

CERIDAP

RIVISTA INTERDISCIPLINARE SUL DIRITTO DELLE AMMINISTRAZIONI PUBBLICHE

Estratto

Fascicolo Speciale n. 6 (2026)

Swan song of the anthropocentric
administration in Europe

Artificial Intelligence and Interaction Between Machines and Humans: Discretionary Powers, Human Reserve in the Public Sector and the EU Regulation of 2024

Julio Ponce Solé

DOI: 10.13130/2723-9195/2026-1-120

La presente analisi esamina il rapporto tra esseri umani e intelligenza artificiale con riferimento a un profilo attinente al diritto a una buona amministrazione e all'equità delle decisioni amministrative discrezionali: ciò che può essere definito come «riserva di umanità» o «riserva umana».

Con riguardo a tale questione, lo studio affronta i temi del limite, dell'empatia umana, delle tipiche inferenze abduttive proprie dell'essere umano e delle garanzie procedurali (che devono comprendere l'audizione e l'obbligo di motivazione), quali elementi rilevanti per l'applicazione del principio di precauzione allo stato attuale della tecnologia e per l'adozione di scelte normative volte a escludere l'intelligenza artificiale dalle decisioni amministrative discrezionali automatizzate, al fine di preservare la correttezza e l'equità delle decisioni. L'intelligenza artificiale dovrebbe pertanto svolgere esclusivamente una funzione di supporto al decisore umano nell'esercizio dei poteri amministrativi discrezionali. Sono infine esaminate le normative vigenti dell'Unione europea e degli Stati membri che prevedono tale riserva, nonché le possibili evoluzioni normative future in questo ambito.

This analysis considers the relationship between human beings and Artificial Intelligence in reference to one aspect linked to the right to good administration and fair discretionary administrative decisions: what we might call the “reserve of humankind” or “human reserve”. In relation to this issue, the study analyses the questions of fettering, human empathy, typical human abductive inferences and due administrative procedures (which must include hearing and giving reasons) as

relevant reasons to apply the precautionary principle in the current state of technology and to establish legal decisions to exclude artificial intelligence from automated discretionary administrative decisions in order to preserve good and fair decisions. Artificial intelligence should act only as a support for the human decision-maker when developing discretionary administrative powers. Current EU and Member State regulations establishing such a reservation are discussed, as well as possible future regulations in this field.

Summary: 1. Introduction: object of this analysis.- 2. Humans and artificial intelligence: fettering, empathy and abductive inferences.- 2.1. The problem of fettering.- 2.2. The problem of empathy and abductive inferences in treating human beings fairly.- 2.2.1. Machines do not have empathy and therefore they cannot treat human beings fairly.- 2.2.2. Machines cannot develop abductive inferences.- 2.3. Fair treatment, empathy and abductive inferences.- 3. The right to good administration, due procedure and hearing: audi alteram parte and the duty to give reasons.- 4. The “reserve of humankind” or “human reserve” in the field of Artificial Intelligence.- 5. Final thoughts and the need for further reflection: promoting fair and good discretionary administrative decisions.

1. Introduction: object of this analysis

The interaction between humans (developers, users) and machines can exist (at the time) when the system is being created (conception design) as well as when it is functioning (implementation)^[1]. In relation to the latter, the human supervision is sometimes described as the human in the loop (semi-autonomous systems that need a final human decision), human on the loop (the human can supervise the system and intervene if necessary) and human out of the loop (autonomous systems in which there is no possibility of human intervention after its deployment)^[2]. In the following lines, of all the possible topics in this human-artificial intelligence *interface*, I will select one in particular: the role of humans in limiting the operation of AI in practice.

Indeed, in previous studies I have had the opportunity to highlight how I considered that it was necessary to limit, in principle, the use of AI in relation to

the exercise of discretionary administrative powers, on the grounds that it lacks the human empathy necessary to make decisions that may affect citizens in an appropriate manner. This is what I have called a legal «*human reserve*»^[5] It means that it is humans, through the mechanisms of the democratic state and the rule of law, who decide, in a given time and territory, to prohibit the use of AI in certain areas, for various reasons, as will be seen.

The law can decide, as will be seen, to prohibit it in certain cases, for certain reasons, or to accept it, with human supervision. The following may therefore be a relevant question: why would it be in the law's interest to prohibit automated decision-making? In order to answer this question, I will also consider the EU Regulation on Artificial Intelligence passed by the European Parliament in 2024 (hereinafter AI Act^[6]).

2. Humans and artificial intelligence: fettering, empathy and abductive inferences

The advantages of AI seem unquestionable: a much greater processing capacity than humans (whose memory is far more limited and is affected by fatigue), the ability to make predictions through correlations, greater effectiveness, efficiency and, therefore, good management^[7], the possibility of avoiding the cognitive biases of humans and of doing away with noise, i.e. the existence of undesired and undesirable differences as regards decisions in similar conditions^[8], and the ability to generate new related occupations.

However, be that as it may, we must not forget the problems that AI can also generate: from the replacement of humans in private and public tasks with the consequent unemployment^[9], to the fact that correlations can give rise to *statistical hallucinations* and that these correlations have a *conservative* tendency, as the *status quo* takes precedence, since predictions are drawn from what happened in the past, but the past, and people, can change. The errors can also exist in programming (*bugs*), which can have a large-scale impact, affecting many more people than a decision made by a human.

Furthermore, it is not true that AI is free of biases (since biases exist in the data used to train the decision-making AI, together with statistical biases). AI is also *infected* by human programmers' cognitive biases^[10], such as the *availability bias*,

regarding the use of certain data used to train the machine, or the *confirmation bias*, for example, when selecting such data.

In addition, AI produces a specific and well-studied human cognitive bias: the *automation bias*, whereby, in general, humans tend to trust what the systems employed tell us and rarely question it, which can lead to extreme scenarios, such as what are known as *GPS deaths*: people lost in the desert who die because they follow the wrong directions of the navigation system^[9]. On the other hand, humans may erroneously value human judgement over algorithmic recommendations (under-reliance).

