Journal of Business and Economics, ISSN 2155-7950, USA
June 2018, Volume 9, No. 6, pp. 497-506
DOI: 10.15341/jbe(2155-7950)/06.09.2018/004

http://www.academicstar.us

Machine Learning Algorithms in Administrative Decision-Making

Paola Savona
(Department of Law, Lumsa University, Italy)

Abstract: Machine learning algorithms play a significant role in the digital economy. They suggest products and services to clients, select friends and news, give navigation advice to drivers, make translations. Moreover, learning algorithms are increasingly used to make important decisions about individuals. Companies, for example, rely on machine learning to approve loan, evaluate investments, calculate insurance risks, evaluate workers' performance or select people to hire. Governments use it to detect terrorists and prevent future attacks, target citizens or places for police scrutiny, select tax payers for audit, detect frauds, grant or deny visas, and more. The influence of machine learning in administrative decision-making might rapidly grow in the near future. The paper analyses opportunities and risks involved in relying on learning algorithms to support or to make administrative decisions with the aim of understanding the challenges that the use of those tools poses to the core principles of the rule of law.

Key words: algorithms; machine learning; data mining; big data; administrative decision-making; black box
JEL codes: K230

1. Introduction

Algorithms shape the world we live in. They recommend us books, films and music; suggest us friends and news; help us in translations; give navigation advise to drivers; make our houses and cities smart (Domingos, 2015). They guide our actions, alter our behavior by influencing our choices in everyday life, and thereby determine the success of products and services (Latzer et al., 2017).

Furthermore, algorithms are growingly used to make important decisions about individuals. From credit scoring to recruitment, the rating of universities to the evaluation of workers' performance, from the approval of financial transactions to the diagnosis of diseases, activities once entrusted to human beings are now performed by (or with the support of) computer systems (Rieke, Robinson & Yu, 2014; Citron & Pasquale, 2015; O'Neil, 2016; Kroll et al., 2017). According to Balkin, we are rapidly moving from the age of the Internet to an Algorithmic Society, i.e., to "a society organized around social and economic decision making by algorithms, robots, and AI agents; who not only make decisions, but, in some cases, also carry them out" (2017, p. 5).

Big data is the rough material, the "new oil" of the algorithmic society (Mayer-Shönberger & Cukier, 2013; Zeno Zencovich & Codiglione, 2016; The Economist, 2017). It is, at the same time, the fuel that runs such society, and the product of its operations (Balkin, 2017). Data processing in fact produces more data that in turn can be used by algorithms to improve their performance. As Balkin notes, varying Kant's famous statement, "algorithms

Paola Savona, Ph.D., Assistant Professor in Administrative Law, Department of Law, Lumsa University; research areas/interests: big data, data protection, automated decision-making, transparency, risk regulation. E-mail: p.savona@lumsa.it.

without data are empty; data without algorithms are blind" (2017, p. 6).

Indeed the huge amounts of data increasingly generated by multiple sources would be useless — simply noise - without powerful computational tools capable of handling and analyzing them. The very value of big data lies in analytics and its ability to turn data into useful information and valuable knowledge (Executive Office of the President, 2014; Clegg, 2017). Data mining techniques, in particular, as a subset of big data analytics, use machine learning algorithms to detect unexpected correlations and novel patterns in data, with the main purpose of predicting future trends, processes or behaviors (Mayer-Shönberger & Cukier, 2013; Domingos, 2016; Coglianese & Lehr, 2017).

The paper analyses opportunities and risks involved in relying on machine learning algorithm to support or to make administrative decisions, with the main aim of understanding the challenges that the use of those tools poses to the core values of the rule of law.

2. What Is Machine Learning?

An algorithm can be generally described as an unambiguous sequence of clear instructions for solving a given problem in a finite amount of time (Steinbock, 2005; Latzer et al., 2017). It is "any well-defined computational procedure that takes some value, or set of values, as *inputs* and produces some value, or set of values, as *outputs.*" (Cormen at al., 2001, p. 10; Kroll et al., 2017).

