Automated decisions and Article No. 22 GDPR of the European Union: an analysis of the right to an "explanation".

by Elena Falletti (Università Carlo Cattaneo, Italy)

Abstract: The aim of this proposal is to analyze the use of algorithms in automated decision procedures concerning subjective positions having legal relevance. In this regard, the first paragraph of Article No. 22 GDPR, entitled "Automated decision-making relating to natural persons, including profiling" establishes that: "The data subject has the right not to be subjected to a decision based solely on automated processing, including profiling, which produces legal effects that concern him or that significantly affects his person". This norm gives rise to important questions concerning the very nature of decision formulation through algorithms. In hypothesis, the juridical nature or juridical effect of the concept of "decision" could be restricted to circumstances such as financial solvency, predictability of motor vehicle accidents. In these cases, the "Right to explanation" refers to the algorithmic decision obtained. For the purposes of this abstract, the most interesting element, which could give rise to a comparative judicial dispute, given the validity of the GDPR in all the countries of the European Union, concerns the provision of par. 2 of Article No. 22, letters a) and c) which refers to the implementation of "appropriate measures" to protect "the rights, freedoms and legitimate interests of the interested party". What is the "appropriate" counterpart used for implementing such measures? It would seem unquestionable that the only element that can satisfy this "appropriateness" is human intervention, i.e. someone who has the necessary authority, ability and competence to modify or revise the decision disputed by the user. Another authoritative source claims that such appropriate measures may also consist of automated systems that control the algorithms, i.e. periodic reviews, introduction of procedures that verify the accuracy of the decision-making process or correct errors, discrimination issues, or inaccuracies from an outdated database in order to avoid the self-learning algorithm based on wrong data and processes. Nevertheless, there are those who express perplexity about the effectiveness of the right to explanation, especially in relation to the difficulty in interpreting the legislative passages that envisage it within the GDPR itself, together with doubts about the mandatory nature of this explanatory right.
This paper focuses on the meaning of right to the explanation of the decision pursuant to Article No. 22 GDPR in order to allow the interested parties to understand the reasons for the decision and possibly to dispute it.

1. Artificial Intelligence and automated decision-making: an overview

Since the 1960s technological evolution has changed the perspective of the need to ensure the protection of the privacy of each individual, and this has begun to affect all citizens of industrialized societies[1], both individually and as a mass society.[2] This is the aftermath of the introduction into the production process of the electronic computer and telematic transmission of information relating to the person, collected in digital archives. This event led to a Copernican revolution in the way of creating, organizing and interrogating databases. Electronic processors allow storage of large amounts of data by caching it in their memory. The information was originally contained on

---

1   S. Rodotà, *Il diritto di avere diritti*, Bari-Roma, 2013, 319 ss.
2   E. Canetti, *Masse und Macht* (trad. it. F. Jesi), Milano, Adelphi, 1981, 17-18.

magnetic tapes, then on digital supports, and is used to perform such data operations as calculation, combination, aggregation, analysis, or data selection, enabling the acquisition of new knowledge. In this case, the use of "artificial intelligence" as a concept is accepted because the computer is considered an electronic prosthesis of human intelligence.[3] However, it has been observed that the difficulty in defining artificial intelligence does not concern the concept of artificiality, but the more ambiguous one of human intelligence.[4] In fact, there may be very different definitions of intelligence, and they are connected with other human characteristics that are difficult to define, such as "consciousness", "learning", "language", "abstraction", "reasoning", "adaptation" and so on.[5]

The most immediate connection between the use of artificial intelligence and the protection of privacy concerns the large-scale accumulation of news and data of the most varied nature in an extremely restricted space. The use of software capable of self-learning has considerably reduced the amount of time required for information searching. Consequently, automatic data processing is also possible at a distance, due to the connection between the main processor and various terminals connected to it through a telematic network. Through this type of data collection (and connection), the person to whom these data refer, and whose sphere of confidentiality is at constant risk of breach[6], completely loses his/her control of the data, since he or she cannot be constantly informed of the passages that the said collection undergoes.[7]

As has been observed, artificial intelligence has already been the subject of theoretical analysis for many years.[8] However, at the same time it could not be applicable due to "*the absence of a large amount of data available*".[9] Nevertheless, theoretical study is still necessary to bring AI studies into the context of the appropriate legal interpretation. Scientific developments in the field of Artificial Intelligence have been significant especially in terms of classification and conceptual profiles. In this regard, one of the most recent opinions on this subject distinguishes between research related to Artificial Intelligence in three areas, related to, albeit separate from: symbolic systems,[10] which follow the original approach of Alan Turing, robotics and neural networks.[11]

The use of automated imitative systems of human intelligence in the management of complex cultural operations has undergone significant development since the 1960s, thanks to the implementation of remote controlled machines in working environments characterized by repetitive actions, such as those of industrial "robots", particularly in the automotive sector, which are adapted

---

3   G. Sartor, *Cognitive Automata and the Law*, 2006, http://ssrn.com/abstract=963760.

4   M. U. Scherer, *Regulating Artificial Intelligence Systems: Risks, Challenges, Competencies, and Strategies*, 29 Harv. J. Law & Tec 353, (2016), 358 ss.; D. C. Brock, *Learning from Artificial Intelligence's Previous Awakenings: The History of Expert Systems*, AI Magazine, 2018, 3, 4 ss.

5   M. U. Scherer, op. cit.

6   S. Wachter, B. Mittelstadt, L. Floridi, *Why a Right to Explanation of Automated Decision-Making Does Not Exist in the General Data Protection Regulation*, International Data Privacy Law, 2017.

7   G. Sartor, *Privacy, Reputation, and Trust: Some Implications for Data Protection*, 2006, http://ssrn.com/abstract=891123.

8   A. Renda, *Artificial Intelligence: Ethics, Governance and Policy Challenges Centre for European Policy Studies,* Bruxelles, 2019; AI Now Insititute at New York University, AI Now Report 2018.

9   G. Finocchiaro, *Intelligenza artificiale e protezione dei dati personali*, Giur. It., 2019, 7, 1670.

10 L. Floridi, *La quarta rivoluzione. Come l'infosfera sta trasformando il mondo*. Trad. it. M. Durante, Milano, 2017, 154 ss.

