-making systems involve now prediction and the criminal justice system is no exception. Police officers, judges, juries, probation officers, and parole boards all make risk-based assessments every single day. Predictive tools which seek to help make these difficult, life-altering decisions more objective and fair have been embraced throughout the criminal justice system[4].

1.1. The use of predictive algorithm in Law enforcement

Red boxes spread across Google maps of the city, highlighting 500-by-500-square-foot location and police rush to the indicated area to spot criminals before the crime ever takes place. It seems that Minority Report

maggio 2018, https://www.corriere.it/cultura/18_maggio_18/cartabia-violante-il-mulino-saggio-cassese-c1288514-5ab0-11e8-be88 f6b7fbf45ecc.shtml (20/03/2019).

[2] «Data mining is a process of discovering various models, summaries, and derived values from a given collection of data»: M. Kantardzic, *Data Mining: Concepts, Models, Methods, and Algorithms*, *John Wiley & Sons*, 2011, 6.

[3] In supervised learning it is provided to the software a number of examples of inputs and the respective output, so the algorithm has to identify and learn the rules that connect them, in order to draw up the output for the new inputs. Classification and regression are examples of supervised models, their goal is to assign new inputs to classes that are already known. The most widely used predictive modeling techniques (supervised) are three: (1) decision trees are classification models that partition data into subsets based on categories of input variables. (2) regression, which finds key patterns in large data sets and is often used to determine how much specific factors influence the movement of an asset. (3) neural networks that handle nonlinear relationships in data, which is increasingly common as we collect more data. The unsupervised learning model works in a different way: in this case the algorithm has to group a variety of inputs into unknown classes, learning on its own the recurrent patterns. See: A. Rezzani, *Big Data Analytics. Il manuale del data scientist*, *Apogeo Education*, 2017; D. Abbott, *Applied Predictive Analytics: Principles and Techniques for the Professional Data Analyst*, *John Wiley & Sons*, 2014; https://www.sas.com/en_gb/insights/analytics/predictive-analytics.html#dmtechnical (13/04/2019).

[4] See *Jurek v. Texas*, 428 U.S. Supreme Court 262, 1976: «Prediction of future criminal conduct is an essential element in many of the decisions rendered throughout our criminal justice system. The decision whether to admit a defendant to bail, for instance, must often turn on a judge's prediction of the defendant's future conduct. And any sentencing authority must predict a convicted person's probable future conduct when it engages in the process of determining what punishment to impose. For those sentenced to prison, these same predictions must be made by parole authorities».

fiction[5] has become, somehow, reality. Predictions are drawn up by algorithms insert in software like Pred Pol, used all around the world, from New York to Milan, and that is basis of “predictive police”.

Predictive police describes a system that analyzes available data to predict either where a crime may occur or who could be involved. Crime often seems random, but it follows patterns. So the question is whether we can build mathematical structures to identify such patterns, and the answer is yes, thanks to machine learning techniques the software make predictions based on probabilistic calculations that gives results from a huge amount of data stored in the software. The goal is to identify those areas where certain types of crimes are frequent and prevent their reoccurrence. The information used by the algorithms change depending on the company that creates the software: PredPol uses years of crime data to establish these patterns and then the algorithm uses near real-time crime data to predict the next property crime. Other systems[6] use even more esoteric data - from the weather to phases of the moon - to arrive at their crime forecasts.

Are these predictive policing systems actually effective? A study conducted by the University of California seems to respond positively, pointing out that in those cities in which PredPol is used, crimes has reduced on average 7,4%[7]. Nevertheless, the results of the study have been criticized by some researchers of the University of Grenoble, because the study was carried out, among others, by the two co-founders of PredPol[8].

[5] Minority Report is a 2002 American science fiction action film directed by Steven Spielberg and based on the short story *The Minority Report* by Philip K. Dick: in the year 2054, a specialized police department (PreCrime), thanks to the foreknowledge provided by the called “pregogs”, which are three psychics, arrest criminals before crimes take place. One of the main themes faced by the movie is the role of preventive government activity in protecting its citizens, in a future state based on technology. In particular such technological developments make the presence of the government nearly boundless.

[6] Risk terrain modeling can be applied to any geographic extent (i.e. local, regional, global; urban, suburban, rural; land, sea). It can be used to analyze almost any topic. The topic data represents the problem or issue that you will analyze(e.g., incident locations of robberies, traffic crashes, or drug overdoses).Use data that is representative of the entire study area (e.g., convenience stores, gas stations, bars, parking lots, schools...). http://www.riskterrainmodeling.com/how-to-rtm.html (15/04/2019). HunchLab it’s probably the more sofphisticated system becouse it's not just about anticipating crime, it's about figuring out the best way to respond. Policing tactics should not only be effective, but also reflect the community's priorities. HunchLab provides features that: (1) align patrol activities with the priorities of the community, (2) intelligently allocate resources to prevent over-policing, and (3) determine which tactics work and which don't. https://www.hunchlab.com/ (15/04/1019). It is important to underline that none of this programs use personally identifiable information of the criminals, in order to «eliminate the possibility for privacy or civil rights violations seen with other intelligence-led or predictive policing models», as it's said in their websites.

