

# Artificial intelligence and law: procedural dilemmas and ethical issues

## Introduction

Far from placing ourselves in the realm of science fiction, when we speak of “artificial intelligence” we refer to the dilemmas and doubts that arise in the present in many aspects of the reality that surrounds us. All of us have seen on television how complex surgical operations are performed without just visible human presence, how are performed so automated calculations on satellite orbits, or, more simply and closely to our daily activities as lawyers, how in a few seconds we can have at our disposal the recent jurisprudence on certain specific matter on which we have to file a lawsuit.

The enormous investment in certain scientific fields continuously produces advances that contribute significantly to modifying the scientific and technological context in a few years until leading to paradigm changes<sup>1</sup> or, more emphatically, to new industrial revolutions.<sup>2</sup> But, in reality, we are not talking about sudden changes without prior foundation. Many decades ago the unfortunate Alan M. TURING, based on the postulates of mathematical logic, gave a lecture in which he raised something as current as the question of whether a machine can think.<sup>3</sup> The novelties are rather in the generalization of the application of digital technologies as a consequence of their continuous transformation and improvement. Its immersion in our daily life, on the one hand, produces an inevitable rapturous effect and, on the other, a certain reaction of fear and mistrust.

If we limit ourselves to focusing our observation on the field of Law, intuitively we can list interesting advantages in the application of digital technologies, which lighten the work, speed up the tasks, increase the accuracy of the results if in Law we are allowed to speak in these terms. But, also a priori doubts arise before the possible “dehumanization” of the work of the jurists, which distances them from the equity assessments in the specific case<sup>4</sup> or that multiplies the effects of highly debatable biases.<sup>5</sup>

<sup>1</sup>KUHN,TS, *The Structure of Scientific Revolutions*, 50th Anniversary edition, Chicago-London, 2012, systematized the idea of progress through scientific revolutions.

<sup>2</sup>The considerations of SCHWAB, K., *The fourth industrial revolution*, Barcelona, 2016, are well known, for which, from the sociological point of view, he speaks of the consequences of digitization: “The most important technological innovations are at hand about to generate a transcendental change in the whole world, something inevitable”.

<sup>3</sup>TURING,AM, “Computing Machinery and Intelligence”, *Mind*, vol. LIX, no. 236, 1950, pp. 433-460; translated into Spanish as ¿Pueden las máquinas pensar?, Oviedo, 2012.

<sup>4</sup>Let us remember how ARISTOTLE, *Ética a Nicómaco*, Madrid, 2008, stated: “... when the law presents a universal case and circumstances arise that are outside the universal formula, then it is fine to the extent what the legislator omits and it is wrong to simplify the correction of this omission, since the legislator himself would have made this correction if he had been present and would have legislated this way if he had known about it. (...) And such is the nature of what is equitable: a correction of the law insofar as its universality leaves it incomplete”(5:10).

<sup>5</sup>The case of COMPAS(Correctional Offender Management Profiling for Alternative Sanctions) is well known: an algorithm used in criminal justice to predict the probability of recidivism based on about one hundred

Volume 8 Issue 5 - 2020

**Lorenzo M bujosa vadell**

Professor of Procedural Law, University of Salamanca, Spain

**Correspondence:** Lorenzo M bujosa vadell, Professor of Procedural Law, University of Salamanca, Spain, Email lbujosa@usal.es

**Received:** August 14, 2020 | **Published:** September 28, 2020

From the procedural perspective, we are interested in all the elements that can remove the decisions of the judge from excessive subjectivization and, consequently, from irrationality and arbitrariness. But, along the path of objectification that digital applications lead us to, we can take forward specific considerations, not taken into account by the programmers and which are neither more nor less than the links between jurisdictional activity and constitutional axiological idea of justice.

In all this, of course, it is essential to attend to the ethical questions involved, which have been put in the forefront by different organizations that are conscious on the problems of the application of artificial intelligence in the judicial process. This confirms the magnitude of the risks that we have to face, trying to go on through the balanced way: taking advantage of the virtues of technological advances in the process and avoiding falling into dramatic violations of guarantees consolidated in our legal systems.