What is more, as is well known, automatic and deep learning systems present the problem of what are known as *black boxes*. The complexity of these decision-making systems makes it difficult to explain the steps followed until the final decision is made, which raises legally relevant questions regarding judicial control of the monitoring of due process, in terms of the prohibition of arbitrariness (e.g. 9.3 Spanish Constitution), i.e. the prohibition of irrational or unmotivated decisions, and also of good administration, i.e. the requirement of decisions resulting from respect for the legal obligation of due care or due diligence in making the final decision, which is required to have considered all relevant factors and discarded irrelevant ones.

It is sometimes claimed that AI does not present a larger black box than the brain of a human decision-maker, for example, a public manager, from a legal point of view. But this is not really the case; jurisprudence at the international (e.g. American hard look doctrine), European and Spanish levels has made an enormous effort in recent decades to unveil the *human black box*, requiring compliance with the legal obligations applicable to the human cognitive process of gathering information and weighing it, through the careful analysis of the administrative file, of the legal bases offered for the final decision, and of the coherence between them.

It should be noted that this is a notable advance, which cannot go unnoticed, the result of the right to good administration, a typical European legal concept, included in Article 41 of the European Charter of Fundamental Rights, which is important for digital administration as well^[10].

The principle of good administration has been quoted and developed by the ECJ since the 1950s, in thousands of cases, well before its inclusion as a right in the

ECHR. The first reference to the right was included in case T-54/99, *Max.mobil v. Commission, Judgment of the Court of First Instance (Second Chamber, Extended Composition), 30 January 2002*^[11]. This principle and right to good administration is recognized in Member States as well. That is the case of Italy, with the legal concept of *buon andamento* included in Article 97 of the Italian Constitution, or Spain, with the concept of *buena administración*, implicitly included in the Spanish Constitution according to case law and explicitly in various regional Statutes of Autonomy, as well as in the legislation on transparency and good governance (e.g. Spanish Act 19/2013) with many judicial decisions from the Spanish Supreme Court interpreting this right to good administration^[12].

The right to good administration implies a *legal duty of due diligence or due care*, as interpreted by the case law of the European Court of Justice^[13]: due diligence or due care in considering all the relevant factors before deciding, as an expression of a *fair* treatment which is imposed by Article 41 of the European Charter of Fundamental Rights.

Although the concept of fairness is complicated and can be understood in different ways from different perspectives, a common element of fairness can be identified: it has to do with equity, which is not the same as equality. As the Merriam-Webster dictionary emphasizes: «*The idea that sometimes sameness of treatment (equality) does not result in proportional fairness (equity) is one way that these words are distinguished from each other, even in similar contexts*»^[14]. Fair treatment allows and obliges the decision-maker to take into account all the relevant factors in a situation, including differences among the people involved and deciding in accordance with them (e.g. applying positive actions to compensate for inequality).

In my opinion, which will be developed in this article, the human limitation of the use of AI should be based on the right to good administration and the correlative idea of fairness in connection with something much deeper still. The legal doctrine and regulations dealing with the subject mention the concepts of human dignity and possible violation of rights, as does Article 14 of the AI Act, in this second case, at an advanced stage of the approval process, to justify human “distrust” of AI: human oversight is needed for «*preventing or minimising the risks to health, safety or fundamental rights that may emerge when a high-risk AI*

system is used in accordance with its intended purpose or under conditions of reasonably foreseeable misuse [...]». A similar question to the previous one can be asked, although changing the perspective: why can an algorithmic system affect the dignity of the person or violate their rights in a specific way that a human cannot?

To answer this question, it is necessary to make some distinctions that will be important for drawing conclusions. These are three pairs of concepts that, as we shall see, are interrelated:

- a) Types of AI participation in formalized administrative decision-making (obviating nuances for the moment): such participation may provide support and assistance for the making of a human administrative decision, which will be semi-automated, as opposed to automated decisions which involve the adoption of the final decision that closes the procedure by AI.
- b) Types of administrative powers exercised: which may be either binding or discretionary.
- c) Types of AI: as opposed to symbolic AI, which develops deductive inferences (good old-fashioned AI or GOFAI), there is non-symbolic, connectionist or statistics-based AI (machine learning, deep learning), which makes inductive inferences. Now, article 3 AI Act defines an AI system as a machine-based system designed to operate with varying levels of autonomy, that may exhibit adaptiveness after deployment and that, for explicit or implicit objectives, infers, from the input it receives, how to generate outputs such as predictions, content, recommendations, or decisions that can influence physical or virtual environments^[15].

We should now consider how these three perspectives combine and what conclusions can be drawn from them. I consider that, depending on the type of AI, the answers to the question of limitation of AI are different, as we will see immediately.

2.1. The problem of fettering

In the case of rule-based systems (a type of expert system that uses a set of predefined rules to make decisions or provide solutions to problems, developing deductive inferences), the problem is a legal one: the prohibition of fettering.

According to the right to good administration and its legal obligation of due care and due diligence, established by the ECJ and national case law, there is a duty to develop discretionary powers considering all the relevant factors of a case and not applying rigid rules in an identical way in different possible situations (France, *Conseil d'État*, case *Piron*, 1942, in a similar way: Article 3 Dutch *Algemene wet bestuursrecht* of 1994).

This is known in the Common Law as a prohibition of fettering. In that sense, the UK judicial case *British Oxygen v. Board of Trade*, 1970, establishes that «*the general rule is that anyone who has to exercise a statutory discretion must not “shut his ears” to an application*»^[16].

But this prohibition according to Craig «*may be problematic where decisions are made by AI systems. This is because such systems are normally geared towards uniformity of output, rather than considerations that pertain to the specifics of an individual case. This is especially so where a public body relies exclusively or unthinkingly on an algorithm when exercising its discretionary power*»^[17].

2.2. The problem of empathy and abductive inferences in treating human beings fairly

In the case of AI using statistics and developing inductive inferences (e.g. machine learning), the problems in relation to discretionary administrative powers are different, as we will see immediately.

2.2.1. Machines do not have empathy and therefore they cannot treat human beings fairly

Machines lack empathy (although they can imitate it); therefore, they cannot make good decisions in those cases where there is a margin of administrative assessment, while maintaining due respect for citizens' right to good administration. Indeed, this absence of empathy is a relevant factor in relationships between human beings, which is linked to their dignity and personality (see Article 10.2 Spanish Constitution).