Machine learning algorithms (or learners) are characterized by the ability to find hidden statistical patterns in their inputs and improve over time as they receive more data (Bostrom, 2014; Surden, 2014; Yeung, 2017). Today, such algorithms are used in a variety of applications such as speech recognition, spam filters, language translation, rout-finders, or recommender systems (Bostrom, 2014). If performing well, they can produce automated results that would require high-order cognitive processes to be reached by a person in a similar situation (Surden, 2014). Machine learning is therefore considered a branch of artificial intelligent, as learning algorithms can produce "intelligent" results in complex tasks.

The term "learning", referred to the machine, does not imply, however, that computer systems artificially replicate the advanced cognitive systems involved in human cognition. As Surden suggests, the term is used in a "functional sense", to indicate how learning algorithms possess the ability to change their behaviour to enhance their performance on some task through experience (2014, p. 90).

Data mining techniques employ machine learning algorithms to search for relationships among attributes in data and identify correlations that are not visible to human eye (Hildebrandt & Koops, 2010; Nissenbaum, 2010; Rubinstein, 2013). This process, also known as "knowledge discovery in data" (KDD), is new (and differs from scientific methodology) insofar as the pattern recognition does not necessarily require the predefinition of hypotheses to be tested on samples (Anderson, 2008). As Mayer-Shönberger and Padova note, "at least to an extent, big data reverses the direction of discovery, using data to foster hypotheses rather than 'prove' existing hypotheses" (2016, p. 315). Analysts "let the data speak": learning algorithms look for models on their own by inferring them from the large amount of data generated by the past experience (Mayer-Shönberger & Cukier, 2013).

The resulting knowledge is based on statistical correlations, which do not necessarily imply a causal relationship (Anderson, 2008; Hildebrandt & Koops, 2010). It is, nevertheless, a very useful type of knowledge as correlations and patterns discovered in data can be exploited to assess the likelihood of future outcomes

(Mayer-Shönberger & Cukier, 2013). As new data becomes available, self-learning algorithms can test the model, refine the patterns and thereby improve their forecasting power (Coglianese & Lehr, 2017).

Data mining can be used for assessing risks and probabilities of uncertain developments in various fields of application. It can be used to analyze data about the physical world — for example in weather forecast or to predict climate change or the flow of traffic — or it can be employed to analyze data about people and predict human behaviors (Cohen, 2013). Thereby it allows new, sophisticated forms of automated profiling, in which individuals are assigned to particular categories (profiles) based on their similarity to members of a comparable class sharing similar clusters of attributes (Nissenbaum, 2010; Bosco et al., 2015). These profiles are then used to make decisions about individuals, often without human intervention (Hildebrand, 2008).

3. Applications in the Public Sector

Data mining is increasingly employed in the private sector to enhance efficiency, increase productivity, and improve decision-making (Rubinstein, 2013). Learning algorithms are widely used, for instance, to predict consumers' behavior, evaluate investments and financial instruments, calculate loan rates, credit scores, and insurance risks, select people to hire.

Governments have also discovered the predictive power of data analysis. Since September 11, 2001 data mining has captured the attention as a promising tool for identifying potential terrorists and pre-empt terrorist attacks (Cate, 2008; Zarsky, 2011). Intelligence agencies have begun collecting, retaining and analyzing information about hundreds of millions of people, searching for suspicious data linkages and behaviors (Rubinstein, Lee & Schwartz, 2008).

Those huge amounts of data have been then exploited, among other things, by creating watch lists of people who fit terrorist profiles, which are used not only in investigation but also in administrative determinations (Steinbock, 2005; Citron, 2008; Korf, 2015). For example, in the U.S. "no fly program", passengers' data is compared against federal government watch lists (the so-called no fly lists), to determine if passengers may pose a security risk. When the computer system generates a match between the name of the passenger and the name of a person included in one of those lists, the passenger is either barred from boarding an aircraft or he/she has to undergo an enhanced screening before boarding (Steinbock, 2005; Citron, 2008). In other words, a computer system replaces (in the first case), or supports (in the second), a human decision directly affecting the freedom of movement of U.S. and non-U.S. citizens.