[7] See: G.O. Mohler, M.B. Short, S. Malinowski, M. Johnson, G.E. Tita, A.L. Bertozzi, P.J. Brantingham, *Randomized Controlled Field Trials of Predictive Policing,* in *Journal of the American Statistical Association,* 2015; http://newsroom.ucla.edu/releases/predictive-policing-substantially-reduces-crime-in-los-angeles-during-months-long-test.

[8] «Nous avons réalisé une analyse critique du système commercial de prédiction policière dénommé Predpol visant à évaluer objectivement l’efficacité des prédictions réalisées. Après avoir effectué un travail de recoupement d’informations, nous avons recueilli les données nécessaires pour répondre à notre problématique. Nos simulations numériques ont pu mettre en évidence que, sur l’étude concernant la prédiction des homicides et délits avec arme à feu à Chicago, les résultats parus dans l’article (Mohler, 2014) ne sont pas probants. En effet, avec un algorithme basique de “prédiction par meilleur rang avec points chauds dynamiques” nous obtenons des résultats équivalents. Une valeur du pAUC (Aire partielle sous la courbe) standardisée donne 0.559±0.01 pour notre étude versus 0.557±0.01 (Mohler, 2014, Fig. 1, numérisée par OPR) pour la leur, mettant in fine en question la plus-value prétendue par G. Mohler et Predpol Inc», I. Benslimanè, *Étude critique d’un système d’analyse prédictive appliqué à la criminalité : Predpol, Université Joseph Fourrier – Grenoble*, 2014. available in https://cortecs.org/wp-content/uploads/2014/10/rapport_stage_Ismael_Benslimane.pdf (11/04/2019).

However, concerns are not just related to the effective results, but refer mainly to the way data are collected and used.

Prediction has always been part of policing. Police officers regularly predict the places and persons involved in criminal activity and seek to deter this pattern of lawbreaking. The novelty in predictive policing techniques lies in the tools that law enforcement can use.

The first development is due to the environmental criminology and the early experiments that studied the geography of crime[9]. Such experiments informed police practice as crime mapping became a way to identify and study patterns of criminal behavior. As data collection and data analysis grew more sophisticated, new predictive techniques and computer-mapping technologies also developed to make use of the information.

This version of predictive policing started to be employed in the mid-2000s[10]: the idea was to use forecasting data analysis tool, that at first consisted in computer-augmented hotspot systems, in order to anticipate, prevent and reduce criminal activity. Everything from data mining to geospatial prediction is – and still is - perceived as a powerful instrument that could be used by law enforcement to deploy resources more effectively and respond to crimes in a more appropriate way.

So, in its first iteration, predictive policing was seen as just an extension of what law enforcement had already been doing for decades and therefore it didn't get particular attention from a privacy perspective[11].

Nevertheless something started to happen when police departments from around the world realize that they could use, in their new policing systems, non-traditional data and a huge amount of information which could be at their disposal. The "Big Data"[12] trend was ongoing and, for example, the social network analysis suddenly was seen as an extremely effective tool in the hand of law enforcement: by discovering whom

[9] The 1980s and early 1990s were a rich period for the geography of crime, with several important books such as D.J. EVANS, D.T. HERBERT, *The Geography of Crime*, *Routledge,* 1989; D.J. EVANS, N.R. FYFE, D.T. HERBERT, *Crime, Policing and Place*, *Routledge*, 1992; D. T. HERBERT, *The geography of urban crime*, *Longman*, 1982. See also D.K, ROSSMO, *Geographic profiling: target patterns of serial murderers*, *Simon Fraser University*, 1987. available in https://core.ac.uk/download/pdf/56371040.pdf: This study had three specific purpose which were, in increasing order of importance: (1) to add the knowledge of serial murder; (2) to examine the geography of serial murder at both macro and microlevels; and (3) to construct a method for determining the most probable area of offender residence from the location of the crimes. The dissertation research has subsequently led to the development pf geographic profiling, an information management strategy for serial violent investigation, at iii.

[10] For example the Santa Cruz Police Department (SCPD) was one of the first in the U. S. to employ predictive policing in its daily operations. The software in use was developed by researchers at the University of California, Los Angeles, and Santa Clara University, with input from crime analysts from SCPD. The program was first implemented in July 2011. In July 2012, the program moved from its experimental phase into full operational use. http://www.cityofsantacruz.com/government/city-departments/city-manager/community-relations/city-annual-report/march-2012-newsletter/predictive-policing.

[11] For example, police departments around the world know that certain events – such as demonstrations, New Year's Eve celebrations, political rallies and other events – can require additional policing efforts. So why not make use of data already on hand to identify potential hot spots and prevent crime before the trouble come to pass?