## The complexity of artificial intelligence

In spite of the frequency with which in general we hear talk of “artificial intelligence”, for the uninitiated it still is a difficult term to define. It is a complex concept due to the breadth and heterogeneity of the elements it contains and, furthermore, because it is unequivocally unstable, due to its constant mobility and evolution. It is, perhaps, the best current example of work in progress of which it is already

factors. Vid. the judgment of the Supreme Court of the State of Wisconsin: *Loomis v. Wisconsin*, 881 NW2d 749 (Wis. 2016), which denies Loomis's claim regarding the violation of the right to due process with all guarantees (due process of law), based, among other reasons, on the fact that the judicial decision to the conviction was based on a risk assessment instrument that took into account gender and race and that was not susceptible to being disputed because it was a commercial secret. In turn, the Supreme Court of the United States denied the writ of certiorari on June 26, 2017.

Vid. also “Would You Trust An Artificially-Intelligent Expert?”, *The National Law Review* July 11, 2020, Volume X, Number 193:

<https://www.natlawreview.com/article/would-you-trust-artificially-intelligent-expert>

From a point of view more critical vid. TASHEA, J., “Courts Are Using AI to Sentence Criminals. That Must Stop Now”, at <https://www.wired.com/2017/04/courts-using-ai-sentence-criminals-must-stop-now/>

necessary to assess its various legal consequences,<sup>6</sup> but we still don't really know where it is going to take us, leaving us with multiple speculations that, in many cases, still fall within the realm of the unreal dreams.

Indeed, we speak of the existence of waves or generations of artificial intelligence, depending on the evolution of the instruments to which it refers and, above all, of the possibilities of action regarding the management of a multitude of data, substituting with advantage to the intelligence of the human being.<sup>7</sup> We can read also on "superintelligence",<sup>8</sup> referred to a continuous comparison with the human biological brain, with the aim of overcoming the capabilities of the human being.<sup>9</sup> Here we find ourselves with the great paradox that we are trying to overcome something that is still largely a mystery to us, despite the great advances in neuroscience.<sup>10</sup> It seems that we are on the historical-cultural path of English Romanticism,<sup>11</sup> when literature specifically alluded to the "creation of a mind".<sup>12</sup> And let's not forget that both the concept of mind and intelligence are also elusive, complex concepts with disputed delimitations.<sup>13</sup>

It is not my intention to intervene in specialized controversies for which I am far from being prepared, but it will be necessary to embrace a definition of artificial intelligence, which even with its defects and insecurities, allows us to move forward towards what interests us at most: the legal-procedural consequences of this new reality, and above all, as we will see immediately, the ethical limits of those consequences. Thus, I accept the apparently broad and

simple definition that has been upheld in the European Union by which it is stated that "The term 'artificial intelligence' (AI) applies to systems that display intelligent behavior, since they are able to analyze their environment and take action—with a certain degree of autonomy—in order to achieve specific objectives".<sup>14</sup> In this perspective, a main issue that concerns us directly stands out: that of generating trust in a human-centered artificial intelligence.

In times in which legal subjectivity has expanded notably, and thus the ability to be part of a process has been recognized by some Courts not only to separate beings, as the Atrato River in Colombia, also to nature itself or to Mother Earth, we are asked whether a robot, endowed with artificial intelligence,<sup>15</sup> it may one day become party to the judicial proceedings too, and even if an artificial being will be able issuing judgments that satisfy claims in a reasonable way. Of course, robotics cannot be confused with artificial intelligence, since we all have gadgets, mechanisms that work in an automated way for example, kitchen robots, but to some of them can be applied complex algorithms with what would already enter our concept.<sup>16</sup>

The European Parliament has ruled on several occasions about this conjunction between artificial intelligence and robotics, based on the idea that "artificial intelligence (AI) and robotics that are transparent and that integrate ethical considerations have the potential to enrich our lives and consolidate our capacities, both at the individual level and for the common good", but which at the same time are susceptible to malicious use that violates or, at least, jeopardizes fundamental rights and, for this, calls for a periodic evaluation of the legislation in order to ensure that it is fit for purpose in relation to artificial intelligence while respecting the fundamental values of the Union.<sup>17</sup>

It is worth remembering -and surprising us once again, for its acuity- that Isaac ASIMOV, as early as 1942, already concerned himself with the close relationship between artificial intelligence and insurmountable ethical guidelines, when formulating his Laws of the robotic:

1. A robot will not harm a human being or, through inaction, allow a human being to be harmed.
2. A robot must comply with the orders given by humans, except for those that conflict with the first law.
3. A robot must protect its own existence to the extent that this protection does not conflict with the first or second law.<sup>18</sup>

<sup>14</sup>Communication from the Commission to the European Parliament, the European Council, the Council, the European Economic and Social Committee and the Committee of the Regions. Artificial intelligence for Europe. Brussels, April 25, 2018 COM (2018) 237 final

<sup>15</sup>Vid. LACRUZ MANTECÓN, ML Robots y personas. Una aproximación legal a la subjetividad cibernética, Madrid, 2020. In an interdisciplinary way, BARRIO ANDRÉS, M. (Dir.), Derecho de los Robots, Madrid, 2018 and ROGEL VIDE, C., (Coord.), Los robots y el Derecho, Madrid, 2018.

