Public Officials and the Design of Algorithms. Lessons from the Italian Experience*

Fulvio Costantino

(Researcher in Administrative Law at The University of Macerata)

ABSTRACT The essay deals with the topic of algorithm design in the public sector. It takes into account the state of the Italian public administration, the national and international debate, and the jurisprudence that has dealt with the issue. Particular emphasis is placed on the contribution that compliance with the principles of transparency and participation can make to the solution of the most relevant problems at a global level.

1. *Introduction*

The Italian public administration, given the high average age and the almost exclusively legal and economic training of its officials, needs important plans for hiring staff, equipped with the skills that the administration is most lacking: the entry of young engineers, architects, sociologists and computer scientists must be promoted[1].

In the face of this need, the official training must be constantly rethought[2]: the question relating to how new human and instrumental resources are intended to modify the *modus operandi* of the administration is particularly relevant[3].

It is a matter not only of conveying the learning of the principles and provisions contained in the discipline, at every level, but also of refining the skill to decline them in administrative activity, in particular in the face of an increased and wider use of techniques and tools that adopt languages and comply to logics that must necessarily be adapted to the pursuit of the public interest.

The problem arises in reference to the ever wider use that even the public authorities make of artificial intelligence[4]. If the digitization of

---

*Article submitted to double-blind peer review.

[1] The current situation is well summarized in S. Cassese (ed.), *Rapporto Aspen Institute Italia, Le riforme della pubblica amministrazione nella XVII Legislatura. I motivi ispiratori, i risultati conseguiti, gli obiettivi da raggiungere*, 2018, available on www.aspeninstitute.it, 37. See M. D'Onghia, *Il fabbisogno del personale pubblico e le politiche di reclutamento tra nuove sfide e riduzione dei costi*, in *Rivista giuridica del lavoro e della previdenza sociale*, 2018, 435, in the light of Legislative Decree no. 75/2017 (regarding the recruitment of public personnel and the organization of offices, and a new regulation of the identification of needs), which in any case points out the problem deriving from the scarcity of economic resources. On this decree see M. Magri, *Il lavoro pubblico tra sviluppo ed eclissi della "privatizzazione" (Commento a d.lg. 25 maggio 2017, n. 75)*, in *Giornale di diritto amministrativo*, 2017, 581. Regarding simply the type of skills that are required by the use of artificial intelligence, see www.bigdata4innovation.it/big-data/big-data-data-science-verso-data-driven-economy-data-monetization, where the indispensable figures in a company are summarized: here reference is made to the *Data Governance Manager* (who could also be the *Chief Data Officer,* who deals with *Governance* strategies and their implementation); to the *Data Owner,* responsible for defining the specific *business* requirements of the data and how they must be addressed towards *business* actions; to the *Data System Owner,* who follows the technical management of *the data governance* systems; to the *Data Quality Manager,* who defines the reference framework for the company *Data Quality*; to the *Data Steward,* who defines and controls the implementation of company policies and *standards* related to data quality; to *Data Users,* who are called to work on data with different levels of authorization, with different tasks to achieve the objectives. This simple list shows well how data analysis requires multiple and specialized skills.
[2] The issue of training is central, but not adequately addressed. Regarding the training of the *elite,* F. Patroni Griffi, *"Élite" politiche e amministrative tra crisi della*

democrazia e riforme dell'amministrazione, Relazione al Convegno "Emergenza classe dirigente" tenutosi a Firenze il 3 novembre 2017, in *Federalismi.it*, 2017. M. Ramajoli, *Quale cultura per l'amministrazione pubblica?,* Relazione al Convegno "Verso le carriere amministrative" tenutosi a Roma il 19-20 gennaio 2017, in *Giornale di diritto ammnistrativo*, 2017, 187. With reference to public management, also from a comparative point of view, see G. Barbetta, *La formazione dei dipendenti pubblici ancora Cenerentola tra esigenze di razionalizzazione e contenimento della spesa*, in *Rivista giuridica del lavoro e della previdenza sociale*, 2013, 875.
[3] On the *modus operandi* of the public administration, for all, R. D'amico, *L'analisi della pubblica amministrazione. Teorie, concetti e metodi. Vol. I La pubblica amministrazione e la sua scienza*, II ed., Milano, Franco Angeli, 2015. With reference instead to how computerization impacts on the work of the administration, see F. Costantino, *Autonomia dell'amministrazione e innovazione digitale*, Napoli, Jovene, 2012.
[4] Two essays in which the issue of the relationship between public administration and artificial intelligence is addressed (it should be noted that the latter feeds on large masses of data, the so-called *big data,* for which the link between *big data* and artificial intelligence is normally dealt with): F. Costantino, *Rischi e opportunità del ricorso delle amministrazioni alle predizioni dei big data*, in *Diritto pubblico*, 2019, 43, and F. Costantino, *Intelligenza artificiale e decisioni amministrative*, in *Rivista italiana per le scienze giuridiche*, 2017, 357; also M. Tresca, *I primi passi verso l'Intelligenza Artificiale al servizio del cittadino: brevi note sul Libro Bianco dell'Agenzia per l'Italia digitale (Relazione al "III Colloquio italo-francese sul Diritto del Web", Roma, 21 giugno 2018)*, in *Rivista di*

administration in Italy so far has been carried out promptly and effectively in the municipalities in which the average age of officials is lower, on the one hand, this makes it essential to adopt hiring logics that allow the public authorities to make use of already trained people who are able to make use of innovations with a view to pursuing the public interest; on the other hand, the change arises for everyone, citizens and officials of the administration, and everyone must be able to exploit without prejudice the benefits that the changes underway allow[5].

Automation is not a very recent phenomenon: have been studied in years in Italy the cases in which it is directly the computer to make a decision, thanks to instructions defined algorithms. Studies on the automated administrative decisions have been published, for years judicial decisions have been adopted related to measures substantially, even if not formally, adopted by the computer – i.e. about exceeding the speed limits[6].

An aspect of particular interest concerns the ability of processors, based on past data, to make decisions, using the future projections that they themselves produce[7]. The issue has become current again in the light of the affirmation of the so-called *big data:* this term refers to large amounts of data, which can be transferred electronically, centralized at private operators[8] and public authorities[9] and used in multiple

ways. Indeed, thanks to the ability of information systems to intertwine and process information, it is possible to obtain new knowledge, of even greater value, to predict events and behaviours, and decide accordingly[10]. *Big data* can operate thanks to the so-called *algorithms*, sets of instructions or rules given to a computer to follow and apply: they define priorities, data classification and association.

Information is the basis for the decisions of the administration, which has a support that allows more effective and effective measures. Many disparate, unrelated data are reconnected, recurrences and sequences are identified. The common process is reversed: you have the data available, and you are called to ask questions. The administration, based on enormous amounts of data relating to the past, can make decisions, making use of the future projections that the computers produce.

Cases of "prediction" have always been carried out by experts: by police officers with regard to the commission of crimes, by social workers with regard to mistreatment, by teachers for the assignment of scholarships; the quality of the algorithmic prediction is now much better than in the "analog" mode.

The prediction, on the other hand, produces the effect of limiting the sphere of action of a subject, both of the recipient of the measure resulting from the prediction, who undergoes a particular treatment on the basis of any behaviour that has not - yet - held, and of the users of the forecast, as influenced by it[11].

It should be added that, among the algorithms, the *Machine Learning* (ML) ones play a particularly significant role: they are designed for learning, so that the computers can modify its rules based on the new data: based on the training data, the machine can learn the rules by itself and build its own decision model; it therefore does not just execute instructions but finds patterns in the data and turns them into

diritto dei media, 2018, 222, www.medialaws.eu; D. U. Galetta – J. C. Corvalàn, *Intelligenza Artificiale per una Pubblica Amministrazione 4.0? Potenzialità, rischi e sfide della rivoluzione tecnologica in atto*, in *Federalismi.it*, 2019.