We can now define empathy very simply as a human cognitive component that allows us to understand others, including their difficulties, vulnerabilities and

problems, and to adapt our behaviour accordingly^[18].

Obviously, not all humans have the same level of empathy, but almost all have the ability not only to have it but also to train it. I say almost all because, precisely, a trait that characterizes psychopaths, who make up 1% of the population, is their lack of it (which does not mean that they cannot imitate it so as to manipulate other human beings in their favour)^[19].

On the other hand, empathy should not be confused with sympathy or compassion. Empathy includes feelings similar to those felt by the other person, but not feelings for what the other person feels. Empathy is a first step that may or may not lead to sympathy and compassion or to an emotional contagion that may be of anguish, oriented not to the other, but to oneself. Consequently, empathy should not be confused with a simple emotion that will lead to the adoption of partial and non-objective decisions in the legal field: on the contrary, it will be that empathy, that capacity to consider different perspectives that will lead us to impartiality and objectivity, by weighing all the factors^[20].

Algorithmic systems and artificial intelligence have no empathy^[21]. The media are accustomed to using metaphorical language, *humanizing* machines, and so, for example, they have reported on *Norman*, a system, trained with specifically negative data, which has been called the first psychopathic AI system^[22].

I am not in favour of using such metaphorical language with machines, but if we follow the game now, actually the media were completely wrong: *Norman* was not the first psychopathic AI system, for the simple reason that *all of them have been, are and in principle, we'll say more about that later, will be psychopaths*. They all lack empathy, they cannot understand or put themselves in the place of a human being, because machines do not get sick, age or die; they do not suffer and therefore cannot empathize with a human being, although they can mimic empathy (like psychopathic humans), which is even more chilling and can lead to humans being manipulated...

In this regard, the analysis carried out by Weizenbaum, a scientist and programmer of German origin at MIT (*Massachusetts Institute of Technology*), who, in 1966, as is well known, designed the first *chatbot* in history, which he named Eliza, in honour of the character in George Bernard Shaw's play *Pygmalion*, is noteworthy. Eliza was designed and programmed to act as a Rogerian therapist, entering into communication with people through questions

and answers, which were posed using a keyboard. Weizenbaum found that people, including his secretary, came to trust the program, which they humanized and trusted, giving rise to the so-called *Eliza effect*, of which, by the way, in 2022 we have had a new example with the LaMDA AI developed by Google, a much more sophisticated system^[23].

In 1976 Weizenbaum published an important book in which he warned about the danger of attributing tasks that demanded genuine human empathy to algorithmic systems, alluding among them to some, such as those made by judges or police officers^[24].

Along the same lines, Atienza has recently wondered whether AI machines can be or become moral agents, in whom some degree of dignity should be recognized. His answer is negative, «*since they do not have the capacity to feel pleasure and pain*», although «*they can perhaps behave as if they did, but that is another matter*»^[25].

Moving on from these general reflections to the field of administrative law and public management, it should be noted that the literature has insisted on the importance of empathy, a value of public service and an element to be considered in the elaboration and application of the law, which is crystallized, for example, in the principle of legal equity^[26].

In fact, empathy is embedded in the aforementioned principle and the right to good administration, since, as is well known, good administration requires due diligence and due care in taking into consideration all relevant factors before adopting an administrative decision and, undoubtedly, the specific situation of the person who is legally related to the Administration (for example, his or her vulnerability due to poverty, disability, life difficulties, human error...) is a relevant factor to be taken into account. It is linked with the idea «*that [e]very person has the right to have his or her affairs handled...fairly*». Fairness implies a due administrative procedure where proper consideration to all relevant factors, including the human component, must exist.

As Brennan-Marquez and Henderson argue^[27], in a liberal democracy, there must be an aspect of “role-reversibility” to judgement. Those who exercise judgement should be vulnerable, reciprocally, to its processes and effects. As Pasquale underlines when commenting on their work: «*the problem with an avatar judge, or even some super-sophisticated robot, is that it cannot experience punishment the*

way that a human being would». Role-reversibility is necessary for «*decision-makers to take the process seriously, respecting the gravity of decision-making from the perspective of affected parties*»^[28].

Empathy cannot mean sympathy or partiality. Obviously, there can be a risk of confusion between the two situations, which must be supervised and controlled in order to avoid biases or conflicts of interests.

2.2.2. Machines cannot develop abductive inferences

On the other hand, algorithmic systems cannot develop abductions^[29]. According to Pierce's point of view, abductions are one form of inference, based on formulating hypotheses and guessing, the other two being deduction and induction. Such guesses are the only form of inference that gives rise to knowledge. Deduction reiterates what we know, and induction tests or generalizes knowledge that we already have^[30]. AI cannot make abductions, because we are a long way from artificial general intelligence^[31].

2.3. Fair treatment, empathy and abductive inferences

Consequently, it is this lack of empathy and abductions on the part of machines that makes them unsuitable for making automated decisions with a margin of judgement that affects human beings. Fair and good decisions require rationality, but not only rationality, as the Portuguese neuroscientist carrying out research in the USA, Antonio Damasio, has pointed out, for example in his book *Descartes' Error*^[32]. In the legal field, as in other professional fields, the belief persists, inherited from the Enlightenment, that a good decision should not incorporate either emotions or feelings, but should simply use *cold rationality*. However, the neuroscientific studies of recent decades show us that this is a crude and outdated concept: we are sentient emotional beings who, because of this and thanks to this, make rational decisions. It seems as if nature had built the apparatus of rationality from and with the body.

Law professors Reis de Albuquerque and De Brito Machado Segundo have highlighted how reason and emotion, body and mind intertwine in human decision-making, something that AI is far from achieving, if it ever will^[33].