[12] Nowadays there are several definitions of Big Data, depending on the topic we associate the term to. Nevertheless it is possible to identify some common features in every definition that has been given, in order to draw up a consensual definition: «Big Data represents the Information assets characterized by such a High Volume, Velocity and Variety to require specific Technology and Analytical Methods for its transformation into Value.» A. DE MAURO, M. GRECO, M. GRIMALDI, *What is Big Data? A Consensual Definition and a Review of Key Research Topics*, *in ResearchGate*, 2014, 8.

criminals were talking to on social media, police officers could start to piece together some very complex criminal networks. Furthermore, big data capabilities make it possible to develop individual profiling based on past criminal activity, current association and other factors that are related to social danger and criminal inclination. That's the reason why law enforcement, as part of a large push toward proactive policing, is switching surveillance and investigation resource to focus on prediction.

One of the biggest test cases took place in New Orleans, where predictive policing, thanks to the sophisticated data mining tools from Silicon Valley’s Palantir, was able to uncover links between gang members, outline extensive criminal histories, and even those individuals who might become future gang members. That was certainly a breakthrough, and the NOPD won acclaim for its policing efforts, but the use of Palantir has also risen serious concerns: first of all, neither the city council members or members of the local community were aware of the collaboration between NOPD and Plantir, so there was no public vetting of the problem. Moreover, it was pointed out that Palantir predictive policing system tended to have an outsized impact on specific sections of the population – such as black people or LGBT communities – reiterating historical biased and discriminatory practices of the police of New Orleans[13].

2. Issues: how predictive analytics impact on Civil Rights

Predictive policing provides tools that have the potential to influence developments in data science in order to improve the effectiveness and efficiency of law enforcement agencies. However, this ultimately depends on humans understanding the limitations and assumptions embedded in the software and ensuring that ineffective and unjust outcomes are avoided. Fist concerns are related to the way data are collected, with specific reference to every individual's right to privacy: the difficulty in obtaining usable, accurate, and clean data to integrate into a predictive system exposes a massive vulnerability. Any data-driven system risks being undermined by bad data. Such information includes flaws, fragmentation, and the internal and external pressures to collect vast amounts of information constantly, instantaneously, and without adequate financial resources to ensure accuracy. Secondly, it is not clear how algorithm works because it does not provide a transparent explanation of the output it processes. This implies that the use of predictive software can undermine the ability for individual officers or law enforcement agencies to give an account of their decisions in important ways. From the type of data that is entered into the software and their processing derive what are the actual police operations, which may prove discriminatory. The assumptions behind predictive technologies are affected by unseen influences that may have unintended and discriminatory consequences. First, the data itself can be the result of biased collection and, in fact, implicit bias has been demonstrated to impact policing decisions on the street. In particular, the targeting

[13] See chapter 3.

of certain areas or certain races creates the impression of higher crime rates in those areas, which then justifies continued police presence there[14].

2.1. Data-gathering and Right to Privacy

The use of techniques that involve the collection of high volume of data, their storage, the cross-checking of database and, more specifically the systematic profiling in the interest of predictive policing, produce potential interference in the right to privacy and the treatment of personal data. In particular, the term profiling refers to the cross-checking, through algorithms, of data collected from various sources , in order to predict the occurrence of crimes and their location (crime hotspot), or the drafting of individual criminal profile ("predictive composite")[15]. Database developed by law enforcement or acquired by data-brokers, social networks and internet are some of the sources that are used in predictive policing. The adverse effects on the protection of human rights connected to predictive policing are considerable: firstly, the identification of categories of subjects with different levels of social danger ("social sorting") implies clear risks of stigmatization and discrimination. Second, as recorded by the European Parliament in 2017, the possibility of false-positive due to the inaccuracy of processed data, cannot be overlooked: «low-quality data and/or low-quality procedures behind decision-making processes and analytical tools could result in biased algorithms, spurious correlations, errors, an underestimation of the legal, social and ethical implications»[16].

It should, however, be pointed out that the techniques under examination should be seen in the broader context of national surveillance for the purpose of maintaining public safety: the right of the individual to privacy of personal data must therefore be balanced with the public interest in security. Furthermore, what should serve as guarantees for the individual, is the so-called pseudoanonymisation, that consist in«the processing of personal data in such a manner that the personal data can no longer be attributed to a specific data subject without the use of additional information, provided that such additional information is

[14] All the issues mentioned will be explored in the following paragraphs 2.1, 2.2 and 2.3 that refer, respectively, to privacy, algorithm transparency and discrimination concerns. Also see H.V. JAGADISH, *The Promise and Perils of Predictive Policing Based on Big Data*, in *The Conversation*, 2015 http://gizmodo.com/the-promise-and-perils-of-predictive-policing-based-on-1742958281 (4/05/2019); R. VAN BRAKEL, P. DE HERT, *Policing, surveillance and law in a pre-crime society: Understanding the consequences of technology based strategies*, in *Researchgate*, 2011, 163 192.

[15] See: A. BABUTA*, Big Data and Policing. An Assessment of Law Enforcement Requirements, Expectations and Priorities*, in *Royal United Services Institute for Defence and Security Studies*, 2017; M. MENDOLA, *One Step Further in the 'Surveillance Society': The Case of Predictive Policing*, *Tech and Law Center*, 2016.