5 A. Banfi and G. Galli, *La digitalizzazione delle pubbliche amministrazioni*, in *OCPI*, 2020, www.osserva-toriocpi.unicatt.it/cpi-Digitalizzazione%20PA(1).pdf.

6 V. Buscema, *Discrezionalità amministrativa e reti neurali artificiali*, in *Il Foro amminstrativo*, 1993, 620; G. Duni (ed.), *Dall'informatica amministrativa alla teleam-ministrazione*, Roma, Istituto Poligrafico Zecca dello Stato, 1992; A. Masucci, *L'atto amministrativo informatico. Primi lineamenti di una ricostruzione*, Napoli, Jovene, 1993; A. Natalini, *Sistemi informativi e procedimenti amministrativi*, in *Rivista trimestrale di diritto pubblico*, 1999, 449; A.G. Orofino, *La patologia dell'atto amministrativo elettronico: sindacato giurisdizionale e strumenti di tutela*, in *Foro amministrativo – C.d.S.*, 2002, 2257.

7 I. Kerr and J. Earle, *Prediction, Preemption, Presumption: How Big Data Threatens Big Picture Privacy,* in *Stanford Law Review Online*, 66, 2014, 72.

8 *State actors*? According to D. Rudofsky, *Modern State Action Doctrine in the Age of Big Data,* in *New York University Annual Survey of American Law,* 2016, 741. The processing of these large amounts of data is now within the reach of more operators than in the past: on this point see S. Fahey, *The democratization of big data,* in *Journal of National Security Law & Policy,* 2014, 325, 332.

9 On public data, see F. Malomo and V. Sena, *Data Intelligence for Local Government? Assessing the Benefits and Barriers to Use of Big Data in the Public Sector*, in *Policy and Internet*, 9, 2017, 9; K. C. Desouza and J. B. Jacob, *Big Data in the Public Sector: Lessons for Practitioners and Scholars*, in *Administration & Society*, 49,

2017, 1043; S. Vydra and B. Klievink, *Techno-optimism and policy-pessimism in the public sector big data debate*, in *Government Information Quarterly*, 36, 2019, 101383; C. Fredriksson, F. Mubarak et al., *Big Data in the Public Sector: A Systematic Literature Review*, in *Scandinavian Journal of Public Administration*, 21, 2017, 39; M. Maciejewski, *To do more, better, faster and more cheaply: using big data in public administration*, in *International Review of Administrative Sciences*, 83, 2016, 120; U. Fantigrossi, *I dati pubblici tra Stato e mercato*, in *Amministrare*, 2007, 291.

10 See McKinsey Global Institute research, *Big Data: The next frontier for innovation, competition, and productivity,* 2011, the 2012 White Paper by researchers *Challenges and Opportunities with Big Data*; B. Franks, *Taming The Big Data Tidal Wave: Finding Opportunities in Huge Data Streams with Advanced Analytics*, New York, Wiley, 2012.

11 See R. H. Sloan and R. Warner, *Algorithms and Human Freedom*, in *Santa Clara High Technology Law Journal*, 2019, 36, 1.

instructions to follow. The use of *machine learning* algorithms is practiced in *web* search, spam filters, reporting systems, ad placement, credit evaluation, fraud detection, stock trading, drug design.

In the public sector, almost all applications currently used are based on algorithms of the first type, and therefore on predefined, clear and explainable instructions[12]. However, for example, the predictive police tools already in use in Italy are based on machine learning[13]. These are more effective solutions to the objectives that are to be pursued but gradually escape human understanding.

The problem, which is the subject of this contribution, is the design of algorithms, from the perspective of the responsibility of public officials.

2. *Section I. Neutrality and skills*

A decisive profile concerns the supposed neutrality and objectivity of the work of the processors, which calls into question the role of the public official, even and above all if he does not have adequate computer skills[14].

In fact, even beyond the aforementioned hypothesis of automated measures[15], he may be in a position to formally take a decision taken substantially by artificial intelligence, but without being able to understand, integrate and correct it.

The problem is that algorithms can either inform decision making or make decisions. If the predictions show to the administration where to focus attention (surveillance, investments, maintenance), the choices do not appear to be *in se* questionable. The predictions thus guide operators even when the decision belongs entirely to human beings (expenditure to be paid or fraudulent; tax return to be audited; travellers to be subjected to further checks). This is an unregulated case: art. 22 of the GDPR in fact prohibits the adoption of a fully automated measure and refers to legal effects or effects that affect in a similar way, and therefore to the decision.

The level of attention on the direct use of IT tools in the adoption of administrative measures is very high; instead the use in the preliminary investigation and decision support is not equally studied.

Nor should we think that the algorithms that probe, intertwine and process *big data* are neutral[16]. The programmer's hand affects how the algorithms operate, and each of them is only one of the possible ones: in this sense the

---

[12] C. Coglianese and D. Lehr, *Regulating by Robot: Administrative Decision Making in the Machine-Learning Era*, in *The Georgetown Law Journal*, 105, 2016, 1147.

[13] R. Coluccini, *La polizia predittiva è diventata realtà in Italia e non ce ne siamo accorti*, 21st November 2018, vice.com/it/article/pa5apm/polizia-predittiva-italia-lombard-i-xlaw-prevedere-crimini- algorithms. The XLAW system, developed by an inspector of the Naples police headquarters, validated by the Federico II and Parthenope Universities of Naples and already adopted by at least 9 police stations, adopts *machine learning* to find criminal models, starting from the reports provided; the software is able to identify new models without the intervention of the programmer, providing geo-referenced predictive *alerts,* as the software identifies "hunting grounds", areas of the cities where crimes occur most (a map of the city is displayed and circles indicate the crimes which will take place in a specific street and at a specific time, with notification to the smartphone of the police officer) (E. Lombardo, *Sicurezza 4P. Lo studio alla base del software XLAW per prevedere e prevenire i crimini*, Venice, Mazzanti Libri ME Publisher, 2019). The data entered are the complaints presented by the victims of thefts and other types of crime, and the consequent judicial measures. Numerous other information is also taken into account, relating to the frequency of criminal acts and the neighborhoods in which they tend to recur, the number of people residing in a certain area of the city, the presence of banks and commercial activities, holidays, timetables (businesses, schools, offices, public transport), commercial and tourist flows, weather conditions, sporting and musical events; the characteristics of the suspect (gender, height, citizenship, distinctive and biometric signs), of the victims, the types of stolen goods. Furthermore, the system creates unforeseen criminal models (drug dealing in homes, baby gangs, theft in construction sites).
It is not currently possible to access the algorithms, as the rights to the source code are held exclusively by the developer and the software is given in free loan to the Central Anti-Crime Directorate of the Ministry of the Interior.

[14] The well-known ruling of the Consiglio di Stato, sect. VI,

8th April 2019, n. 2270, which reformed the equally well-known decision, Tar Lazio, Sect. III *bis*, 10th September 2018, n. 9227, on which, see I. Forgione, *Il caso dell'accesso al software MIUR [Ministero dell'Istruzione, dell'Università e della Ricerca] per l'assegnazione dei docenti (Nota a TAR Lazio, Roma, sez. III* bis*, 14 febbraio 2017, n. 3769)*, in *Giornale di diritto amministrativo*, 2018, 647. It refers to a case in which the official did not intervene in making the decision, but the problem of what happens beforehand and therefore of the ability of the official to understand the algorithm and ask for its correction, to avoid decisions incorrect raised; the idea that the computer takes decisions without the administration being able to intervene to avoid it looks unacceptable.