<sup>16</sup>As SALAZAR, I. says in La revolución de los robots. Cómo la inteligencia artificial y la robótica afectan a nuestro futuro, Gijón, 2019, p. 51, "the robot revolution, of an unprecedented technological era in a completely interconnected world, is happening in record time".

<sup>17</sup>Cf. Resolution of the European Parliament, of February 12, 2019, on a European global industrial policy on artificial intelligence and robotics (2018/2088 (INI)), and, previously, the Resolution of the European Parliament, of February 16, 2017, with recommendations addressed to the Commission on civil law regulations on robotics (2015/2103 (INL)).

<sup>18</sup>ASIMOV, I., "Vicious Circle", published under the title Runaround, in the journal Astounding Science Fiction, in March 1942.

<sup>6</sup>SUSSKIND, R., Online Courts and the Future of the Justice, Oxford, 2019, p. 263, puts it in a striking way: "Scarcely a week passes without news of 'an AI' or a 'robot lawyer' that is outperforming or poised to replace traditional human lawyers in some legal task or other."

<sup>7</sup>Cf. SUSSKIND, R., Online Courts..., op. cit., pp. 264-275.

<sup>8</sup>BOSTROM, N., Superintelligence. Roads, dangers, strategies, 2016, p. 22, tentatively defines it as "any intellect that greatly exceeds the cognitive performance of humans in virtually all areas of concern."

<sup>9</sup>BOSTROM, N., Superintelligence..., op. cit., p. 60: "At present, the computational power of the biological brain still compares favorably with that of digital computers, although leading-edge supercomputers are reaching performance levels that are within the range of plausible estimates of brain processing power. But hardware is improving rapidly, and the ultimate performance limits for hardware are much higher than those for biological computing substrates."

<sup>10</sup>Vid. COBB, M., The Idea of the Brain. The past and the future of Neuroscience, New York, 2020.

<sup>11</sup>The memory of FRANKENSTEIN, by Mary SHELLEY, published on March 1, 2018, a year and a half after the famous year in which there was no summer, is inevitable here.

<sup>12</sup>Worth the examples of KURZWEIL, R., How to create a mind. The secret of human thought, Berlin, 2013, or in a critical sense, FODOR, J., La mente no funciona así. Alcance y límites de la psicología computacional, Madrid, 2003. Also PENROSE, R., La nueva mente del emperador, Madrid, 2019, originally published in 1989.

<sup>13</sup>OLIVEIRA, A., The Digital Mind. How Science is Redefining Humanity, London, 2017, p. 88: "... intelligence is a more elusive concept than had once been thought. Though it is commonly accepted that intelligence is required in order for a human to address any of the problems mentioned in the precedent paragraph, it isn't at all clear that the techniques computers used to solve those problems endowed them with general human-like intelligence. In fact, those problems were tackled with specialized approaches that were, in general, very different from the approaches used by humans". It refers to solving complicated mathematical theorems, designing correctly movement chess, plan activities, etc.

As we will see later, the centrality of the human being will continue to be the fundamental criterion in the face of advances in the application of artificial intelligence in the process.

## Judicial process and artificial intelligence

Of course, the application of information and communication technologies to the process with very varied purposes is not new, all of them aimed at facilitating the procedural activity. In Spain since the 1994 reform, the Organic Law of the Judiciary alluded, for the first time in our legal system, to the possibility of using technical, electronic and computer means for the development of the activity and the exercise of the functions of courts and courts.

