CODES OF CONDUCT AND GUIDELINES (CHAPTER X) *

Asunción Nicolás Lucas
Senior Lecturer in Administrative Law
University of Oviedo

1. *Codes of conduct (Article 95 AIA)*

1.1. *Codes of conduct as a complement to legal regulation for systems that are not considered 'high-risk'*

The European Union recognises the importance of regulating the development, deployment, and use of Artificial Intelligence (AI) in order to protect the fundamental rights of citizens and to ensure safety and ethics in its application. The reasons are that Artificial Intelligence can significantly impact areas such as privacy, non-discrimination, transparency, and accountability, and appropriate regulation is required to address these challenges.

This is why, alongside the desire to create European regulation in this area, [1]

---

* This paper has been written within the framework of the Research Project 'Algorithmic tools for citizens and public administrations', funded by the Spanish Ministry of Science and Innovation (PID2021-126881OB-I00).

[1] Regulation with which the EU aims to lead the development of safe,

which has led to the passing of the Artificial Intelligence Act (AIA), the promotion of voluntary codes of conduct[2] can complement such legal regulation by providing practical guidance and ethical standards for those developing, deploying or using[3] Artificial Intelligence systems. Artificial intelligence-related codes of conduct could be said to be sets of principles, guidelines, and ethical standards designed to guide the development, deployment, and responsible use of Artificial Intelligence systems. Codes of conduct aim to promote ethical and responsible practices in the development and use of AI, as well as to foster public confidence in this emerging technology. This phenomenon is arguably associated with economic globalisation[4] and will serve as the basis for a commitment to the conduct of companies designing and building AI systems.

They would constitute a sample of what we can call self-regulation by the entities, institutions, private organisations, or subjects that intervene in this context, in order to regulate themselves. And in this self-regulation, ethics has a marked role to play. Despite a certain scepticism on the part of jurists, there are certain advantages and contributions to the development of AI ethics. Among

---

trustworthy, and ethical artificial intelligence, as indicated by the European Council (Extraordinary meeting, 1-2 October 2020: https://www.consilium.europa.eu/es/policies/a-digital-future-for-europe/), and to ensure the protection of ethical principles, as requested by the European Parliament. See the European Parliament resolution of 20 October 2020 with recommendations to the Commission on a framework of ethical aspects of artificial intelligence, robotics and related technologies: https://www.europarl.europa.eu/doceo/document/TA-9-2020-0275_EN.html.

[2] This is one of the three options proposed by D. García San José, *Implicaciones jurídicas y bioéticas de la Inteligencia Artificial (AI). Especial consideración al marco normativo internacional*, in *Cuadernos de Derecho Transnacional*, vol. 13, no. 1, 2021, pp. 255 ff. The three options for regulating Artificial Intelligence would be the drafting of international treaties, codes of conduct and judicial activism by international control bodies.

[3] On the public uses of AI in the public sector, L. Cotino Hueso *Los usos de la inteligencia artificial en el sector público, su variable impacto y categorización jurídica*, in *Revista canaria de Administración pública*, no. 1, 2023, pp. 211 ff.

[4] ILO Governing Body, *Overview of global developments and Office activities concerning codes of conduct, social labelling and other private sector initiatives addressing labour issues*, GB.273/WP/SDL/1, 1998, p. 85: https://webapps.ilo.org/public/english/standards/relm/gb/docs/gb273/sdl-1.htm.

other considerations, it would be better adapted to private, business scenarios, corporate organisations, and technical and professional sectors, in that it is conveyed through self-regulation and through ethical rules and codes of conduct to which operators could adhere.

Public ethics has been emphasised for its unique preventive character, as opposed to the more reactive role of law. The development of AI is also about prevention rather than cure. In the face of a criminal or disciplinary reaction from the law, it is necessary to develop compelling, coherent, and generalised alert and vigilance mechanisms. It would be a matter of defining an AI ethics policy that would include codes of ethics or conduct that coexist with the regulatory framework.[5]

For example, in the area of data protection (which already included this practice before the Artificial Intelligence Act), as we will have the opportunity to comment, these are voluntary compliance mechanisms that establish specific rules for data controllers or processors in order to correctly apply the General Data Protection Regulation and, consequently in our country, the Organic Law on the Protection of Personal Data and the Guarantee of Digital Rights. In this case, its subjective scope is indeed somewhat broader since it covers not only private subjects but also public entities; associations and other bodies representing categories of data controllers and data processors; companies or groups of companies; constitutional bodies, Public Administrations, independent administrative authorities, Public Universities, Public Sector Foundations, consortia; bodies that assume supervisory and extrajudicial conflict resolution functions. In short, the aim would be to facilitate the application of the regulations according to the different characteristics of the various sectors of activity of their promoters.[6]

The reference to data protection[7] is not trivial insofar as the AI Regulation

---

[5] L. Cotino Hueso, *Ética en el diseño para el desarrollo de una inteligencia artificial, robótica y big data confiables y su utilidad desde el derecho*, in *Revista Catalana de Dret Públic*, no. 58, 2019, pp. 41-42; D. Terrón Santos, *Limitar la IA desde la ética y el derecho*, in J.L. Domínguez Álvarez y D. Terrón Santos (eds.), *Desafíos éticos, jurídicos y tecnológicos del avance digital*, Iustel, 2023, pp. 175 ff.

[6] https://www.aepd.es/derechos-y-deberes/cumple-tus-deberes/medidas-de-cumplimiento/codigos-de-conducta.

[7] Data are one of the elements to be taken into account by predictive algorithms, together with the algorithmic formula and computational capacity. Precisely the use of data leads A. Huergo Lora, *Una aproximación a los algoritmos desde*

itself justifies the consistency of its provisions with the Charter of Fundamental Rights of the European Union and existing secondary EU law on data protection, consumer protection, non-discrimination, and gender equality, without prejudice, of course, to the General Data Protection Regulation [Regulation (EU) 2016/679] and the Criminal Data Protection Directive [Directive (EU) 2016/680].

Among other things to consider, codes of conduct can address a wide range of issues, including *ethics and accountability* issues: they should set out ethical principles that AI developers and users should follow, to ensure that their actions are ethical and responsible. This may include respect for human rights, fairness, and transparency in designing and operating AI systems. *Non-discrimination*: They should focus on avoiding unfair discrimination in AI systems, ensuring that they do not perpetuate bias, or prejudice based on protected characteristics such as race, gender, religion, or sexual orientation. [8]

---

el Derecho Administrativo, in A. Huergo Lora (ed.), *La regulación de los algoritmos*, Aranzadi, 2020, page 52, states that 'it is not surprising that data protection rules are the first legal response to the phenomenon'. S. Civitarese Matteucci, *Umano troppo umano. Decisioni amministrative automatizzate e principio di legalità*, in *Diritto pubblico*, 2019, no. 1, pp. 5 ff., underlines that the General Data Protection Regulation deals with the right to data protection, but not with other legally relevant perspectives, such as the normative basis for the administration to act in an automated way.

It is also true, as stated by Huergo Lora, op. cit., p. 56, that 'technology has made it easier for users to record all their activities, which has many advantages (think of the activity bracelets that allow measuring sports practice or physical activity in general, the cards and electronic payments in general, which facilitate the control of expenditure and avoid having to file thousands of paper documents, or geolocation and its multiple utilities, etc.). With few exceptions, citizens collaborate voluntarily and create large amounts of data. An economic model has been established in which this data is useful not only for citizens but also for companies, and it is often the potential of this data that has led companies to create the applications that users use'.

[8] On the risk of discrimination of algorithms, see Huergo Lora, op. cit. Huergo Lora, op. cit. page 84, when he states that 'in the case of algorithms, in which, by definition, the programmer does not define the profile, but rather it is elaborated by the algorithm from the available data, discrimination may be the consequence of using, to 'train' the algorithm, data that should not be taken into consideration because it is prohibited by Article 14 EC, such as sex, for example. If we use the sex of the subjects as one of the elements that can be used to produce a predictive

Transparency and explainability: should promote transparency in developing and deploying AI systems, allowing users to understand how they work and why they make certain decisions. This may involve clear documentation of the algorithms used and disclosure of the data and processes involved in decision-making. *Privacy and data protection*: They should ensure the protection of privacy and personal data in the context of AI, ensure compliance with data protection laws and regulations, and take measures to minimise the risk of privacy abuses or breaches. *Security and reliability*: with the aim of preventing accidents, failures, or malicious manipulations that could cause harm to individuals or society at large. *Inclusion and accessibility*: In the design of AI systems, ensuring that they are accessible to all people, including those with disabilities or in disadvantaged situations.

Although Europe has just approved and published its Artificial Intelligence Regulation, until all of its content is definitively applied, months and years will pass (depending, as we will see in another chapter, on the part of the Regulation to which we refer), and AI continues to advance unchecked. To avoid this possible lack of control, the United States and Europe proposed to present a voluntary code of conduct for companies developing this type of technology,[9] which companies from other countries such as Canada, the United Kingdom, Japan, and India tried to join, and which could provide a *regulatory bridge* until the practical application of this EU regulation, thereby giving the public confidence (in the words of the Vice-President of the European Commission for digital and competition, Margrethe Vestager). At the beginning of September 2023, representatives of Canada, France, Germany, Italy, Japan, the United Kingdom, and the United States reached

---

profile, the algorithm will arrive at a result that gives preferential treatment to one sex over the other, which is contrary to Article 14.

On this risk of discrimination, see also A. Soriano Arnanz, *Decisiones automatizadas y discriminación: aproximación y propuestas generales*, in *Revista General de Derecho Administrativo*, no. 56, 2021; M. Moreno Rebato, Inteligencia artificial *(Inteligencia artificial (Umbrales éticos, Derecho y Administraciones Públicas),* Civitas, 2021.

[9] Representatives of the United States and the European Union held a bilateral meeting in Sweden in May 2023 for this purpose; in this regard, *the US and the EU will propose a voluntary code of conduct for generative artificial intelligence*, in El País: https://elpais.com/tecnologia/2023-05-31/eeuu-y-la-ue-propondran-un-codigo-de-conducta-voluntario-para-la-inteligencia-artificial-generativa.html#).

Countries such as China, which also has plans for regulation, including a safety inspection of artificial intelligence tools, may remain on the sidelines.

an agreement with a position close to the Japanese position, halfway between the *laissez-faire* of the United States and European ethical regulation. Finally, at its last meeting (30 October 2023), the G7 announced the creation of an international code of conduct to minimise the risks of artificial intelligence within the framework of the Hiroshima artificial intelligence process, to which we will devote a section to this study. However, as we know, this code is not binding, so that large technological giants are not bound by it.

In a way, these parallel paths respond to the need to find a firm regulation of AI but also to try to give a *de facto* prior response to the risks posed by this not-so-new technology, as well as to the fears raised in all sectors, including personalities who move precisely in this world.[10]

Without dwelling too much on this topic, since the advantages and disadvantages of AI have been recurrently analysed in any exposition on this subject, both from an ethical and legal point of view,[11] Artificial Intelligence presents a transformative potential in various areas of society, but it also brings with it certain risks[12] and challenges that need to be carefully addressed. Some

---

[10] A group of 350 executives, researchers, and engineers with expertise in this technology signed an open letter whose title already says a lot about the risk of this technology: *Mitigating the risk of extinction [for humanity] of AI should be a global priority along with other risks on a societal scale, such as pandemics and nuclear war.* Among others, Sam Altman (CEO of OpenAI), Demis Hassabis (Google DeepMind), Dario Amodei (Anthropic), Geoffrey Hinton, and Yoshua Bengio (AI researchers and considered the godfathers of the modern AI movement) signed the document. Sam Altman had told the US Senate that he understood that people were anxious about how AI could change the way we live but that we needed to work together to identify and manage the risks and downsides so that we can also enjoy the tremendous benefits that AI will provide.