[16] European Parliament, *Report on fundamental rights implications of big data: privacy, data protection, nondiscrimination, security and law-enforcement*, 2016/2225(INI), *Committee on Civil Liberties, Justice and Home Affairs*, 2017, par. M.

kept separately and is subject to technical and organizational measures to ensure that the personal data are not attributed to an identified or identifiable natural person»[17].

Nevertheless, the pseudoanonymisation – as well as in the case of anonym data- does not guarantee absolute and definitive anonymity of personal data[18], as there is the possibility of a re-identification of individuals through the comparison of various types of anonym data. The consequence is clear: at present, the re-identification process that deliver data interfere with individuals' information privacy rights. Also the European Court of Human Rights has pointed out the risks of stigmatization linked to the storage of personal data and information. In particular, the historical judgment of the Court of 4 December 2008 in case *S. and Marper v. The United Kingdom*, highlights that «the mere storing of data relating to the private life of an individual amounts to an interference within the meaning of Article 8 [...] stemming from the fact that persons […] who have not been convicted of any offence and are entitled to the presumption of innocence, are treated in the same way as convicted persons.[...] the perception that they are not being treated as innocent is heightened by the fact that their data are retained indefinitely in the same way as the data of convicted persons, while the data of those who have never been suspected of an offence are required to be destroyed»[19].

Despite the alleged anonymity of the data and their collection, the use of big data, as demonstrated, is likely to cause violations of individual rights related to the potential re-identification of their holders. That happens because database crossings or re-identification techniques can trace the information back to a single individual or a criminal profile. This is the reason why predictive policing techniques should be conducted with «greater algorithmic accountability and Transparency»[20].

[17] Directive (EU) 2016/680 of the European Parliament and of the Council, art. 3, par. 5. The directive is about the protection of natural persons with regard to the processing of personal data by competent authorities for the purposes of the prevention, investigation, detection or prosecution of criminal offences or the execution of criminal penalties, and on the free movement of such data, and repealing Council Framework Decision 2008/977/JHA : https://eur-lex.europa.eu/legal-content/ENIT/TXT/?uri=CELEX:32016L0680&from=FR.

[18] P.G. Della Morte, *Big data e protezione internazionale dei diritti umani. Regole e conflitti*, 2018, 156 ss.

[19] In England, Wales and Northern Ireland, since 2004, any individual arrested for any recordable offence has had a DNA sample taken and stored as a digital profile in the National DNA Database. Even if the individual was never charged, if criminal proceedings were discontinued, or if the person was later acquitted of any crime, their DNA profile could nevertheless be kept permanently on record. The majority of the Council of Europe member states allow the compulsory taking of fingerprints and DNA samples in the context of criminal proceedings; however the United Kingdom (specifically, England, Wales and Northern Ireland) was the only member state that expressly permitted the systematic and indefinite retention of such DNA profiles. In contrast, DNA samples taken in Scotland when individuals are arrested must be destroyed if the individual is not charged or convicted. The case involved two claimants from Sheffield, England: Mr. S. and Michael Marper. Mr S. was arrested on 19 January 2001 at the age of eleven and charged with attempted robbery. His fingerprints and DNA samples were taken. He was acquitted on 14 June 2001. Michael Marper was arrested on 13 March 2001 and charged with harassment of his partner. His fingerprints and DNA samples were taken. The charge was not pressed because Marper and his partner became reconciled before a pretrial review had taken place. *S. and Marper v. the United Kingdom*, CEDU, ric. 30562/04 and 30566/04, 2008. See also: I. Voinamotoc, *La génétique et l'article 8 de la CEDH: la généalogie de Marper c Royaume Uní dans le droit international*, in J. Casadevall, G. Raimondi, E. Fribergh, P. Titiun, P. Kempees, J. Darcy, *Mélanges en l'honneur de Dean Spielmann: liber amicorum Dean Spielmann*, *Oisterwijk*, 2015, 399 ss.

[20] European Parliament, *Report on fundamental rights implications of big data: privacy, data protection, nondiscrimination, security and law-enforcement*, cit., par. 8.

2.2. Data-processing and algorithm transparency

The value of transparency for predictive policing is extremely significant. On one hand, it is important that police officers can understand the information they receive, so they can act on it appropriately, and judges can determine whether or not an officer's reliance on that information was reasonable. On the other hand, the local community can hold police officers to account for their strategy, but when the decision has been made by a predictive policing algorithm, this open dialogue and debate is particularly difficult. In this regard, it is questionable whether public authorities should be required to include appropriate information about the use of algorithmic decision-making tools in order to encourage such information to be provided publicly and proactively – and set expectations for the private companies that provide these programs[21]. Transparency is essential for predictive policing to be used in an effective, legal, and ethical way that does not eviscerate the reasonable suspicion standard[22]. Nevertheless, such transparency surrounding the use of predictive policing algorithms is widely lacking. Some companies, claim the right to keep the code to themselves, powering their algorithms as trade secrets[23].