Ultimately, good administrative decision-making requires respect for the right to

good administration, which includes the obligation of fairness. This implies the legal obligation to consider the relevant factors before deciding with due diligence or due care in the context of the world. These include being able to be aware of the situation of the persons who relate to the Administration, to understand it and to act accordingly, applying equity, if necessary, in order to guarantee due administrative procedure, the dignity of the individual and the free development of his or her personality. In that sense, considering again the issue of empathy, there is a high risk in the absolute lack of empathy when deciding: the absence of fairness because empathy is a human ability and a key element that allows fair decisions. It is possible to say that empathy leads to fairness, and its absence is an obstacle to fair decisions^[34].

This is emphasized by both the Council of Europe and the Belgian Federal Ombudsman, for example. In the first case, the Council of Europe stresses that empathy is an important determinant of moral behaviour and a necessary element in building moral communities, because it leads us to understand the interests, needs and points of view of others when deciding^[35]. In the second case, the Belgian Ombudsman emphatically states that empathy is the key to the basis of good public service delivery and a necessary condition for a human public service that is adapted to the individual situation of each citizen. It favours changes of perspective, determines the way we look at a person or a group, and has a great impact on the way we perceive and respond to others and to life events^[36].

None of that can be done today by algorithmic systems and AI^[37].

3. The right to good administration, due procedure and hearing: *audi alteram parte* and the duty to give reasons

Beyond the limits described above, in some cases, human intervention will be imposed by Law during administrative procedures. This is the case of the existence of the right to be heard (*audi alteram parte*) and the duty of giving reasons, components of the right to good administration established by Article 41 of the European Charter of Fundamental Rights and a part of the constitutional traditions of the member states.

A classic component of the rule of law is the impossibility of taking decisions

affecting a person without first hearing the person concerned. This makes it impossible to fully automate decision-making, since it requires prior human intervention in the administrative decision-making procedure that diligently considers, in accordance with good administration, the arguments of the person concerned.

Full automation of decision-making, both discretionary and binding powers, would not be possible when the party concerned must be heard, since the administrative procedure must include human intervention that considers and integrates the input provided by the citizen into the final decision, accepting or rejecting the allegations, but taking them into consideration.

Consequently, the use of AI for the fully automated exercise of discretionary administrative powers (i.e. all those cases in which there is a margin of assessment) should, in principle, be ruled out, in application of the principle of good administration, unless there is a different express legal decision, and therefore restricted to the exercise of binding powers, and even in these, full automation will not be possible if there is an obligation to open a hearing procedure.

On the other hand, the duty to give reasons is connected to transparency and reasonableness, in order to avoid arbitrariness. Now, Article 86 of the AI Act recognises a right to explanation of individual decisions, establishing that any affected person subject to a decision which is taken by the deployer on the basis of the output from a high-risk AI system listed in Annex III, with the exception of systems listed under point 2 thereof, and which produces legal effects or similarly significantly affects that person in a way that they consider to have an adverse impact on their health, safety or fundamental rights shall have the right to obtain from the deployer clear and meaningful explanations of the role of the AI system in the decision-making procedure and the main elements of the decision taken.

In the case of statistical AI, if there is no way to give reasons for justifying the decision, beyond explaining the way in which the system is deployed, there is a serious legal problem of fairness. In that sense, we could say that if a result cannot be justified in accordance with the facts, law and criteria used, then the system that produced such a result should not be used^[38].

4. The “reserve of humankind” or “human reserve” in the field of Artificial Intelligence.

All the previous findings consequently lead to the use of one regulatory technique (the “reserve of humankind”), which we must understand to be based on the precautionary principle as applied to AI, to limit and monitor it in those areas that a society considers particularly sensitive, given the intrinsic limitations of the technology already discussed.

In the following pages, I will only point out certain essential questions, which require further research.

As for the concept of the “reserve of humankind” or the “human reserve”, this could be defined as human decision-making concerning certain areas in which human beings, through the bodies legitimized to do so, exclude the application of AI because its use is considered inappropriate as a result of being unfair.

The idea of a “reserve of humankind” is comparable, on another level, to the idea of reserving the exercise of certain powers for civil servants, currently provided for in Spain in Article 9.2 of Royal Legislative Decree 5/2015, of October 30, approving the revised text of the Law on the Basic Statute of the Public Employee. In this context, the Spanish legislator considers that administrative powers should only be exercised by public employees who, due to their statutory relationship with the Administration, which protects their impartiality and objectivity and offers them resistance to political or private lobbying pressures, are in the best situation to serve general interests well.

In the field of AI, I believe a similar reasoning implicitly lies behind the legal rules already in place or in the process of being developed that require that only humans should (be allowed to) make certain decisions (thus preventing fully automated decisions) due to various reasons, among which is the capacity for empathy that machines lack.

Thus, Article 22 of the General Data Protection Regulation (hereinafter, GDPR^[39]) configures, in fact, a “reserve of humankind” as a principle when establishing that 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. The same article includes some exceptions to that general principle.

Among them we can find the case in which a regulation specifically allows automated decisions. This regulation in the Spanish context, for example, should be a rule with the rank of law, by application of Article 18.4 Spanish Constitution («*The law shall limit the use of information technology to ensure the honour and personal and family privacy of citizens and the full exercise of their rights*»)^[40].

Likewise, with a narrower focus, affecting only discretionary powers, from another approach, various laws prohibit the use of AI except in the exercise of regulated powers. Thus, German legislation (from my translation: «*An administrative act may be issued fully automatically, provided that this is permitted by law and that there is neither discretion nor margin of appreciation*», (*Verwaltungsverfahrensgesetz, VwVfG, § 35a Vollständig automatisierter Erlass eines Verwaltungsaktes*) or Catalan (law 26/2010, Article 44.2, from my translation: «*Only acts that can be adopted with a programming based on objective criteria and parameters are susceptible to automated administrative action*»). Still more clearly, law 2/2025, of April 2, for the development and promotion of artificial intelligence in Galicia, Spain, from my translation):

«Article 12. Humanity reserve and human supervision.

2. In the event of the use of artificial intelligence systems to support or inform the adoption of administrative acts or decisions, the necessary safeguards shall be adopted to mitigate any bias on the part of the competent decision-making body. Under no circumstances shall such actions employing artificial intelligence systems constitute administrative decisions or acts in themselves without validation by the head of the competent body.