The visa issuing procedure to enter United States is based on data mining as well. In 2013 the Department of State launched the "Kingfisher Expansion" (KFE) visa vetting system for conducting interagency counterterrorism screening of all visa applicants. Consular officers are now required to submit visa applications to the National Counterterrorism Center, which in turn uses an automated process to compare multiple fields of information drawn from the application against intelligence community and law enforcement agency databases in order to identify terrorism concerns. When such process results in a "red-light" hit, the visa is denied (Wasem, 2015; Assistant Secretary for Consular Affair Department of State, 2016).

Likewise, in Australia data mining is intensively used for border control and national security purposes. In 2016, the Australian Government announced the decision to establish a new automated Visa Risk Assessment (VRA) tool to assess terrorism and criminal threats. According to the Minister for Immigration and Border Protection, "the VRA capability will consolidate a wide range of immigration and border information in real time

enabling broad ranging threat identification and automated risk profiling" (Dutton, 2016).

In Europe too, predictive data analysis has become part of the current strategies to counteract terrorism, especially after terrorist attacks in Paris in November 2015. The most notable example is probably the EU Directive 2016/681 *on the use of passenger name record (PNR) data for the prevention detection, investigation and prosecution of terrorist offences and serious crime* of 27 April 2016. The Directive does not foresee the creation of "no fly lists". However, it provides that PNR data of people travelling in the Member States shall be processed for the purpose of "carrying out an assessment of passengers prior to their scheduled arrival in or departure from the Member State to identify persons who require further examination by the competent authorities (…) in view of the fact that such persons may be involved in a terrorist offence or serious crime" (article 6, paragraph 2, letter a). The Directive further allows the analysis of PNR data for the purpose of updating or creating new criteria to be used in such assessment "in order to identify any persons who may be involved in a terrorist offence or serious crime" (article 6, paragraph 2, letter c). In other words, it seems that PNR data of people travelling in Europe may be processed both to develop profiles of potential terrorists, and to apply those profiles to people whose data matches with them (European Data Protection Supervisor, 2011).

As also the PNR Directive shows, data processing is currently used not only for counteracting terrorism, but, more in general, to detect and prevent crime (so-called predictive policing). In several countries, police departments apply computer modelling to historical crime data and other kind of information, to forecast where a crime is likely to occur, or detect people at risk of committing it (Joh, 2014; Crawford & Schultz, 2014).

Applications of machine learning techniques have also crossed the security and law enforcement context. The U.S. Environmental Protection Agency (EPA), for instance, relies on machine-learning algorithms to predict environmental exposure to chemicals, and to detect compounds that should be subject to stricter regulation (Cuellar, 2016; Coglianese & Lehr, 2017). Data mining techniques are also being used in many countries to predict tax underreporting, detect frauds, prevent corruption, and evaluate the eligibility for public benefits (Citron, 2008; Coglianese & Lehr, 2017). In the United States, automated risk assessment tools are even used in sentencing as a way to calculate the probability that a convicted person commits another crime in the future (Supreme Court of Wisconsin, *State of Wisconsin* v. *Erik L. Loomis*, 13 July 2016).

4. Machine Learning in Administrative Decision-Making

As shown above, machine learning systems are presently used not only for triggering investigation (e.g., in law enforcement, fiscal assessment, fraud control) but also as the sole basis for administrative determinations (the prohibition of boarding on airplanes, the denial of visa). It can reasonably been expected that applications of data mining and machine learning techniques in administrative decision-making will rapidly grow in the future (Cuellar, 2016).