[11] R. López de Mantaras y P. Messeguer González, *Inteligencia artificial*, Catarata-CSIC, 2017.

[12] Following, for example, A. Soriano Arnanz, *Decisiones automatizadas y discriminación: aproximación y propuestas generales*, in *Revista General de Derecho Administrativo,* no. 56, 2021, who speaks of singularity, autonomy, protection of personal data, biases, errors, discrimination, private and not public control of the code and data, opacity and impossibility of access to the source code of the algorithm. A relationship with biases can also be seen in A. Cerrillo i Martínez, *La inteligencia artificial y el control de sus posibles sesgos,* in M. Villoria Mendieta (ed.), *Ética pública en el siglo XIX,* INAP, 2021, pp. 93 ff.

of the principal risks of Artificial Intelligence include: *inequality and discrimination*, AI algorithms can reflect and amplify existing biases in the data they are trained on. This can lead to discriminatory decisions in areas such as recruitment, credit allocation, or criminal justice, disproportionately affecting certain groups. *Privacy and security,* the massive data collection necessary to train AI systems, raises privacy concerns. This data may be vulnerable to security breaches and misuse if not adequately protected. *Technology dependency,* increasing automation driven by AI may lead to greater reliance on technology, which could have negative repercussions in the event of system failures or cyber-attacks. *Job displacement*, the implementation of AI may lead to the automation of jobs, which could result in the displacement of workers in specific industries. This will require effective retraining and retraining programmes. *Ethics and liability*, the autonomy of AI systems raises ethical challenges, such as who is responsible for erroneous or harmful decisions made by algorithms without adequate human oversight. *Superintelligence,* in the long term, the development of superintelligent AI raises questions about control and security, as well as how to ensure that the goals of such intelligence are compatible with human values.

It is, therefore, essential to address these risks through appropriate policies and regulations, as well as by promoting ethical[13] and transparent research in

---

[13] On the role of ethics in the development of AI and its relation to the development of codes of conduct, see also the papers of the *XVIII Congreso de la Asociación Española de Profesores de Derecho Administrativo*, entitled *El Derecho Administrativo en la era de la Inteligencia Artificial.* https://www.aepda.es/AEPDAEntrada-3987-XVIII-CONGRESO-DE-LA-ASOCIACION-ESPANOLA-DE-PROFESORES-DE-DERECHO-ADMINISTRATIVO.aspx.

S. Tavares da Silva, in *El Derecho Administrativo en la era de la inteligencia artificial*, speaks of the role of ethics in identifying principles applicable to artificial intelligence and thus overcoming human imperfection or respecting a series of values that are in accordance with the dignity of the human person, as stated by J.A. Hernández Corchete, in *¿Un nuevo estatuto jurídico para el ciudadano?* although there must be an integration between law and ethics in the regulation of Artificial Intelligence, as stated by A. Mantelero, in *Retos y regulación de la Inteligencia Artificial: la toma de decisiones en los asuntos públicos y la administración de justicia,* or by A. Huergo Lora, in *De la digitalización a la inteligencia artificial: ¿evolución o revolución?* who points out the dimension of moral principles as the basis of any normative regulation, although this goes further because it provides a sanctioning system that distinguishes the legal the field of AI. The goal should be to harness the benefits of AI while mitigating its potential adverse effects.

1.2. *EU Regulation of codes of conduct (Article 95 AIA)*

In the sense of what we have discussed in previous paragraphs, Article 95 AIA states as follows:

'1.The AI Office [14] and the Member States shall encourage and facilitate the drawing up of codes of conduct, including related governance mechanisms, intended to foster the voluntary application to AI systems, other than high-risk AI systems, of some or all of the requirements set out in Chapter III, Section 2 taking into account the available technical solutions and industry best practices allowing for the application of such requirements.

2.The AI Office and the Member States shall facilitate the drawing up of codes of conduct concerning the voluntary application, including by deployers, [15] of specific requirements to all AI systems, on the basis of clear objectives and key performance indicators to measure the achievement of those objectives, including elements such as, but not limited to:

---

from the ethical, which is why AI calls for legal regulation and not just ethical principles. Principles that were born within ethics but have ended up being juridified and integrated into legal rules, as stated by S. de la Sierra, in *El ejercicio de potestades mediante Inteligencia Artificial (The exercise of powers through Artificial Intelligence*). On the contribution of ethics, see the paper by J.L. Bermejo Latre, *La aplicación de la inteligencia artificial en la actividad formal e informal de la Administración*.

[14] According to Article 3 (47) AIA, 'AI Office' means the Commission's function of contributing to the implementation, monitoring and supervision of AI systems and general-purpose AI models, and AI governance, provided for in Commission Decision of 24 January 2024; references in this Regulation to the AI Office shall be construed as references to the Commission.

[15] According to Article 3 (4) AIA, 'deployer' means a natural or legal person, public authority, agency or other body using an AI system under its authority except where the AI system is used in the course of a personal non-professional activity. These persons have to take appropriate technical and organisational measures to ensure that the systems are used in accordance with the accompanying instructions for use; they will have to entrust human supervision to natural persons who have the necessary competence, training, and authority; and they will ensure that the input data are relevant and sufficiently representative for the intended purpose of the AI system (Article 26 AIA, although referring to high-risk AI systems).

(a) applicable elements provided for in Union ethical guidelines for trustworthy AI;[16]

(b) assessing and minimising the impact of AI systems on environmental sustainability, including as regards energy-efficient programming and techniques for the efficient design, training and use of AI;

(c) promoting AI literacy, in particular that of persons dealing with the development, operation and use of AI;

(d) facilitating an inclusive and diverse design of AI systems, including through the establishment of inclusive and diverse development teams and the promotion of stakeholders' participation in that process;

(e) assessing and preventing the negative impact of AI systems on vulnerable persons or groups of vulnerable persons, including as regards accessibility for persons with a disability, as well as on gender equality.

3. Codes of conduct may be drawn up by individual providers or deployers of AI systems or by organisations representing them or by both, including with the involvement of any interested stakeholders and their representative organisations, including civil society organisations and academia. Codes of conduct may cover one or more AI systems taking into account the similarity of the intended purpose of the relevant systems.

4.The AI Office and the Member States shall take into account the specific interests and needs of SMEs, including start-ups, when encouraging and facilitating the drawing up of codes of conduct'.

Artificial Intelligence (AI), this rapidly evolving set of technologies, can generate, as we have said, many economic and social benefits in multiple sectors and social activities (think of sectors as relevant as climate change, the environment, health, the public sector, finance, mobility, agriculture), but also many risks for specific individuals or larger groups or collectives or society as a whole. Hence, the need to develop such Artificial Intelligence (AI) related *codes of conduct* for low to moderate risk AI systems, which are often approached through the prism of ethical principles and guidelines that promote the responsible development and deployment of the technology. These ethical codes and principles seek to provide a framework for the responsible development of AI, including in low-risk systems, promoting values such as

---

[16] This document proposes, among other things, to have organisations that can certify that AI systems are transparent, accountable, and fair. Organisations that may issue certifications, also referred to in this AI Regulation, conformity assessments, notified body certifications or market surveillance organisations.

transparency, fairness, inclusiveness, and respect for human rights. By adopting and applying these ethical standards, a safer and more ethical approach to integrating AI into various areas of society can be fostered.

1.2.1. *What characteristics should codes of conduct have?*

According to this article, the AI Office and the Member States shall encourage and facilitate the elaboration of these codes of conduct with the corresponding governance mechanisms.

Secondly, these are voluntary texts. The exact requirements set out in Chapter III, Section 2 for 'high risk' AI systems will apply to 'non-high risk' systems. Requirements concerning the quality of the data sets used, technical documentation and recording, transparency, communication of information to users, human oversight, robustness, accuracy, and cybersecurity apply to high-risk AI systems. Many leading-edge operators were already implementing the minimum requirements proposed by the preparatory work developed by a high-level expert group on AI and endorsed by the Commission in its 2019 Communication on human-centric AI. The problem may arise because there are systems that are not high risk but that do have a high impact, and the application of these principles will be left to voluntary compliance through these codes of conduct.[17]

Mention is made in this article; however, it is a specific requirement for applying the Union's ethical guidelines for reliable AI.[18] In these Guidelines,

---

[17] L. Cotino Hueso, *Los usos de la inteligencia artificial...*, op. cit., 2023, pp. 222-223, mentions among others the prosecution of fraud, money laundering, taxes, social security, labour inspection, traffic sanctions or similar that 'happen to be especially conflictive. In my opinion, this is an unacceptable option, given the enormous advance of these uses of public AI, which generate so much impact with total opacity'. See also B. Olivares Olivares, *Implicaciones de la normativa sobre protección de datos en el desarrollo de la inteligencia artificial por la administración tributaria: la gobernanza de los datos*, in S. Moreno González (ed.), *Nuevas tecnologías disruptivas y tributación*, 2021, Thomson Reuters-Aranzadi, pp. 180 ff.

[18] HLEG, *Ethics guidelines for trustworthy AI*, 2019(https://digital-strategy.ec.europa.eu/en/library/ethics-guidelines-trustworthy-ai). As stated by S. Muñoz Machado, *Prólogo*, *El Cronista del Estado Social y Democrático de Derecho*, vol. 100, 2022, p. 8, 'artificial intelligence systems must have human supervision and control. Special attention must be paid to prevent artificial intelligence systems from causing

the Expert Group had developed seven non-binding ethical principles for AI that were intended to help ensure the trustworthiness and ethical basis of AI. These seven principles were: human action and oversight; technical soundness and safety; privacy and data management; transparency; diversity, non-discrimination and equity; social and environmental well-being; accountability.

Other specific requirements included: environmental sustainability; energy efficiency from a programming point of view, as well as techniques for designing, training, and using AI; promotion of literacy, in particular with people involved in the development, operation, and use of AI; inclusive and diverse AI equipment and system designs, assessing the harms that AI systems may have on vulnerable people or groups, for example in terms of accessibility for people with disabilities or gender equality.

These Codes of Conduct may be developed by individual AI system providers; by AI system deployers; by organisations representing AI systems; or by both. Any other interested party or their representative organisations (e.g., civil society organisations and academia), in any case, private and non-public sector entities, may also participate in their elaboration.

They may relate to one or more AI systems, considering the similarity of purpose pursued by them. They should pay particular attention to the specific interests and needs of small-scale providers and start-ups (SMEs).

AI systems associated with products that the Regulation does not consider as high risk, and which are therefore not obliged to comply with the requirements set out in the Regulation must nevertheless be safe once placed on the market or put into service. To contribute to this objective, Directive 2001/95/EC of the European Parliament and the Council on general product safety would apply as a safety net.[19]

On the other hand, according to the European Data Protection Board (ECDC, Guidelines 1/2019),[20] a draft code of conduct should include a

---

harm, aggravating existing harm, or harming people. Human dignity and physical and mental integrity are particularly protectable'. Cf. also M. Moreno Rebato, *La propuesta de Reglamento de la Unión Europea sobre inteligencia artificial y las Directrices éticas para una inteligencia artificial fiable: una oportunidad para la Administración Pública Española*, in G. Vestri (ed.), *La disrupción tecnológica en la Administración Pública: retos y desafíos de la inteligencia artificial*, Thomson Reuters-Aranzadi, 2022, pp. 67 ff.