[15] The debate now also extends to discretionary measures: see recently S. Bianchini and L. M. Lucatuorto, *Discrezionalità e contemperamento degli interessi nei processi decisionali dell'amministrazione digitale*, in *Ciberspazio e diritto*, 2009, 41 and M. Asprone, S. Gaudini, *L'automazione dell'attività amministrativa tra vincolatività e discrezionalità*, in *Nuova rassegna*, 2011, 145. On the subject, again recently, P. Otranto, *Decisione amministrativa e digitalizzazione della p.a.*, in *Federalismi.it*, 2017, and F. Patroni Griffi, *La decisione robotica e il giudice amministrativo*, in www.giustizia-amministrativa.it, 2018. The issue is not new at all: dealt with it, for example. A. Usai, *Le prospettive di automazione delle decisioni amministrative in un sistema di teleamministrazione*, in *Diritto dell'informazione e dell'informatica*, 1993, 163 and the philosophers G. Caridi and S. Pellecchia, *Un modello operativo per l'automazione di procedure giuridiche*, in *Rivista internazionale di filosofia del diritto*, 1985, 3, 37.

[16] S.K. Glaberson, *Coding over the cracks: Impoverished Algorithms: Predictive Analytics and Child Protection*, in *Fordham Urban Law Journal*, 2019, 46.

relationship between scientist and public decision maker becomes crucial. An effect of desubjectivization occurs, so that the work of machines is perceived by users as more objective than it is.

An interesting essay examined the use of algorithms for predictive policing[17] and identified the assumptions from which they (mostly implicitly) move: that the data used do not reflect accurately the reality; that the future is like the past; that the variables omitted are irrelevant; that the algorithms are neutral; that the data justifiably discriminate; that the place of possible offenses plays a primary role; that police deployment is the primary intervention; that the implementation of algorithms happens in a perfect way; that the change of police deployments prevents crime; that attention to crime is always appropriate. Even without the need to discuss them all, it is clear how the choices made upstream, on the data to be entered, on the measures to be taken, on the philosophy behind the algorithm determine the effects. The French National Bar Council (CNB) states: «we must avoid the obsession with efficiency and predictability that motivates the use of the algorithm that leads us to design categories and rules no longer in consideration of our ideal of justice, but so that they are more easily 'coded'».

The supposed technological neutrality[18] is an excellent expedient to avoid the problem of the democratic *accountability* of the decision maker. Also for these reasons, some "prejudices" underlying the algorithms can never be completely removed, also because they are not necessarily explicit, and not necessarily for malice, but generally for unconsciousness: as an effect, the values of the so-called data scientists are replicated, with the risk of perpetuating visions of society resulting from stereotypes and prejudices[19].

Rand Hindi (entrepreneur and data scientist) expresses his point of view on mutual prejudices: "Artificial intelligence makes fewer mistakes than humans, but it makes mistakes that humans would not have made"[20]. In particular, humans

are afraid of the possibility of a machine error that humans would not normally have made.

3. *Defensive Administration*

The context in which the issue of the decision induced by the computer takes place is therefore the responsibility of the official: in its concrete application, there is the concern of the public administrators not to be held accountable for their alleged improper actions. Traditionally it leads to adapt to what is elsewhere or by others decided[21].

A strong tendency of public decision-makers to deference to decisions, or even simply to the predictions of processors, was found - the so-called "*Automation bias*"[22]: operators of automated systems tend to trust the responses of such systems, even in circumstances where their judgment or experience would have led to a different conclusion. People assume that algorithmic decisions are made on the basis of indisputable rigid sciences, operating at a level beyond human capabilities, and they are afraid of making mistakes: after all, it is in the nature of

---

[17] L. Bennett Moses and J. Chan, *Algorithmic prediction in policing: assumptions, evaluation, and accountability*, *Policing and Society*, 28, 2018, 806-822.

[18] See also S. Valentine, *Impoverished Algorithms: Misguided Governments, Flawed Technologies, and Social Control*, in *Fordham Urban Law Journal*, 2019, 46, 364.

[19] But see J. Kleinberg, J. Ludwig, S. Mullainathany and C. R. Sunstein, *Discrimination in the age of algorithms*, in www.papers.ssrn.com/sol3/papers.cfm?abstract_id=3329669.

[20] R. Brauneis and E.P. Goodman, *Algorithmic Transparency for the Smart City*, in *Yale Journal of Law and Technology*, 2018, 103. The UE *White Paper On Artificial Intelligence - A European approach to excellence and trust* COM(2020), 65 final, refers to the accident fatal of a Tesla car that did not identify the truck in the middle of

the road, while a driver who observes the road carefully cannot make the same mistake.

[21] S. Battini and F. Decarolis, *L'amministrazione si difende*, in *Rivista trimestrale di diritto pubblico*, 2019, 293 takes up the issue of responsibility, already addressed under a more strictly legal profile, in Id., *Responsabilità e responsabilizzazione dei funzionari e dipendenti pubblici (Intervento al Convegno "De la responsabilité à la responsabilisation des fonctionnaires", Tours (Francia), 20-21 novembre 2014)*, in *Rivista trimestrale di diritto pubblico*, 2015, 53; see also. L. Cominelli, *Cittadini e amministrazioni: burocrazie disfunzionali e diffusione dei rimedi*, Relazione al convegno "...il più vicino possibile al cittadino...", Trieste, 29 e 30 settembre 2009, in *Diritto e società*, 2010, 3, 124; R. Cavallo Perin and B.Gagliardi, *Status dell'impiegato pubblico, responsabilità disciplinare e interesse degli amministrati*, in *Diritto amministrativo*, 2009, 53.
The first contribution is particularly relevant for our purposes: up to now the issue of the so called *defensive administration* (the risk that the official does not take decisions or conducts useful for the pursuit of the public interest, and instead assumes others, or remains inert, for fear that individual disutilities may result from his actions) has not been adequately examined. In this regard, the limitation of liability before the Corte dei Conti for willful misconduct or gross negligence is not enough to reduce the perceived risk (on this point, see also E. D'Alterio). The perception of risk is largely the result of uncertainty about the boundary between lawful and unlawful, so that there is a tendency to postpone decisions, ask for opinions, make decisions "which are also respectful of the rules deemed inapplicable, because one day someone could decide to apply them": it is then correctly observed that the first reform should be the reduction of regulatory chaos.

[22] About Automation bias in public sector, see C. Hall, *Challenging Automated Decision-making by Public Bodies: Selected Case Studies from Other Jurisdictions*, in *Judicial Review*, 25, 2020, 8; the report T. Snow, *Decision-making in the Age of the Algorithm*, 2019, in www.nesta.org; J. Cobbe, *Administrative law and the machines of government: judicial review of automated public-sector decision-making*, in *Legal Studies*, 39, 2019, 636.

---

algorithmic systems to engage in complex and potentially indecipherable calculations and calculate on a scale of factors and at a speed that man cannot replicate.

Following the indication of the machine may be tempting for public officials, as they may be considered exempt from liability, acting as defined, and indeed a failure to comply could expose them to a possible conviction for damage (consider the case in which the computer prescribes to carry out maintenance work on some infrastructures to avoid damage, in the event that the administration remains inactive and an accident occurs). The phenomenon is exacerbated in the context of the use of force, in which case there is both the risk of annulment of the measure and, above all, of negative external consequences. The effect achieved is the weakening of the judgment capacity of the public decision-maker, with contextual de-responsibility and de facto demotion.