According to Coglianese and Lehr, decisions such as the license of aircrafts or pilots, the order to shutdown a pipeline at risk of imminent accident, the antitrust review of a proposed merger, the ban on the use of toxic chemical compounds, the order to stop a financial transaction which is the result of insider trading, could be taken by computer systems by means of machine learning systems, potentially without any human intervention (2017). Little imagination is required to conceive other cases in which learning algorithms could support or replace administrative decision-making. They could be employed, for instance, in licensing novel drugs or food, deciding on asylum requests, selecting public officers, deciding whether to admit a student to a university, evaluating the

financial reliability of a firm contracting with a public body. In short, machine learning systems could be used whenever an agency has to make a predictive assessment of future facts or behaviors, or has to decide despite incomplete information.

The reliance on machine learning seems promising for making government, like the private sector, smarter (Coglianese & Lehr, 2017). Learning algorithms could significantly reduce time and costs of administrative action thereby enhancing its efficiency. They could also improve decision-making under uncertainty since they make it possible to produce new insights and findings not visible to the human eye. Moreover, with the private sector increasingly relying on machine learning to make decisions, the use of the same analytical tools by public authorities might prove necessary in order to regulate economic activity more effectively (Coglianese & Lehr, 2017). However, contrary to what some legal scholars think, an increased automation of decision-making would not necessarily increase the accuracy, fairness and transparency of administrative decisions.

As a matter of fact, algorithms are not infallible. The rate of error is particularly high whenever legal rules, which are to some extent always vague and ambiguous, are translated to the binary code of computers. In such a process, performed by programmers who lack legal expertise, nuances of meaning may be lost or distorted, and as a result the rule altered (Citron, 2008; Perry & Smith, 2014). Furthermore, as machine learning algorithms analyze large amounts of unverified data, a certain lack of precision is physiological: the process of working on all available data, instead of on selected and controlled samples, necessarily implies a loss of accuracy in the outcomes (Mayer-Shönberger & Cukier, 2013). Finally, predictive analytics may generate false positives (individuals or objects wrongly labeled as dangerous). Indeed, as shown above, the forecasting power of machine learning comes from its ability to discover statistical correlations, which do not necessarily imply a causal relationship. However, as statisticians are well aware, when different complex factors interact with each other, there are always variables that cannot be predicted, so the prospect of a high number of false positives is inherent in the process (Steinbock, 2005; Cate, 2008; Hildebrand & Koops, 2010).

The U.S. “no fly program” is a good example of how frequent failures can occur in systems using data mining techniques: since 2005 several hundred of thousands of unsuspecting people, including small children and well-known figures such as Senator Edward Kennedy, have been prevented from boarding their scheduled flights while their names were mistakenly included on watch lists (Citron, 2008; Cate, 2008). The same program also shows how errors can be difficult to detect and correct. Unlike human officers who can assess the soundness of a decision to be taken and eventually modify it, computer systems have no autonomous capacity of self-correction. Even “intelligent” algorithms do not have such ability: learning algorithms can improve their performances over time as more data becomes available, but they can neither verify the data that they process, nor rectify a programming error.

These limits of algorithmic decision-making cannot be overcome by simply providing the human review of the results of data processing, since the intervention of human operators does not necessarily protect against computer’s errors (Citron, 2008). In fact, due to the widely shared belief that numbers cannot be wrong, officials hardly take a decision different from the one that is suggested by the machine (Citron, 2008; Conseil d’Etat, 2014; Clegg, 2017). The wrong belief in the infallibility of numbers thus makes the difference between decisions taken by computer systems an decisions made with their support evanescent.

Algorithms are not even necessarily fair. They are the product of the beliefs, fallibilities and biases of the person who created them (Barret, 2016). Moreover, since machine learning systems infer correlations from data generated by the past experience, they may replicate old prejudices and pre-existing inequalities encoded in the

data that they process (Barret, 2016; Conseil d'Etat, 2014; Article 29 Working Party, 2017). As a consequence, algorithms can be intrinsically (and unintentionally) discriminatory; and data mining as all profiling may have (and actually have had in many applications) a disparate impact on affected individuals or groups (Executive Office of the President 2014; Citron & Pasquale, 2014; Yeung, 2017). Such discriminatory effects may remain unnoticed if algorithms are not regularly tested and their outcomes controlled.