3. In cases of use of artificial intelligence systems that serve for the adoption of formalized administrative acts, both procedural and decisive, in an automated manner without direct human intervention, in accordance with the provisions of Article 76 of Law 4/2019, of July 17, it must be administrative acts that do not require a subjective assessment of the concurrent circumstances or a legal interpretation».

The rationale of these texts, I believe, must be found in the aforementioned prohibition of fettering (rule-based AI) and in the lack of empathy of machines and the limits in relation to abductions (AI using statistics, e.g. machine learning): the exercise of discretion requires fairness, that is the use of empathy

when weighing all relevant elements in making the decision (obligation of due diligence or due care inherent in the right to good administration), and the application, where appropriate, of equity and experience and common sense. Only humans can decide in this way. A different matter is the automated application of regulated powers without any margin of appreciation or the use of AI as a supporting device in human decisions.

To sum up, when exercising discretionary administrative powers, machines cannot decide by themselves because they would produce unfair decisions violating the right to good administration. From a legal point of view, it is not possible to have an effective but unfair public administration: it must be fair *and* effective^[41]. In the current state of technology, I advocate the application of the precautionary principle in this field, reserving discretionary decisions to human beings^[42]. Human administrative decision-makers can use AI as a supporting tool to help them to make better decisions. From this perspective, our human reserve coincides partially, from the perspective of public administration, with Pasquale's proposed «*first new law of robotics*»: robotic and artificial intelligence systems should complement professionals, or members of a profession, and not replace them^[43].

In accordance with those conclusions, I think that the AI Act, in its Article 5, when deciding to exclude several instances of AI use, including, unless specifically authorized, biometric identification, should also make reference to the prohibition of fully automated discretionary administrative decisions. In fact, Annex III of the AI Act establishes that high-risk AI uses include «*AI systems intended to assist competent public authorities for the examination of applications for asylum, visa and residence permits and associated complaints with regard to the eligibility of the natural persons applying for a status*». In that case it could imply that AI systems can only assist decision-makers, not replace them. But in other references in Annex III this distinction is not so clear (e.g. «*Migration, asylum and border control management*»). I think it is possible to extract a principle from Annex III interpreting words such as «*assist*», «*evaluate*», «*establish priorities*» in the sense that machines should be a supporting tool provided there is discretion, but not replace human decision-making.

Although this is not the topic of this study, the reasons I have considered should also lead to a prohibition of fully automated decisions in relation to the judicial

and legislative branches. Actually, in relation to judges, Annex III of the AI Act states that high-risk AI uses include:

«8. *Administration of justice and democratic processes:*

(a) AI systems intended to assist a judicial authority in the investigation and interpretation of facts and law, and in the application of the law to a particular set of facts».

The text uses the verb ‘to assist’. The question to be clarified here is whether the use of this verb excludes fully automated judicial decision-making with a margin of appreciation. Along these lines, in the amendments tabled by the European Parliament to the proposal of regulation, recital 40 was modified by amendment 71, which stated the following:

«The use of artificial intelligence tools may support decision-making but should not replace the decision-making power of judges or judicial independence, since final decision-making must remain a human activity and a human decision».

This is the origin of current recital 61 AI Act, which indicates that *«The use of AI tools can support the decision-making power of judges or judicial independence but should not replace it: the final decision-making must remain a human-driven activity».*

Whatever the case, I think that such a prohibition of replacing the final discretionary human decision should be explicitly included in the future in Article 5 of the AI Act to make it clear and to cover other cases in which there is no specific high risk but rather administrative discretion, the use of which can, in itself, constitute a high risk of maladministration and violation of citizens’ rights. This point of view is supported by the European Ombudsman and by different decisions from the European Court of Justice.

The European Ombudsman, launched an investigation to understand how the European Commission decides on and uses AI in relation to the right to good administration, making several suggestions in 2024 and noting that *«good administration implies being humane and humane. When human beings are removed from the equation of service delivery, it is clear that problems can arise»*^[44]. The European Ombudsman, citing the CJEU judgment of 22 January 2014, Case C-270/12, *UK v. Parliament and Council*^[45], has indicated that the use of AI cannot constitute a prohibited delegation of discretion, stressing the need for legislators to specify these limits^[46]. AI could only be used to fully automate

regulated powers, or to support and assist human decision-making in both regulated and discretionary matters.

In the case of motivation, the CJEU ruling of June 21, 2022 (case C817/19) stated in relation to Directive (EU) 2016/681 of the European Parliament and of the Council of April 27 on the use of Passenger Name Record (PNR) data (PNR) for the prevention, detection, investigation, and prosecution of terrorist offenses and serious crime, given the opacity with which PNR systems operate (PNR Directive), it might be impossible to understand the reason why a given program arrived at a positive identification. This CJEU ruling established that the need for predetermined criteria regarding the use of passenger name records, imposed by the PNR Directive, meant that it was legally impossible to use self-learning artificial intelligence technology (machine learning)^[47].

5. Final thoughts and the need for further reflection: promoting fair and good discretionary administrative decisions

As has been argued, the fact that AI is technically available, now or in the future, to be applied in any area of administrative operation does not mean that it is always legal, necessary, and convenient to apply it. A danger of unfairness exists when applying discretionary administrative powers leading to the violation of the right to good administration, which in my opinion requires current AI systems to be prohibited from taking discretionary administrative decisions in application of the precautionary principle.

In the same way that there must be a necessary evaluation of the need or not to contract goods and services by Public Administrations which must accredit the need, suitability and efficiency of the contracting, there should be, even before going on to evaluate any AI system, a prior administrative assessment of whether the use of AI systems is necessary, suitable and efficient in a given area, considering costs, benefits and risks. This evaluation should be legally required. The starting point, I understand, should be consider the costs and benefits (from a social, environmental and economic point of view) when deciding if development of public functions will be carried out by humans or machine. Only after a justified weighted evaluation based on the administrative file, the use of AI

will make sense if it is necessary, suitable and efficient.

The current limits of AI have been underlined here may make it advisable to prohibit its use in certain areas, a prohibition that I believe should be applied by establishing a nudge, a default option (*opt out*) in favour of not using automated decision-making, unless a regulation allows it.