In this sense, the reliance on artificial intelligence would not represent anything other than the most recent declination of the so-called attitude of the so called *defensive administration*, very widespread. An attitude that well aspires to the illusory prospect of elimination of discretion[23], which would benefit from the obfuscation of responsibility that occurs in "multi-agent" structures and, in this case, from the intertwining between human and technical operators, since the algorithmic decision, even if made by man, it is however a co-decision[24].

The phenomenon, to be handled effectively, requires the adoption of various instruments related to the work of civil servants: the recruitment of personnel able to design and use these systems; the training of administrators, especially local ones, to provide them with the tools, first of all cultural, to intervene, which prevent the seduction of a demotion disempowering; training to face unjustified automation errors; the investment in policies aimed at the efficiency (and the evaluation of the relative results) of the administration, to overcome an approach that has contented itself with controlling the mere legality of the administrative action and has thus promoted de-responsibility[25].

Two issues, however, appear to be central: the first, relating to the *ex ante* design of the algorithms by the administration; the second, the *ex post* control on the use of the algorithm. We will concentrate on the first profile, following the instructions coming from official documents and judicial rulings.

4. *Documents and the quest for a strategy*

The documents, whether they are normative sources or guidelines, address the issue of the design of artificial intelligence with a view to preserving human decision.

The great attention paid to the guidelines affecting the sector appears to be relevant: the globality of the phenomenon, the speed of innovation and the need not to stifle progress make the adoption of legal rules difficult, especially in light of the strong competition between regulations on the subject[26]. The aversion to norms also derives from the inevitability of interdisciplinarity, since the legislator, alone, could approve rules out of context and ineffective[27].

The Guidelines on the protection of individuals with regard to the processing of personal data in a world of Big Data of the Council of Europe, dated January 2017, drawn up on the basis of the Convention on the Protection of Individuals with regard to Automatic Processing of data personnel (so-called Convention 108, 1981, amended in 2010), clearly indicate, with reference to decisions, that the autonomy of human intervention in the decision-making process must in any case be

---

[23] On the subject of the reduction of discretion, see G. Tropea, *La discrezionalità amministrativa tra sempli-ficazioni e liberalizzazioni, anche alla luce della legge n. 124/2015*, in *Diritto amministrativo*, 2016, 107.

[24] G. Avanzini, *Decisioni amministrative e algoritmi informatici. Predeterminazione, analisi predittiva e nuove forme di intelligibilità*, 2018, Editoriale Scientifica, Napoli, 40.

[25] The risk is that the official does not take decisions or conducts useful for the pursuit of the public interest, and instead assumes others, or remains inert, for fear that individual disutility may result from those choices. The "defensive" official distorts for individual purposes conduct

that should be characterized by collective ends, not much differently from the corrupt official. While the latter pursues an individual gain with collective loss, the defensive officer forgoes a collective gain to avoid the risk of an individual loss (…). However, this is a less visible, more subtle and, for this reason, perhaps more dangerous form of illegitimacy. It does not consist of acts, decisions or even omissions referable to specific obligations to act. Instead, it consists of more indefinite delaying techniques, of apparent choices, of delays allowed by the legal system, of formal acts that hide substantial inertia, etc. All this (…) is not recorded by the "radar" of the more traditional administrative law of legality. Different tools are needed to intercept it, which are precisely those of the administrative law of efficiency (...) (Battini and F. Decarolis, *L'am-ministrazione si difende*, 2019).

[26] The French CNIL, the institution responsible for the ethical and social issues raised by the evolution of digital technologies, in particular algorithms and artificial intelligence, in addressing the question on the regulation of algorithms, states that: it would be too early to impose rules which would necessarily prove unsuitable and destined to be rapidly obsolete by the technical developments which are now proceeding at an immeasurable pace compared to that of legal invention.

[27] See also, about the use of soft law, R. Hagemann. J. Huddleston Skees and A. Thierer, *Soft Law for Hard Problems: The Governance of Emerging Technologies in an Uncertain Future*, in *Colorado Technology Law Journal*, 2018, 17, 37.

preserved; that one cannot rely on results of the data or on decontextualized information without regard to the circumstances relating to data; that in the event of a significant impact on individual rights, or in any case legal effects, the human decision-maker must give reasons for the motivation underlying the processing, including the consequences for it[28].

The European Union adopted in April 2019 *Ethics Guidelines for Trustworthy AI*. The document is interesting, because the European Union is able to adopt rules[29]. The reliability of AI is considered the effect of three components: legality, ethics and robustness from a technical and social point of view. On the one hand, the statement underlines how the legal profile is fundamental, and indeed it is placed in the first place, although the flourishing of ethical guidelines - and the same decision of the Union to make use of them - could have led to think that the legal profile should not be privileged; on the other, it is stated that it is not sufficient, if not accompanied by the other two components. On the other hand, however, the principles are unlikely to be effective if not translated into rules; they also make references to legal institutions which are commonly implemented in national and supranational systems in other sectors.

The guidelines deal with respect for human autonomy, damage prevention, fairness and explicability: "provisions" are invoked (only to verify which ones and of what nature) to avoid possible damage to democracy, the rule of law, distributive justice. Human intervention and surveillance, technical robustness and security, confidentiality and data *governance*, transparency, diversity-non-discrimination-equity, social-environmental well-being, *accountability are invoked* (page 5). They refer to participation, as the involvement of stakeholders (page 6). A *checklist* is suggested to be followed.

The reference to participation, controls, transparency, motivation, that are legal instruments, are evident.

Also interesting are the European Parliament document of February 2017, "*Recommendations to the Commission on Civil Law Rules on*

robotics", which places the need for human control of automated and algorithmic decision-making processes[30], the Communication "*Artificial intelligence for Europe*" adopted in April 2018[31], as well as the most recent "*White Paper on artificial intelligence*" (February 2020).

Already in the 2018 document, the development of principles relating to the responsibility of machines and the regulation of the algorithms used are proposed, as well as the establishment of an Agency with regulatory powers both on Artificial Intelligence and on the algorithms adopted. A form of liability, a robot registration system, compulsory insurance and a guarantee fund for damage caused by uninsured robots (par. 59), as well as strict forms of liability are proposed (par. 53 and 54).

The main document to examine is the "*White Paper on Artificial Intelligence - A European approach to excellence and trust*" (COM 20202 65): it lists all the requirements that could be established by law. The chapters (§D) concern: data-based training (it is necessary to verify that the safety rules of the EU are respected, for example that the data sets are sufficiently large and cover all relevant scenarios to avoid dangerous situations, that the outcomes do not lead to discrimination for gender, ethnicity and other, and that the privacy regulations are respected); the storage of data for a reasonable time (to keep track of the programming algorithm and training methods, of data used to train the AI systems, including its features and how the data have been chosen) and in some cases, keep the data themselves, to allow tracing and verifying actions and decisions, and therefore to verify and access, while protecting confidential information. This would also be useful for operators, who would know the rules to respect[32].

The communication focuses on the requirements of robustness and precision (*sub* d): algorithms must be developed in a responsible way, with an adequate *ex ante* risk assessment and must behave according to forecasts. In this sense, it must be ensured that errors can be adequately managed, that the system is resistant

---

[28] https://rm.coe.int/16806ebe7a. See A. Mantelero, *Regulating big data. The guidelines of the Council of Europe in the context of the European data protection framework*, in *Computer Law & Security Review*, 33, 2017, 584.