Above all, algorithms are not transparent. Though they make everything visible they are in themselves invisible (Hildebrand, 2009; Richards & King, 2013; Rodotà, 2014). Algorithms, indeed, are commonly described as black boxes: one knows what goes in, on one side, and one sees what comes out the other, but one does not know what goes on between the two (Mayer-Shönberger & Cukier, 2013; Pasquale, 2014; Rouvroy, 2016).

Opacity of machine learning is due to three main reasons: first, algorithms are normally covered by intellectual property rights; second, they are not comprehensible to the majority of people, who lack the technical skill necessary to read a computer code; third, experts and even the analysts who programmed a self-learning algorithm might find it difficult to explain the way in which it operates because the internal logic of the algorithm might be altered as it learns on training data (Burrel, 2016; Lepri et al., 2017; Conseil d'Etat, 2017). In other words, in certain instances analysts too cannot look inside the black box to understand how the transformation from inputs into outputs occurs (Coglianese & Lehr, 2017).

The black box nature of machine learning makes its application in administrative decision-making, although promising, extremely problematic. Decisions supported or made (the difference, as explained, is not as relevant as it might seem) by machine learning systems result from opaque processes of transformation of inputs into outputs instead of transparent procedures allowing control and participation of affected individuals. Data mining in fact provides essentially no due process: there is no notice, no opportunity to be heard, no confrontation with evidence, no giving reasons, but only a result (Steinbock, 2005; Citron, 2008). Algorithmic decisions are inherently unpredictable by the addressees, since machine learning generates unexpected, not intuitive, knowledge. They are difficult, sometimes impossible, to challenge and to review. Indeed, even if the algorithm is disclosed, its logic is accessible only to a small number of experts, and in some cases not even to them. Finally, since decisions based on self-learning systems are not predictable, the injuries that they might cause are not foreseeable. As a result, damage arising by such decisions might encounter difficulties in being compensated at least under those regimes of civil liability requiring to prove the fault of the public administration (Balkin, 2015).

5. The Legal Framework

Since the use of machine learning algorithms in administrative decision-making may undermine the principles of legal certainty, due process, judicial review and liability of public bodies (i.e., the core principles of the rule of law), it should be regulated in order to provide limits and adequate safeguards.

In the United States there is no general law concerning data processing and automated decision-making. Government data mining mostly occurs without a statutory or otherwise regulatory framework, and without legal guarantees for affected individuals (Cate, 2008).

In Europe the situation is partially different. The Data Protection Directive 95/46/EC (DPD) prohibited decision-making based exclusively on automated profiling, granting (with some exceptions) the right to every person "not to be subject to a decision which produce legal effects concerning him or significantly affects him and which is based solely on automated processing of data intended to evaluate certain personal aspects relating to

him, such as his performance at work, creditworthiness, reliability, conduct, etc." (article 15, par. 1). It also provided a right to know the logic involved in automatic processing article 12, par. 1, let. a).

The EU General Data Protection Regulation 2016/679 (GDPR), which applies from 25 May 2018, broadens the protection against automated decision-making. Article 22, paragraph 1, prohibits decisions "based solely on automated processing", which produces legal effects concerning an individual, or "similarly significantly" affect him or her, whether or not they include profiling (Article 29 Working Party, 2017).

The scope of this provision is still narrow, since it refers to fully automated decision-making, and may be easily circumvented by including a merely formal human intervention in the decision process, with no influence on the outcomes of that process (Petkova & Boehm, 2017). Accordingly, the Article 29 Working Party in its Guidelines on automated decision-making has warned that the prohibition cannot be avoided "by fabricating human involvement": the oversight of the decision should be "meaningful" and "be carried out by someone who has the authority and the competence to change the decision" (2017, p. 10).