It is sure that it took a long time to digitize the activity of the administration of justice, but the Spanish Law 18/2011, of July 5, regulating the use of information and communication technologies in the Administration of Justice made certainly a big step. With this Statute, the electronic judicial office, the electronic judicial file, the electronic records were introduced, ... Already the original version of the Civil Procedure Law of January 7, 2000, had included, among the means of evidence (art. 299.2), the instruments that allow to archive and know or reproduce words, data, figures and mathematical operations carried out for accounting or other purposes, relevant to the process.<sup>19</sup> Later, Law 19/2015, of July 13, on administrative reform measures in the field of the Administration of Justice and the Civil Registry, introduced electronic auctions. And from the entry into force of Organic Law 13/2015, of October 5, we have in Spanish criminal procedural legislation a broad regulation on the use of various electronic devices for the prosecution of both traditional crime and the complexities of cybercrime.

All this may arise threats to the integrity of fundamental rights in a due process, but it can be said that in the use of any of these technological applications there is any system of “strong” artificial intelligence, in the sense that there were “machines self-conscious”,<sup>20</sup> much less with autonomous decision-making capacity. It is usual to speak of different degrees of artificial intelligence: strong, weak, specific, general, as is typical of a broad and plural concept. This vagueness leads us to continual challenges, which we must try to overcome.

Thus, for the moment the replacement of the human being by the thinking machine seems to be out the scenery. It seems that there is not yet an application of strong and general artificial intelligence, with the ability to apply general knowledge to the tasks that are posed to it. In this way we would place ourselves before the fascinating debate about machine learning and making decisions according to that learning obviously, this will depend on what we understand by “learning”.<sup>21</sup> Anyway it is indisputable that weak artificial intelligence is already among us, based not on the recognition of independent

<sup>19</sup>For a complete and critical view, vid. BUENO DE MATA, F., *La prueba electrónica y proceso 2.0.*, Valencia, 2014.

<sup>20</sup>SUSSKIND, R., *Online Courts...*, op. cit., p. 265: “The term, ‘intelligence’, for many observers suggests perhaps that the latest systems are in some sense actually ‘conscious’. In the philosophical jargon of AI, a system that is conscious would be an exhibit of what is known as ‘strong AI’”.

<sup>21</sup>LÓPEZ MORATALLA, N., *Inteligencia artificial ¿Conciencia artificial?*, Madrid, 2017, p. 52: “Machines built with the so-called cognitive technology, simulating the human brain, could have Machine Learning or Deep Learning. In a simplified way, it can be said that, in the first case, they simulate the functioning of the neuron, working on one layer or several layers. In the second, they would simulate complex neuronal networks”.

patterns that, in turn, are integrated, but on machine learning, with the processing of large amounts of data and with true deductive reasoning from inductive learning.<sup>22</sup>

As a consequence of all this, SUSSKIND tells us about disruptive technologies for the lawyer of tomorrow, in the sense that they fundamentally challenge and change the operation of a given sector.<sup>23</sup> In our area we can say that there is a trend in access to justice<sup>24</sup> that changes the way, at least partially: from face-to-face and physical to virtual and electronic, even it has much more facilities to obtain information, and we can also affirm that some decisions based on objective criteria can be automated.<sup>25</sup> Possibly a clear example could be the decisions regarding the order for payment procedure (procedimiento monitorio).

Out of these exceptional cases of automation, it seems that it is more realistic to limit ourselves to the cases in which artificial intelligence is used as an auxiliary instrument for the judge,<sup>26</sup> although it is not reasonable to disregard the instruments that allow great predictability in judicial decisions,<sup>27</sup> with the problems that this entails, such as implicit biases both in the configuration of the algorithm and in the selection criteria of the data that are entered for the decision.

Already in 1998 SARTOR and BRANTING proposed an analysis of the judicial applications of artificial intelligence<sup>28</sup> and pointed to a continuous trend of expansion towards more complex areas that imply important challenges, moving towards a greater influence on the judicial decision. It started from the examination of the automatic preparation of complementary judicial documents, to observe the creation of a support and information supply system for an adequate adoption of judicial decisions and the application of artificial intelligence systems that can help regulate judicial discretionality, the

<sup>22</sup>Cf. LÓPEZ MORATALLA, N., *Inteligencia artificial...*, op. cit., p. 74.

<sup>23</sup>SUSSKIND, R., *El abogado del mañana. Una introducción a tu futuro*, 2nd ed., Madrid, 2020, p. 77.

<sup>24</sup>Vid. SOLAR CAYÓN, JI, *La inteligencia artificial jurídica. El impacto de la innovación tecnológica en la práctica del Derecho y el mercado de servicios jurídicos*, Cizur Menor, 2019, pp. 306-313.