11 N. Nilsson, *The Quest for Artificial Intelligence: A History of Ideas and Achievements*, Cambridge University Press, 2010, 13, web https://ai.stanford.edu/~nilsson/QAI/qai.pdf; I. M. Cockburn, R. Henderson, S. Stern, *The Impact of Artificial Intelligence on Innovation,* National Bureau of Economic Research, Cambridge (USA), 2018, 9 ss.

in logistics or transport as well.[12] Legal experience has accepted programs that facilitate the solution of complex questions by using "Expert systems".[13] Such programs are able to interact with a "question/answer" model, or with questions about specific cases, and the program is able to provide an appropriate answer according to a decision tree model,[14] allowing the reconstruction of the logical steps between the question posed to the system (a database, for example) and its answer and *vice versa*.[15] However, it should be noted that this model, although implemented with some commercial success, is now outdated and no longer represents the innovative reference point on the subject.

Even if the fortune of such systems has by now faded,[16] they maintain a specific academic interest, in particular as regards the use of assistance systems for human decision-making, such as in the formulation of medical diagnoses or in the argumentative support in jurisprudential elaboration.[17] In this regard, common law jurisprudence has offered a first positive assessment of the use of these systems in judicial proceedings. In fact, in relation to a legal case about the family asset management carried out through multiple trusts, the Jersey Royal Court, in the United Kingdom,[18] affirmed that the use of expert programs in the disclosure of large quantities of documents is more efficient frome the point of view of the cost profile, accuracy of research and analysis of relevant documents compared to traditional manual verification.[19]

Neural systems and their "learning approach" currently play a central role in research in the field of artificial intelligence, related to the ability to learn from the experience accumulated in the reference databases. They have overcome the symbolic logic that reacted in a predefined way to specific stimuli as happens with symbolic systems. A significant further step concerns the ability of the algorithms to react to the stimuli to which they are subjected so as to develop accurate methods for the prediction of particular events, be they physical or logical.[20]

If the "Q/A" model does not pose significant problems relating to the protection of the person, or to the respect of the principle of transparency in automated decision-making methods, these problems

---

12 I. M. Cockburn, R. Henderson, S. Stern, The Impact of Artificial Intelligence on Innovation, National Bureau of Economic Research, Cambridge (USA), 2018, 9 ss.

13 In this article, the reference is to "legal expert systems". See K. D. Ashley, *Artificial Intelligence and Legal Analytics*, Cambridge University Press, Cambridge, 2017, 8, where there is a quotation of R. E. Susskind according to whom *"expert systems are computer applications that contain representations of knowledge and expertise … which they can apply – much as human beings do – in solving problems, offering advice and undertaking a variety of other tasks. In law, the idea is to use computer technology to make scarce expertise and knowledge more widely avalaible and easily accessible"*.

14 N. Bostrom, *Superintelligence. Paths, Dangers, Strategies*, Oxford, 2014, 7ss; S. Hénin, *AI. Intelligenza Artificiale tra incubo e sogno*, Milano, 2019, 43; E. Chrysler, *Using Decision Tree Analysis to Develop an Expert System. Information Systems Education Journal*, 2006, 4 (55). http://isedj.org/4/55/. ISSN: 1545-679X.

15 M. Tegermark, *Life 3.0. Being Human in the Age of Artificial Intelligence*, New York, 2017, 66 ss.

16 K. D. Ashley, op. cit.

17 N.Aletras, D. Tsarapatsanis, D. Preoţiuc-Pietro, V. Lampos, *Predicting judicial decisions of the European Court of Human Rights: a Natural Language Processing perspective*. PeerJ Computer Science, 2016, 93; L. Bennett Moses, J. Chan, *Using Big Data for Legal and Law Enforcement Decisions: Testing the New Tools*" [2014] UNSWLawJl 25; (2014) 37 ss.

18 In the matter of the A and B Trusts [2018] JRC 068 (10 April 2018),available on http://www.bailii.org/je/cases/UR/2018/2018_068.html

19 The Court wrote: *"parties in disputes of this kind should consider using artificial intelligence systems to identify relevant documents unless such systems could clearly be shown not to be reliable. The use of such systems does not cause any issue of principle; indeed, if parties were to attempt solely to analyse documents using exclusively or extensive manual processes alone, such a methodology runs a significant risk of being disproportionate and in breach of the overriding objective"*. In the matter of the A and B Trust, [2018] JRC 068, cit.

20 M. Cockburn, R. Henderson, S. Stern, op. cit.

arise more clearly in the machine learning model implemented with artificial neural networks (henceforth ANN).[21] These imitate what happens in biological neurons through the exchange of information through the synapses, i.e. the procedure that allows the human brain to learn[22]. The best-known model[23] works as follows: an input layer is recognized, connected to sensors that perceive the digitized information to be processed (these can be images, texts, numbers, operations), then the information goes through different hidden intermediate layers, where the calculation operations and learning is carried out, and an output layer that transmits the results of the processing.
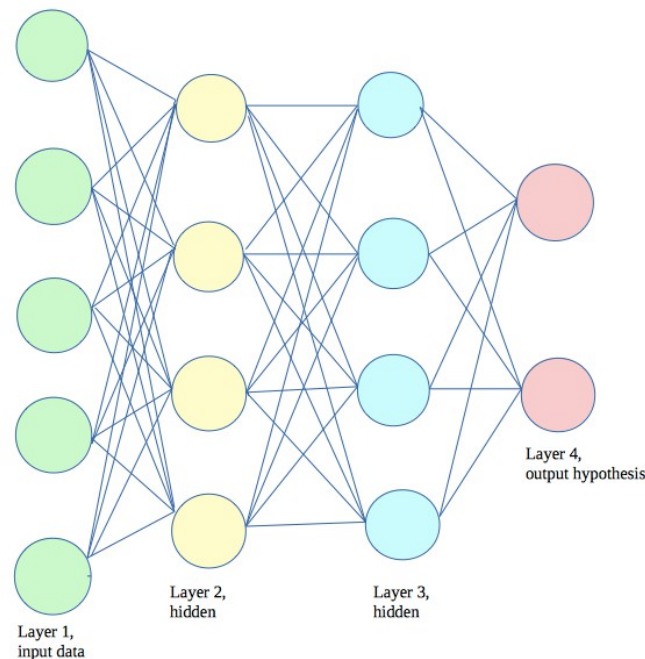

Table n. 1: Artificial neural network. In this graphic, the first layer represents the inputs. Hidden layers represent the computational units and the relative relationships that give as a result the hypothesis shown by the output layer.