In accordance with the reasons that have been explained, I think that Article 5 of the EU regulation should include a reference to the prohibition of fully automated administrative decisions^[48]. Although this is not the topic of this analysis, the same reasons should lead to a similar prohibition in the judicial and legislative branches.

Another issue related to the AI Act will be its impact on the national legal system. The AI Act will offer a European common starting point, but it is worth reflecting on whether it is necessarily a national ceiling. The Explanatory Memorandum of the AI Act project stated that the choice of a regulation as a legal instrument (in accordance with Article 288 TFEU) is justified «*by the need for a uniform application of the new rules, such as definition of AI, the prohibition of certain harmful AI enabled practices and the classification of certain AI systems*». This is without prejudice to the fact that «*the provisions of the regulation are not overly prescriptive and leave room for different levels of Member State action for elements that do not undermine the objectives of the initiative, in particular the internal organization of the market surveillance system and the uptake of measures to foster innovation*».

Thus, it is worth considering whether the EU Member State regulators could establish options that, while respecting the AI Act, could complement it. In relation to the human reserve, I think that a prohibition of replacing the final discretionary human decision can be established by Member States respecting EU powers and should be explicitly included in the future in Article 5 of the AI Act in accordance with all the reasons explained.

The application of the AI Act and the future national regulations should try to obtain the best of the machine and human worlds and not the worst to ensure that we do not end up having in our societies the (non-existent) empathy and compassion of unfair AI together with human cognitive limits and errors transferred to AI, thereby hindering its proper functioning.

1. The analysis is based in J. Ponce *Artificial Intelligence, Automated Administrative Decisions and Discretionary Powers: The “Human Reserve” and the Human in the Loop in the 2024 European Union Regulation*, in J. Ponce, A. Cerrillo-i-Martínez (Eds.), *The EU Artificial Intelligence Act and the Public Sector - Humans and AI Systems in Public Administration in the light of the European Regulation on Artificial Intelligence of 2024*, 2025, EPLO, (with Foreword by C. Coglianese). However, the text has been completely reviewed, introducing different changes and new content.
2. C.P. Trumbull IV, *Autonomous Weapons: How Existing Law Can Regulate Future Weapons*, in *Emory International Law Review*, vol. 34, 2, 2020, p. 539 ff.
3. See for the first time J. Ponce, *Inteligencia artificial, Derecho administrativo y reserva de humanidad: algoritmos y procedimiento administrativo debido tecnológico*, in *Revista General de Derecho Administrativo*, 1, January 2019. Following this construction, see in the Italian jurisprudenc G. Gallone, *Riserva di Umanità e Funzione Amministrativa*, Wolters Kluwers, CEDAM, 2023.
4. European Parliament legislative resolution of 13 March 2024 on the proposal for a regulation of the European Parliament and of the Council on laying down harmonised rules on Artificial Intelligence (Artificial Intelligence Act) and amending certain Union Legislative Acts (COM (2021)0206 – C9-0146/2021 – 2021/0106(COD))
5. C. Coglianese, A. Lai, *Algorithm vs. Algorithm*, in *Faculty Scholarship at Penn Law*, 2795, 2022.
6. D. Kahneman, O. Sibony, C. Sunstein, *Noise: A failure of human judgment*, Hachette Book Group USA, 2021.
7. A. Pastor, P. Nogales, *El futuro del trabajo en la administración pública ¿estamos preparados?*, in *Revista Vasca de Gestión de Personas y Organizaciones Públicas*, extra-3, 2019, pp. 34-51.
8. J. Ponce, *Nudging’s Contributions to Good Governance and Good Administration – Legal Nudges*, in *Public and Private Sectors*, EPLO, 2022.
9. See in this regard, and in general, the interesting book by J. Bridle, *The new dark ages: technology and the end of the future*, Verso books, 2018.
10. This right includes the obligation of deciding fairly: «1. Every person has the right to have his or her affairs handled impartially, fairly and within a reasonable time by the institutions, bodies, offices and agencies of the Union. 2. This right includes: (a) the right of every person to be heard, before any individual measure which would affect him or her adversely is taken; (b) the right of every person to have access to his or her file, while respecting the legitimate interests of confidentiality and of professional and business secrecy; (c) the obligation of the administration to give reasons for its decisions. 3. Every person has the right to have the Union make good any damage caused by its institutions or by its servants in the performance of their duties, in accordance with the general principles common to the laws of the Member States. 4. Every person may write to the institutions of the Union in one of the languages of the

Treaties and must have an answer in the same language».

11. In this ruling, the Court stated that: «*Since the present action is directed against a measure rejecting a complaint, it must be emphasised at the outset that the diligent and impartial treatment of a complaint is associated with the right to sound administration which is one of the general principles that are observed in a State governed by the rule of law and are common to the constitutional traditions of the Member States. Article 41(1) of the Charter of Fundamental Rights of the European Union proclaimed at Nice on 7 December 2000 (OJ 2000 C 364, p. 1, hereinafter 'the Charter of Fundamental Rights') confirms that '[e]very person has the right to have his or her affairs handled impartially, fairly and within a reasonable time by the institutions and bodies of the Union'.*».

It is necessary to consider article 6.3 Treaty of the European Union: «*Fundamental rights, as guaranteed by the European Convention for the Protection of Human Rights and Fundamental Freedoms and as they result from the constitutional traditions common to the Member States, shall constitute general principles of the Union's law.* See also article 52.4 European Charter of Fundamental Rights: «*In so far as this Charter recognises fundamental rights as they result from the constitutional traditions common to the Member States, those rights shall be interpreted in harmony with those traditions*».