[19] https://eur-lex.europa.eu/legal-content/ES/ALL/?uri=CELEX:32001L0095.

[20] https://www.edpb.europa.eu/sites/default/files/files/file1/edpb_guidelines_201901_v2.0_codesofconduct_en.pdf.

justificatory memorandum describing in detail the purpose of the code, its scope, the supporting documentation; the legal standing of the promoter; the material scope of application, which data processing operations it applies; the territorial scope of application; in the case of transnational codes, the supervisory authority; the supervisory mechanisms; as well as consultations with stakeholders and be in conformity with national law.

1.2.2. *Background on the subject*

Without wishing to be exhaustive, several relevant moments can be pointed out in the path of AI towards ethical control and the development of codes of conduct.

At the international level, several relevant milestones linking ethics and AI can be pointed out. Thus, in 2016-2017, the Institute of Electrical and Electronics Engineers (IEEE) produced the first policy document on ethically aligned design, developing a set of ethical guidelines for designers and developers of AI systems, including consideration of human safety and well-being, transparency and accountability, social inclusion and non-discrimination, fairness and equity, and ethical collaboration.

In 2017, the ACM (*Association for Computing Machinery*) held a conference on AI, ethics and society.[21] Also in January 2017, the so-called Asilomar Principles of Artificial Intelligence,[22] were developed, including, for example, the safety of AI systems throughout their lifetime, transparency in failures, judicial transparency (satisfactory and auditable explanation by a competent human authority), accountability of designers and developers of AI systems, repositories of the moral implications of their use, misuse and actions; the design of AI systems so that their goals and behaviours can be aligned with human values, human dignity, rights, freedoms and cultural diversity; privacy; shared benefit; and shared prosperity.

At the same international level, one of the first institutions to concern itself with these issues was the OECD (Organisation for Economic Co-operation and Development), which approved a Recommendation on Artificial Intelligence in 2019,[23] including principles that promote artificial intelligence

---

[21] https://www.acm.org/code-of-ethics.

[22] https://futureoflife.org/open-letter/ai-principles/.

[23] https://legalinstruments.oecd.org/en/instruments/OECD-LEGAL-0449.

with respect for human rights and democratic values. Indeed, it set out a series of ethical principles for AI, referring to inclusion, human welfare and human rights, transparency and accountability, robustness and security, privacy and data protection, international collaboration, and governance.

The previous year, in 2018, the Toronto Declaration on the Protection of Equal Rights and Non-Discrimination in Machine Learning Systems had been agreed.

At the United Nations level, led by the International Telecommunications Union, a global and inclusive dialogue on AI is also being promoted, especially in connection with the sustainable development goals set out in the 2030 Agenda.

At the European level, the path is directed towards the use of ethics as an inspiration to develop a *made in Europe* brand, although it is true that in the globalised world in which we live, AI knows no borders, and there is intense competition in this area with other large regions and countries in the world, such as the USA or China, where ethical standards are certainly much lower, although this should not slow down European development.[24]

Succinctly, this body states the following: first, AI must serve people and the planet, driving inclusive growth, sustainable development, and well-being. Second, AI systems must be designed to respect the rule of law, human rights, democratic values, and diversity and incorporate appropriate safeguards - for example, allowing for human intervention where necessary - to ensure a just and equitable society. Third, AI systems must be governed by transparency and responsible disclosure to ensure that people know when they are interacting with them and can object to the outcomes of that interaction. Fourth, AI systems must operate robustly, reliably, and securely throughout their lifetime, and potential risks must be assessed and managed at all times. Fifth, organisations and individuals developing, deploying, or managing AI systems should be accountable for their proper functioning in line with the above principles.

[24] HLEG, *Ethics guidelines for trustworthy AI*, 2019 (https://digital-strategy.ec.europa.eu/en/library/ethics-guidelines-trustworthy-ai).

On the risks of applying ethical codes in certain countries, such as the Republic of Korea, and in general on the application of ethical principles to artificial intelligence, A.M. Porcelli, *La inteligencia artificial y la robótica: sus dilemas sociales, éticos y jurídicos*, in *Derecho global. Studies in Law and Justice*, vol. 6, no. 16, 2020: https://www.scielo.org.mx/scielo.php?script=sci_arttext&pid=S2448-51362020000300049. As the author states, 'intelligent systems must be developed in a way that puts people at their centre and thus earns the trust of the public. This implies that

We could point to the European Commission's funding of the *Robolaw*, The *Regulating Emerging Robotic Techologies in Europe: Robotics facing Law and Ethics*,[25] as a basis for the elaboration of a first code of conduct, which sanctions a voluntary ethical code of conduct as a basis for the regulation of the social, environmental and human health impacts of robotics, a code that would determine compliance with ethical standards by researchers, professionals, users and designers.

Along this path, the Declaration on the impact of the development of big data on the protection of individuals with regard to the processing of their personal data in the EU was formulated as early as 2014.[26] AI systems need data to develop, and data availability depends on the legislation regulating its use, which can obviously be more or less permissive. European legislation is understood to be, in this sense, less permissive in relation to the US or China, which will inevitably go hand in hand with higher ethical standards in the European case.

In 2015, the European Data Protection Supervisor produced his Opinion 4/2015,[27] in which he recalls that 'the ethical dimension of data processing needs to be taken into account'. In the same vein, see Opinion 7/2015 or 8/2016, as well as the decision to set up an advisory group on ethics in December 2015.[28]

On 19 April 2016, the European Commission published the document Advancing the Internet of Things in Europe,[29] highlighting the need for legal regulation of digital technologies due to their rapid evolution.

In turn, the European Parliament has the expert group Panel for the Future of Science and Technology, which in 2016 produced several documents on

---

artificial intelligence applications must not only comply with the law, but must also respect ethical principles and ensure that their implementation avoids unintended harm... The European Union is clear that a new ethical and legal framework that responds to the reality of artificial intelligence needs to be complemented and dictated. However, more must be offered than just a list of ethical principles...'.

[25] https://cordis.europa.eu/project/id/289092/reporting.

[26] E. Gil, *Big data, privacidad y protección de datos*, Boletín Oficial del Estado, 2015.

[27] https://www.edps.europa.eu/sites/default/files/publication/15-09-11_data_ethics_en.pdf.

[28] https://www.edps.europa.eu/data-protection/our-work/publications/ethical-framework/ethics-advisory-group_en?etrans=en.

[29] https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX%3A52016SC0110.

ethical, legal, and political aspects and impacts of AI in various areas. The European Commission is also advised by the *European Group on Ethics in Science and New Technologies* (EGE),[30] which has published several documents on ethical principles and guidelines. For example, in 2018, the Declaration on Artificial Intelligence, Robotics, and Autonomous Systems, which reflects on a series of ethical principles; the Communication on Artificial Intelligence for Europe (2018), which sets out a roadmap for drawing up ethical guidelines; and the Coordinated Plan on Artificial Intelligence (2018), which emphasises AI *made in Europe*. Also, in 2017, the European Parliament passed several resolutions on robotics[31] (including an annex calling for a code of conduct for robotics engineers,[32] a code of ethics for research ethics committees or on big data,[33] a code of ethics for research ethics committees or on big data).

---

[30] https://research-and-innovation.ec.europa.eu/strategy/support-policy-making/scientific-support-eu-policies/european-group-ethics_en.

[31] European Parliament resolution of 16 February 2017 with recommendations to the Commission on Civil Law Rules on Robotics (2015/2103(INL)). This text states that 'Robotics research activities should be conducted in accordance with the precautionary principle, anticipating potential safety impacts of outcomes and taking due precautions, proportional to the level of protection, while encouraging progress for the benefit of society and the environment'.

[32] A few years earlier, in 2008, the European Commission issued the *Recommendation of 7 February 2008 on a code of conduct for responsible nanosciences and nanotechnologies research* (Document 32008H0345), a code that will be voluntary for the interested parties, although an express call is made to the Member States to encourage its voluntary adoption, and the aim is to use it in commercial relations on an international scale. It also responds to the safe, ethical and effective framework we have been talking about.

[33] European Parliament resolution of 14 March 2017 on fundamental rights implications of big data: privacy, data protection, non-discrimination, security and law-enforcement (2016/2225(INI)). In this regard, see also A. Boix Palop, *Los algoritmos son reglamentos: la necesidad de extender las garantías propias de las normas reglamentarias a los programas empleados por la Administración para la adopción de decisiones*, in *Revista de Derecho Público: Teoría y Método*, no. 1, 2020, pp. 223-270; A. Cerrillo i Martínez, *El impacto de la inteligencia artificial en el Derecho administrativo. ¿Nuevos conceptos para nuevas realidades técnicas?*, in *Revista General de Derecho Administrativo*, no. 50, 2019; o L. Cotino Hueso, *Riesgos e impactos del big data, la inteligencia artificial y la robótica. Enfoques, modelos y principios de la respuesta del Derecho*, in *Revista General de Derecho Administrativo*, no. 50, 2019.

In the same year, the European Council called for awareness of the urgency of addressing new trends, including issues such as artificial intelligence, while ensuring a high level of data protection, as well as digital rights and ethical standards.

In December 2018, the European Economic and Social Committee adopted *the Opinion on 'Trust, privacy and security for consumers and businesses in the internet of Things (IoT)'*,[34] highlighting the need for legal regulation, as well as the risks affecting security, privacy, intimacy and the ethical or contrary to fundamental human rights nature of this technology.

Two years earlier, the General Data Protection Regulation (GDPR) had been published. Already, this Regulation referred to codes of conduct as guidelines for the implementation of appropriate measures demonstrating the controller's or processor's compliance with its obligations, in particular with regard to the identification of the risk related to the processing, its assessment in terms of its source, nature, likelihood, and severity and the identification of good practices to mitigate the risk. Such codes of conduct could, in particular, set out the obligations of controllers and processors, taking into account the likely risk to the rights and freedoms of natural persons arising from the processing. Adherence or non-adherence to such codes of conduct could even be considered as an attenuating or aggravating circumstance for the purpose of determining the level of fines for administrative offences in this area. Apart from the fact that such adherence can also be used as an element to demonstrate the controller's compliance with its obligations and will be taken into account when carrying out data protection impact assessments. They would also provide a guarantee in the framework of the transfer of personal data to third countries or international organisations.

These codes of conduct referred to in the GDPR would enable control of the compliance of data controllers or processors with their obligations. They must be approved by the supervisory authority, which will give its opinion on whether the draft code submitted to it, as well as a possible modification or extension, is in compliance with the Regulation and will approve it if it considers the guarantees offered by the document to be sufficient. If the Committee's opinion is positive, it will submit it to the Commission, which may decide employing 'implementing acts' that the adopted code of conduct shall have general validity within the Union. It will then be publicised and filed

---

[34] https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:52018IE1038.

in a register kept by the Committee. A posteriori, and without prejudice to the functions and powers of the supervisory authority, there may be Bodies accredited by the supervisory authority itself to supervise a posteriori compliance with the codes of conduct. This same supervisory authority, within its respective territory, shall encourage the elaboration of codes of conduct, elaborate and publish the requirements for the accreditation of the supervisory bodies of these codes, and carry out the accreditation of these bodies.