Each layer is connected with all the others and each connection has its own weight, however the reconstruction of the logical path followed by the learning units is confusing and difficult to interpret. Although they have generated much enthusiasm and are an integral part of the AI development debate, the ANNs are (still) too orderly (given that, unlike biological brains, they prefer elegance and the power of mathematics), too simple, too limited in the number of components and too dry (because they ignore many other biological factors such as temporal factors of synaptic sequences),[24] when compared with a human brain.

The "Expert system" presents a justification tree logic of its decisions through logical paths that can always be reconstructed linearly, while the new frontier of AI uses neural logic, modeled on the

---

21 M. Hénin, op. cit., 50 ss.; M. A. Boden, *L'intelligenza artificiale,* Bologna, 2016, 79 ss.

22 M. Hénin, op. cit., M. A. Boden, op. cit.

23 M. Hénin, ult. op. loc. cit.

24 G. Finocchiaro, *Intelligenza artificiale e diritto. Intelligenza artificiale e protezione dei dati personali,* Giur. It., 2019, 7, 1670I

functioning of the synapses of the brain, which however cannot reconstruct the path of the logical steps of its results.

This factor highlights the difference between the models of expert systems (now used residually), compared to neural networks. The latter collect and store data from past experiences and process them in invisible procedures: the so-called "Black box", where matrix calculations are performed, and which cannot be reconstructed in their logical path. Following the mentality of the jurist, one could compare this difference between the traditional approach of the *id quod plerumque accidit* with the quantum approach of the uncertainty principle. In fact, matrix calculations performed by AI in black boxes are not limited to repeating the experiences already stored, but rather they recombine them, with the risk of perpetuating the bias and possible discriminatory elements present in the black box[25]. Therefore, it is necessary in the training or learning phase for the neural network to collect objective and truthful data.

2. Automated decisions and Article No. 22 GDPR: the right to an explanation

The Article No. 22 GDPR (General Data Protection Regulation, Regulation EU/679/2016) approach is wider than the one formerly established by Article No. 15.1 of the Data Protection Directive,[26] which is no longer in force. This new perspective concerns: specific information to the data subject, the right to obtain human intervention, the right to express the point of view of the data holder, the right to obtain an explanation of the decision reached, and the right to challenge the decision.[27]

Indeed, Article No. 22 GDPR[28] explicitly establishes the individual right of the data subject not to "be subjected to a decision based solely on automated processing, including profiling,[29] which

---

25 M. Pasquinelli, *How a Machine Learns and Fails – A Grammar of Error for Artificial Intelligence*, Spheres (http://spheres-journal.org/how-a-machine-learns-and-fails-a-grammar-of-error-for-artificial-intelligence/), 2019, 9 ss.

26 Article No 15 of the "Directive 95/46/EC of the European Parliament and of the Council of 24 October 1995 on the protection of individuals with regard to the processing of personal data and on the free movement of such data" stated that "Automated individual decisions 1. Member States shall grant the right to every person not to be subject to a decision which produces 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. 2. Subject to the other Articles of this Directive, Member States shall provide that a person may be subjected to a decision of the kind referred to in paragraph 1 if that decision: (a) is taken in the course of the entering into or performance of a contract, provided the request for the entering into or the performance of the contract, lodged by the data subject, has been satisfied or that there are suitable measures to safeguard his legitimate interests, such as arrangements allowing him to put his point of view; or (b) is authorized by a law which also lays down measures to safeguard the data subject's legitimate interests". A. Roig, "*Safeguards for the right not to be subject to a decision based solely on automated processing (Article 22 GDPR)*", in European Journal of Law and Technology, Vol 8, No 3, 2017, 2.

27 A. Roig, cit.

28 Article 22 Automated individual decision-making, including profiling 1. The data subject shall have the right not to be subject to a decision based solely on automated processing, including profiling, which produces legal effects concerning him or her or similarly significantly affects him or her. 2. Paragraph 1 shall not apply if the decision: (a) is necessary for entering into, or performance of, a contract between the data subject and a data controller; (b) is authorised by Union or Member State law to which the controller is subject and which also lays down suitable measures to safeguard the data subject's rights and freedoms and legitimate interests; or (c) is based on the data subject's explicit consent. 3. In the cases referred to in points (a) and (c) of paragraph 2, the data controller shall implement suitable measures to safeguard the data subject's rights and freedoms and legitimate interests, at least the right to obtain human intervention on the part of the controller, to express his or her point of view and to contest the decision. 4. Decisions referred to in paragraph 2 shall not be based on special categories of personal data referred to in Article 9(1), unless point (a) or (g) of Article 9(2) applies and suitable measures to safeguard the data subject's rights and freedoms and legitimate interests are in place.

29 According to Article No. 4 of the GDPR, the term "profiling' means any form of automated processing of personal data consisting of the use of personal data to evaluate certain personal aspects relating to a natural person, in

produces legal effects that concern him or that significantly affects his person".[30] This refers to Automated Decision Making (henceforth ADM), through which a duly programmed IT system can produce a significant decision for the subjects involved based exclusively on the algorithmic evaluation of the personal data of the profiled subjects/users without the aid of human intervention.[31] At the base of this subtraction (a kind of habeas corpus, of self-determination concerning one's digital personality), there is the need to protect individual fundamental rights, whose protection is recognized by the GDPR in Whereas No. 71 and in Article No. 22. Although this is important, it does not seem to be sufficient in the light of the coordinated reading of the exceptions contained in the paragraphs following the same Article No. 22 GDPR, as will be seen later. However, it is a significant, albeit extensible, barrier against decisions formed without human intervention[32] that affect individual (human) rights. The attempt to find a balance between the centrality of the human factor and the need for decision-making transparency with the attempt to avoid obstacles to technological innovation, and the need for improvement of decision-making procedures, especially administrative ones, towards which it is necessary to verify the exercise of power and its related judicial review, is evident.[33] Scholars observed that similar warranties should also be applied in the private sector, where they appear inefficient regarding those who suffer an automated decision.[34] However, a solution could be found applying broadly, and with horizontal effect to the Google Spain decision (already enforced regarding the right to be forgotten).[35] The effect of such an extension of this approach means that the automatic indexing of the results of online searches by the Google search engine could fall under the application of the first paragraph of Article No. 22 GDPR.