[29] See N. A. Smuha, *The EU Approach to Ethics Guidelines for Trustworthy Artificial Intelligence*, in *Computer Law Review International*; 20, 2019; 97; S. Larsson, (Accepted/In press), *On the Governance of Artificial Intelligence through Ethics Guidelines*, *Asian Journal of Law and Society*, 2020, in www.portal.research.lu.se/portal/files/79693973/Larsson_18_May_2020_On_the_Governance_of_Artificial_Intelligence_through_Ethics_Guidelines.pdf.

[30] Discussed by A. Ponce, *A Law on Robotics and Artificial Intelligence in the EU?*, ETUI Research Paper - Foresight Brief of the 2nd September 2017, https://dx.doi.org/10.21-39/ssrn.3180004.

[31] Artificial Intelligence for Europe, COM (2018), 137 final. See A. Ponce, *Artificial Intelligence: A Game Changer for the World of Work*, ETUI Research Paper - Foresight Brief of the 5th June 2018, in www.dx.doi.org/10.2139/ssrn.3198581.

[32] The communication also proposes "proactive transparency regarding capabilities, limitations, purpose, operating conditions and the level of accuracy expected. Citizens should be informed of the interaction with artificial intelligence, unless it is evident. The information must be objective, concise and easy to understand and appropriate to the specific context".

to attacks and that mitigation measures are taken. With reference to human control, it is intended to prescribe the involvement of human beings for high-risk applications. The possible solutions are that the result does not become effective before it has been reviewed and validated, or that it is effective but a subsequent human control is guaranteed, or that the system can be monitored in real time with the possibility of intervention or that some operational constraints are imposed on the system in the design phase[33].

Another issue concerns the recipients of the obligations (§ E): the latter must be shared in particular among the developers and the people who apply the algorithms and therefore have to control. It is obviously essential that economic operators are held accountable even if they are not established in the Union.

It indicates the need for a preliminary assessment of compliance (§ F) with testing, inspection and certification procedures, in addition to controls on algorithms and data for high-risk cases, although not all requirements can be verified ex ante; in addition, AI systems can evolve, so repeated evaluations are required; furthermore, it is necessary to ensure that any anomalies can be remedied. This does not exclude *ex post* controls by the national authorities, possibly also entrusted by third parties. The establishment of a voluntary labelling system is envisaged (§ G).

In terms of governance (§ H), the need not to overlap the competences and to have a structure that allows the exchange of information, practices, the provision of opinions and advice, based on a network of national and sectoral authorities, is reiterated: a committee of experts is also mentioned; the participation of stakeholders must be guaranteed; conformity assessments could be assigned to bodies designated by member states that carry out independent assessments.

If we discuss the measures that Italy is taking or intends to take, we must take into account that in the Commission European asked for the June 2019 the adoption of a strategy for AI, but Italy, unlike other countries, did not meet the deadline[34].

The Italian government's strategy for AI, which is expected (a draft circulated for consultation in July 2019[35]), appears to be based on the principles of anthropocentrism, reliability and sustainability of AI. It is stated that the aim is to counter the risks of exacerbating discrimination and social and territorial imbalances potentially deriving from an unconscious use of AI. To this end, an inter-ministerial and multi- stakeholder control room is set up[36], as well as an Italian Institute for AI (I3A)[37].

---

[33] With regard to biometric data and remote identification, through facial recognition systems in public places, which may involve risks for fundamental rights, the processing is possible only for reasons of significant public interest, on the basis of EU law, in compliance with principle of proportionality, of data protection and with the provision of protective measures: in the case of judicial police activities, without prejudice to the other guarantees, the processing must be strictly necessary.

[34] Coordinated Plan on Artificial Intelligence, COM (2018) 795 final, eur-lex.europa.eu/legal-content/EN/TXT/HT-ML/?uri=CELEX:52018DC0795&from=IT.

[35] www.mise.gov.it/images/stories/documenti/Proposte-per-una-strategia-italiana-2019.pdf.

[36] In order to ensure: *1. a harmonious, effective and evolutionary implementation of the national strategy on AI; 2. monitoring and evaluation of its economic and social impact; 3. a constant discussion on the subject with the scientific and business community.* The Strategy also aims, although it is not clear with which resources, to *increase investments, public and private, in AI and related technologies, strengthen the ecosystem of research and innovation in the field of AI* (research centers), *strengthen the educational offer at every level to bring AI to the service of the workforce, exploit the potential of the data economy, real fuel for AI, support the adoption of AI-based digital technologies* (network), *consolidate the regulatory and ethical framework that regulates the development of AI* (cybersecurity, co-regulation, forms of certification), *promote awareness and trust in AI among citizens.*

[37] The strategy addresses in priority areas (industry and manufacturing, agri-food (monitoring crops and soil conditions, forecasting harvests in relation to environmental and meteorological conditions), tourism and culture (profiling and personalizing services), infrastructures and energy networks, smart cities and mobility (parking, traffic, self-driving vehicles, lighting, public transport, monitoring of buildings, bridges), health and social security (assisted surgery, nursing, medical imaging, health records), where the role of the public regulator and administrator is evident ; but the "public administration" itself is also explicitly indicated among the sectors (document management, virtual assistants and chat bots, predictive analysis of business risks and support for the examination of incentive applications, fight against tax evasion and crimes on the web) and the document underlines how artificial intelligence allows to simplify bureaucratic processes. For these purposes, the more interesting actions to be undertaken are homogenizing, centralizing (the aim is to define standard formats for PA data, both those produced by it and those collected from other sources, and for the training of AI systems; possibility of changing the legal status of PA data so that they can belong to the State rather than to the single administration that has got them), sharing data (Promotion of Data Sharing Agreements, standard contracts in which the parties undertake to manage the supply and management of data according to shared rules; introduction of data sharing obligations in strategic sectors with a high potential of collective interest; pilot projects that exploit Data Trusts oriented towards social and environmental sustainability, in which the organizations in possession of the data temporarily entrust its management to certified third parties for the pursuit of a public interest) and relaunching the public administration to make public policies more efficient. With particular reference to efficiency, the protection of legality is a subject of great attention, for the fight both against tax evasion, computer crimes (on the web, cyberattacks generated by AI, theft of personal information), and against "traditional" forms of crime, in particular mafias and terrorism (it also deals with pre-commercial procurement and innovative challenges). In this regard, it is indicated the purpose to introduce artificial intelligence in public administration at all levels and to

The Use of AI by Public Administration

5. *Judicial indications*

More instructions can be found also in judicial reasoning. Recent Italian cases (Consiglio di Stato, sec. VI, n. 2270 of 8th April 2019 and 13rd December 2019, n. 8472)[38] that concern the assignment of the place of employment to teachers[39] - although can be traced back to the classic issue of the computerized administrative provision, on which the doctrine questioned itself some decades ago - enunciated principles that can be useful.

The Consiglio di Stato in these rulings correctly claims that the public administration

must be able to exploit the potential of the so-called digital revolution and underlines the promised benefits in terms of efficiency and neutrality (7.1[40]), also in public services (8.1). It rightly points out that the claim of neutrality is questionable in the design of algorithms: the adoption of predictive models and criteria, the choice of data, the interpretation and formulation of judgments are the result of precise choices and values.

With respect to this process, the Consiglio di Stato admits that it is often difficult to obtain the necessary transparency (7.2).

The problem, however, is that the Consiglio di Stato has before it a fairly simple case (for which perhaps there was no need for this so correct and deep examination) of *software,* in which data were entered and produced a decision.

As the Consiglio di Stato states, the usefulness of standardized procedures is evident, in the case of large quantities of requests, of certain and objective data, when the absence of any discretionary appreciation data occurs; it responds to principles of efficiency and economy, of good performance, and produces benefits in terms of speed, diligence and impartiality.