<sup>25</sup>Article 42 of the aforementioned Law 18/2011 refers to automated judicial actions: “In the event of automated action, the definition of the specifications, programming must be previously established by the State Technical Committee of the Electronic Judicial Administration, maintenance, supervision and quality control and, where appropriate, the audit of the information system and its source code.

The systems will include the management indicators established by the National Commission for Judicial Statistics and the State Technical Committee for Electronic Judicial Administration, each one within the scope of their powers”. NIEVA FENOLL, J., *Inteligencia artificial y proceso judicial*, Madrid, 2018, p. 33-41, assesses what it calls “automatable decisions, attending to” the most obvious aspects, which in fact in some places are already being automated.

<sup>26</sup>SUSSKIND, R., *Online Courts...*, op. cit., p. 265: “instead, my attention here is on ‘weak AI’ -functionally, these systems seem to be doing some of the work of lawyers but without the cognitive states enjoyed by humans, such as our self-awareness and emotional satisfaction”.

<sup>27</sup>For a historical perspective of these advances, vid., PALIWALA, A. (Ed.), *A history of legal informatics*, Zaragoza, 2010. A current perspective in SAN MIGUEL CASO, C., “The techniques of judicial prediction and its impact on the process”, CONDE FUENTES, J., and SERRANO HOYO, G., (Dirs.), *Digital justice in Spain and the European Union*, Barcelona, 2019.

<sup>28</sup>Vid. SARTOR, G., ad BRANTING, K., (Eds.), *Judicial Applications of Artificial Intelligence*, Dordrecht, 2010.

adequate selection and application of the doctrine of the precedents and, ultimately, the computer support to reach the “best” judicial resolution.<sup>29</sup>

On the other hand, the existence of computer programs that introduce the most advanced artificial intelligence systems should not lead us to the error of thinking that they will be directly applied in daily jurisdictional activity. Furthermore, there is a well-founded complaint about the lack of use of the systems already available in daily judicial management,<sup>30</sup> which is not necessarily a consequence of the digital divide, but rather of the conservative trend of work in courts and tribunals.

## The relevance of ethical issues

The profound challenge caused by the application, to some extent, of artificial intelligence to some aspects of the judicial process runs the risk of producing a mirage or enthralling, for which we underestimate the negative consequences and insecurities we may incur. For this reason, as we have already seen, we insist on placing the human being at the center and on articulating a series of ethical rules that must be respected to prevent all this problem from getting out of hand.

In many of the computer applications the key to their effectiveness is found in the handling of an exorbitant amount of data, some of them close to privacy,<sup>31</sup> which already places us before some specific questions from the constitutional point of view.<sup>32</sup> Thus, using the well-known methodology of human rights generations, VASAK speaks of the third generation,<sup>33</sup> in which PEREZ LUÑO places the consequences of the technological revolution.<sup>34</sup>

However, a preliminary step must be taken into account, by which it is intended to offer a series of moral reflections, and even ethical principles, that could serve as a guide to the legislator. In this way, we return to the European Union, which already has a long tradition in what has been known since the Stockholm Program as “E-Justice” or “electronic justice”.<sup>35</sup> In recent years has highlighted a

<sup>29</sup>Specially the last chapter, in which TARUFFO, M., “Judicial decisions and Artificial Intelligence”, SARTOR, G., ad BRANTING, K., (Eds.), *Judicial Applications ...*, op. cit., pp. 207-220. The Italian jurist says (p. 212) “If one considers the evident features of complexity, variability, flexibility and discretion that are typical of court decisions, any approach aimed at interpreting the legal reasoning according to logical rules and models may appear as doomed to failure”. And he continues (p. 214): “The so-called ‘easy cases’ are often not easy enough to be standardized in terms of computerized models. A fortiori such a standardization is extremely difficult or impossible to achieve in the ‘average’ judicial case, let alone in hard cases. It seems, therefore, that the type of computerization we are considering may be useful but only in a relative narrow area of judicial practice”.

<sup>30</sup>Cf. SUSSKIND, R., *El legislador de mañana...*, op. cit., p. 147, where he speaks, precisely for that reason, of “disappointing progress.”

<sup>31</sup>Vid. MAYER-SCHÖNBERGER, V., and CUKIER, K., *Big Data. The revolution of big data*, Madrid, 2013, and HOFFMANN-RIEM, W., *Big Data. Challenges also for the Law*, Madrid, 2018.