The following paragraphs of the same article contain the remarkable exceptions to the first paragraph. First of all, the second paragraph states that the prohibition of automated decision-making does not apply in cases in which such ADM procedures are necessary for the purposes of concluding and executing a contract between the interested party and the data controller or if the person concerned has given his/her consent to ADM. This case has been inserted in order to ensure that the interested party has greater awareness of the automatic procedure to which he or she is

---

particular to analyse or predict aspects concerning that natural person's performance at work, economic situation, health, personal preferences, interests, reliability, behaviour, location or movements"

30 E. Troisi, *AI e GDPR: Automated Decision Making, la protezione dei dati e il diritto alla 'intellegibilità´ dell'algoritmo,* 20 European Journal of Privacy Law & Technologies, 2018/1, http://www.ejplt.tatodpr.eu/Article/Archive/print_html?idn=2&ida=64&idi=-1&idu=-1&print=1#2; R. Messinetti, *La tutela della persona umana versus l'intelligenza artificiale. Potere decisionale dell'apparato tecnologico e diritto alla spiegazione della decisione automatizzata,* Contratto e Impresa, 2019, 861; F. Costantini, *Profilazione e "automated decision making" in ambito lavorativo nella giurisprudenza italiana,* Lavoro nella Giur., 2019, 11, 984; G. Noto La Diega, *Against the Dehumanisation of Decision-Making. Algorithmic Decisions at the Crossroads of Intellectual Property, Data Protection, and Freedom of Information, Journal of Intellectual Property,* Information Technology and E-Commerce Law, 2018, 1, 16 ss.

31 E. Troisi, op. cit.; A. Pajno, M. Bassini, G. De Gregorio, M. Macchia, F. P. Patti, O. Pollicino, S. Quattrocolo, D.Simeoli, P. Sirena, *AI: profili giuridici. Intelligenza Artificiale: criticità emergenti e sfide per il giurista,* BioLaw Journal – Rivista di BioDiritto, n. 3/2019, 217 ss.

32 A. Pajno, M. Bassini, G. De Gregorio, M. Macchia, F. P. Patti, O. Pollicino, S. Quattrocolo, D.Simeoli, P. Sirena, AI: profili giuridici, cit.

33 A. Pajno, M. Bassini, G. De Gregorio, M. Macchia, F. P. Patti, O. Pollicino, S. Quattrocolo, D.Simeoli, P. Sirena, AI: profili giuridici, cit.

34 S. Wachter, B. Mittelstadt, L. Floridi, *Why a Right to Explanation of Automated Decision-Making Does Not Exist in the General Data Protection Regulation,* International Data Privacy Law, 2017, Vol. 7, No. 2

35 A. Pajno, M. Bassini, G. De Gregorio, M. Macchia, F. P. Patti, O. Pollicino, S. Quattrocolo, D.Simeoli, P. Sirena, AI: profili giuridici, cit.

subjected, with the consequence that the automated decision is to be considered illegitimate in all cases in which the interested party has not expressly allowed it, directly or through a contractual agreement where the circumstance is explicit.[36] In light of the behavioral experience of individuals, it is not difficult to hypothesize that the exception relating to the manifestation of consent contained in the prohibition of the first paragraph could take the form of usually superficial users of the ADM services through their formal manifestation of consent. Indeed, it is a necessary act for the purpose of using the proposed service, and people could be unaware in relation to the adverse consequences in case of refusal, as happened with the manifestation of consent to the collection of personal data.[37]

In the event that automated personal processing is allowed, either by law or by express consent from the party subjected to this decision, the data controller has the duty, pursuant to paragraph 3 of the aforementioned Article No. 22 GDPR, to take appropriate measures to protect his or her rights, freedoms and legitimate interests. In particular, the data controller is in any case obliged to guarantee him/her the right to obtain human intervention, as well as to allow the subordinate subject to express his/her opinion and to contest its decision.[38] Some American scholars speak of "right to a human decision".[39]

Article No. 22 GDPR expressly refers to treatments concerning "decision-making", in the sense that algorithms produce and apply a decision having effects on the person concerned. This means that treatment of mere conservation and organization of personal data, which do not undergo evaluation or analysis, should be excluded. However, this circumstance has an extremely limited scope, as the collection and classification of the data itself involves an evaluation by the algorithm that will produce a result. Therefore, all those data treatments that are not typically inferential, in which the use of algorithms is limited, i.e., to the conservation and/or organization of personal data without involving any analytical and/or evaluation steps of those data. In an automated decision-making treatment, on the other hand, the data are collected (or subjected to the algorithm) precisely because this, by performing a given calculation and by applying mathematically defined deductive rules to the data-set program, analyzes them to arrive at a specific 'solution', which is a 'decision-making' output.

In my opinion, this rule should be interpreted extensively because, from a practical point of view, the case has significant scope given that in digital practice the mechanisms of scoring are used for the purposes of e.g. selection among the aspirants of a job position, the predictability of motor accidents concerning a potential insured person, or the solvency of possible clients of a bank account holder to the aforementioned mathematically formalized decision-making procedure.[40] These programs are based on assessments performed by algorithms on the data statistics related to the work or debt history of potential candidates, applying certain predefined guidelines,[41] therefore the evaluation-decision element with legal effects on the subject is an essential part.

---

36 E. Troisi, op. cit.

37 *I. A. Caggiano, Il consenso al trattamento dei dati personali nel nuovo Regolamento europeo. Analisi giuridica e studi comportamentali,* Osservatorio del diritto civile e commerciale, 2018, 67 ss.; L. Gatt, R. Montanari, I.A. Caggiano, *Consenso al trattamento dei dati personali e analisi giuridico-comportamentale. Spunti di riflessione sull'effettività della tutela dei dati personali,* Politica del diritto, 2017, 363 ss.

38 E. Troisi, op. cit.

39 A. Z. Huq, *A Right to a Human Decision* (May 3, 2019). Virginia Law Review, Vol. 105; U of Chicago, Public Law Working Paper No. 713. Available at SSRN: https://ssrn.com/abstract=3382521; T. Wu, *Will Artificial Intelligence Eat the Law? The Rise of Hybrid Social-ordering Systems*, Columbia Law Review, 2019, 2023.

40 B. Goodman, S. Flaxman, *European Union regulations on algorithmic decision-making and a "right to explanation*", AI Magazine, 2017, C. O' Neil, *Armi di distruzione matematica. Come i big data aumentano la disuguaglianza e minacciano la democrazia,* Firenze, 2017.

A very interesting element, which could give rise to comparative legal dispute with the enforcement of the GDPR in all the countries of the European Union, concerns the provision of par. 2 of the Article No. 22 GDPR, letters a) and c).[42] It refers to the implementation of "suitable measures" to protect the rights, freedoms and legitimate interests of the data subject, especially the right to obtain human intervention by the data controller, to express their opinion and to challenge the decision.