The technical reason of such lack of transparency is linked to the black box phenomenon, and so, in many cases where algorithm-based technologies are deployed, the police officers involved in operating the programs won't have any comprehension of how the algorithm works. Even for computer programmers, algorithmic decision making can be difficult to decipher: having access to the algorithm used does not mean there will be a greater degree of transparency because explaining how data is then used and how the algorithm works remain an incredibly difficult question.

If no one can explain how a decision has been made, it is impossible to challenge it properly[24].

This is a serious problem given that, where the state restricts a person's rights, there must be legal basis for that infringement in order to protect people from being treated differently based on arbitrary or illogical factors, such as where they live or their race. In fact, the underlying data is usually collected from public sources, and law enforcement officers have a significant amount of freedom in acquiring this information, which means that they can obtain the predictions without facing any legal standard. That's the reason why

[21] R. Brauneis, E.P. Goodman, *Algorithmic Transparency for the Smart City*, In *SSRN*, 2017.

[22] *See* T. Z. Zarsky, *Transparent Predictions,* in *U. Ill. L. Rev.*, 1503, 2013 (describing the increased implementation of data mining in government processes, the lack of transparency therein, and proposing corrections).

[23] E.E. Joh*, The Undue Influence of Surveillance Technology Companies on Policing,* in N.Y.U. L. Rev., 2017*,* https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2924620: «police department may rely increasingly on big data tools, they do not create them. The police are costumers who contract with private vendors» which guard their algorithms as trade secret.

[24] The House of Lords Select Committee on Artificial Intelligence concluded that «achieving full technical transparency is difficult, and possibly even impossible, for certain kinds of AI [artificial intelligence] systems in use today […]. We believe it is not acceptable to deploy any artificial intelligence system which could have a substantial impact on an individual's life, unless it can generate a full and satisfactory explanation for the decisions it will take». Oswald and Urwin, Written evidence submitted to the House of Lords Select Committee on Artificial Intelligence, 2018 Available at: http://data.parliament.uk/writtenevidence/committeeevidence.svc/evidencedocuument/science-and-technology-committee/algorithms-in-decisionmaking/written/69002.html (27/ 03/ 2019).

investment in policing should focus on the development of programs and algorithms that are able to reduce biased approaches, work to address wider and underlying issues of bias that permeate the criminal justice system. The development and trial of such programs, as well as their evaluation, should be supervised by independent experts and conducted entirely independently. Furthermore, considering the impact that such policing systems have on human rights, the outcome should made public[25].

However, while transparency can aid in preventing discrimination and stigmatization- such as through the programmer's choice of variables, the methods of data collection, or police reliance on the programs- is also unlikely to correct flailing in application. So, in order to prevent a severe impact on the protection of individual rights it is necessary for predictive policing programs to be implemented in a transparent way, but that is not enough. If the use of these algorithms is transparent, but does not lead to the correction of encoded bias in the data or the use of "dirty" information, transparency is fairly hollow as an institutional principle.

2.3. Bias and potential discriminatory practices

As police officers use the information as developed by predictive algorithms more often, judges will begin to expect this kind of hard data and may begin to reject the current subjective, experiential, or anecdotal evidence that officers currently rely upon[26]. However, neither police activity or judgments based purely on pre-crime predictions is likely to happen any time soon. That because, first of all, the law requires police officers and judges to act on facts that are specific to the case under investigation and therefore the probabilistic and general nature of the information used by big data may not be able to provide this specificity. Also, predictive algorithms may use factors that courts can't consider in their judgments, such as the race or the sexual orientation of the subject[27]. Moreover, the underlying data that the algorithms process may be in itself biased and so, using algorithms wouldn't actually increase accuracy but merely reinforce years of discriminatory policing: basically, if the underlying data is discriminatory, then the results that are based on that data will be discriminatory, and the supposedly "neutral" algorithms will be doing nothing more than tightening up the existing bias in the criminal justice system. Discriminatory searches, stops, arrests, and convictions will become the basis of the city's predictive algorithms, which creates two distinct problems. The first lies on the fact that crime data is notoriously incomplete. Certain crimes - like

[25] Police departments should tell the public which predictive systems they use, by what criteria they chose them and how they evaluate them. Moreover, each law enforcement agency shall allow individuals that report police misconduct or defendants to access documentation relevant to the individual's stops, adopt specific measures for preventing unauthorized access or release of Data and comply with the public disclosure laws of the State.

[26] A.G. Ferguson, *Crime Mapping and the Fourth Amendment: Redrawing "High-Crime Areas"*, in *Hastings L. J.*, 2011, 221-22: «If the officer did not base his decision on specific data about a specific crime problem in a specific area, or if the data relied upon did not demonstrate a specific and relevant crime problem, then reliance on this information should not be considered».