After stopping briefly at the codes of conduct referred to in the GDPR and continuing with the list of relevant milestones, the *White Paper on Artificial Intelligence*,[35] was published on 19 February 2020, with a twofold objective. On the one hand, to promote the adoption of AI. On the other hand, what is of more interest to us at the moment is addressing the risks associated with specific uses of this technology. AI must be reliable and safe; it must respect fundamental rights, and the ethics applied to AI. As stated in this text, 'Artificial Intelligence is developing fast. It will change our lives by improving healthcare (e.g. making diagnosis more precise, enabling better prevention of diseases), increasing the efficiency of farming, contributing to climate change mitigation and adaptation, improving the efficiency of production systems through predictive maintenance, increasing the security of Europeans, and in many other ways that we can only begin to imagine. At the same time, Artificial Intelligence (AI) entails a number of potential risks, such as opaque decision-making, gender-based or other kinds of discrimination, intrusion in our private lives or being used for criminal purposes'. And it is in avoiding these risks that the EU and the member states must act, both from a regulatory point of view and in the elaboration of these codes of conduct, that we are analysing through the application of ethical principles.

In the same year, the European Parliament also did a great deal of work in this area, passing a series of resolutions in October on issues such as ethics. These included the *European Parliament resolution of 20 October 2020 with recommendations to the Commission on a framework of ethical aspects of artificial intelligence, robotics, and related technologies* (2020/2012(INL)),[36] in which it specifically

---

[35] European Commission, *White Paper On Artificial Intelligence - A European approach to excellence and trust*, COM (2020) 65 final. This text identifies options to achieve the dual objective of promoting the adoption of AI and addressing the risks linked to specific uses of this new technology on EU values and fundamental rights. The *White Paper* follows HLEG *Ethics guidelines for trustworthy AI* (2019).

[36] https://www.europarl.europa.eu/doceo/document/TA-9-2020-0275_EN.html.

recommends that the Commission propose legislative measures to harness the opportunities and benefits of AI while ensuring that ethical principles are protected. The Resolution includes the text for a legislative proposal for a Regulation on ethical principles for the development, deployment, and use of artificial intelligence, robotics, and related technologies, which would later become the *Proposal for a Regulation laying down harmonised rules on Artificial Intelligence* (European Commission). This proposal was subject to an impact assessment examined by its Regulatory Scrutiny Committee, which issued a favourable report on 21 March 2021. Four options with different degrees of regulatory intervention were assessed. One of them, which is relevant to the issue at hand, is Option 3+, which already refers to *codes of conduct* applicable to *non-high risk* AI systems. This option would contribute to reducing the risk that fundamental rights are violated[37] or that the security of individuals could be endangered.

Moving to the national level, the Spanish government also acted along these lines by approving in November 2020 the National Artificial Intelligence Strategy,[38] in an attempt to lead the development and integration of AI into the productive fabric, the economy, and society, emphasising the implications of AI with ethics and morality, as it states on page 4 as follows: '... the central issue underlying the development and spread of AI is the ethical question. The widespread use and management of data through the actions of algorithms and autonomous systems have multiple ethical and moral implications that require

---

[37] In this regard, see J.F. Sánchez Barrilao, *Derecho constitucional, desarrollo informático e inteligencia artificial: aproximación a la propuesta del Parlamento Europeo a favor de una regulación sobre robótica*, in J. Valls Prieto (ed.), *Retos jurídicos por la sociedad digital*, 2018, Aranzadi, pp. 21 ff.; M.A. Presno Linera, *Derechos fundamentales e inteligencia artificial en el Estado social, democrático y digital de Derecho*, in *El Cronista del Estado Social y Democrático de Derecho*, no. 100, 2022, pp. 48 ff.; D. García San José, *La Europa de los derechos ante los avances científicos y tecnológicos*, Tirant lo Blanch, 2021.

[38] https://portal.mineco.gob.es/es-es/digitalizacionIA/Paginas/ENIA.aspx. The promotion of Artificial Intelligence constitutes one of the main elements of the *Agenda España Digital 2025*, presented in July 2020, in its line of action 9 on 'Data Economy and Artificial Intelligence', based on the work already started in the context of the Agenda for Change, presented in February 2019, the support programmes for Digital Enabling Technologies and the Spanish Strategy for R&D&I in Artificial Intelligence presented in March 2019. It is also in line with the commitment made with the rest of the EU countries to position the EU as a world leader in this area.

processes and control mechanisms that protect our values, principles, and rights. The drafting of a Charter of Digital Rights and the creation of AI ethics monitoring mechanisms are two of the initiatives included in this Strategy in order to reach a consensus on an appropriate framework for this technological development. The 6th strategic axis (p. 20) is dedicated to the establishment of an ethical and regulatory framework that reinforces the protection of individual and collective rights in order to guarantee social inclusion and welfare based on a series of ethical principles: inclusion or non-discrimination; social welfare; sustainability and ecological responsibility. Along these lines (although with a relatively low level of concreteness) is the so-called Charter of Digital Rights, promoted by this document, although it lacks any legal value,[39] or the attempt

---

[39] The Charter would be an example of a soft law instrument, like the codes of conduct we are analysing, although it can serve to interpret the rules that make up the legal system, as stated by C.I. Velasco Rico, in the paper presented at the XVIII Congress of the Spanish Association of Administrative Law Professors on *Administrative Law in the Age of* Artificial *Intelligence* , held in Vigo in January 2024, entitled *Marco regulatorio de los sistemas algorítmicos y de inteligencia artificial: el papel de la Administración (Regulatory framework of algorithmic and artificial intelligence systems: the role of the Administration*). On the other hand, M.A. Presno Linera, op. cit, pp. 50-51, gives as an example of precisely the opposite, of a Charter with *legal value*, Law number 27/2021 of 17 May, of the *Portuguese Charter of Human Rights in the Digital Age*, 'which includes, in addition to the protection of classic rights, such as the freedoms of expression, demonstration, association or participation, in the digital world, and of recognising recent rights, such as the right to be forgotten and protection against abusive geolocation, the use of artificial intelligence and robots: 1. The use of artificial intelligence shall be guided by respect for fundamental rights, ensuring a fair balance between the principles of explainability, security, transparency, and accountability, taking into account the circumstances of each individual case and establishing processes to avoid bias and discrimination. 2. Decisions taken by means of algorithms that have a significant impact on the addressees shall be communicated to the interested parties, shall be subject to appeal and shall be auditable under the terms provided for by law. The principles of beneficence, non-maleficence, respect for human autonomy and justice, as well as the principles and values enshrined in Article 2 of the Treaty on European Union, namely non-discrimination and tolerance, shall apply to the creation and use of robots (Article 9)'. On the Charter, see also L. Cotino Hueso (ed.), *La Carta de Derechos Digitales*, Valencia, Tirant Lo Blanch, 2022; S. de la Sierra Morón, in the same work, *Una introducción a la Carta de Derechos Digitales*, pp. 27 ff. The Charter is accessible at the

to promote national and international forums for dialogue, awareness and participation in AI aimed at fostering dialogue between government, science, social partners, the private sector and civil society.

1.2.3. *Principles to which codes of conduct should be subject*

The European AI Regulation suggests the development of ethical codes of conduct[40] for providers of non-high-risk AI systems. These codes of conduct[41] should address ethical principles[42] and ensure compliance with the standards set out in the regulation itself, which, at the same time, encourages all stakeholders, including industry, academia, civil society, and standardisation organisations, to take these ethical principles into account in the development of voluntary standards and best practices. To ensure the effectiveness of these voluntary codes of conduct, they should be based on clear objectives and key performance indicators against which the achievement of these objectives can be measured.

Although we can also understand the statement by Boix Palop (2022), when he concludes that, curiously, AI systems are subject to greater control

---

following link: https://www.lamoncloa.gob.es/presidente/actividades/Documents/2021/140721-Carta_Derechos_Digitales_RedEs.pdf.

[40] On the relationship between ethics and codes of conduct, see L. Cotino Hueso, *Ética en el diseño para el desarrollo de una inteligencia artificial, robótica y big data confiables y su utilidad desde el derecho*, in *Revista Catalana de Dret Públic*, no. 58, 2019, pp. 29 ff. The ethics application would be oriented towards answering questions such as the following: What risks, dangers, and impacts could these new technologies have? What risks, errors, discriminations? Could the AI be strong enough to attack human beings? Could we consider it a threat? Would it be capable of violating human controls? Can the AI be manipulated? With absolute respect for the principle of transparency, can AI have a conscience, can there be artificial moral agents, and can responsibility and rights be declared for non-humans?

[41] On which the European Committee on Artificial Intelligence, in its role of providing advice and assistance to the Commission and the Member States to facilitate the consistent and effective implementation of the AI Regulation, may issue written recommendations and opinions.

[42] A. Boix Palop y L. Cotino Hueso (ed.), *Derecho Público, derechos y transparencia ante el uso de algoritmos, inteligencia artificial y big data,* in *Revista General de Derecho Administrativo*, no. 50, 2019; F. Longo, *Administración pública con valores. Instrumentos para una gobernanza ética*, INAP, 2015.

and are required to comply with more rigorous principles when they are used by private economic agents than by public administrations,[43] emphasising, for example, the principle of transparency, which we will discuss below, although it is not exclusive to our legal system.

However, what principles could fit into these instruments? After analysing the documents referred to in the previous pages, I will draw up a list of principles that I believe should be taken into account when drawing up these codes of conduct:

a. *Transparency and information communication to users:* Providers should ensure transparency in developing, deploying, and operating their AI systems. AI systems should be traceable and explainable[44] and make individuals aware that they are communicating or interacting with an AI system, and those responsible for deployment should be informed about the capabilities and limitations of the AI system and individuals about their rights. This could also include providing clear information so that users can understand how the system works and how decisions are made. The information provided has to be sufficient for users to correctly interpret and use the output information, even if this may cause conflicts of interest (e.g., between the administration using an AI system and the contractor that created it[45]), or behavioural

---

[43] A. Boix Palop, *Transparencia en la utilización de inteligencia artificial por parte de la Administración*, in *El Estado Social y Democrático de Derecho*, no. 100, 2022, pp. 90 ff.

[44] In this regard, cfr. A. Cerrillo i Martínez, *El impacto de la inteligencia artificial en el derecho administrativo, ¿nuevos conceptos para nuevas realidades técnicas?*, in *Revista General de Derecho Administrativo*, no. 50, 2019.

[45] A. Huergo Lora, op. cit.: 'the fact that the algorithmic model has been elaborated by a company (contractor of the Administration) and is protected as a business secret does not authorise the Administration to deny access to the information. If it does not do so, i.e. if it then finds itself in a conflict between compliance with the contract and respect for the rights of the interested party in the procedure, its obligation is clearly to respect the latter, even if this entails, where appropriate, compensation costs against the contractor for breach of contract'. Quoting S. Civitarese Mateucci, *Umano troppo* umano, page 28, 'it is quite evident that, if the administration acquires a service in order to make use of it in its own activity of issuing decisions, the question of intellectual property loses importance'. Although by a very lateral connection with this procurement issue, vid. A.D. Berning Prieto, *El uso de sistemas basados en inteligencia artificial por las Administraciones Públicas: estado actual de la cuestión y algunas propuestas ad futurum para un uso responsable*, in

modifications of the recipients of the algorithms used by such an AI system, because if citizens know the criteria used by the algorithm they may alter their behaviour to try to cheat the algorithm, adjusting it more to the criteria of the algorithms than to the objectives that the applicable rules are trying to achieve. However, we can also point to a problem in a possible lack of transparency, due to the leadership of the private sector in the development of applications based on the use of algorithms.[46]

In addition, algorithms can be controllable by judges and courts. Until now, in the absence of EU regulation, the transparency, auditability, or neutrality of artificial intelligence systems has only been addressed in a very generic way in Article 23 of Law 15/2022 of 12 July on equal treatment and non-discrimination. This article is no more than a mere declaration of intent in determining that public administrations will encourage algorithms involved in decision-making to 'take into account criteria of minimisation of bias, transparency, and accountability, whenever technically feasible' or that they will 'prioritise transparency in design', as well as 'promote the use of an ethical and reliable Artificial Intelligence that respects fundamental rights'.