What appropriate counterpart is useful for implementing such measures? It would seem indisputable that the only element that can satisfy this appropriateness is human intervention, that is the interlocutor, not specifically identified by the GDPR, who has the necessary authority, capacity and competence to modify or revise the decision disputed by the user.[43] For the purpose of identifying this figure, reference could be made to the entity itself which issued the automated decision that has effects on the legal situation (and therefore on the life) of the subject submitted, or directly to the judicial authority. Other scholars claim that appropriate measures may also consist of automated systems that control algorithms, i.e. periodic reviews, introduction of procedures that verify the accuracy of the decision-making process. At the same time, these procedures could correct errors, discrimination, inaccuracies from an outdated or inaccurate database in order to avoid the self-learning algorithm based on wrong data and processes.[44]

Nevertheless, there are those who express perplexity about the effectiveness of the right to explanation, especially in relation to the difficulty in interpreting the legislative passages provided by the GDPR itself, together with doubts about the mandatory nature of this explanatory right; especially in consideration of the complexity of the subject, which is neither easily accessible, nor suitable for users, i.e. anyone who somehow interacts, often unknowingly, with algorithms.[45]

With regard to the configurability of the the right to explanation of the interested party of, it is observed that there is a non-univocal reference between the already mentioned "appropriate measures" expressly established, but not better specified, by Article No. 22, par. 3 and the "adequate guarantees" mentioned in "Whereas" No. 71 of the GDPR Regulation, which should include the "specific information to the interested party and the right to obtain human intervention, to express their opinion, to obtain an explanation of the decision reached after such assessment and to challenge the decision".

Given that such "whereas" is not of a binding nature, and is used by operators as an interpretative tool, a debate has arise about the interpretation of the two combined statements: between, on the one hand, those who maintain the systematic interpretation of the GDPR, and recognize only one right of access and information, but not a right to explanation,[46] and on the other hand, those who claim that the right to explanation is also included.[47]

---

41 G. Malgieri, G. Comandè, *Right to legibility of automated decision-making, in International Data Privacy Law*, 2017, vol.7, n.4; E. Troisi, op. cit.

42 E. Pellecchia, *Profilazione e decisioni automatizzate al tempo della "Black Box Society": qualità dei dati e leggibilità dell'algoritmo nella cornice della "responsible research and innovation"*, Le Nuove leggi civili commentate, 2018, 5, 1209-1235; B. Goodman, S. Flaxman, European Union regulation, cit.,

43 E. Pellecchia, op. cit.

44 Guidelines on Automated individual decisionmaking and Profiling for purposes of Regulation 2016/679, written by the Data Protection Working Party nominee according to Article No 29 Directive EC/1995/46. Last version of 6th February 2018, recently replaced by the European Data Protection Board. E. Pellecchia, op. cit.

45 S. Wachter. B. Mittels*tadt, C. Russell*, Counterfactual explanations without opening the black box: automated decisions and the GDPR, 2017, 23 ss.

46 S. Wachter, B. Mittelstadt e L. Floridi, *Why a right to explanation of automated decisionmaking does not exist in the general data protection regulation*, cit.

47 A. D. Selbst, J. Powles, *Meaningful information and the right to explanation,* in 7 International Data Privacy Law, (2017), 233.

In any case, the European Data Protection Committee recognizes that the maximum transparency guaranteed by high information standards may not be sufficient if the information received is not comprehensible to the recipient, given the complexity of the automated decision-making processes. Therefore it would be appropriate to provide not so much an explanation on how the algorithm works, but complete information communicated in a simple way about the elements used and on the relative weight in the procedure, in order to allow the interested parties to understand the reasons for the decision and possibly to dispute it.[48] On this point, in the doctrine there are those who lean towards a distinction between the decisions that can be explained *ex ante* (especially in reference to a decision-making system that carries out evaluations, and therefore decisions, on a standard model), or *ex post* (in the case of decisions relating to specific cases).[49]

3. AI, privacy, non-discrimination and access to challenging a decision

In this regard, it must be emphasized that the development of artificial intelligence has changed the perspective under which the concept of informed consent (on which the entire normative framework of the GDPR is based) is studied. In fact, the use of automated algorithms transforms all personal data, without distinction. Indeed, the algorithms allow the coordination of "sensitive" data (for which the consent of the entitled person is required) with non-sensitive data, reconstructing the entire individual profile of each person, thereby damaging both his/her privacy (causing the total invasiveness into the private sphere of that person), and his/her dignity (making him/her "naked" and "transparent" in front of those who come into possession of his/her revised data), and subjecting them to danger of serious discrimination.

On this point the Italian Privacy Authority has expressed its concern, stating that:

Personal data collection is the "engine" of artificial intelligence, and (at the same time, since the data protection discipline, although technologically neutral), it is the most advanced regulatory area and most capable of governing the complexity of the digital society, respecting the dignity of the person. (...) Algorithms are not neutral calculation syllogisms, but human opinions structured in mathematical form which, as such, reflect, to a greater or lesser extent, the pre-understandings of those who design them, risking to turn algorithmic discrimination into social discrimination. With respect to these risks, the guarantees enshrined in the new legal framework with regard to automated decision-making processes are important, ensuring their contestability and transparency of logic, and demanding, at least as a last resort, the filter of human intervention, to counter the unconditional delegation to the blind determinism of technology.[50]

Among the many possible cases in the medical, social and technological fields, there are areas of human life where the application of artificial intelligence software can have discriminatory effects,

---

48 E. Pellecchia, op. cit.

49 S. Wachter, B. Mittelstadt, L. Floridi; *Why a Right to Explanation of Automated Decision-Making Does Not Exist*, cit.