There is also a brief reference to the discretionary activity (in particular technical), which could benefit from the efficiency: a previous ruling had led the technical rule back to an administrative rule, and stated that discretionary power has space only in the software programming phase[41].

For the Consiglio di Stato, the coordinates given for the legitimacy of the administrative action are the full knowledge of the criteria applied and the imputability of the decision to the legal body that holds the power, which must be able to verify the logic and legitimacy of the choice and of the results. For these reasons, full knowledge implies transparency, which also applies to the rules expressed in a language other than the legal one and which must make the authors, the procedure, the decision, priorities assigned and relevant data understandable

---

equip law enforcement, investigative and inspection authorities with new technological tools that *make it easier to detect anomalies, allowing for example, to focus investigators' attention on high-risk targets for tax evasion or avoidance and to identify cybercriminals more quickly.* However, there are no particular indications regarding how to develop artificial intelligence. In the absence of a specific strategy, however, it has been published a general Plan *"2025 - a strategy for technological innovation and digitalization of the country"*, a very general framework, that is also about IA: it includes the establishment of an *ethical AI LAB-EL* to develop ethical guidelines for a proper use of artificial Intelligence. Furthermore, point 8 provides that artificial intelligence must be able to "guide public decision-makers towards increasingly conscious choices, efficiently managing a series of administrative procedures, especially if they are repetitive and with low discretion".

[38] A. Mascolo, *Gli algoritmi amministrativi: la sfida della comprensibilità*, in *Giornale di diritto amministrativo*, 2020, 366. There are tribunals who have argued that integral automation is illegitimate, as it is in contrast with transparency, impartiality, obligation to motivate, right of defense; according to other rulings, the IT tools must only assume a serving function, otherwise there would be a violation of the preliminary investigation principle (Avanzini, 86, lists the rulings).

[39] Some teachers criticized their assignment far from their place of residence: the recruitment procedure, which had started with an application to be presented exclusively electronically on a platform set up by the Ministry (MIUR), responsible for data collection and processing, had already been dealt with by T.A.R. Lazio, 3742 and 3769 of 2017: see E. Prosperetti, *Accesso al software e al relativo algoritmo nei procedimenti amministrativi e giudiziali. Un'analisi a partire da due pronunce del Tar Lazio*, in *Diritto dell'informazione e dell'informatica*, 2019, 979, concerning the relationship between right of access and industrial protection of *software*. Access to the assignment criteria and to the source code of the software was requested, pursuant to art. 22 of Law 241 of 1990. The MIUR initially refused and then described the operation of the algorithm, commissioned to a company, without however providing technical data, as protected as an intellectual work "and in any case not subject to access, as it cannot be qualified as administrative document". The TAR has ordered MIUR to issue copies of the software source codes to the applicants. The software code is considered a document, a representation of activities of public interest and it is not enough to indicate the operating criteria, also because it was in doubt that it did not respect the collective agreement of the school staff.

It has been noted that, more than the transparency of the software, it would have been necessary to display its design criteria, leaving the administration with the task of verifying whether they had been respected and, if anything, allowing an assessment action, but without asking the display of the source code, the result of intellectual property.

[40] The paragraphs cited are contained in ruling no. 4872.

[41] "The technical rule behind each algorithm still remains a general administrative rule, built by man and not by the machine, to be then (only) applied by the latter, even if this happens exclusively. As such (…) it must be subject to the general principles of administrative activity; it cannot leave discretionary application spaces (of which the electronic processor lacks), but must reasonably foresee a definite solution for all possible cases, even the most unlikely (and this makes it partly different from many general administrative rules); administrative discretion, if certainly cannot be delegated to the software, is therefore to be found at the time of processing the digital tool; (...) also by means of constant tests, updates and ways of improving the algorithm (especially in the case of progressive learning and deep learning)".

externally. Industrial secrecy and privacy of companies are not relevant when they put their tools at the service of public administrations.

The Consiglio di Stato also refers to the GDPR: articles 13 and 14 in fact require that the notice of an automated decision-making process is given to the interested party and that in the case of a fully automated process «significant information must be provided on the logic used, as well as the importance and expected consequences of this treatment for the 'interested'»; art. 15 instead grants a right to receive information relating to the existence of any automated decision-making processes, even if it has already started and even if it has produced a decision. Finally, art. 22 affirms the right not to be subjected to automated decisions without human involvement that produce legal effects or similarly affect the individual. The Robotics Charter, approved in February 2017 by the European Parliament, is also mentioned.

The principles indicated by the Consiglio di Stato are the principle of knowledgeability (referred to art. 42 of the European Charter of Fundamental Rights), also including the principle of comprehensibility, which entails the right to receive "significant information on the logic used"; the principle of non-exclusivity of the algorithmic decision pursuant to art. 22 Reg. (*human in the loop*); based on recital no. 71 of Regulation 679/2016 must be guaranteed that there is no algorithmic discrimination, so the factors that lead to inaccuracies in the data must be rectified, the risk of mistakes must be minimized and the security of personal data guaranteed, with the preferable rectification of data in "entrance"[42].

Another relevant case is the judgment of appeal number 30 of 2nd January 2020: the Consiglio di Stato, reforming the ruling of the T.A.R. that had granted access to the algorithm, considered that the administrative court was wrong in not having allowed the corporation (Cineca) to appear as a counter-interested party, to protect its copyright, with respect to the application for access to the software developed for the recruitment of school managers for state educational institutions. A malfunctioning of the software that managed the test was reported: it would not have respected the instructions provided by the Ministry on the website, among other things not allowing the automatic or manual saving of the answers. The Consiglio di Stato decided that exposing the source code would cause serious damage to both the confidentiality requirements of the competition procedure (the exposition of formulas and instructions needed to decrypt and manage questions and answers, as well as all the data entered by the candidates of the competition, would produce a violation of the privacy and of the regularity of the competition), and the right of the owner of the algorithm to protect the software design, object of intellectual and industrial property, ex Articles 24, paragraph 6, lett. d), Law no. 241/90, 5 *bis*, paragraph 2, lett. c) Legislative Decree no. 33/2013 and 53, paragraph 5, lett. d), Legislative Decree no. 50/2016.

The indications thus received from Italian courts refer to the need for transparency, the refusal of fully automated procedures, respect for copyright.

6. *Section II. Transparency*

From the point of view of the design of the algorithms it is now possible to verify the impact of the legal principles mentioned so far, starting from transparency.

The appeal for transparency to satisfy the request of private and public subjects for access to algorithms, databases, logic of decisions is shared: the 2016 report of the White House on IA[43] states that many ethical issues related to AI can be addressed through greater transparency and calls for developing systems that are transparent and inherently able to explain to

---

[42] The administrative judge of first instance had been completely opposed to the use of algorithms: the Lazio Regional Administrative Court (sez. III, 28/05/2019, n.6688) stated that "a real administrative activity was lacking, having entrusted the entire procedure to an impersonal algorithm: the high number of subjects and territorial areas involved in an administrative procedure cannot legitimize its devolution to a completely impersonal computer or mathematical mechanism that lacks the evaluative capacity of individual concrete cases, as the necessary guarantee procedural investigation is lacking, in particular when the effect is the adoption of measures that affect legal positions of private subjects: the articles 2, 6,7,8,9,10 of the law 7.8.1990 n. 241 would not be respected. The principles of participation, transparency and access would be harmed. There is a flaw in the motivation, with the frustration of the procedural guarantees, as it does not allow to understand the logical process followed by the administration". T.A.R. Puglia 806 and 896/2016 is also referred to, to claim that "the Administration must prepare, together with tools for simplifying document flows in the event of mass insolvency procedures, parallel traditional administrative procedures which can be activated in an emergency, in case of incorrect functioning of the computer systems set up for the physiological forwarding of the application".
At the same time, the Lazio Regional Administrative Court, Sect. III *bis*, 8312/2016 and Section III *bis*, 8902/2018, state

that the IT procedures applied to administrative procedures must be in a necessarily serving position with respect to them, to avoid technical problems hindering relations between the private sector and the Public Administration and between Public Administrations. Based on this principle, IT procedures must never replace the role of the person in charge of the procedure.
[43] Preparing for the future of artificial intelligence, 31 obamawhitehouse.archives.gov/sites/default/files/whitehouse_files/microsites/ostp/NSTC/preparing_for_the_future_of_ai.pdf.