In addition, we may find that specific AI systems intended to interact with natural persons or to generate content may carry specific risks of impersonation or counterfeiting, regardless of whether they are classified as high risk or not. Therefore, the use of such systems should, in certain circumstances, be subject to specific transparency obligations without prejudice to the requirements and obligations applicable to high-risk AI systems. In particular, natural persons should be notified that they are interacting with an AI system unless it is apparent from the circumstances and context of use and informed when they are exposed to an emotion recognition system or a biometric profiling system. This information and these

---

Revista de Estudios de la Administración Local y Autonómica, no. 20, 2023, pp. 179-181; I. Gallego Córcoles, *La contratación de soluciones de inteligencia artificial*, in E. Gamero Casado (ed.), in *Inteligencia artificial y sector público: retos, límites y medios*, 2023, Tirant Lo Blanch, pp. 503 ff.

[46] C. Ramió, *Inteligencia Artificial y Administración Pública. Robots y humanos compartiendo el servicio público*, Catarata, 2018; C. Velasco Rico, *Vigilando al algoritmo. Propuestas organizativas para garantizar la transparencia*, in Puentes Cociña/Quintiá Pastrana (eds.), *El Derecho ante la transformación digital: oportunidades, riesgos y garantías*, Atelier, 2019, pp. 73 ss.; I. Martín Delgado, *La aplicación del principio de transparencia a la actividad administrativa algorítmica*, in E. Gamero Casado (ed.), *Inteligencia artificial y sector público: retos, límites y medios*, Tirant Lo Blanch, 2023, pp. 131 ff.

notifications should be provided in formats accessible to persons with disabilities. In addition, users who use an AI system to generate or manipulate images, audio files, or videos that bear a striking resemblance to real people, places, or events and could mislead a person into thinking that they are authentic should disclose that they have been artificially created or manipulated by labelling the AI-generated content accordingly and indicating its artificial origin.

In addition, providers shall ensure that AI systems intended to interact with natural persons are designed and developed in such a way that such persons are informed that they are interacting with an AI system, except in situations where this is obvious from the circumstances and context of use, as stated in Art. 50 'where it is obvious from the point of view of a reasonably well-informed, observant and circumspect natural person'. This obligation does not apply to AI systems authorised by law for the purpose of detection, prevention, investigation or prosecution of criminal offences, unless such systems are available to the public to report a criminal offence.

According to Article 50 (7) AIA, 'The AI Office shall encourage and facilitate the development of codes of good practice at Union level to promote the effective implementation of obligations relating to the detection and labelling of artificially generated or artificially manipulated content. The Commission shall be empowered to adopt implementing acts to adopt such codes of practice... If it considers that the code is inadequate, the Commission may adopt an implementing act specifying common rules for enforcing those obligations ...'. These codes of practice are again referred to in Article 56 of the Act, which states that they should be drawn up in order to contribute to the correct application of the standard, also taking into account international approaches. These codes should ensure up-to-date information concerning market and technological developments; contain a summary of the content used for training; identify the type and nature of systemic risks at the Union level, including, if possible, their origin; measures, procedures, and modalities of assessment and management of these same risks. In addition to providers of general-purpose AI models, national authorities, civil society organisations, industry, academia, and other interested parties, as well as downstream providers and independent experts, may be involved in developing these Codes of Practice. These Codes will be subject to regular monitoring and evaluation by the AI Office and the Committee and should be finalised by 2 May 2025 at the latest, and the Commission itself may adopt a code of practice and make it generally valid within the Union by means of an implementing act, which may also establish standard rules for the fulfilment of the obligations of providers of general purpose and general purpose AI models with systemic risk (in accordance with Articles 53 and 55 of the Regulation).

b. *Literacy:* AI systems must equip providers, deployers, and affected persons with the necessary concepts to make informed decisions to ensure proper compliance and correct implementation. These literacy measures, together with the introduction of appropriate follow-up actions, could help to improve working conditions and consolidate the innovation path of reliable AI in the EU. The EESC will support the Commission in promoting AI literacy tools, public awareness, and understanding of the benefits, risks, safeguards, rights, and obligations in relation to the use of AI systems. There will also have to be codes of conduct promoting this literacy among those involved in developing, managing, and using AI. In this regard, Article 4 of the text of the AI Regulation adopted by the European Parliament states that 'Providers and deployers of AI systems shall take measures to ensure, to their best extent, a sufficient level of AI literacy of their staff and other persons dealing with the operation and use of AI systems on their behalf, taking into account their technical knowledge, experience, education and training and the context the AI systems are to be used in, and considering the persons or groups of persons on whom the AI systems are to be used'.[47]

c. *Diversity and non-discrimination:* Codes of conduct should ensure that AI systems do not perpetuate or amplify existing discrimination. Providers should work to avoid unfair bias and discrimination based on protected characteristics such as gender, race, religion, etc. That is, artificial intelligence systems should not unfairly discriminate against individuals or groups on the basis of these characteristics. They must promote equal access, gender equality, and cultural diversity while avoiding discriminatory effects and unfair bias prohibited by national or EU law. They also need to be developed in an inclusive manner,

---

[47] Along the same lines, Strategic Axis 2 of the National Artificial Intelligence Strategy, entitled *Promoting the development of digital skills, boosting national talent and attracting global talent*, points out that 'it is necessary to increase the technical training in AI of the working population, both the user public and specialists, in order to facilitate access to new quality jobs and meet the challenges of the labour market of the future...' (p. 18). This literacy would, according to this document, involve promoting education and training in digital skills for the entire population throughout their lives, with an ethical, humanist and gender perspective. On literacy, in this case, of legal professionals, M. Barrio Andrés et al., *La visión de los expertos*, in *Informe Break the limits*, Aranzadi-La Ley, 2023.

involving relevant stakeholders such as business and civil society organisations, academia, research bodies, trade unions, and consumer protection organisations. At the same time, in order to know whether or not discrimination exists, it is essential to be able to access the source code of programmes and to have access to the algorithms that configure them.[48]

d. *Beneficence.* AI should be developed for the common good and benefit of humanity, improve individual and collective well-being, generate prosperity, wealth, sustainability, increase citizen empowerment, equitably distribute economic, social and political opportunities, protect the democratic process and the rule of law; provide low-cost, high-quality common goods and services; data literacy and representativeness; achieve sustainable development goals.

e. *Non-maleficence.* AI has a preventive functionality; it must not do physical, psychological, financial, or social harm. This applies both when developing and configuring AI as well as for its outcomes. It alludes to the criteria of precaution and active responsibility. The assets that would be at risk if these principles were not complied with would be fundamental rights,[49] the democratic process, minorities, or the environment.

f. *Fairness.* The use of AI for distributing resources and equitable access to resources must be achieved; discrimination of any kind must be eliminated, and bias and stigmatisation must be avoided. Shared benefits and shared prosperity would be generated. However, this implies high accountability standards and effective redress or remedy if harm occurs.

g. *Privacy and data protection:* AI providers are expected to respect the privacy and data protection of users. They should implement measures to ensure the security and confidentiality of information used by their systems. AI systems should be developed and used in accordance with privacy and data protection standards insofar as they process data that meet high standards in terms of quality and integrity. It is also true that, while data protection regulations

---

[48] M. Barrio Andrés, *Luces y sombras del Estado algorítmico de Derecho*, en *Derecho Digital e Innovación*, no. 5, 2020; J. Jiménez López, *Oscuridad algorítmica en el sector público*, in G. Vestri (ed.), *La disrupción tecnológica en la Administración Pública. Retos y desafíos de la inteligencia artificial*, Thomson Reuters-Aranzadi, 2021, pp. 41 ff.

[49] R. Martínez Martínez, *Inteligencia artificial, derecho y derechos fundamentales*, in T. de la Cuadra Salcedo and J.L. Piñar Mañas (eds.), *Sociedad digital y derecho,* BOE, 2018, pp. 259 ff., has expressed that AI can be a space open to utopia but also a gateway to a dystopian world.

provide rules for the use of personal data, algorithmic predictions are mostly made based on non-personal data; it is not necessary to know the name and surname of the target subjects, but only the notes that may be relevant. Correlations are sought from non-personal data (anonymisation) and profiles are created.

In connection with this data protection, insofar as data may be subject to cross-border exchange, the Commission may formulate initiatives, also on a sectoral basis, aimed at facilitating the reduction of technical obstacles to such data exchange for the development of AI, including in relation to the data access infrastructure and the semantic and technical interoperability of different types of data.

h. *Quality and accuracy:* AI systems must be accurate and reliable. Providers should strive to ensure the quality of their models and algorithms, avoiding inaccurate or biased results.

i. *Collaboration with authorities and risk assessment:* Providers should collaborate with regulatory authorities and participate in assessing risks associated with their AI systems, especially those considered high risk.

j. *Training and competence:* Suppliers should ensure that those involved in developing and deploying AI systems have adequate training and competence. This can help reduce risks associated with misuse of the technology.

k. *Responsibility:* Developers and users of artificial intelligence systems should take responsibility for the impact of their actions and decisions and take steps to mitigate any potential harm. They should know how to prevent self-learning systems from training with malicious data.

l. *Human oversight:* AI systems shall be designed and developed in such a way that they can be effectively monitored by natural persons during the period they are in use in order to prevent or minimise risks to health, safety, or fundamental rights. AI systems must be developed and used as tools in the service of individuals, with respect for human dignity and personal autonomy, and must be operated in such a way that they can be adequately controlled and monitored by human beings (Art. 14 AIA). [50] Vigilance necessarily implies that those who

---

[50] At other levels, this human surveillance is already regulated in Article 22 of the General Data Protection Regulation, Article 41 of Spanish Law 40/2015, of 1 October, on the Legal Regime of the Public Sector, or in Article 96 of Spanish Law 58/2003, of 17 December, on General Taxation, albeit somewhat sparingly. The Spanish Charter of Digital Rights also refers to the issue in Article XXV, paragraphs 2 and 3, in relation to the exercise of both regulated and discretionary administrative

supervise must have the precise knowledge and skills, including the necessary literacy, to identify the biases of AI systems and their own biases and be able to reverse them. Human intervention must be effective in the sense that it can interfere with the functioning of the AI system and even stop it if necessary.

With AI, we voluntarily cede part of our decision-making power to machines, so human autonomy should be promoted; 'humans should always decide what decisions to make'. There has to be human supervision of machines.[51] Human beings must be able to deviate from the decisions made by machines. However, it can also be dangerous because human beings get used to the routine, and it can be easier for them to let the machine decide for them and let themselves go.

m. *Accuracy, robustness, cybersecurity*: AI systems shall be resilient to errors, failures, and inconsistencies that may arise in the AI systems themselves or the environment in which they operate, mainly due to their interaction with humans or other systems. AI systems should be developed and operated in

---

powers. Likewise, one can consult what has been called the first regulation of AI in Spanish law, article 23 of Law 15/2022, of 12 July, on equal treatment and non-discrimination, although it only contains a generic reference to *accountability*: it already mentions some of the fundamental principles we have been talking about (transparency, avoidance of discrimination, application of ethical rules and respect for fundamental rights).