50 Italian Privacy Authority Report, A. Soro, "*Proteggere i dati per governare la complessità*", Rome, il 10.7.2018 available on https://www.garanteprivacy.it/documents/10160/0/Relazion%2B2017%2B-%2BDiscorso%2Bdel %2BPresidente%2BAntonello%2BSoro+&cd=9&hl=it&ct=clnk&gl=it: Non soltanto (...) i dati personali sono il "motore" dell'intelligenza artificiale, ma anche e soprattutto perché la disciplina di protezione dati, pur tecnologicamente neutra, è il settore normativo più avanzato e maggiormente capace di governare la complessità della società digitale, nel rispetto della dignità della persona. (…) Gli algoritmi non sono neutri sillogismi di calcolo, ma opinioni umane strutturate in forma matematica che, come tali, riflettono, in misura più o meno rilevante, le precomprensioni di chi li progetta, rischiando di volgere la discriminazione algoritmica in discriminazione sociale. Rispetto a questi rischi, risultano importanti le garanzie sancite dal nuovo quadro giuridico in ordine ai processi decisionali automatizzati, assicurandone la contestabilità e la trasparenza della logica, ed esigendo, almeno in ultima istanza, il filtro dell'uomo, per contrastare la delega incondizionata al cieco determinismo della tecnologia

that are not "neutral"[51], in relation to sensitive cases such as the verification of possession of the requirements for obtaining a residence permit or citizenship[52] and granting or denying "universal credit", a form of last resort welfare benefit.[53] These categories show how concepts like inclusion and exclusion, circumstances of eligibility, and "requirements of conduct" differentially affect entitlement for various groups of both citizens and migrants[54]", and delegating the decision connected to these subjects apparently seems to be immune to negative (non-neutral) effects on people involved. Whether or not this procedure may seem neutral and correct, it could actually represent the dehumanization of a process of an essential status attribution to the people involved, which is not always easy to access and obtain, leaving them in the grip of a confrontation.

On this point, we could refer to recent information collected in relation to the British experience[55], on the basis of data collected by a public institution, the Department for Work and Pensions (henceforth DWP). Although Brexit (i.e. the procedure concerning the exit of the United Kingdom from the European Union) has not yet seen the conclusion of an agreement accepted by both parties, the data referred to were collected on May 2019, therefore still under the enforcement of EU/679/2016 Regulation. Here, specific reference is made to the United Kingdom, given the experience in the automation of serial decisions, starting as early as the mid-2000s, with the automated implementation of credit injunction procedures, such as Money Claim Online and Possession Claim Online.

In the experience of the United Kingdom, it was precisely Brexit that made it necessary to involve automation of decisions in order to examine all the requests for the granting of settled status, or British citizenship, which came into force in view of the United Kingdom leaving the European Union. As of 30 September 2019, the total number of applications was 1,860,200, of which 1,524,500 were concluded. Among these, 61% were accepted, while 38% were granted the "pre-settled status" and 2% of applications were rejected on the grounds of "suitability".[56] The criticism of this type of approach concerns the substance of the status, i.e. the distinction between "settled status", "pre-settled" status and citizenship, with significant consequences for possible future policy changes.[57]

"Universal Credit" (here in after "UC", established by the Welfare Reform Act 2012) is a social benefit consisting of a payment to help with living costs[58] to people who are on "a low income or out of work."[59] In order to assess the difficulties of using automatic decision systems for the provision of social benefits to vulnerable people, a survey made by the National Audit Office remarks that only 54% of applicants were able to make a claim online without help, most of them needed to verify their identity online, and the Assisted Digital Support providend by the DWP is

---

51 C. Sunstein, *Algorithms, Correcting Biases*, 12.12.2018, https://ssrn.com/abstract=3300171, 5 ss.

52 S. L. Harris, *Between citizen and migrant: the risks of ambiguity, uncertainty and opaque systems,* King's College News Centre, 16.10.2019 https://www.kcl.ac.uk/news/between-citizen-and-migrant-the-risks-of-ambiguity-uncertainty-and-opaque-systems

53 S. Howes, K. M. Jones, *COMPUTER SAYS 'NO!' Stage one: information provision*, London-Glasgow, 2019, 5.

54 L. Morris, *Reconfiguring Rights in Austerity Britain: Boundaries, Behaviours and Contestable Margins*, Journal of Social Policy, 2019, 272; I. Shutes, *Work-related conditionality and the access to social benefits of national citizens, EU and non-EU citizens.* Journal of Social Policy, 2016, 4.

55 On this point, let me reference E. Falletti, *E-Justice, Esperienze di diritto comparato*, Milan, 2008 pp. 57 ff.

56 Home Office, EU Settlement Scheme Statistics, September 2019, Second edition Experimental Statistics, 9 October 2019.

57 S. L. Harris, Between citizen and migrant, cit.

58 See https://www.gov.uk/universal-credit

59 S. Howes, K. M. Jones, COMPUTER SAYS 'NO!', cit.

not always sufficient.[60] The report itself underlines that "overall 81% of sanctions that go to tribunal are overturned, suggesting decisions are not always correct".[61]

Despite the presence of these difficulties, UC is granted through an online account with which claimants manage their UC application and communicate with the DWP. However, applicants are unaware that UC is a decision based system,[62] given that the procedure features a sort of stream of letters, informal chats with a work coach or payment statement, making a proper formal decision difficult to identify.[63] Similarly, the rules for calculating payments are unclear,[64] so, for claimants it could be difficult to challenge the DWP statement denying them benefits. The National Audit Office Report underlines that the use of a "fully automated risk analysis and intelligence system on fraud and error" has not been developed "sufficiently to understand and assess fraud and error or to provide staff with effective reporting to enable them to identify potential fraud".[65]

This is a relevant issue because it is under direct effect of the Article 22 GDPR, since Brexit had not been achieved at the time of the occurrence of these facts (May 2019). Disatisfied claimants were directed to call the UC helpline for an explanation even though DWP is obliged to provide a written statement of reasons.[66] However, this not solve the issue because the system allowed the claimant to use a helpline, that did not have full access to data calculation of awards, so the claimant's doubts on his /her request might not have a complete explanation.[67] Indeed, as reported,[68] the EWS case studies highlight that even when the DWP did provide written justification to a request by the claimant, the answer did not comply with the law and/or was not helpful to the claimant.

According to Regulations 7(3)-(4)[69] and 51[70] of The Universal Credit, Personal Independence Payment, Jobseeker's Allowance and Employment and Support Allowance (Decisions and Appeals) Regulations 2013, claimants can obtain from the authorities "written reasons for a decision within a month of being notified of the decision, and that they must be provided with these reasons within 14 days of their request or as soon as reasonably practical".[71] This regulation is stricter than Article No. 22 GDPR which is vaguer and includes no specific provision regarding methods and times for explanation to claimants.

