THE CERTIFICATION AS A MECHANISM FOR CONTROL OF ARTIFICIAL INTELLIGENCE IN EUROPE

By

Carlos Galán, PhD
Doctor in Computer Science
Law Degree and Lawyer specialized in ICT Law
Professor of the Carlos III University of Madrid (Spain)
President of the Legal Technology Agency
CCN Advisor
Member of the European Artificial Intelligence Alliance
Member of the Cyberpolitical Working Group of the Elcano
Royal Institute

.

Summary / Abstract:

The use of Artificial Intelligence (AI) is one of the most significant technological contributions that will permeate the life of western societies in the coming years, providing significant benefits, but also highlighting risks that need to be assessed and minimized. A reality as disruptive as AI requires that its technology and that the products and services supported by it provide enough guarantees of its proper functioning. The present work analyzes and makes a proposal for the application of the mechanisms of Certification of Conformity to the maintenance of the afore mentioned guarantees.

----

1. INTRODUCTION : THE RELIABILITY OF ARTIFICIAL INTELLIGENCE

On July 27, 1987, the one that these lines subscribe defended before the Doctoral Thesis Court of the Polytechnic University of Madrid the work "*A model for the understanding of problems proposed in natural language for a limited domain*". As can be collected, the work was the result of an investigation related to one of the classic paradigms of Artificial Intelligence (AI): the automatic recognition of natural language.

Since 1987, the recognition of natural language, like the rest of the competing domains of AI, has evolved greatly.

After a few years of "desert crossing", practically empty of the spectacular results that had been predicted, the AI now enjoys a magnificent present, which will surely result in an extraordinary future [1].

We do not believe to sin of chimerics if we affirm, with complete emptiness, that AI will be one of the backbone of society in the coming years; able to modify our habits and alter our relationship with the surrounding world [2]. It is not just about machines that play (and win) against humans, autonomous vehicles or robots that help in such tasks, etc .: we are talking about a new ecosystem that will integrate intelligent elements, capable of reasoning as you would a human being (sometimes, better and faster). The speed of data processing and the volume handled, traditional obstacles to AI developments, have collapsed dramatically. At present, it is already possible to handle in real time terabytes of information (which has made possible the emergence of new tools such as machine learning or deep learning) and, in general, the adoption of the methods, procedures and tools of the AI in multiple and varied scenarios: from the autonomous vehicle, to health care, through domestic or service robots, education, cybersecurity [3], personalized medicine, the fight against climate change, the best use of resources natural or transport infrastructure; all in an economic scenario that could exceed 38 billion dollars in 2025 [4].

The day is not far away when intelligent machines, beyond the performance of routine activities (which would be part of the so-called *weak AI*), will be able to show behaviors previously reserved for the human being: the ability to reason autonomously, to learn from the past ... and to make decisions (what has been called *hard AI*).

When one of my students asks me: "Professor: can the machines think?", after sketching a smile reminiscent of Turing [5] or Minsky [6], I reply, emphatically: "A machine, as you call them, could do the same as an y human being, who, by the way, is but also a machine. Or not?" [7].

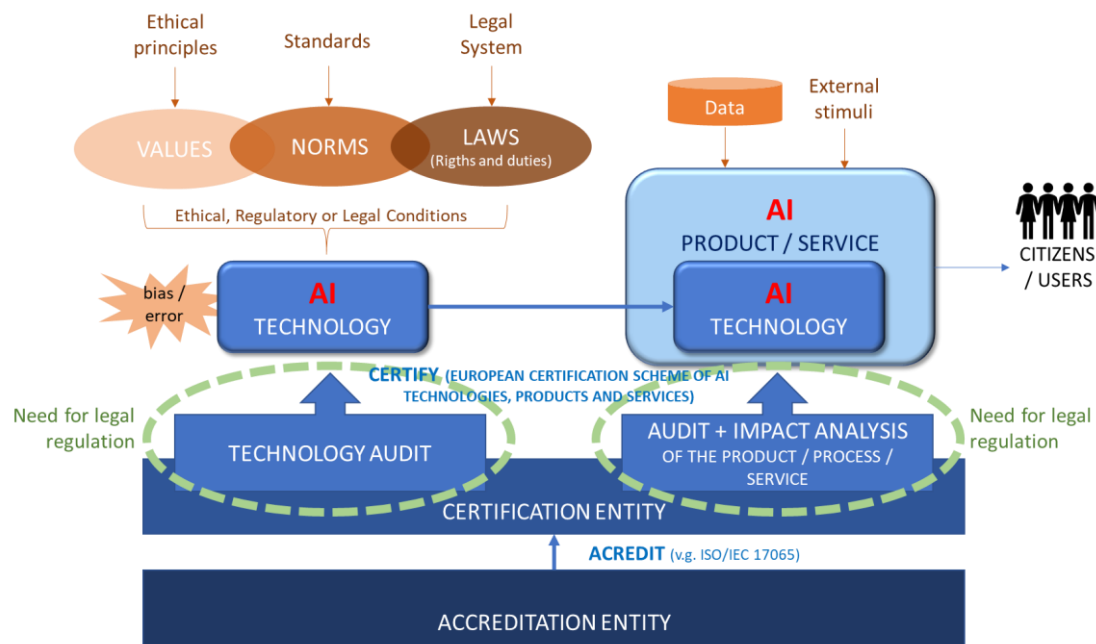Faced with this future, closer than it may seem, it should be prepared.