It is also true that Article 14 AIA is among the provisions regulating the principles applicable to *high-risk* AI systems, but Article 95 AIA, which is the subject of this commentary, refers precisely to the application of the same principles.

[51] J. Ponce Solé, *Inteligencia artificial, Derecho administrativo y reserva de humanidad: algoritmos y procedimiento administrativo debido tecnológico*, in *Revista General de Derecho Administrativo*, no. 50, 2019; by the same author, *Reserva de humanidad y supervisión humana de la Inteligencia artificial*, in *El Cronista del Estado Social y Democrático de Derecho*, no. 100, 2022, pp. 58 ff. ; also by the same author, *Seres humanos e inteligencia artificial: discrecionalidad artificial, reserva de humanidad y supervisión humana*, in E. Gamero Casado (ed.), *Inteligencia artificial y sector público: retos, límites y medios*, Tirant Lo Blanch, 2023, pp. 195 ff.; J.A. Plaza, *Lecciones de ética para máquinas que piensan y toman decisiones*, in *El País: Retina*: https://elpais.com/retina/2017/12/19/innovacion/1513661054_305253.html. Also see Y. de la Cueva, *¿Quién vigila al algoritmo*, in *El Notario del siglo XXI,* no. 87, 2019(https://www.elnotario.es/opinion/opinion/9637-quien-vigila-al-algoritmo); and G. Vestri, *La inteligencia artificial ante el desafío de la transparencia algorítmica: una aproximación desde la perspectiva jurídico-administrativa*, in *Revista Aragonesa de Administración Pública*, no. 56, 2021, pp. 368 ff.

such a way as to be robust in the event of problems and resilient to attempts to alter their use or operation to enable their unlawful use by third parties and minimise unintended harm (Article 15 AIA). To address the technical aspects of how to measure appropriate levels of accuracy and robustness as well as any other relevant performance parameters, the Commission, in cooperation with relevant stakeholders and organisations, such as metrology and benchmarking authorities, shall, as appropriate, encourage the development of benchmarks and measurement methodologies.

Technical solutions to address potential AI-specific vulnerabilities shall include measures to prevent, detect, combat, resolve, and control attacks that attempt to manipulate the training data set (data poisoning), or pre-trained components used in training (model poisoning), input information designed to cause the model to make a mistake (adversarial examples or model evasion), confidentiality attacks or model flaws.

n. *Social and environmental well-being*: AI systems must be developed and used in a sustainable and environmentally sound manner for the benefit of all human beings while monitoring and assessing the long-term effects on people, society, and democracy.

o. *Other principles:* democracy,[52] dignity, non-discrimination, equality, justice, human freedom, privacy, education, access to information, robustness, reliability, reproducibility (i.e., that results are consistent across different

---

[52] Couldn't the algorithms used by AI systems break democracy? Couldn't a kind of indoctrination be found through the internet? This could be deduced from reading the following text: 'Algorithmic predictions, even if they are or could be used in the activity of public administrations, have been developed mainly because they are useful for the private sector. Thus, for example, a company (or even a political party) that contracts an advertising space on Facebook, not only gets (unlike what happens when a space is contracted on a billboard, or in a print media, or radio) that the ad is seen only by some users (those selected by the algorithmic model), but it can get that not all these users see the same ad, but that it is adapted to their profile. For example, users with more environmentalist tendencies will see an advertisement of the party focused on that part of its programme, while others will receive a message more adapted to their preferences'; cfr. A. Huergo Lora, *Una aproximación a los algoritmos desde el Derecho Administrativo*, in A. Huergo Lora (ed.), *La regulación de los algoritmos*, Aranzadi, 2020, p. 48. By the same author and in the same sense, *Administraciones Públicas e Inteligencia Artifical: ¿más o menos discrecionalidad?* in *El Cronista del Estado Social y Democrático de Derecho*, no. 96-97, 2021, pp. 78 ff.

situations, computational frameworks and input data), provision for fallback plans in case of problems (that change, e.g., procedure or need human operator action); intuition (to adapt to human intentions and to be able to cooperate with the human being); intelligibility (the person sharing space with the machine has to understand what the machine is going to do); adaptability (to the human person with whom it cooperates); fitness for purpose.

1.2.4. *A model of conduct: the Hiroshima Code*

As we have already mentioned in previous pages, the United States and Europe proposed, prior to the approval of the European Regulation on Artificial Intelligence, to present a voluntary code of conduct for companies developing this type of technology,[53] which they tried to get companies from other countries such as Canada, the United Kingdom, Japan, and India to join, and which could provide a *regulatory bridge* until the practical application of this EU regulation, thereby giving the public confidence. This initiative is part of international discussions that have been ongoing for years on the surveillance of artificial intelligence within the OECD, the Global Partnership on Artificial Intelligence, the EU-US Council on Trade and Technology, as well as EU digital associations. This negotiation at the international level, in which the EU also participates, is consistent with the legally binding rules coming out of the EU and the recently adopted Artificial Intelligence Regulation.

In the same perspective, and after arduous negotiations, the group of countries that make up the G7 (United States, Japan, Germany, United

---

[53] We refer to what has been said in previous pages of this text. On this code and the previous meetings that gave rise to it, see T. de la Quadra Salcedo Fernández del Castillo, *Inteligencia artificial, Administraciones Públicas y Derecho. Una visión comparada de un Derecho en construcción*, in *XVIII Congreso de la Asociación Española de Profesores de Derecho Administrativo*, on *El Derecho Administrativo en la era de la Inteligencia Artificial*. https://www.aepda.es/AEPDAEntrada-3987-XVIII-CONGRESO-DE-LA-ASOCIACION-ESPANOLA-DE-PROFESORES-DE-DERECHO-ADMINISTRATIVO.aspx, who points out other milestones that can be related to this Code, such as the UNESCO Recommendation of 2021 on the ethics of artificial intelligence, or the document of the Chinese Ministry of Science and Technology on Ethical Standards for the new AI Generation, as well as the Ibero-American Charter of Principles and Rights of Digital Environments, of 2023.

Kingdom, France, Italy, and Canada) announced at their meeting on 30 October 2023 the creation of an international code of conduct to minimise the risks of artificial intelligence within the framework of the Hiroshima artificial intelligence process. This code will ask large companies in the sector to commit to taking measures to curb the huge social impact of this emerging technology, to create risk management systems, and invest in cybersecurity controls. As we will discuss in more detail, other rules include transparency, privacy, and the protection of intellectual property.

The purpose of this code,[54] together with further approved principles, will be to complement existing international legislation in the field of AI. Its main objective will be to promote safe and trustworthy AI worldwide, but it will only provide voluntary guidance for the actions of organisations developing state-of-the-art AI systems, including basic models and generative AI systems.

The organisations, which will base their approach on risk, can be from academia, civil society, the private and/or the public sectors.

The text sets out the agreed principles in eleven sections, although the list of actions is not exhaustive, it is a living document (revisable and updatable as necessary) and applicable to all stages of the lifecycle to cover, as and when applicable, the design, development, deployment, and use of advanced AI systems.

As we have pointed out, compliance with the approved code will be voluntary and complementary to international regulations, and the organisations that develop these AI systems should, at the same time, develop effective supervision mechanisms for their systems, self-evaluation mechanisms, to achieve respect for the rule of law, human rights, procedural guarantees, diversity, equity and non-discrimination, democracy and human

---

[54] The text is available at: https://digital-strategy.ec.europa.eu/es/library/hiroshima-process-international-code-conduct-advanced-ai-systems. Although they are not exactly the same thing, 'codes of conduct' can be found in other areas, although not directly related to Artificial Intelligence, but as a set of principles to be taken into account, for example the Google Supplier Code of Conduct(https://about.google/intl/ALL_en/supplier-code-of-conduct/) or the Codes of Good Governance (for example, that of the Comisión Nacional del Mercado de Valores, relating to the good governance of listed companies: https://acrobat.adobe.com/id/urn:aaid:sc:EU:2ec0f4f6-5cc8-4b49-81f6-63661679beeb?viewer%21megaVerb=group-discover).

protagonism in the design, development, and implementation of advanced AI systems.

The principles or measures that these organisations should accept in the approved code are grouped into eleven lines of action which, as we have said, are neither closed nor, much less, unchangeable:

1. Take appropriate measures throughout the development of advanced AI systems, including prior to and throughout their deployment and placement on the market, to identify, evaluate, and mitigate risks across the AI lifecycle. This includes employing diverse internal and independent external testing measures [...]. Testing and mitigation measures, should, for example, seek to ensure the trustworthiness, safety and security of systems throughout their entire lifecycle so that they do not pose unreasonable risks. In support of such testing, developers should seek to enable traceability, in relation to datasets, processes, and decisions made during system development. These measures should be documented and supported by regularly updated technical documentation. Attention should be devoted to the following risks as appropriate: chemical, biological, radiological, and nuclear risks; Offensive cyber capabilities; Risks to health and/or Safety; Risks from models of making copies of themselves or 'self-replicating' or training other models; Societal risks, as well as risks to individuals and communities; Threats to democratic values and human rights; or Risk that a particular event could lead to a chain reaction with considerable negative effects that could affect up to an entire city, an entire domain activity or an entire community.

2. Identify and mitigate vulnerabilities, and, where appropriate, incidents and patterns of misuse, after deployment including placement on the market. Organizations should use, as and when appropriate commensurate to the level of risk, AI systems as intended and monitor for vulnerabilities, incidents, emerging risks and misuse after deployment, and take appropriate action to address these. Organizations are encouraged to consider, for example, facilitating third-party and user discovery and reporting of issues and vulnerabilities after deployment such as through bounty systems, contests, or prizes to incentivize the responsible disclosure of weaknesses. Organizations are further encouraged to maintain appropriate documentation of reported incidents and to mitigate the identified risks and vulnerabilities, in collaboration with other stakeholders. Mechanisms to report vulnerabilities, where appropriate, should be accessible to a diverse set of stakeholders.

3. Publicly report advanced AI systems' capabilities, limitations and domains of appropriate and inappropriate use, to support ensuring sufficient transparency, thereby contributing to increase accountability. Organizations should make the information in the transparency reports sufficiently clear and understandable to enable deployers and users as appropriate and relevant to interpret the model/system's output and to enable users to use it appropriately.

4. Work towards responsible information sharing and reporting of incidents among organizations developing advanced AI systems including with industry, governments, civil society, and academia.

5. Develop, implement and disclose AI governance and risk management policies, grounded in a risk-based approach – including privacy policies, and mitigation measures. Organizations should establish policies, procedures, and training to ensure that staff are familiar with their duties and the organization's risk management practices.

6. Invest in and implement robust security controls, including physical security, cybersecurity and insider threat safeguards across the AI lifecycle.

7. Develop and deploy reliable content authentication and provenance mechanisms, where technically feasible, such as watermarking or other techniques to enable users to identify AI-generated content. The provenance data should include an identifier of the service or model that created the content, but need not include user information. Organizations should also endeavor to develop tools or APIs to allow users to determine if particular content was created with their advanced AI system, such as via watermarks.

8. Prioritize research to mitigate societal, safety and security risks and prioritize investment in effective mitigation measures. This includes conducting, collaborating on and investing in research that supports the advancement of AI safety, security, and trust, and addressing key risks, as well as investing in developing appropriate mitigation tools. Organizations commit to conducting, collaborating on and investing in research that supports the advancement of AI safety, security, trustworthiness and addressing key risks, such as prioritizing research on upholding democratic values, respecting human rights, protecting children and vulnerable groups, safeguarding intellectual property rights and privacy, and avoiding harmful bias, mis- and disinformation, and information manipulation. Organizations also commit to invest in developing appropriate mitigation tools, and work to proactively manage the risks of advanced AI systems, including environmental and climate impacts, so that their benefits can be realized.