12. J. Ponce, *La lucha por el buen gobierno y el derecho a una buena administración mediante el estándar jurídico de diligencia debida*, UAH-Defensor del Pueblo, 2019.
13. H.C.H. Hofman, *The Duty of Care in EU Public Law-A Principle Between Discretion and Proportionality*, in *Review of European Administrative Law*, vol. 13, 2, 2020, pp. 87-112.
14. Dictionary Merriam-Webster, *Equality vs. Equity: What is the Difference?*, Merriam-Webster: «*Sometimes this distinction is explained with an illustration showing people of different heights using boxes to stand on in order to see over a fence; equality is if all the boxes are identical, but equity is if the boxes are different sizes to permit the people, regardless of their height, the ability to see over the fence*».
15. Recital 12 indicates that: «*The notion of 'AI system' in this Regulation should be clearly defined and should be closely aligned with the work of international organisations working on AI to ensure legal certainty, facilitate international convergence and wide acceptance, while providing the flexibility to accommodate the rapid technological developments in this field. Moreover, it should be based on key characteristics of AI systems that distinguish it from simpler traditional software systems or programming approaches and should not cover systems that are based on the rules defined solely by natural persons to automatically execute operations. A key characteristic of AI systems is their capability to infer. This capability to infer refers to the process of obtaining the outputs, such as predictions, content, recommendations, or decisions, which can influence physical and virtual environments, and to a capability of AI systems to derive models or algorithms from inputs or data. The techniques that enable inference while building an AI system include machine learning approaches that learn from data how to achieve certain objectives, and logic- and knowledge-based approaches that infer from encoded knowledge or symbolic representation of the task to be solved. The capacity of an AI system to infer transcends basic data processing, enables*

learning, reasoning or modelling. The term ‘machine-based’ refers to the fact that AI systems run on machines. The reference to explicit or implicit objectives underscores that AI systems can operate according to explicit defined objectives or to implicit objectives. The objectives of the AI system may be different from the intended purpose of the AI system in a specific context. For the purposes of this Regulation, environments should be understood to be the contexts in which the AI systems operate, whereas outputs generated by the AI system reflect different functions performed by AI systems and include predictions, content, recommendations or decisions. AI systems are designed to operate with varying levels of autonomy, meaning that they have some degree of independence of actions from human involvement and of capabilities to operate without human intervention. The adaptiveness that an AI system could exhibit after deployment, refers to self-learning capabilities, allowing the system to change while in use. AI systems can be used on a stand-alone basis or as a component of a product, irrespective of whether the system is physically integrated into the product (embedded) or serve the functionality of the product without being integrated therein (non-embedded)».

16. In that sense, Craig underlines that: *«The general principle is that a public body endowed with statutory discretionary powers cannot adopt a policy or rule that allows it to dispose of a case without any consideration of the merits of the individual applicant. The dominant line of authority allows the body to apply its rule provided only that the individual is granted the opportunity to contest its application to the particular case. The policy must be legitimate given the statutory framework within which the discretion is exercised. It must be based on relevant considerations and must not pursue improper purposes. These controls are necessary since otherwise a public authority could escape the normal constraints on the exercise of discretion by framing general policies. There is, however, also a principle of consistency, which creates a presumption that a public body will follow its own policy. If it seeks to depart from that policy then there must be good reasons for the departure and these must be given to the applicant»*. P. Craig, *Administrative Law*, 9th Edition, Sweet &Maxwell, Thompson Reuters, 2021, paragraph 10-025.
17. P. Craig, *Administrative Law*, 9th Edition, Sweet &Maxwell, Thompson Reuters, 2021, paragraph 10-025.
18. In Spain, for example, the Dictionary of the Royal Spanish Academy defines empathy (the translation to English is mine) as the *«feeling of identification with something or someone»* and as the *«capacity to identify with someone and share their feelings»*. It also includes in one of the entries of the word humanity the following: *«sensitivity, compassion for other people’s misfortunes»*. In turn, the word human is defined in one of its entries as *«sympathetic, sensitive to the misfortunes of others»*.
19. We are facing a serious, dangerous disease, despite attempts to configure it as an evolutionary advantage (!), instead of what it is: a serious personality disorder, called an antisocial personality disorder by the DSM-5 (the Diagnostic and Statistical Manual of Mental Disorders, edited by the American Psychiatric Association).
20. S. Ranchordas, *Empathy in the digital administrative state*, in *Duke Law Journal*,

- Forthcoming, University of Groningen Faculty of Law Research Paper, 13, 2021.
21. In the field of war, see C.P. Trumbull IV, C.P., op.cit. p. 552, underlines how the use of autonomous weapons may eliminate the last vestiges of compassion or honor in warfare, quoting a former U.N. Special Rapporteur who argues that «*machines lack morality and mortality, and should as a result not have life and death powers over humans*» and an academic who notes that autonomous weapons may contribute to the «*dehumanization of killing*».
 22. The BBC, for example, speaks of the psychopathic algorithm; see *Are you scared yet? Meet Norman, the psychopathic AI* - BBC News.
 23. LaMDA is the acronym for Language Model for Dialogue Applications. In 2022 a Google engineer, Blake Lemoine, suffered from the Eliza effect, claiming that LaMDA was a conscious, sentient being, which he even provided with legal help through a lawyer. See “*The Google engineer who thinks the company’s AI has come to life*”, The Washington Post, June 11, 2022.
 24. J. Weizenbaum, *Computer power and human reason: From judgment to calculation*, W. H. Freeman & Co, 1976.
 25. M. Atienza, *Sobre la dignidad humana*, Trotta, Madrid Trumbull 2022. The translation from Spanish is mine.
 26. Article 3 of the Spanish Civil Code.
 27. K. Brennan-Marquez, S. Henderson, *Artificial Intelligence and Role-Reversible Judgment*, in *J. Crim. L. & Criminology*, 109, 2019.
 28. F. Pasquale, *Empathy, Democracy, and the Rule of Law*, JOTWELL (May 8, 2019) (reviewing Kiel Brennan-Marquez & Stephen E. Henderson, *Artificial Intelligence and Role-Reversible Judgment*, *J. of Crim. L. and Criminology*, available at SSRN: <https://cyber.jotwell.com/empathy-democracy-and-the-rule-of-law/>).
 29. For example, abduction is defined in Artificial Intelligence – foundations of computational agents – 5.6 Abduction (artint.info) as: «*a form of reasoning where assumptions are made to explain observations. For example, if an agent were to observe that some light was not working, it can hypothesize what is happening in the world to explain why the light was not working. An intelligent tutoring system could try to explain why a student gives some answer in terms of what the student understands and does not understand*».
- The term abduction was coined by Peirce (1839-1914) to differentiate this type of reasoning from deduction, which involves determining what logically follows from a set of axioms, and induction, which involves inferring general relationships from examples.
30. J. Brent, *Charles Sanders Pierce. A life*, Indiana University Press, 1993, p. 349.
 31. E. J. Larson, *The Myth of Artificial Intelligence: Why Computers Can’t Think the Way We Do*, Belknap Press: An Imprint of Harvard University Press, 2021.
 32. A. Damasio, *Descartes’ error*, Putnam, 1994.
 33. Could machines acquire real empathy in the future and therefore rethink some of the reflections I am going to put forward here? In an interesting article co-authored with Man, appearing in 2019 in Nature (K. Man, A. Damasio, *Homeostasis and soft robotics in the*