And this is where an element of the greatest importance comes into play, in whose formulation the present work finds its genesis: admitting that an intelligent machine can modify its behavior according to its internal logic and exogenous factors, who assures us that its behavior goes to develop within what we usually call ethical principles *ethical principles* [8] and, in any case, respecting the current legal system? What does the law have to say in this situation? How do you say it?

Faced with this challenge, there are two possibilities not only not exclusive, but complementary: incorporating a set of rules to the AI technology construction systems capable of adjusting the behavior of AI products or services to the ethical-legal rules of application; or ensuring, via audit, that such technologies, products or services conform to a previously defined model of ethical-legal behavior.

Both answers are, as we say, complementary. The present work deals with the second one: the certification of conformity (based on independent audits) as a mechanism of ethical-legal control of AI technologies, products or services.

2. THE PROPOSED MODEL
The figure shows a conceptual scheme of the proposed model.

The proposed AI Certification model in Europe

2.1 AI Technology, Products and Services

We will call **AI Technology**, in general and without a thoroughness, to the set of methods, procedures, tools and results, scientific or technological, that constitute the basis for building AI Products or Services, that is, products or services that have "ability to interpret correctly external data, learn from such data and use that learning to achieve specific objectives and tasks through flexible adaptation " [9].

The AI Technology is therefore configured as the pieces with which we will build solutions supported, to a greater or lesser extent, in AI techniques and whose intelligent behavior will be given by the greater or lesser importance or magnitude of the AI technology involved in the product or final service.

The **AI Products/Services**, meanwhile, are those that may be used by their recipients, "consumers" or end users, whether *general purpose* (addressed to all) or *specific purpose* (directed to a sector of the population or the industry), and whose configurator components include elements belonging to AI Technology. For example, an autonomous vehicle (*AI Product* paradigm) will have an inference engine (AI Technology) in charge of processing the stimuli from the outside (and from the interior, if applicable) in accordance with the provisions of the AI software used, attending to the signals it receives from the sensors and, surely, to the experience accumulated based on the previous processing of thousands or millions of data.

Therefore, not all the constituent elements of an AI Product must be, necessarily, AI Technology. On the contrary, the usual thing is that in an AI Product/Service coexist elements developed and implemented using what we can call *traditional technologies* with components developed based on AI Technology.

2.2 Configuration elements of AI Technology and AI Products/Services

While "traditional technology" is based on the design and implementation of *deterministic algorithms*, that is, procedures in which the same results are obtained for the same input data; AI models (especially those included in the so-called *hard AI*) use *heuristic procedures*, that is, computational models that not only take into account the input data, but also the accumulated experience and *knowledge* derived from such experience, which It facilitates that, in the face of the same stimuli, different responses can be generated. This is what we will call the *adaptive capacity* of AI, a quality that brings these systems closer to human reasoning.

This heuristic behavior is what gives AI its greatest benefits and also its most significant risks.

That an intelligent machine (an AI system, in short) is able to derive its autonomous operation towards unethical or illegal behaviors only depends on the fact that its developers have incorporated a certain type of reliable and contrasted elements into the AI Technology that prevent such drift.

Here is the essence of the problem: What are these elements capable of making an AI system work properly in accordance with the ethical and legal principles assumed by the user community to which it is addressed? And how to incorporate such principles into the system? [10].

There are several aspects that can (and should) direct the design and development of AI Technology, and that should be taken into account by their manufacturers. The most significant are those shown in the following table.