9. Prioritize the development of advanced AI systems to address the world's greatest challenges, notably but not limited to the climate crisis, global health and education. These efforts are undertaken in support of progress on the United Nations Sustainable Development Goals, and to encourage AI development for global benefit. Organizations should prioritize responsible stewardship of trustworthy and human-centric AI and also support digital literacy initiatives that promote the education and training of the public, including students and workers, to enable them to benefit from the use of advanced AI systems, and to help individuals and communities better understand the nature, capabilities, limitations, and impact of these technologies.

10. Advance the development of and, where appropriate, adoption of international technical standards.

11. Implement appropriate data input measures and protections for personal data and intellectual property. Appropriate measures could include transparency, privacy-preserving training techniques, and/or testing and fine-tuning to ensure that systems do not divulge confidential or sensitive data.

These eleven principles will provide organisations that develop, deploy or use advanced AI systems with guidelines to promote the security and reliability of the technology, mitigate risks and misuse, detect the existence of vulnerabilities, encourage responsible information sharing, incident reporting, and investment in cybersecurity, among other aspects. They were drawn up on the basis of surveys of interested parties [55] and this code was followed by the Bletchley Declaration [56] in which 29 countries, including Spain, participated in the idea of supporting an inclusive international network of scientific research on AI security that complements existing multilateral collaboration and forums. Subsequent to it is also the Executive Order on Safe, Secure, and Trustworthy Development and Use of Artificial Intelligence. [57]

The first of the principles set out in this code already refers to testing by the organisations creating these Artificial Intelligence systems, which should be carried out through *controlled test sites*, which are also referred to in the Regulation. [58]

---

[55] Cfr. https://digital-strategy.ec.europa.eu/en/news/commission-gathers-views-g7-guiding-principles-generative-artificial-intelligence.

[56] Available at: https://www.gov.uk/government/publications/ai-safety-summit-2023-the-bletchley-declaration/the-bletchley-declaration-by-countries-attending-the-ai-safety-summit-1-2-november-2023. For more information see also C.I. Velasco Rico, *Marco regulatorio...*, op. cit., p. 33.

[57] Available in: https://www.whitehouse.gov/briefing-room/presidential-actions/2023/10/30/executive-order-on-the-safe-secure-and-trustworthy-development-and-use-of-artificial-intelligence/.

[58] According to Article 57 (1) AIA, 'Member States shall ensure that their competent authorities establish at least one AI regulatory sandbox at national level, which shall be operational by 2 August 2026. That sandbox may also be established

In order to avoid regulatory fragmentation in the Union, the Commission may also adopt *implementing acts* specifying the detailed arrangements for the establishment, development, implementation, operation, and supervision of these controlled test sites, which will ensure that *potential providers are facilitated through the learning outcomes of the AI controlled test sites, to comply with the conformity assessment obligations under this Regulation and the voluntary application of the codes of conduct.* In this respect, reference is made to the commentary on Article 97 AIA.

2. *Guidelines (Article 96 AIA)*

European Union Guidelines are non-binding guidelines and documents addressed to different subjects with the aim of facilitating the implementation of other EU legislation. Although non-binding,[59] these guidelines play an important role in implementing and enforcing EU law. Guidelines can take various forms, such as, for example: *Commission practical guidelines* setting out good practices and procedures for risk prevention, providing a frame of reference for Member States and other stakeholders to improve their processes and ensure compliance with EU law; *Council recommendations*,[60] as proposals addressed to Member States or other EU institutions, to suggest specific actions, but without imposing legal obligations, to guide policy and legislative actions in various areas; *Commission communications*, as documents used to inform, clarify or interpret specific aspects of EU law or policy, which may address technical, political or administrative issues, and which seek to provide a coherent and uniform interpretation of EU rules; or *EU social partner agreements*, i.e. pacts between employers' organisations and trade unions at European level, which may influence EU labour law and social policies, and which, although

---

jointly with the competent authorities of other Member States. The Commission may provide technical support, advice and tools for the establishment and operation of AI regulatory sandboxes'.

[59] We already have more previous examples of Guidelines, such as the HLEG *Ethics guidelines for trustworthy AI* (2019), cited above.

[60] It is true that Recommendations (covered by Article 288 TFEU) are one of the non-binding forms that EU acts can take, along with opinions. Because of their non-binding nature, they also have no legal consequences if they are not complied with, but they can provide guidance on the interpretation or content of EU law.

not directly binding, may be implemented by directives or decisions of the Member States.

In short, the EU Guidelines, although not legally binding (another example of soft law), are key tools for the practical and uniform application of EU law, providing guidance and setting standards that help Member States and other stakeholders to comply with the obligations and objectives of the European Union. However, they do not have the same scope as a delegated act since, unlike delegated acts, they are not legally binding. They will, of course, need to be updated when deemed necessary (we should not forget that we are dealing with rapidly evolving AI systems) at the request of Member States or the AI Office or the initiative of the Commission itself. In issuing these Guidelines, the Commission will pay particular attention to the needs of SMEs (along the same lines of protecting these small businesses, as seen in other articles of the Regulation), including start-ups, local public authorities, and sectors most likely to be affected by this Regulation.

Along these lines, Article 96 AIA introduces Guidelines that may be adopted by the Commission on the application of the Regulation. These may concern the implementation of the requirements and obligations referred to in Articles 8 to 15 and Article 25, i.e., they may concern the requirements for high-risk AI systems relating to their compliance, their management system, data and data governance, technical documentation, record keeping, transparency and communication of information to deployers, human oversight, accuracy, robustness, and cybersecurity, as well as responsibilities along the AI value chain in relation to the obligations of suppliers and deployers of high-risk AI systems. In these aforementioned scenarios, those Guidelines shall take into account the generally recognised state of the art in AI, as well as the relevant harmonised standards and common specifications referred to in Articles 40 and 41, or harmonised standards or technical specifications to be established in accordance with Union harmonisation law.

The Guidelines may also address prohibited practices referred to in Article 5 on the practical implementation of the provisions related to substantial modifications, transparency obligations of providers and users of AI systems intended to interact directly with natural persons, or which are in general use, or generate synthetic audio, image, video or text content, or involve emotion recognition or a biometric categorisation system, as well as systems that generate or manipulate images or audio or video content or generate or manipulate text.

These Guidelines may also act by providing detailed information on the relationship between this Regulation and the list of Union harmonisation legislation set out in Annex I, as well as other relevant Union legislation, including as regards consistency in its application, as well as on the application of the definition of AI system set out in Article 3 (1) AIA.

We can still make one last reference to guidelines that the Commission can approve following this same Article 96, but which are not included in the same article but in other paragraphs. I am referring specifically to Article 99 concerning penalties and other enforcement measures, such as warnings or non-financial measures, which may be established by the Member States, measures applicable to the infringements contained in the Regulation, and which must take into account the Guidelines issued by the Commission.

Alternatively, the reference to the same Guidelines in Article 6 (5), which should be adopted by 2 February 2026, after consultation of the European Committee on Artificial Intelligence, to specify the practical implementation of that Article, thus concerning the rules for the classification of high-risk AI systems, as well as a comprehensive list of practical examples of use cases of high-risk and non-high-risk AI systems.

3. *Conclusions*

The US consultancy Gartner estimates that the AI market could be worth $127 billion by 2025, up from $2 billion in 2015. The United States and China will lead the way in terms of investment.

Moreover, although there are voices such as the Swedish philosopher from Oxford University, Nick Bostrom, who anticipates a 90% chance that between 2075 and 2090, there will be machines as intelligent as humans, or Stephen Hawking, who predicted that machines would ultimately surpass humans in less than 100 years, the truth is that far from making us obsolete, AI will make us more efficient and allow us to carry out actions that we would never have been able to do due to their complexity.

The EU has indeed come a long way in drafting the text that has been the subject of partial commentary in these lines. However, there is still a long way to go before the regulation of AI is complete. This regulation perhaps develops in more detail some of these principles that we have had the opportunity to discuss in parallel work between jurists, technicians, and other specialists in the development of AI systems.

The fact that we have assessed different principles that codes of conduct should contain, based on the various documents we have analysed, leads us to conclude that it is difficult for all of these principles, or at least the most important ones, to become the necessary catalyst for the response expected from such instruments. Although codes of conduct are important tools to guide behaviour and practices within organisations, their effectiveness as drivers of the desired response depends on multiple factors beyond the principles they contain. The diversity of contexts and the need for effective implementation and genuine commitment are some of the key challenges that must be overcome for these instruments to be truly impactful.

The use of codes of conduct for 'low risk' AI systems is a crucial measure to promote the ethical and responsible development of this technology. As we have discussed, this is because they promote trust in AI systems by setting clear standards for their design, implementation, and use. These codes aim to protect rights and values and should ensure that AI systems respect human rights, privacy, and other fundamental values. This is essential even for low risk applications, as any technology can have significant ethical implications. They will promote transparency by requiring disclosure of information about how AI systems work, what data they use, and how they make decisions, which is crucial for users to understand and trust these technologies.

While low-risk AI systems may have limited impact, establishing ethical codes from the outset can help prevent more serious ethical or legal problems in the future as technology evolves. In this way, they will prevent future risks. They should incorporate different perspectives - they should be developed collaboratively and include a variety of perspectives, including those that can be brought to the attention of experts in ethics, technology, human rights, and society at large. This ensures that a wide range of ethical and societal concerns are addressed.

In summary, codes of conduct are an essential tool to ensure that even AI systems considered low risk are developed and used ethically and responsibly, thus promoting technological innovation in harmony with fundamental human values. Nevertheless, they are not enough; they must be the necessary complement to the EU legislation we are now discussing and international action on Artificial Intelligence as well.

Finally, the analysis of the Guidelines that the Regulation allows the Commission to draw up places us before a non-binding instrument with no normative value, which does not aspire to replace Community policies or current or future EU regulatory actions nor to prevent their introduction. We must also consider them as living documents, susceptible to fundamental revision and updating in accordance with the evolution of technology, as well as other elements such as the social and economic environment. Furthermore, why not also use it as reference documents for other countries around us or for future conferences between States related to Artificial Intelligence?

4. *References*

M. Barrio Andrés and others, *La visión de los expertos*, in *Informe Break the limits*, Fundación Aranzadi, La Ley, 2023.

M. Barrio Andrés, *Luces y sombras del Estado algorítmico de Derecho*, in *Derecho Digital e Innovación*, no. 5, 2020.

J.L. Bermejo Latre, *La aplicación de la inteligencia artificial en la actividad formal e informal de la administración*, in *XVIII Congreso de la Asociación Española de Profesores de Derecho Administrativo*, 2024. Full text available in: https://www.aepda.es/AEPDAEntrada-3987-XVIII-CONGRESO-DE-LA-ASOCIACION-ESPANOLA-DE-PROFESORES-DE-DERECHO-ADMINISTRATIVO.aspx.

A.D. Berning Prieto, *El uso de sistemas basados en inteligencia artificial por las Administraciones Públicas: estado actual de la cuestión y algunas propuestas ad futurum para un uso responsable*, in *Revista de Estudios de la Administración Local y Autonómica,* no. 20, 2023, pp. 165-185.