<sup>32</sup>GARCÍA COSTA, M.F., “Perfiles constitucionales de la justicia electrónica”, in GÓMEZ MANRESA, M.F., and FERNÁNDEZ SALMERÓN, M., *Modernización e innovación en la Administración de Justicia*, CizurMenor, 2019, pp. 23-35.

<sup>33</sup>VASAK, K., *Pour les droits de l’homme de la troisième génération*, Strasbourg, 1979.

<sup>34</sup>PÉREZ LUÑO, A.E., *Human rights in technological society*, Madrid, 2012.

<sup>35</sup>Stockholm Program - An open and safe Europe serving and protecting the citizen [Official Gazette No. C 115 of May 4, 2010]. Vine. SENÉS MOTILLA, C., (Coord), *Presente y futuro de la e-Justicia en España y en la Unión Europea*, Cizur Menor, 2010, and DE LA OLIVA SANTOS, A.; GASCÓN

concern about the ethical dimension of artificial intelligence, as outlined in the Communication from the European Commission, on April 8, 2019, which takes as headings that of “Building trust in artificial intelligence centered on the human being”<sup>36</sup> and this text highlights that “Artificial Intelligence is not an end in itself, but a means that should serve people with the ultimate goal of increasing their well-being. For this, the reliability of artificial intelligence must be guaranteed”.

Thus, we find “reliability” as a key concept, which must be reinforced with a series of essential requirements that are formulated as requirements for the application of artificial intelligence:

- Human intervention and supervision
- Technical strength and security
- Privacy and data management
- Transparency
- Diversity, non-discrimination and equity
- Social and environmental well-being
- Accountability

We are in a general perspective, and therefore prior to the specific application of artificial intelligence to the process, but the conclusion is far from minor: “The ethical dimension of artificial intelligence is not a luxury or an accessory: it must be an integral part of the development of artificial intelligence”.

Obviously, these very delicate matters should not be separated from the hands of the experts, and for this reason, a “European Group on Ethics in Science and New Technologies” was set up, which acts as an advisor to the European Commission. Among his works, we are interested in the “Declaration on artificial intelligence, robotics and ‘autonomous’ systems”, issued on March 9, 2018,<sup>37</sup> since it starts from some basic ethical principles and democratic assumptions and, therefore, of an evident constitutional scope:

- Human dignity
- Autonomy
- Responsibility
- Justice, equity and solidarity
- Democracy
- Rule of law
- Physical and mental safety and integrity
- Data protection and privacy
- Sustainability

Also a high-level expert group on artificial intelligence was set up as the center of dialogue between the government of the European Union and the people involved in the future of artificial intelligence. This Group of Experts took up the central idea of reliability and on April 8, 2019 proposed some “Ethical Guidelines for a reliable artificial

INCHAUSTI, F.; AGUILERA MORALES, M., (Coords.), *La e-Justicia en la Unión Europea. Desarrollos en el ámbito europeo y en los ordenamientos nacionales*, CizurMenor, 2012.

<sup>36</sup>Brussels, 8.4.2019 COM (2019) 168 final.

<sup>37</sup>[http://ec.europa.eu/research/egp/pdf/egp\\_ai\\_statement\\_2018.pdf](http://ec.europa.eu/research/egp/pdf/egp_ai_statement_2018.pdf)

intelligence “.<sup>38</sup> These guidelines are based on three components that must be satisfied throughout the entire cycle of the system:

- a. artificial intelligence must be lawful, that is, comply with all applicable laws and regulations;
- b. it must be ethical, so as to guarantee respect for ethical principles and values; and
- c. it must be robust, both from a technical and social point of view, since artificial intelligence systems, even with good intentions, can cause accidental damage.

More recently, in an even more institutional way, the European Commission used the White Paper technique to open a larger debate on artificial intelligence.<sup>39</sup> Again, the approach is broad, therefore it does not refer specifically to the application of artificial intelligence to justice, but with considerations much closer to the object of our concern, since, on the one hand, we speak of the existence of risks to fundamental rights, especially the protection of personal data and privacy and non-discrimination.<sup>40</sup> It is stated that such risks may be the result of defects in the general design of artificial intelligence systems or the use of data that can be biased without prior correction (for example, a system is trained using only or mainly data relating to men, and this translates into worse results in relation to women). But there is also another dimension of risks to safety and the effective functioning of the civil liability regime.<sup>41</sup>

It is important to underline the need of human supervision, as it helps to ensure that an artificial intelligence system does not undermine human autonomy or cause other adverse effects. In short, it is a matter of trying to compensate for some particular characteristics of some of the artificial intelligence technologies, such as opacity, complexity, unpredictability and partially autonomous behavior, which can complicate the effective application of the regulations on fundamental rights.