---

60 A. Morse, Rolling *Out Universal Credit. Report by the Comptroller and Auditor General,* London, 2018, 72.
61 A. Morse, Rolling Out Universal Credit.
62 S. Howes, K. M. Jones, *COMPUTER SAYS 'NO!',* cit.; M. Monaghan, J. Ingold, *Policy Practitioners'Accounts of Evidence-Based Policy Making: The Case of Universal Credit*. Journal of Social Policy, 48 (2), 2019. pp. 351-368; A. Morse, Rolling Out Universal Credit, cit. 60 ss.
63 S. Howes, K. M. Jones, COMPUTER SAYS 'NO!', cit.
64 Difficulties are encountered regarding the calculation of childcare and housing costs (S. Howes, K. M. Jones, COMPUTER SAYS 'NO!', cit.).
65 A. Morse, Rolling Out Universal Credit, cit.
66 S. Howes, K. M. Jones, COMPUTER SAYS 'NO!', cit., 15.
67 S. Howes, K. M. Jones, COMPUTER SAYS 'NO!', cit., 15. Nevertheless, according to a Governance Social Research at least 16% of the claimants does not provided with correct information on the phone (Governement Social Research, Department for Work and Pensions, Claimant Service and Experience Survey 2017/18, Annual Survey, London, 31 January 2019, 22.
68 New research highlights a catalogue of issues with universal credit, 19.7.19, https://www.nurseryworld.co.uk/News/article/new-research-highlights-a-catalogue-of-issues-with-universal-credit
69 https://www.legislation.gov.uk/ukdsi/2013/9780111531556/regulation/7
70 https://www.legislation.gov.uk/ukdsi/2013/9780111531556/regulation/51
71 S. Howes, K. M. Jones, *COMPUTER SAYS 'NO!',* cit., 15. The authors specify that "the information that is subsequently provided should be adequate so that a person can understand the decision sufficiently well to be able to bring a challenge against the decision if it is wrong".

4. Comparative case law on automatic decisions and the right to obtain an explanation under GDPR

The Italian administrative jurisprudence has expressed itself in its highest assembly, i.e. the Council of State, on a matter arising from the use by the Ministry of Education of an algorithm that defines the assignment of secondary school teaching staff to the roles of teaching located throughout the national territory. The Council of State intervened following appeals on which the administrative courts of first instance had pronounced, with contradictory sentences, due to the difficulty of framing an innovative subject with in the traditional legal categories. In particular, the possibility that the algorithm (even if instructed by the human element, and therefore preset according to precise axiological indications) can ensure compliance with administrative procedural guarantees has given rise to debate.

The Council of State has clarified that the technical rule on the basis of which the algorithm works remains a general administrative rule, built by humans and not by the machine to be applied only by the machine itself. Therefore this rule must comply with these requirements:[72]

a) even if expressed in mathematical terms, this rule has full legal and administrative validity and therefore must be subject to the principles of publicity and transparency, and therefore must adhere to Article No. 22 GDPR, and be reasonable and proportionate;

b) it cannot leave discretionary application spaces, but must reasonably provide a precise solution for all possible cases. In this way, the administrative discretion (which cannot be delegated to the software) can be traced back to the processing of the algorithmic instrument, in obedience to Article No. 22 GDPR;

c) it is assumed that the state administration performs an *ex ante* role of mediation and composition of the interests of the parties involved, by means of constant tests, updates and improvement of the algorithm;

d) the possibility must be provided for the judge to manage, assessments on a human level on assessments made directly by automatic means, to allow the verification of the correctness of the automated procedure in all its phases.

In France, the Cour de Cassation confirmed the decision of the *Tribunal de Grande Instance* of Lyon on the annulment of the decision of the "*Comité d'hygiène, de securité et des conditions de travail (CHSCT)*" of the Crédit mutuel du Sud-Est which had appointed an expert consultant for the implementation of an artificial intelligence program (the Watson operating system of IBM) to follow the customers. According to what was stated in the decision of 12 April 2018, the implementation of this artificial intelligence program does not demonstrate the change in working conditions, as far as health and safety are concerned, in light of the art. L. 4614-12-2 ° of the French Code du travail. Indeed, the Cour de Cassation shares the assumption of the TGI that the use of this technology will cause "minor consequences in the direct working conditions of employees whose tasks will be facilitated". According to the employer, the Watson application will help employees manage the numerous emails, allowing them to re-orientate the keywords to the matters or treat them in order of priority or respond appropriately and propose different possible solutions. However, the CHSCT was concerned that this cognitive technology project entailed the risk of a

---

72 A. Pajno, M. Bassini, G. De Gregorio, M. Macchia, F. P. Patti, O. Pollicino, S. Quattrocolo, D.Simeoli, P. Sirena, AI: profili giuridici, cit.

redistribution of employee roles within an agency and therefore a significant change in working conditions.[73]

More recently, the French Conseil Constitutionnel expressed its opinion on the legitimacy of a law which extended the possibility for the public administration to resort, albeit exceptionally, to decisions capable of producing legal effects on individuals. These decisions would ben based on automatic processing of personal data. The same rule authorized automated decisions in the event that:

a) the activity of the algorithm did not concern sensitive data;

b) an administrative remedy was possible;

c) adequate information was provided concerning the use of the algorithms themselves.

The constitutional criticalities concerned a possible conflict with Article No. 21 of the French Constitution regarding the attribution of administrative powers, especially in relation to the ability to learn algorithms that could have caused the application of rules different from those set. The Conseil constitutionnel ruled out these profiles of unconstitutionality, stating that the guarantees of freedom and individual rights had been observed. These guarantees included the limitation of the use of such typologies of decisions, the provision of the conditions that legitimized them, and the possibility for the addressee to obtain the explanation in intelligible and detailed terms of the algorithmic procedure provided for by Article No. 22 GDPR.

In the United Kingdom, the High Court of England and Wales[74] has rejected a lawsuit concerning the use of automated facial recognition in relation to possible violations of privacy and data protection. The facts from which the suit arose concern a facial recognition technology that is instantly capable of capturing the biometric features of those who pass within the range of vision of police surveillance cameras, and it is able to compare such data with biometric features of people present in the police archive. Faced with this technology, the parties' defenses are opposed: on the one hand, the claimant, an association defending human rights, states that this technology is illicit under three profiles: 1. it is contrary to Article No. 8 of the ECHR, 2. it violates data protection pursuant to the Data Protection Act 1998 and 2018 and 3. it violates the anti-discrimination legislation. On the other hand, the defendant, the South Wales police authority, says it is extremely useful in crime prevention and investigations.