A. Boix Palop, *Los algoritmos son reglamentos: la necesidad de extender las garantías propias de las normas reglamentarias a los programas empleados por la Administración para la adopción de decisiones*, in *Revista de Derecho Púbico: Teoría y Método*, no. 1, 2020, pp. 223-270.

A. Boix Palop, *Transparencia en la utilización de inteligencia artificial por parte de la Administración*, in *El Estado Social y Democrático de Derecho*, no. 100, 2022, pp. 90-105.

A. Boix Palop and L. Cotino Hueso (eds.), *Derecho Público, derechos y transparencia ante el uso de algoritmos, inteligencia artificial y big data*, in *Revista General de Derecho Administrativo*, no. 50, 2019.

A. Cerrillo i Martínez, *El impacto de la inteligencia artificial en el Derecho administrativo. ¿Nuevos conceptos para nuevas realidades técnicas?*, in *Revista General de Derecho Administrativo*, no. 50, 2019.

A. Cerrillo i Martínez, *La inteligencia artificial y el control de sus posibles sesgos*, in M. Villoria Mendieta (ed.), *Ética pública en el siglo XIX*, INAP, Madrid, 2021, pp. 93-112.

S. Civitarese Mateucci, *'Umano troppo umano'. Decisioni amministrative automatizzate e principio di legalità*, in *Diritto pubblico*, 2019, vol. 1, pp. 5-41.

L. Cotino Hueso*, Ética en el diseño para el desarrollo de una inteligencia artificial,* *robótica y big data confiables y su utilidad desde el derecho*, in *Revista Catalana de Dret Públic*, no. 58, 2019, pp. 29-48.

L. Cotino Hueso, *Riesgos e impactos del big data, la inteligencia artificial y la robótica. Enfoques, modelos y principios de la respuesta del Derecho*, in *Revista General de Derecho Administrativo*, no. 50, 2019.

L. Cotino Hueso (ed.), *La Carta de Derechos Digitales*, Valencia, Tirant Lo Blanch, 2022

L. Cotino Hueso, *Los usos de la inteligencia artificial en el sector público, su variable impacto y categorización jurídica*, in *Revista canaria de Administración pública*, no. 1, 2023, pp. 211-242.

Y. de la Cueva, *¿Quién vigila al algoritmo?*, in *El Notario del siglo XXI*, 2019, no. 87.

I. Gallego Córcoles, *La contratación de soluciones de inteligencia artificial*, in E. Gamero Casado (ed.) *Inteligencia artificial y sector público: retos, límites y medios*, 2023, Valencia, Tirant Lo Blanch, pp. 503-567.

D. García San José, *La Europa de los derechos antes los avances científicos y tecnológicos*, Valencia, Tirant lo Blanch, 2021.

E. Gil, *Big data, privacidad y protección de datos*, Boletín Oficial del Estado, 2015.

J.A. Hernández Corchete*, ¿Un nuevo estatuto jurídico para el ciudadano?*, in *XVIII Congreso de la Asociación Española de Profesores de Derecho Administrativo*, 2024. Full text available in: https://www.aepda.es/AEPDAEntrada-3987-XVIII-CONGRESO-DE-LA-ASOCIACION-ESPANOLA-DE-PROFESORES-DE-DERECHO-ADMINISTRATIVO.aspx.

A. Huergo Lora, *Una aproximación a los algoritmos desde el Derecho Administrativo*, in A. Huergo Lora (ed.) *La regulación de los algoritmos*, Cizur Menor, Aranzadi, 2020, pp. 23-87.

A. Huergo Lora, *Administraciones Públicas e Inteligencia Artifical: ¿más o menos discrecionalidad?*, in *El Cronista del Estado Social y Democrático de Derecho*, no. 96-97, 2021, pp. 78-95.

A. Huergo Lora, *De la digitalización a la Inteligencia Artificial: ¿evolución o revolución?*, in *XVIII Congreso de la Asociación Española de Profesores de Derecho Administrativo*, 2024. Full text available in: https://www.aepda.es/AEPDAEntrada-3987-XVIII-CONGRESO-DE-LA-ASOCIACION-ESPANOLA-DE-PROFESORES-DE-DERECHO-ADMINISTRATIVO.aspx.

J. Jiménez López, *Oscuridad algorítmica en el sector público*, in G. Vestri (ed.), *La disrupción tecnológica en la Administración Pública. Retos y desafíos de la inteligencia artificial*, 2021, Thomson Reuters-Aranzadi, pp. 41-66.

R. López de Mantaras and P. Messeguer González, *Inteligencia artificial*, Madrid, Catarata-CSIC, 2017.

F. Longo, *Administración pública con valores. Instrumentos para una gobernanza ética*, Madrid, INAP, 2015.

A. Mantelero, *Retos y regulación de la Inteligencia Artificial: la toma de decisiones en los asuntos públicos y la administración de justicia*, in *XVIII Congreso de la Asociación Española de Profesores de Derecho Administrativo*, 2024. Full text available in: https://www.aepda.es/AEPDAEntrada-3987-XVIII-CONGRESO-DE-LA-ASOCIACION-ESPANOLA-DE-PROFESORES-DE-DERECHO-ADMINISTRATIVO.aspx.

I. Martín Delgado, *La aplicación del principio de transparencia a la actividad administrativa algorítmica*, in E. Gamero Casado (ed.) *Inteligencia artificial y sector público: retos, límites y medios*, Valencia, Tirant Lo Blanch, 2023, pp. 131-194.

R. Martínez Martínez, *Inteligencia artificial, derecho y derechos fundamentales*, in T. de la Quadra Salcedo and J.L. Piñar Mañas (eds.), *Sociedad digital y derecho*, BOE, 2018, pp. 259-278.

M. Moreno Rebato*, Inteligencia artificial (umbrales éticos, Derecho y Administraciones Públicas)*, Civitas, 2021.

M. Moreno Rebato, *La propuesta de Reglamento de la Unión Europea sobre inteligencia artificial y las Directrices éticas para una inteligencia artificial fiable: una oportunidad para la Administración Pública Española*, in G. Vestri (ed.), *La disrupción tecnológica en la Administración Pública: retos y desafíos de la inteligencia artificial*, Thomson Reuters-Aranzadi, 2022, pp. 67-81.

S. Muñoz Machado, *Prólogo*, in *El Cronista del Estado Social y Democrático de Derecho*, no. 100, 2022, pp. 5-13.

B. Olivares Olivares, *Implicaciones de la normativa sobre protección de datos en el desarrollo de la inteligencia artificial por la administración tributaria: la gobernanza de los datos*, in S. Moreno González (ed.), *Nuevas tecnologías disruptivas y tributación*, Thomson Reuters-Aranzadi, 2021.

J.A. Plaza, *Lecciones de ética para máquinas que piensan y toman decisiones*, in *El País Retina*. Full text available in: https://elpais.com/retina/2017/12/19/innovacion/1513661054_305253.html.

J. Ponce Solé, *Inteligencia artificial, Derecho administrativo y reserva de humanidad: algoritmos y procedimiento administrativo debido tecnológico*, in *Revista General de Derecho Administrativo*, no. 50, 2019.

J. Ponce Solé, *Reserva de humanidad y supervisión humana de la Inteligencia artificial*, in *El Cronista del Estado Social y Democrático de Derecho*, no. 100, 2022, pp. 58-67.

J. Ponce Solé, *Seres humanos e inteligencia artificial: discrecionalidad artificial, reserva* *de humanidad y supervisión humana*, in E. Gamero Casado (ed.), *Inteligencia artificial y sector público: retos, límites y medios*, Valencia, Tirant Lo Blanch, 2023, pp. 195-225.

A.M. Porcelli, *La inteligencia artificial y la robótica: sus dilemas sociales, éticos y jurídicos*, en *Derecho global. Estudios sobre Derecho y Justicia*, vol. 6, no. 16, 2020.

M.A. Presno Linera, *Derechos fundamentales e inteligencia artificial en el Estado social, democrático y digital de Derecho*, in *El Cronista del Estado Social y Democrático de Derecho*, no. 100 2022, pp. 48-57.

T. de la Quadra Salcedo Fernández del Castillo, *Inteligencia artificial, Administraciones Públicas y Derecho. Una visión comparada de un Derecho en construcción*, in *XVIII Congreso de la Asociación Española de Profesores de Derecho Administrativo*, 2024. Full text available in: https://www.aepda.es/AEPDAEntrada-3987-XVIII-CONGRESO-DE-LA-ASOCIACION-ESPANOLA-DE-PROFESORES-DE-DERECHO-ADMINISTRATIVO.aspx.

C. Ramió, *Inteligencia Artificial y Administración Pública. Robots y humanos compartiendo el servicio público*, 2018, Catarata.

J.F. Sánchez Barrilao, *Derecho constitucional, desarrollo informático e inteligencia artificial: aproximación a la propuesta del Parlamento Europeo a favor de una regulación sobre robótica*, en J. Valls Prieto (ed.), *Retos jurídicos por la sociedad digital*, Cizur Menor, Thomson Reuters/Aranzadi, 2018, pp. 21-76.

S. de la Sierra Morón, *Una introducción a la Carta de Derechos Digitales*, in L. Cotino Hueso (ed.), *La Carta de Derechos Digitales*, Valencia, Tirant Lo Blanch, 2022, pp. 27-52.

S. de la Sierra, *El ejercicio de potestades mediante Inteligencia Artificial*, in *XVIII Congreso de la Asociación Española de Profesores de Derecho Administrativo*, 2024. Full text available in: https://www.aepda.es/AEPDAEntrada-3987-XVIII-CONGRESO-DE-LA-ASOCIACION-ESPANOLA-DE-PROFESORES-DE-DERECHO-ADMINISTRATIVO.aspx.

A. Soriano Arnanz, *Decisiones automatizadas y discriminación: aproximación y propuestas generales*, in *Revista General de Derecho Administrativo*, no. 56, 2021.

S. Tavares da Silva, El Derecho Administrativo en la era de la inteligencia artificial, comunicación, in XVIII Congreso de la Asociación Española de Profesores de Derecho Administrativo, 2024. Full text available in: https://www.aepda.es/AEPDAEntrada-3987-XVIII-CONGRESO-DE-LA-ASOCIACION-ESPANOLA-DE-PROFESORES-DE-DERECHO-ADMINISTRATIVO.aspx.

D. Terrón Santos, *Limitar la IA desde la ética y el derecho*, in J.L. Domínguez Álvarez y D. Terrón Santos (eds.), *Desafíos éticos, jurídicos y tecnológicos del avance digital*, Madrid, Iustel, 2023, pp. 175-190.

C.I. Velasco Rico, *Vigilando al algoritmo. Propuestas organizativas para garantizar* *la transparencia*, in B. Puentes Cociña and A. Quintiá Pastrana (eds.*), El Derecho ante la transformación digital: oportunidades, riesgos y garantías*, 2019, Atelier, pp. 73-90.

C.I. Velasco Rico, *Marco regulatorio de los sistemas algorítmicos y de inteligencia artificial: el papel de la Administración*, in *XVIII Congreso de la Asociación Española de Profesores de Derecho Administrativo*, 2024. Full text available in: https://www.aepda.es/AEPDAEntrada-3987-XVIII-CONGRESO-DE-LA-ASOCIACION-ESPANOLA-DE-PROFESORES-DE-DERECHO-ADMINISTRATIVO.aspx.

G. Vestri, *La inteligencia artificial ante el desafío de la transparencia algorítmica: una aproximación desde la perspectiva jurídico-administrativa*, in *Revista Aragonesa de Administración Pública*, no. 56, 2021, pp. 368-398.