Likewise, from an ethical perspective, from outside the European Union it may be interesting to cite the well-known ASILOMAR Principles on artificial intelligence, approved within the Future of Life Institute, at the meeting of more than one hundred experts from different disciplines, met in Pacific Grove (California) in January 2017.<sup>42</sup> Also in a generic way, a long series of requirements that artificial intelligence systems must meet are listed. From a judicial process point of view, we are mainly interested in the following:

“Security: Artificial intelligence systems should be secure throughout their operational life and verifiability where possible.

Transparency in system failures: If an artificial intelligence system causes harm it should be possible to determine the cause.  
Transparency in decisions: Any intervention of an autonomous

<sup>38</sup><https://op.europa.eu/es/publication-detail/-/publication/d3988569-0434-11ea-8c1f-01aa75ed71a1>

<sup>39</sup>White Paper on artificial intelligence - An European oriented approach to excellence and trust. Brussels, February 19, 2020 COM (2020) 65 final

<sup>40</sup>“The use of artificial intelligence can affect the values on which the EU is founded and cause the violation of fundamental rights, such as freedom of expression, freedom of assembly, human dignity, the absence of discrimination based on reason of sex, race or ethnic origin, religion or creed, disability, age or sexual orientation, and, in its application in certain areas, the protection of personal data and private life, the right to effective judicial protection and a fair trial, or consumer protection”.

<sup>41</sup>“AI technologies can present new security risks for users when integrated into products and services.”

<sup>42</sup>Vid. <https://futureoflife.org/ai-principles/>

system in a decision should be accompanied by a satisfactory and controllable explanation by a competent human authority.

(...) Respect for values: Highly autonomous artificial intelligence systems should be designed so that their goals and behaviors can be aligned with human values throughout their operations.

Human values: Artificial intelligence systems should be designed and operated to be compatible with the ideals of human dignity, rights, freedoms and cultural diversity  
Personal privacy: People should have the right to access, manage and control the data that are generated by the application of artificial intelligence systems”.

Although all these criteria and requirements are formulated in a general way, they are naturally applicable to the various modalities that artificial intelligence may adopt in the administration of justice. It is essential that such instruments are also reliable and that the necessary precautions are taken with regard to security, transparency, privacy, non-discrimination and human dignity. If there are doubts yet regarding any of these items, it would be better to discard the advantages that may derive from their use, as there are too many risks to be able to affirm with justification that respect for the human being continues to be at the center of all this. It seems clear, however, that this requires a precise regulatory development, which is to be determined.<sup>43</sup>

## Open conclusions

All the previous considerations, the preparatory documents, the doctrinal reflections, the continuous technological advances can only lead us to open conclusions. We are at a time when we ask ourselves questions, dilemmas, needs, but also demands. It is not, therefore, a totally nebulous and undefined situation. Although it is clear that it is an object of study in continuous evolution, we have some unavoidable ethical references which should materialize more and serve as boundaries for legislative developments.

The realistic view of the application of artificial intelligence to the process focuses today on the consideration of methods of support for the judge, for example allowing rapid selection of case law. However, in some decisions we are already approaching the application of instruments that, at least partially, may involve a certain substitution of the judge in some of its decisions, such as in automated payment requirements; in the evaluation of electronic evidence (block chain); in assessment of probabilities of recidivism. All this should not hide from us the persistence of serious problems that are waiting a reasonable resolution: the question of the neutral construction of the algorithm; the problem of biases in the data that are being entered into the computer system; the protection of data (big data) and privacy, or even the issue of autonomy in learning and its control.

## Acknowledgments

None.

## Conflicts of interest

Author declares there are no conflicts of interest.

<sup>43</sup>LATORRE, JI, *Ética para máquinas*, Barcelona, 2019, pp. 206-207: “Foolishness is rampant. Artificial intelligences advance without agreed ethical codes. No politician talks about what to program, what criteria for good and evil govern the actions of a drone, a robot or the traffic manager of a city. This error is not new. Humans have not been able to establish consensual ethical standards for previous advances”. In this sense, COECKELBERGH, M., *AI Ethics*, Cambridge (Massachusetts), 2020, pp. 144-165, refers to “what needs to be done and other questions that politicians must answer”.