In the first complaint against automated facial recognition, the judges considered the possible violation of Article No 8 of the ECHR against those who had been scanned unknowingly, simply because they walked a certain route. For some time the jurisprudence, both of the European Court of Human Rights and of the English courts, has been occupied by scientific technologies in the investigative field, the use of which interferes with the protection of private life, as in the case of the collection of genetic data or fingerprints. On this point the S. and Marper v. United Kingdom decision is decisive: on that occasion the Court of Strasbourg declared that the collection of biometric data represents a violation of private life and consists of an illegal supplement pursuant to Article No. 8 of the ECHR, given the intrinsically private provenance of the personal information collected. However, in the case in question we refer to the collection of data concerning the scanning of the face, without any direct contact with the body of the person under examination. Therefore, according to the applicants this technology violates Article No. 8 of the ECHR both in

---

73 Cour de cassation, civile, ch. sociale, arrêt du 12 avril 2018 https://www.legalis.net/jurisprudences/cour-de-cassation-civile-ch-sociale-arret-du-12-avril-2018/

74 Bridges, R (On Application of) v The Chief Constable of South Wales Police [2019] EWHC 2341 (Admin) (04 September 2019).

the light of the first paragraph, as at the moment there is no appropriate legislation in the United Kingdom that guarantees the privacy of subjects who are subjected to such investigations, and in light of the second paragraph, because the current British legislative framework (that is the Police and Criminal Evidence Act 1984 and the Data Protection Act in its two versions of 1998 and 2018) is insufficient and presents the appointment with the unjustified technology under the perspective of its use in a democratic society, not being the same interference required by law.

The judges reject this reconstruction using the powers attributed to the police by Common Law (in particular with reference to the previous case In R (Catt) against the Association of Chief Police Officers [2015] AC 1065) and recognize the legitimacy of research and personal information regarding maintaining public order, preventing and combating criminal activities. Furthermore, they specify that the collection and conservation of biometric data observed through cameras and facial recognition programs does not assume the characteristics of intrusiveness prohibited by the Common Law and the protection of Habeas Corpus, because there is no physical threat . However, some scholars stated that the collection and conservation of such data can still violate the psychological integrity of a person, and therefore his/her right to privacy, in the absence of informed consent.

The second complaint concerns the violation of the Data Protection Act 1998 and of the Data Protection Act 2018, which implements the European Union legislation currently in force (pending the resolution of the Brexit node). On this point the judges reject the reference of the plaintiffs to the decision of the Court of Justice of European Union Breyer against Bundesrepublik Deutschland (C-582/14), which adopts a broad interpretation of the dynamic IP address as personal data. In fact, the English judges argue that the anonymisation of personal data does not preclude the identification of the person to whom those data refer.

As regards the use of facial recognition for the purposes of "law enforcement", the Court rejected the argument of the defendant, according to which only the personal data of those who were on the checklists were processed in a sensitive manner. On this point, the Sect. 35 (8) (b) of the Data Protection Act 2018 can be correctly read as applicable both to the biometric data contained in the list of persons to be checked, and to the biometric data of the members of the public. In fact, this is only possible if each person is uniquely identified. This conclusion is supported by the inclusion of biometric data in the GDPR (Regulation EU/679/2016) and in the EU/68/2016 Directive on data protection, and the fact that this legislation is incorporated in British law by the Data Protection Act 2018. Otherwise there would be a breach of the Data Protection Act 2018.

Finally, concerning the third reason, anti-discrimination, with regard to the possible occurrence of so-called "False positives", the Court rejected the request, stating that the parties did not provide sufficient evidence in this regard. On this point the doctrine noted that a periodic revision of such cases is necessary in the light of the Equality Act 2010 and the Human Rights Act 1998.

The same court decided a case[75] regarding the use of algorithms to decide the attribution of funding and grants to students. Specifically, the Court allowed the "Office for Students"[76] (hereinafter OfS) to publish the registration refusal for Barking and Dagenham College. These decisions were

---

75 Barking & Dagenham College, R (on the application of) v The Office for Students [2019] EWHC 2667 (Admin) (11 October 2019)

76 established by the Higher Education and Research Act 2017 ('HERA') with effect from 1 January 2018. Its functions under that Act include establishing and maintaining a register of English higher education providers (s. 3(1)). The information contained in the register must be made publicly available (s. 3(9)) (Barking & Dagenham College, R (on the application of) v The Office for Students [2019] EWHC 2667 (Admin))

grounded on three indicators: student completion rates, qualification outcomes and graduate employment. According to the Office for Students, the Barking and Dagenham College performed badly and refused registration under specific rules.[77] The college argued that OfS gave little or no weight to contextual factors, and that the detail of the algorithms used to create these metrics was not shared with the college until after the decision to refuse registration had been taken. This case is very similar to the decisions established by the Italian Council of State under the software use perspective. In this case, the Court cannot pursue an "interim relief" because the students had a specific right to know regarding the status of their data used in the (automatic) decision-making process.[78]

The judgment affirmed that such "interim relief" would not be available, primarily because it was felt that current students had a "right to know" the status of their provider. But we also get to see some of the arguments that would underpin a full appeal; these were primarily focused on the quality and applicability of data used in OfS' decision-making process. According to scholars, "(W)hile on the surface the funding algorithms provided a direct line between the data and the decision, the reality was softened by the ability of the Higher Education Funding Council for England (Hefce) to retrospectively adjust funding should the initial data hypothesis prove to be incorrect or unwarranted". But this seems an *ex post* adjustment of a disfunctional funding management. What is important here is to understand how algorithms work(ed) in granting students funding, especially in terms of transparency and complexity. In this case, "the failure of the OfS to disclose the detail of the algorithms at a suitably early stage in the registration process arguably places the provider at a distinct disadvantage."[79] This case underlines, on the one hand, that the public authorities, in this case the Office for Students" had a responsibility (and a liability) to assign student funding properly, in a transparent way; and on the other hand, the fact that the higher education providers needed to manage algorithms fairly given the critical and fundamental role attributed to them by the regulation itself.[80]

---

77 These rules were explained at par. 9 of (Barking & Dagenham College, R (on the application of) v The Office for Students [2019] EWHC 2667 (Admin)). These rules are not of interest regarding automatic decision, so they were not specifically analyzed.

78 A. Youell, *Could your data end up in the dock?*, 13.10.19, available on https://wonkhe.com/blogs/could-your-data-end-up-in-the-dock/

79 A. Youell, Could your data end up in the dock?, cit.

80 A. Youell, cit.