[27] In this regard, some scholars argue that many of the risk prediction factors currently in use in sentencing decisions may be unconstitutional because they rely directly or indirectly on race or other suspect classes. S.B. Starr, *Evidence-Based Sentencing and the Scientific Rationalization of Discrimination*, in *Stanford Law Review,* 66, 2014, 803.

murder, burglary, and auto theft- tend to be frequently reported to authorities, while other crimes - such as sexual assault, domestic violence, and fraud - tend to be underreported[28]. As to deficiency, the fragmented nature of crime data on the national and local level makes reliance on it arguable. The second problem is related to the fact that when an algorithm determines whether a neighborhood is a "high crime area", it will have a skewed interpretation of the frequency of crimes in different areas. This will lead to focus the intervention of law enforcement in certain arears then elsewhere, which will create a self-fulfilling prophecy as more individuals are stopped, searched, arrested, and thus convicted in those areas. It is called "ratchet" effect[29]: if certain factors are already perceived as leading to higher levels of criminal activity, a predictive algorithm will lead authorities to conduct and authorize more searches on areas and suspects who meet these factors, leading to more arrests that are linked to those circumstances. This will create high rate of police operations toward certain individuals or areas of a city – usually the ones where minorities and immigrants live – and disproportionately small amounts of data about other individuals or areas[30].

All of these objections can be overcome only if the algorithms and software used by the big data analyses are made more transparent so that it is possible to evaluate the underlying processes and the standards being used.

3. Case study: The investigation on the New Orleans Police Department

The Department of Justice investigated the New Orleans Police Department (NOPD) twice. The first investigation began in 1996 focusing on a wide-range of police misconduct, but it ended without a consent decree because the NOPD committed to turn over a new leaf and reform itself[31].

[28] For example the DOJ's Bureau of Justice Statistics says in its National Criminal Victimization Survey (NCVS) that in 2016, «U.S. residents age 12 or older experienced 5.7 million violent victimizations». A majority of those crimes, some 58 percent the DOJ says, were never reported to police. https://www.bjs.gov/content/pub/pdf/cv16.pdf; Also see: https://ec.europa.eu/eurostat/statistics-explained/index.php?title=Crime_statistics (14/ 04/ 2019).

[29] B.E. Harcourt, *Against prediction: profiling, policing, and punishing in an actual age,* in *University of Chicago Press*, 2006, 145 ss.

[30] B.E. Harcourt, *Against Prediction: Sentencing, Policing, and Punishing in an Actuarial Age,* in *University of Chicago Public Law & Legal Theory Working Paper,* 94, 2005, p. 28 : «And the fact is, given the paucity of reliable information on natural offending rates, law enforcement relies heavily on arrest, conviction, and supervision data in deciding how to allocate resources. This, in turn, accelerates the imbalance in the prison population and acts like a ratchet [...]. The reason, in essence, is that when we profile, we are essentially sampling more from a higher-offending population. Instead of sampling randomly – which would net a proportional representation of the offending population – we are sampling in greater numbers from the pool of higher offenders, and thereby skewing our sample results. Somewhat counter-intuitively, the only way to produce a prison population that mirrors the offending population is to sample randomly from the general population – to engage in essentially random searches, or random audits, or random policing. Barring that arbitrariness, our results will be distorted. What the ratchet effect does is to disproportionately distribute criminal records and criminal justice contacts with terrible effects on the profiled population. Disproportionate criminal supervision and incarceration reduces work opportunities, breaks down families and communities, and disrupts education. It contributes to the exaggerated general perception in the public imagination and among law enforcement officers of the criminality of the targeted group».

[31] A. Johnson Jr., *What the Studies Said,* in *New Orleans Magazine,* 2011. http://www.myneworleans.com/New-Orleans-Magazine/May-2011/WHAT-THE-STUDIES-SAID/. In particular it was said that «The federal government delayed their "takeover"

However, In 2010 at the invitation of the Mayor of the city, the Department reopened its investigation of NOPD reviewing records between 2005 and 2011 pointing out that the measures taken by NOPD , inter alia, addressed bias-based profiling and other discriminatory policing on basis of race, national origin, and LGBT status[32]. Indeed, in 2009 arrest data provided by NOPD indicates that, against the detention of 500 African-American males under the age of 17, only eight white males of the same age were taken into custody. The same situation was experienced by females in this same age group. In terms of arrest rates, for both African-American males to white males, and African-American females to white females, it was nearly 16 to 1[33]. Furthermore, the report documented evidence of "dirty data.". In particular, it observed that the arrests and Field Interview Cards[34] data perpetrated by the NOPD where characterized by disparities and inconsistencies, and it questioned NOPD policies that encouraged unwarranted and potentially privacy-violating data collection[35]. In order to support criminal investigation, strategic homicide-reduction strategies, and to obtain indictments by employing, in 2012, the NOPD started a partnership with Palantir[36]. Leaving aside the lack of transparency of the collaboration[37], that has led to the annulment of the agreement in 2018, what is interesting to underline is that the system's analysis throws back similar racial disparities and other biases of the NOPD's practices and policies. In particular, the documents highlighted that the system associated violent or gang crimes to «overwhelmingly young, African American, male,

and gave the opportunity to reform itself» over the next decade. «Last week the Department of Justice gave the NOPD a "vote of confidence" in how the department was managed and how its officers treat citizens».