The general prohibition of decision-making based solely on data processing may be derogated if the decision is necessary for entering into a contract between the data subject and a data controller, if there is the explicit consent of the data subject, or if the decision is authorised by Union or Member State law "which also lays down suitable measures to safeguard the data subject's rights and freedoms and legitimate interests" (article 22, par. 2, lett. a), c) b). The Article 29 Working Party, in light of the main risks associated to algorithmic decisions, has stated that "controllers should carry out frequent assessments on the data sets they process to check for any bias, and develop ways to address any prejudicial elements, including any over-reliance on correlations." (2017, p. 17). It has also recommended the use of systems that audit algorithms and regular test the accuracy of automated decision-making.

Stricter conditions are required by the GDPR for laws that authorise automated decision-making based on the processing of special categories of data (i.e., personal data revealing racial or ethnic origin, political opinions, religious or philosophical beliefs, or trade union membership, genetic and biometric data, data concerning health or person's sex life or sexual orientation). In those cases processing must be "necessary for reasons of substantial public interest" and the law "shall be proportionate to the aim pursued, respect the essence of the right to data protection and provide for suitable and specific measures to safeguard the fundamental rights and the interests of the data subject" (article 9, par. 2, let. g).

The GDPR has also strengthened transparency obligations in relation to decision-making based on automated processing. The individual has the right to be informed about the existence of automated decision-making, the logic involved, and about "the significance and the envisaged consequences of such processing" (article 13, par. 2, let. f) and 14, par. 2, let. g).

The GDPR, as the DPD, applies also to processing of personal data by public authorities. Consequently, any administrative decision based solely on automated processing, which has a legal effect on individuals, shall be prohibited unless it is expressly authorised by Union or Member State law, which in turn would lay down suitable safeguards for the individuals concerned. Although this prohibition covers only administrative decisions entirely made by computer systems (including those in which human intervention is only apparent) which affect natural persons (the protection of entities is outside the scope of GDPR), its impact on Member States' legal systems will be significant, as the use of algorithms in administrative decision-making is at present mostly unregulated.

6. Conclusion

The choice to rely on sophisticated algorithms to support or replace administrative decisions should never be, as it actually is in most countries, left to the discretion of the public administration but it should be granted by the legislative power.

The law authorizing the use of automated systems should provide for proper verification and audit mechanisms in order to grant transparency, accountability and the "technological due process" (Citron, 2008; Perry & Smith, 2014; Crawford & Schultz, 2014). A case-by-case analysis of risks and benefits involved in using machine learning systems should be carried out, including also the alternatives to such use (Zarsky, 2011). Relying on learning algorithms might prove to be a rational option, for instance, in risk regulation when agencies have to decide despite incomplete information, as, for example, when an agency has to assess whether a new product on the market might harm people's health or the environment. In cases where scientific knowledge is particularly uncertain and contested, data mining techniques might offer novel and useful patterns to the public authority, which would otherwise have to make blind decisions.

As far as personal data is concerned, data mining should be authorized only if it is strictly necessary in order to fulfill an important public interest, and provided that its efficacy is proven (which is highly disputed, for instance, in relation to measures against terrorism). Decisions based on automated profiling should be avoided, since profiling human behavior undermines the dignity of the individual, whose uniqueness cannot be reduced to the sum of his or her data.

Law may not ignore causes and reasons. Decisions affecting fundamental rights of the individual may not be left to machines. Human judgment is not perfect, and it may be influenced by biases and mistakes, but it also incorporates something an algorithm will never be able to learn: emotion, empathy and common sense.

References

Anderson C. (2008). "The end of theory: The data deluge makes the scientific method obslote?", *Wired Magazine*.

Assistant Secretary for Consular Affairs Department of State (2016). Crisis Of Confidence: Preventing Terrorist Infiltration Through U.S. Refugee And Visa Programs, Written Statement before the United States House Of Representatives Committee On Homeland Security Hearing On February 3.

Balkin J. M. (2017). "The three laws of robotics in the age of big data", Yale Law School, Public Law Research Paper No. 592, available online at: https://ssrn.com/abstract=2890965.

Balkin J. M. (2015). "The path of Robotics Law", *California Law Review Circuit*, Vol. 6, pp. 45-60.