Aspects exogenous to AI Technology	Legal norms	Set of regulations pertaining to the applicable legal system.
	Ethical principles	Set of behavior guidelines, generally accepted by the target community of AI products / services [11].
	Standards	Set of technical standards, usually from international organizations [12], aimed at normalizing the construction or use of products, processes or services [13].
Aspects endogenous to AI Technology	Bias	Undesirable quality - usually caused by an incorrect design or an insufficient or inadequate choice of learning data - that can make AI systems produce biased results.
	The error	Faults - usually not deliberate - in the design or development of technology.

All of them conform the ethical, normative or legal elements that condition the construction of the AI Technology. If one is not present or is improperly implemented, the resulting AI Technology will not have the proper guarantees to ensure that its operation conforms to the legal or ethical models desired by the user community of the final products or services.

Unless it has ad hoc barriers (which is not usually frequent), the behavior of the AI Technology, good or bad, acceptable or not acceptable, is transferred, whether we want it or not, to the AI Product/Service in which it is inserted. Since this product or service is precisely the one used by end users, it is at that moment that the suitability of the basic technology used is evidenced, and when, unfortunately, remediation is usually difficult - or impossible, even - if such behavior has moved away from the lawful or desirable.

2.3 Ex-ante control and ex-post control

Thus, it seems logical to think that, before an AI Product/Service enters the market, it is necessary to ensure that its final behavior will be as expected. This assurance necessarily involves guaranteeing the suitability of the product/service itself, as well as the technology used for its construction. In other words: although it constitutes the most important verification, it is not enough to guarantee that an AI Technology is adequate from the ethical and legal points of view, it is also necessary to verify that the AI Product/Service in which it is "inserted" is equally acceptable after such inclusion [14].

The verification of this suitability can be done in two temporary moments: before or after its construction. Both types of control are not exclusive.

The following table shows the essential characteristics of both types of control [15].

Ex-ante control: Control measures <u>prior </u>to the development or implementation of the AI Technology, Product or Service.	Security **from the design.** Essential activities included: - Conception. - Planning - Development - Tests.	Verification of the conformity of the solution with: • The applicable legal system. • Ethical principles. • Technical standards. • Impact analysis on society (target community of products / services)[16].
Ex-post control: Control measures <u>subsequent </u>to the development or implementation of the AI Technology, Product or Service.	Safety **in the operation.** Essential activities included: - Integration - Acquisition. - Deployment - Exploitation. - Publication. - Conservation - Access. - Interconnection.	Satisfaction of the requirements of: • Transparency and intelligibility of the systems. • Possibility of access and verification. • Explainability , traceability and accountability (responsibility).

Even though it is always preferable, **Ex-ante control** is not easy to implement. As we have indicated in the previous table, it demands from the manufacturers of the AI Technology, Products or Services (and, where appropriate, from their distributors) the submission to a discipline of action that must impregnate not only the conception of the systems but its complete life cycle: planning, design, acquisition, construction, deployment, exploitation, publication, conservation and access or interconnection with them.

On the other hand, an Ex-ante control also requires reliable mechanisms, capable of transferring to the design and development of the underlying AI Technology the legal, ethical and regulatory requirements that ensure that, in the end, an adequate, legal product is obtained and ethically acceptable [17].

In this regard, the work of the High-Level Group of Experts on Artificial Intelligence of the European Commission (AIHLE) has been directed, which has insisted in pointing out, rightly, that AI must be centered on the human being: "*AI must be developed, deployed and used with an "ethical purpose", based on fundamental rights, social values and ethical principles of beneficence (doing good), non-maleficence (not harming), human autonomy, justice and the explainability* [18]. *All of this is crucial to achieve reliable AI.* " [19].

From the European Economic and Social Committee it has been pointed out that AI must respond to minimum security standards (internal and external) that allow it to function properly and without causing damage to users or recipients of its action, even proposing, the elaboration of a uniform and universal code of ethics for the development, deployment and use of AI, so that throughout its operating process AI systems are compatible with the principles of human dignity, integrity, freedom , privacy, cultural and gender diversity and fundamental human rights [20][21].

Furthermore, in our opinion, so that an AI Technology, Product or Service can be used by its recipients, it is not only essential that it be in accordance with ethical principles or a specific legal system; It will also be necessary to carry out an **impact analysis** of such an AI object on the society in which it intends to be inserted. This is especially important when the deployment of the object AI in question can have a significant impact on the rights and freedoms of people or in the established socio-economic order [22].