[32] More to the point, in the report it's said that: «We find reasonable cause to believe that there is a pattern or practice of unconstitutional conduct and/or violations of federal law with respect to discriminatory policing. NOPD personnel at all levels of the Department not only acknowledged that the community perceives racial and ethnic profiling as a significant problem, but some also expressed their own belief that such discriminatory conduct occurs. Both bias and the perception of bias erode citizens' inclination to trust and cooperate with law enforcement, impeding effective and safe policing. Although both community members and officers told us that this dynamic is clearly at work in New Orleans, the Department has failed to respond with systems to prevent, detect, and respond to discriminatory policing, and to ensure that police officers are conducting themselves in accordance with constitutional guarantees of equal protection. The Department's inadequate policies and training in conducting proper stops, searches, and arrests increase the likelihood that officers, without sufficient understanding of how to identify and articulate suspicion based on behavior and other permissible factors, will instead rely on inappropriate stereotypes and bias in their decision-making.» See: U.S. Department of Justice Civil Rights Division, Investigation of the New Orleans Police Department, 2011. https://www.justice.gov/sites/default/files/crt/legacy/2011/03/17/nopd_report.pdf.

[33] See *supra note* 6, at ix

[34] A Field Interview (FI) card is a method of documenting informal police contacts during the course of patrol. They may also be known as contact cards, information or interview cards, or any other name, depending on the jurisdiction.

[35] For example, the report noted that in 2009, when NOPD arrest data was compared to national averages, «the level of disparity for youth in New Orleans is so severe and so divergent from nationally reported data that it cannot plausibly be attributed entirely to the underlying rates at which these youth commit crimes, and unquestionably warrants a searching review and a meaningful response from the Department.». Additionally, the Department of Justice expressed concerns regarding omissions of essential information noting that the NOPD «policies and practices for complaint intake do not ensure that complaints are complete and accurate, systematically exclude investigation of certain types of misconduct, and fail to track allegations of discriminatory policing». See *supra note* 31, at ix and xvii.

[36] Palantir Gotham integrates and transforms data, regardless of type or volume, into a single, coherent data asset. As data flows into the platform, it is enriched and mapped into meaningfully defined objects — people, places, things, and events — and the relationships that connect them. https://www.palantir.com/palantir-gotham/, (09/04/2019).

[37] In fact, the program escaped public notice, partly because Palantir established it as a philanthropic relationship with the city through Mayor Mitch Landrieu's signature NOLA For Life program. Thanks to its philanthropic status, as well as New Orleans' "strong mayor" model of government, the agreement never passed through a public procurement process. https://www.theverge.com/2018/2/27/17054740/palantir-predictive-policing-tool-new-orleans-nopd (09/04/2019).

undereducated, and underemployed», which is the same population that was improperly targeted by NOPD practices and misreported in NOPD data[38]. Even though this correlation could be explained in other ways, the persistent use of biased practice and distortions in NOPD data suggests some level of attribution.

This case study demonstrates that when dirty data is fed into a predictive system, which should be "neutral", it can be easily contaminated by them and ingrain biases.

That happens because there is no independent authority[39] that, potentially, could address and control the policies and activities of the department as well as the subsequent data produced through these practices. But identifying unlawful and biased practices is not enough. The data collection, analysis, and use practices must be reformed as well in order to apply restrictions or prohibitions on the use of the historical data generated by unlawful and bias practices and that because it is necessary to avoid the perpetuation of such practices through the systems that rely on dirty-data.

4. Conclusion

The last twenty years have seen an extensive use of data-driven policies, practices, and technologies in the public sector, in order to reduce the reliance on subjective factors, and to react in a more objective way to social, economic and political issues. However, increasing reliance on data presents serious risks to fairness, equity, and justice, if a careful monitoring is not given to the practices underlying the creation, auditing and maintenance of data.

The case study I mentioned in the previous paragraph, highlights the risks and consequences associated with overconfidence on unaccountable and potentially biased data to address sensitive issues like public safety. The NOPD case shows that, in absence of an independent control, the "dirty-data" that are collected through illegal police practices are still used for law enforcement and other purposes, which could create deep-rooted consequences that will permeate throughout the criminal justice system and society more widely.

First of all, since it seems unlikely that police departments, in absence of explicit requirements and incentives, will self-monitor and reform those activities that create biased data, it is needed an external supervision by an independent authority[40]. If big data and predictive analysis are a very useful tool for the

[38] S. SHIRMER, *Deploying Palantir Gotham in New Orleans,* 2014*,* https://assets.documentcloud.org/documents/4344815/Nola-hc3-Final-20140403 (12/04/2019); A. WINSTON, *Palantir has secretly been using New Orleans to test its predictive policing technology,* in *Verge,* 2018, https://www.theverge.com/2018/2/27/17054740/palantir-predictive-policing-tool-new-orleans-nopd (25/03/2019); Palantir Techs., *NOLA murder reduction: technology to strategies*, 2014, https://www.documentcloud.org/documents/4344816-NOLA-Murder-Reduction-White-Paper.html (12/04/2019): describing the Palantir and City of New Orleans partnership to identify «individuals exhibiting the highest predictors of violence».