Barret L. (2016). "Deconstructing data mining: Protecting privacy and civil liberties in automated decision-making", *Georgetown Law Technology Review*, pp. 153-159.

Bosco F. et al. (2015). "Profiling technologies and fundamental rights and values: regulatory challenges and perpsectives from European data protection authorities", in: Gutwirth S., Leenes R. and de Hert P. (Eds.), *Reforming European Data Protection Law*, Dordrecht, pp. 3-33.

Bostrom N. (2014). *Superintelligence*, Oxford.

Burrel J. (2016). "How the machine 'thinks': Understanding opacity in machine learning algorithms", *Big Data & Society*, pp. 1-12.

Cate F. H.(2008). "Government data mining: The need for a legal framework", *Harvard Civil Rights-Civil Liberties Law Review*, Vol. 43, pp. 435-489.

Citron D. K. (2008). "Technological Due Process", *Washington University Law Review*, Vol. 85, pp. 1249-1313.

Citron D. K. and Pasquale F. (2014). "The scored society: Due process for automated predictions", *Washington Law Review*, Vol. 89, pp. 1-33.

Clegg B. (2017). *Big Data*, London.

Coglianese G. and Lehr D. (2017). "Regulating by Robot: Administrative decision making in the machine learning era", University of Pennsylvania Law School, Research Paper No. 17-18, available online at:

http://scholarship.law.upenn.edu/faculty_scholarship/1734.

Cohen J. E. (2013). "What Privacy is for", *Harvard Law Review*, Vol. 126, pp. 1904-1933.

Conseil d'Etat (2014), Etude annuelle 2014, Le numérique et les droits fondamentaux, La Documentation Française.

Conseil d'Etat (2017). Etude annuelle 2017 - Puissance publique et plateformes numériques: accompagner l'«ubérisation», La Documentation Française.

Cormen et al. (2001). *Introduction to Algorithms*, Cambridge, Massachusetts.

Crawford K. and Schultz J. (2014). "Big data and due process: Toward a framework to redress predictive privacy harms", *Boston College Law Review*, Vol. 55, pp. 92-128.

Cuellar M. F. (2016). "Cyberdelegation and the administrative state", Draft, available online at: https://ssrn.com/abstract=2754385.

Data Protection Working Party (2017). "Guidelines on Automated individual decision-making and Profiling for the purposes of Regulation 2016/679", adopted on 3 October, WP 251.

Domingos P. (2015). *The Master Algorithm*, New York.

Domingos P. (2016). "Then mythes about machine learning", available online at: http://medium.com/@pedromdd/ten-myths-about-machine-learning-d88b48334a3.

Dutton P. (2016). "New visa capability to enhance national security", 12 May, available online at: http: www.peterdutton.com.au.

Economist (2017). "Fuel of the future — Data is giving rise to a new economy", May 6, available online at: http://www.economist.com/news/briefing/21721634-how-it-shaping-up-data-giving-rise -new-economy.

European Data Protection Supervisor (2011). Opinion on the Proposal for a Directive of the European Parliament and of the Council on the use of Passenger Name Record data for the prevention, detection, investigation and prosecution of terrorist offences and serious crime, 25 March.

Executive Office of the President (2014). Big Data: Seizing Opportunities, Preserving Values, Report.

Hildebrandt M. (2009). "Who is profiling who? Invisible visibility", in: Gutwirth S. et al. (Eds.), *Reinventing Data Protection?*, Dordrecht, pp. 239-252.

Hildebrandt M. and Koops B. J. (2010). "The challenges of ambient law and legal protection in the profiling era", *Modern Law Review*, pp. 428-60.

Hildebrandt M. (2008). "Defining profiling: A new type of knowledge?", in: Hildebrandt M. and S. Gutwirth et al. (Eds.), P*rofiling the European Citizen*, Dordrecht, pp. 17-45.

Joh E. E. (2014). "Policing by numbers: Big data and the fourth amendment", *Washington Law Review*, Vol. 89, pp. 35-68.