users the reasons for their results; the UK House of Lords AI Committee stated that[44], in order to gain confidence in IA, the implementation of any system that could have a substantial impact on an individual's life is acceptable only if it can generate a complete and satisfactory explanation for the decisions it will make; the European Parliament, in its 2016 report on AI states that it should always be possible to provide the *rationale* behind any decision made with the help of AI that can have a substantial impact on the life of a person or more people and translate it into a form understandable to human[45].

It can be very difficult to know what algorithms are used by the administrations, and their degree of interference in the decision-making process, despite the transparency obligations[46]. Besides, it emerges that the need for transparency implies the need of intelligibility.

The regulation[47] 2016/679 / UE53, so-called GDPR, entered into force in May 2018, provides for the right to information, guarantees and limits relating to the processing of data, to prevent the processing of personal data may arise discrimination [...] or any other significant economic or social damage (cons. 75)[48]. These rights are also provided for in cases where the use of algorithms is permitted by law, consent, etc.

The guarantees include: the right to receive from the data controller significant information on the logic underlying the decision, on the importance and the expected consequences for the data subject; the right to obtain human intervention; the right to express one's opinion and, finally, the right to challenge the decision of the computer.

There are many exceptions to the limits and guarantees: measures which make use of artificial intelligence are authorized, if required by law, in the presence of a public interest (with a formulation that allows a broad appeal to this hypothesis), or in compliance with the obligation of secrecy, or by virtue of the consent of the interested party; in these cases the Regulation simply advocates compliance with adequate guarantees.

The legislative decree n. 101 of 2018, of adaptation of the Italian law to the GDPR, confirms that the processing of data is allowed for reasons of public interest even in the absence of legal provisions or regulations: the need for *ad hoc* rules is in any case considered the main solution[49]; the range of relevant interests capable of justifying the treatments appears in any case wide.

Therefore, if the possibility of using algorithms even without consent is indisputable, we must ask ourselves whether it is really possible to satisfy the demand for transparency. Decisions may be adopted taking into account a number such a large number of data and parameters that is virtually impossible to reconstruct *ex post* the logical process, and then the motivation of the prediction. The more powerful the algorithm, the more opaque it becomes (as it computes more variables and processes more data). The phenomenon appears even more problematic in the areas of security and public order, where secrecy and lack of motivation in decisions happen more frequently. Algorithms are inherently dynamic, change quickly, which makes them even more difficult to understand.

We are facing a crisis of transparency and understandability. The paradigm is even reversed: human beings become knowable (reified by algorithms) and machines become opaque.

Yet, contrary to common opinion, it has been noted that decisions remain more transparent in the case of machines than in humans: the code of human learning is not written and we have very little control over the data entered into humans for their learning[50].

---

[44] *AI in the UK: ready, willing and able?*, § 105, in www.publications.parliament.uk/pa/ld201719/ldselect/ldai/100/100.pdf.

[45] *European Artificial Intelligence (AI) leadership, the path for an integrated vision*, 2018, 35, in www.europarl.europa.eu/RegData/etudes/STUD/2018/626074/IPOL_STU(2018)626074_EN.pdf.

[46] Also R. Brauneis, E.P. Goodman, *Algorithmic Transparency for the Smart City,* in *Yale Journal of Law & Technology,* 2018, 111, reports that their access requests did not, however, make it possible to know the predictive algorithms. Some administrations did not respond, others claimed not to be subject to the access regulations, others opposed a previously made confidentiality commitment. The problem stems from the fact that often they simply did not have this data.

[47] See for example S. Wachter and B. Mittelstadt, *A right to reasonable inferences: re-thinking data protection law in the age of big data and AI,* in *Columbia Business Law Review,* 2019. Information level: www.cnbc.com/2019/05/23/gdpr-one-year-on-ceos-politicians-push-for-us-federal-privacy-law.html; https://martechtoday.com/the-united-states-finally-starts-to-talk-about-data-privacy-legislation-219299, www.fortune.com/2018/11/29/federal-data-privacy-law/, zdnet.com/article/gdpr-usa-microsoft-says-us-should-match-the-eus-digital-privacy-law/.

[48] See C. Castets-Renard, *Accountability of Algorithms in the GDPR and Beyond: A European Legal Framework on Automated Decision- Making*, in *Fordham Intellectual Property, Media & Entertainment Law Journal*, 2019, 30, 91. See on the American debate on regulation of data collection and use F. Pasquale, *Data informed duties in AI development*, in *Colllege of Law Review*, 2019, 119, 1917.

[49] Art. 2, which introduces in the legislative decree 30 June 2003, n. 196 the art. 2 *ter,* 2 *sexies,* 2 *quinquiesdecies*; art. 22.

[50] See G. Noto La Diega, *Against the Dehumanisation of Decision-Making*, in *Journal of Intellectual Property, Information Technology and Electronic Commerce Law*, 3, 2018, 9: A. Vestby and J. Vestby, *Machine Learning and the Police: Asking the Right Questions*, in

---

Furthermore, with regard to the lack of understanding of the operation of machines, we must take into account that, although we depend on many technologies (airplanes, drugs, medical interventions), we do not understand their operation. Intelligibility does not mean for everyone and not even for those subjected to the measure, and not even for officers, but certainly for experts.

In the case of machine learning, who receives the recommendation does not know even how the program arrived at that conclusion, as the algorithms recognize the recurrences by learning themselves how to weight the different components of the entered data. The creators of the algorithms know the data in general terms, they may not know the weight that has been attributed to particular inputs. The principle of the explainability of algorithms implies a list of the factors taken into account and an explanation of the weight given to each factor[51], but it does not solve the problem.

We must probably build a different notion of transparency: it is possible to predict the behaviour of the machine in the event that it operates on the basis of pre-programmed logic rules. In the case of neural networks, the algorithm adapts and develops new solutions to emerging and dynamic situations: a different approach to the predictability of the behaviour of the machine would consist in accepting the impossibility of monitoring every step of the machine, but asking that whatever action takes place, the machine operates within the limits of the regulatory framework. It would be a compromise between predictability (comprehensibility, transparency) and reliability[52].

A lack of transparency and participation in the framework that is given back to us by European legislation results. In the first sense, it is not established that algorithms should be published; in the second sense, both in a defensive logic they are not subject to review by external and independent researchers, and in a collaborative logic there is no *notice and comment procedure* that subjects them to public scrutiny.

With reference to overcoming opacity, the application of the institution of whistle-blowing was hypothesized: the whistle blower, to prevent forms of prejudice and discrimination, could do light on the algorithm, its impact, the data on

which it was built[53]: the institute has given good results in cases of intellectual property, or information asymmetry, in cases in which public bodies rely on private companies.

Three types of transparency are conceivable: transparency on the underlying data[54], on the method, on the outcome. The interest must be directed on the first and the second ones: the first transparency is on training, testing and operational data; the second on how the results come from data. The rule behind results more important than the motivation of the measure.