[39] About the necessity of such authority see chapter 4.

[40] The term independent authorities indicates those public entities or entities, established by law, which exercise mainly administrative functions in areas considered sensitive or of high technical content (competition, privacy, communications, etc.)such

repression of crimes, it emerges the necessity to find a balance between the need for effective law enforcement and crime prevention and the rights of the individual[41]. However, since privacy law has been generally unable, so far, to deal with the concept that «thousands of small acts of data gathering - each individually unharmful, authorized by the user, or gathered by different parties - may in their total, quantitative volume create a privacy violation[42]» a new notion of "quantitative privacy" might take hold considering the development of new technologies and big data. That's the reason why, in my opinion, it is necessary to draw up a specific regulation[43] for this particular investigation technique, as well as an independent authority that monitor, in the first place, the acquisition of data[44], authorize its use and to verify, in the long term, police operations. Furthermore, where appropriate, such authority should impose penalties in the event that the rules of law are not observed or when detect practices which are discriminatory – or in any way detrimental – to the individual's constitutional rights.

It is therefore appropriate to draw up technical measures that are able to guarantee algorithmic accountability.

and transparency in order to avoid negative consequences with regard to the right to privacy and discrimination issues[45].

It is therefore appropriate to draw up technical measures that are able to guarantee algorithmic accountability and transparency in order to avoid negative consequences with regard to the right to privacy and discrimination issues. Only the respect of these safeguards can remove, or at least reduce, the compliance of predictive policing practices with the protection of civil rights. That would legitimate the use, by law enforcement, of powerful tools that undoubtedly help the preservation of public safety.

as to require a particular position of autonomy and independence against the Government, in order to ensure greater impartiality (cd. neutrality) with respect to the interests involved. See G. Falcon, *Lezioni di diritto amministrativo*, *Cedam,* 2016. About the character of independence see G. Napolitano, *Autorità indipendenti e agenzie amministrative*, in M. Clarich, G. Fonderico, *Dizionario di diritto amministrativo*, *Il Sole 24 Ore*, 2007, 87 ss.; R. Chieppa, G.P. Cirillo*, Le autorità amministrative indipendenti*, *Cedam*, 2010, 38.

[41] For example the art. 26, paragraph 1 of the italian Privacy Code states that sensitive data can be processed only with the written consent of the data subject and prior authorization of the Guarantor. However, in paragraph 4 of art 26 it is explained that in some cases, as in one of defensive investigations, sensitive data can be processed without consent, but with the authorization of the Data Protection Authority. Moreover, art 27 specifies that the processing of judicial data by private individuals or public economic entities is allowed only if authorized by express provision of law or measure of the Guarantor specifying the relevant public interest purposes of processing, the types of processed data and executable operations.

[42] K. Miller, *Total Surveillance, Big Data, and Predictive Crime Technology: Privacy's Perfect Storm*, *Journal Technology of Law and Policy*, 2014, 105, 127.

[43] Even the most sophisticated predictive systems will not drive to a police reform without regulatory and institutional changes. And so, scrutiny and balances are needed to mitigate police discretionary power.

[44] Such scrutiny should seek to make a selection of the Data that could be recorded and later used by the software, that's because it is impossible to distinguish between potential and effective problematic policies or practices: an evaluation of this kind can be done only in the long term.

[45] Apart from the case of the New Orleans Police Department that I reported in chapter 3, another example of discrimination practices is demonstrated by a 2016 ProPublica investigation on predictive policing software, which shows that offender-based modelling algorithms were likely to misidentify low-risk black defendants as high risk and high-risk white defendants as low risk. J. Angwin, J. Larson, S. Mattu, L. Kirchner, *Machine Bias. There's software used across the country to predict future criminals. And it's biased against blacks*, 2016

But there's more, predictive analytics it's not a blunt instrument for targeting criminals, it could be used to identify social needs and economic problems that affect those areas with a high rate of crime: indeed, in some cases we could face with environmental risk factors that have a political nature and for that cannot be solved only with the uptake of police. That is the «promise of bright data»[46].

Predictive systems can operate and provide benefits on different levels: on one hand, it would facilitate law enforcement to define the critical areas, allocate resources as effectively as possible at any given moment, intervene at operational level with initiatives aimed at preventing and eradicating criminal phenomena, and constantly measure the results that have been achieved; on the other hand, it would help local administrations to discover the scale of the phenomena and their nature, so they can draw up more efficient policies and measures in the field of criminality and public safety and monitor the outcome. And finally, it would provide the citizens a more specific and objective information about the standard of safety of the city and advices on the best preventive behaviour that they should adopt.

[46] A.G. Ferguson, *The Rise of Big Data Policing: Surveillance, Race, and the Future of Law Enforcement*, 2017, 5-6.