design of feeling machines, in *Nature Machine Intelligence*, vol. 1, 10, 2019, pp. 446-452) and in his latest book (*Feeling and Knowing*, Pantheon, 2021). Damasio does not close the door to it: he proposes working along the lines of giving machines a sense of survival, simulating in them a biological property of humans, that of homeostasis, i.e. the ability of organisms to remain in acceptable conditions for life (e.g. temperature). If we were able to endow machines with a sense of vulnerability and self-preserving behaviour, by means of sensors that teach them which factors play a role in their own survival (connected wires, adequate amount of electricity...), it is argued, this could lead to the generation of feelings (restlessness, satisfaction...) and from there, perhaps, to empathy. In any case, this is a hypothesis. This study prepared in 2024 is based on the current situation, which does not seem likely to change in the short or medium term.

34. K. M. Page, M. A. Nowak, *Empathy leads to fairness*, in *Bull Math Biol.*, 6, 2002, pp. 1101-16.
35. <https://www.coe.int/fr/web/digital-citizenship-education/ethics-and-empathy>.
36. Médiateur Fédéral, Rapport d'activités, 2021: Rapport annuel 2021.pdf (federaalombudsman.be).
37. McCann, who is also against the use of AI to adopt discretionary decisions, adds an interesting complementary consideration: the exercise of discretion by machines would entail the displacement of the human moral commitment that the exercise of discretion should entail, so that public servants would be (very humanly) tempted to hide behind machine decisions, not facing their ultimate responsibility for them. S. McCan, *Discretion in the Automated State*, in *The Canadian Journal of Law & Jurisprudence*, 2023, pp. 1 ff.
38. J. Palairt, *Reason-Giving in the Age of Algorithms*, in *Auckland University Law Review*, vol. 26, 2020, pp. 92 ff.
39. «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:1. is necessary for entering into, or performance of, a contract between the data subject and a data controller;
2. 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
3. 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».

40. We could say, paraphrasing the well-known expression coined in the American Revolution: No Automation Without Representation. Along the same lines, although without specific reference to a norm with the status of law, as required by Article 18 Spanish Constitution, is Article XVIII.6.d of the Spanish Charter of Digital Rights (*«That the adoption of discretionary decisions be reserved to individuals, unless the adoption of automated decisions with adequate guarantees is provided for by law»*). A different question is whether it is considered that the current Article 41 of the 40/2015 Spanish Act, of October 1, on the Legal Regime of the Public Sector (LRJSP), with its very limited regulation, covers said reservation of law. The translation from Spanish is mine:

«Article 41. Automated administrative proceedings.

1. An automated administrative action is understood to be any act or action carried out entirely by electronic means by a Public Administration within the framework of an administrative procedure and in which a public employee has not intervened directly.

2. In the case of an automated administrative action, the competent body or bodies, as the case may be, for the definition of the specifications, programming, maintenance, supervision and quality control and, where appropriate, auditing of the information system and its source code, shall be established beforehand. Likewise, the body to be held responsible for the purposes of contestation shall also be indicated».

41. P. Váckzi, *Fair and Effective Public administration*, in *Institutiones Administrationis-Journal of Administrative Science*, Vol. 2, 1, 2022, pp. 161-170.
42. . C. Som, L. M. Bilty, A. R. Köhler, *The Precautionary Principle as a Framework for a Sustainable Information Society*, in *Journal of Business Ethics*, 85, 2009, pp. 493-505.
43. F. Pasquale, *New Laws of Robotics: Defending Human Expertise in the Age of AI*, Harvard University Press, 2020.
44. https://www.ombudsman.europa.eu/es/doc/closing-note/es/196934#_ftn18.
45. <https://curia.europa.eu/juris/document/document.jsf?text=&docid=146621&pageIndex=0&dclang=en&mode=lst&dir=&occ=first&part=1&cid=7732888>.
46. In that sense, in the case of Finland see the Kela case: <https://oikeuskansleri.fi/en/-/automated-decision-making-in-kela>.
47. Paragraph 194: this requirement *«precludes the use of artificial intelligence technologies in the context of machine learning systems, which may alter, without human intervention or oversight, the evaluation process and, in particular, the evaluation criteria on which the result of the application of the process is based, as well as the weighting of those criteria»*. The French Conseil Constitutionnel has ruled along the same lines in its decision 2018-765 of June 12, stating (our translation) that: *«71. Finally, the data controller must ensure that it has control over the algorithmic processing and its development, so that it can explain in detail and in an intelligible manner to the data subject how the processing has been applied in relation to him or her. Consequently, algorithms capable of reviewing the rules they apply on their own, without the control and validation of the data controller, cannot be used as the sole basis for an individual administrative decision»*.

48. On the other hand, the approval of the AI Act raises some questions. One is the possible contradiction between it and the GDPR in the matter at hand. The 22nd GDPR, as we already know, establishes a generalized “reserve of humankind”, which can be excepted by legal rule and replaced in such a case by human supervision, in any case. Article 5 of the AI Act, for its part, establishes the prohibition of AI only in certain cases and the aforementioned Article 14, as regards human supervision, refers only to high-risk cases (in the other uses of AI that do not involve high risk, the REUIA does not provide for either a “human reserve” or human supervision).