7. *Participation*

Coming to participation, the design of the algorithms concerns above all the balance between public interests, to understand if, when, how and at what price they can be framed in predefined schemes.

Programming a computer requires, even when sensitive interests are involved, an abstract evaluation and therefore a balance of the interests involved *ex ante.*

This poses a crucial problem of reliability and correctness of the algorithms: if they are not well calibrated and controlled, prejudices can be produced to the legal situations of the recipients of the measures, and the processors will be able to replicate any mistake for a potentially infinite number of times, until it is (and always assumed that is) corrected[55].

It therefore becomes necessary first of all to decide when to use it, with what methods and to evaluate what risks of prejudice the administration is willing to accept[56].

The art. 13 of the Italian law 241 of 1990 (Scope of application of the rules on participation) establishes that the provisions about participation do not apply to public administrative action that is directed towards the issue of measures having a normative, general administrative, planning or programming function. In the Italian case, we are facing a strong cultural delay: a rulemaking procedure, characterized by transparency and participation, would be useful for the adoption of algorithms. Inclusive mechanisms of collective decision-

Policing, 2019, 1.

[51] See A. Deeks, *The judicial demand for explainable artificial intelligence*, in *Colllege of Law Review*, 2019, 119, 1829.

[52] A. Deeks, N. Lubell and D. Murray, *Machine Learning, Artificial Intelligence, and the Use of Force by States*, in *Journal of National Security Law and Policy*, 2019, 10, 21.

[53] S. K. Katyal, *Private Accountability in the Age of Artificial Intelligence,* in *UCLA Law Review*, 2019, 54.

[54] We have to focus on data rather than on the software and programmer, according to S. Yanisky-Ravid and S. K. Hallisey, *Equality and Privacy by Design: A New Model of Artificial Intelligence Data Transparency via Auditing, Certification, and Safe Harbor Regimes*, in *Fordham Urban Law Journal*, 2019, 46, 428.

[55] J. Kleinberg, J. Ludwig, S. Mullainathany and C. R. Sunstein, *Discrimination in the age of algorithms*, 33.

[56] A. E. Waldman, *Power, Process, and Automated Decision-Making*, in *Fordham Law Review*, 2019, 88, 613, proposes that regulators, assisted by independent academic experts, audit algorithmic decision-making code for its adherence to social values.

The Use of AI by Public Administration

making, involving stakeholders and civil society, can improve the correctness and validity of algorithms[57].

The ability to dialogue with the parties, generally private corporations, which manage computer systems and, for example, offer predictive services, must be developed, so that these services are built and adapted according to logic that respects rules, principles and objectives[58].

Many predictive products are in fact developed by private companies. Administrations, especially local ones, do not have the skills and resources to generate and manage predictive models for the allocation of relevant public services such as public order, safety, food safety, transport, so they need to negotiate with businesses, universities, ONG.

The risk, in these cases, is that public entities, not only in our legal system, are unable to regulate and supervise the development and implementation of these tools: some studies have found that, in cases of outsourcing, the registers relating to algorithmic procedures were missing, the administration did not ask private operators to implement transparency measures, the contractors opposed the industrial secrecy exception[59], so that some administrations have even secretly used predictive software.

So, if in a Texas case of teachers fired as a result of applying an algorithm[60], the remedy suggested by the court in the case of a conflict between industrial property and transparency is that the school district should find some other policy that does not rely on secret algorithms, not disclose industrial secrets, in other cases some companies have published the source code of their software, in response to requests for greater transparency. Some US administrations have tried to disengage from third parties and have adopted tools developed *in house* or in collaboration with universities[61].

In this sense, if not autonomously, at least the

intervention of associations, research centres and universities could be a solution to recover part of the information asymmetry.

8. *Trade secret protection and Liability*

Trade secret protection should not be a limiting factor, because – as Coglianese noted - the government's goal should not be classified as a trade secret and the mathematical form should be dictated by the government; the results of the verification and validation procedures and the data to ensure accuracy in most cases should be fully releasable without the need to disclose any protected source code or other trade secrets; the information can always be examined by a court behind closed doors, thus protecting any trade secrets or confidential business information. Independent peer reviews may be conducted under nondisclosure agreements; administrations could draw up contracts to clarify what information should be treated confidentially and what information can be disclosed[62].

Regarding liability, there is European product safety legislation but there is no equivalent to protect services, on AI-based services. The "*Report on the safety and liability implications of Artificial Intelligence, the Internet of Things and robotics*" (COM 2020 64) states that product safety legislation should place specific preconditions for addressing the risks of incorrect data being built and mechanisms to ensure data quality over time. In addition, transparency requirements would be needed to overcome the opacity of algorithm-based systems. Those who suffer damage due to artificial intelligence must enjoy the same protection of those who suffer damage due to other technologies. A risk-based approach could be useful for some artificial intelligence applications, considering the sector (i.e. health, transport, energy) and intended use, in particular for reasons of safety, consumer rights and fundamental rights.

The autonomy of the machine can produce unintended and unforeseen harmful outcomes; the regulation on product safety requires manufacturers to take into account the use of products for their life cycle and to give information and warnings to users, but it could be imposed a new evaluation of the product capable of self-learning, notification to the authorities and measures to prevent risks.

In this sense, the provisions on safety and liability are complementary: there is strict liability for defective products (just prove damage and causation); in the case of AI, it is difficult to trace human behaviour, so it could be

---

[57] K. J. Strandburg, *Rulemaking and inscrutable automated decision tools*, in *Colllege Law Review*, 2019, 119, 1851, focuses on explanation's role in decision-criteria development, which it analogizes to rulemaking. On art. 13, M. C. Romano, *Ipotesi sottratte all'applicazione delle norme sulla partecipazione*, in *L'azione amministrativa*, A. Romano (ed.), Torino, Giappichelli, 2016, 436.

[58] The topic is addressed, among others, by W.S. Isaac, *Hope, Hype, and Fear: The Promise and Potential Pitfalls of Artificial Intelligence in Criminal Justice,* in *Ohio State Journal of Criminal Law,* 2018, 543.

[59] R. Brauneis and E.P. Goodman, *Algorithmic Transparency of the Smart City,* in *Yale Journal of Law & Technology,* 2018, 103.

[60] Hous. Fed'n of Teachers, Local 2415 v. Hous. Indep. Sch. Dist., 251 F. Supp. 3d, 1168 (S.D. Tex. 2017). See C. Coglianese, and S. Valentine.

[61] R. Brauneis, and E.P. Goodman, *Algorithmic Transparency for the Smart City*, in *Yale Journal of Law & Technology*, 2018.

[62] C. Coglianese and D. Lehr, *Transparency and Algorithmic Governance*, in *Administrative Law Review*, 2019, 71, 1.

eased the burden of proof, for example in the event of a violation of cybersecurity obligations, or it could be provided strict liability, also because the separation between products and services is not clear-cut, even for software.

With a view to responsibility, the GDPR imposes, echoing the environmental impact assessment, an impact assessment on data protection, to be applied to automated and profiling treatments, large-scale data processing, data obtained from the systematic surveillance of areas accessible on a large scale.

9. *Conclusions*

The analysis carried out highlighted the need for greater transparency and participation in the design of the algorithms. The moment of the *ex post* analysis is not the subject of this contribution, but many of the remarks reported clearly how an extensive use of algorithms will not allow an easy verification of the correctness of the work. Rather, the judges will have to examine the basic choices, the instructions given, the training. That is why we need a well-trained and prescient civil servants.

The Use of AI by Public Administration