Korff D. (2015). "Passenger name records, data mining & data protection: The need for strong safeguards", The Consultative Committee of the Convention for the Protection of Individuals With Regard To Automatic Processing of Personal Data, Council of Europe, T-PD(2015)1, 15 June.

Kroll et al. (2017). "Accountable Algorithms", *University of Pennsylvenia Law Review*, Vol. 165, pp. 633-705.

Latzer M. et al. (2017). "The economics of algorithmic selection on the internet", in: Bauer J. M. and Latzer M. (Eds.), *Handbook on the Economics of the Internet*, Cheltenham.

Lepri B. et al. (2017). "The tyranny of data? The Bright and dark sides of data-driven decision-making for social good", in: T. Cerquitelli, D. Quercia and F. Pasquale (Eds.), *Transparent Data Mining for Small and Big Data*, Dordrecht, pp. 3-24.

Mayer-Shönberger V. and Cukier K. (2013). *Big Data*, London.

Mayer-Shönberger V. and Padova Y. (2016). "Regime change? Enabling big data through Europe's new data protection regulation", *The Columbia Science & Technology Law Review*, Vol. 17, pp. 315-335.

Nissenbaum E. (2010). *Privacy in Context*, Stanford.

O'Neil C. (2016). *Weapons of Math Destruction: How Big Data Increases Inequality and Threatens Democracy*, New York.

Pasquale F. (2014). *The Black Box Society: The Secret Algorithm Behind Money and Information*, Harvard.

Perry M. and Smith A. (2014). " iDecide: the legal implications of automated decision-making", speech delivered at the University of Cambridge – Cambridge Centre for Public Law Conference 2014: Process and Substance in Public Law, Cambridge, 15-17 September, available online at: http://www.fedcourt.gov.au/ digital-law.

Petkova B. and Boehm F. (2017). "Profiling and the essence of the right to data protection", available online at: https://ssrn.com/abstract=2911894.

Richards N. M. and King J. H. (2013). "Three Paradoxes of Big Data", *Stanford Law Review Online*, Vol. 66, pp. 153-159.

Rieke A., Robinson D. and Yu H. (2014). *Civil Rights, Big Data and Our Algorithmic Future- A September 2014 Report on Social Justice and Technology*, available online at: https://bigdata.fairness.io.

Rodotà S. (2014). *Il mondo nella rete*, Roma-Bari.

Rouvroy A. (2016). “Of data and men: Fundamental rights and freedoms in a world of big data”, Council of Europe, Directorate General of Human Rights and Rule of Law, Strasburg.

Rubinstein I. S. (2013). “Big data: The end of privacy or a new beginning?”, *International Data Privacy Law*, pp. 1-14.

Rubinstein I. S., Lee R.D and Schwrtz P. M. (2008). “Data Mining and Internet Profiling: Regulatory and Technological Approaches”, *The University of Chicago Law Review*, Vol. 75, pp. 261-285.

Steinbock D. J. (2005). “Data matching, data mining, and due process”, *Georgia Law Review*, Vol. 40, pp. 1-84.

Surden H. (2014). “Machine learning and law”, *Washington Law Review*, Vol. 89, pp. 87-115.

T. H. Cormen et al. (2001). *Introduction to Algorithms* (2d.ed.), London.

Wasem R. E. (2015). “Immigration: Visa Security Policies, Congressional Research Service 7-5700”, 7 November, available online at: https:www.crs.gov.

Yeung K. (2017). “Algorithmic regulation: A critical interrogation”, King’s College London Law School Research Paper No. 2017-27, available online at: https: //ssrn.com/abstract=2972505.

Zarsky T. Z. (2011). “Governmental data mining and its alternatives”, *Penn State Law Review*, Vol. 116, pp. 285-330.

Zeno Zencovich V. and Codiglione G. (2016). “Ten legal perspectives on the big data revolution”, in: Di Porto F. (Ed.), *Big Data e concorrenza*, Concorrenza e Mercato, Vol. 23, pp. 15-29.