In view of the complexity to transfer ethical and legal principles to AI Technology (and, consequently, to the products or services in which it is incorporated), it seems necessary to contemplate, in addition, an **Ex-post control**, capable of determine, with a high degree of confidence, the goodness of an AI Product/Service, already built, and before its commercialization or deployment.

2.4 Audit and Certification of Conformity

As is known, an audit is: "*A systematic, independent and documented process that seeks to obtain objective evidence and its evaluation to determine the extent to which the audit criteria are met*." [23]. In In other words: an audit indicates the extent to which a certain object is in accordance with the provisions of the reference (reference) standard in question.

This type of Ex-post controls allow us to focus our attention on the observation of the behavior of the audited object, rather than on ensuring that such an object has been designed or constructed according to certain rules. It is, therefore, an evaluation based on the evidence that emerges from the audited object analysis, when it is subjected to an adequate battery of controls, and that will allow us to decide if such object is within the margins admitted by the standard reference. In other words and bringing it closer to our purpose: if your behavior is aligned with the ethical and legal expectations that direct the community in which you intend to integrate.

If this **Compliance Audit** is satisfactory, the Conformity Assessment Entity (by itself or through an authorized third party) will issue a **Certificate of Conformity** that will display, *erga omnes*, the conformity of the technology, product or service with the standard (certification scheme) that has been taken as a reference [24].

The use of certifications, based on independent audits, as an element of exhibiting the conformity of a particular product or service is well known. Compliance with SOG-IS MRA [25], the Trust Services certification model [26], the GDPR compliance certification schemes, or the security certification models sponsored by the so-called *Cybersecurity Act* [27] are good examples. In Spain, the most widespread model for assessing the conformity of the security of information systems (especially aimed at

public sector entities and private companies providing services to those) is the National Security Framework (ENS) [28], which includes the Certification Compliance with the ENS based on the overcoming of periodic compliance audits by the information systems concerned; or the model of Evaluation and Certification of Information Technology Security [29].

In our case, to make this certification mechanism possible, it is necessary to have two previous elements:

> 1. A **Scheme of Evaluation and Certification of the Conformity of the AI Technologies, Products or Services** [30].

> 2. An **European standard** (of a regulatory type, preferably) that regulates the previous scheme, its development, application and updating; and to determine the actors involved in its deployment: Accreditation Entities and Conformity Assessment Entities, essentially [31].

It is evident that the greatest difficulty of this model is found in the first of the points cited and that is where, in our opinion, the first works should focus.

At the end, in our opinion, only that AI Technology, Product or   Service that is duly certified should be deployed in Europe.

3. CONCLUSIONS AND SUBSEQUENT ACTIONS

From all of the above, we can draw some conclusions and propose subsequent actions:

• Since AI is an indispensable element for the development of society, it is necessary to ensure, to the extent possible, that the AI Technology, the AI Products and the AI Services conform to the provisions of the legal regulation that may be applicable, maintaining scrupulous respect for universally accepted ethical principles.

• The previous guarantee can only be achieved with an adequate combination of two types of controls: Ex-ante and Ex-post, before and after, respectively, of the design, development, implementation or deployment of the technology, products or services involved.

• Ex-ante control, although it presents undoubted and logical advantages, is complex and subject, in any case, to the goodwill of designers, developers or implementers.

• Ex-post control, materialized in the form of independent audits and Certifications of Conformity, is configured as a practical and effective mechanism to achieve the objectives pursued.

• The implementation of the AI Compliance Audits requires a European Certification Scheme for AI Technologies, Products and Services [32]; as well as a list of Conformity Assessment Entities that have the required technical capabilities, independence and Impartiality due, and have been duly authorized by an Accreditation Entity [33].

• All of this must be perfectly regulated, through the corresponding European regulations [34].

• AI Technologies, Products or Services can be built within the EU or come from third countries. The above regulation should apply to any technology, product or service of

this type that is intended to be deployed or marketed in Europe. Only technology, certified AI products or services should benefit from public aid [35].

• The Certification Schemes of AI Technologies, Products and Services must be incorporated into the plans derived from the European Strategy of R&D in Artificial Intelligence.

Reiterating the importance of AI in the future of western societies and the urgent need to ensure its proper deployment, we would like to end with a sentence taken from the aforementioned Communication from the Commission to the European Parliament, the European Council, the Council, the European Economic and Social Committee and the Committee of the Regions, on Artificial Intelligence for Europe: "*Our way of addressing the question of AI will define the world in which we will live*."

[1] This is also supported by the EU in its documents: Communication from the Commission to the European Parliament, the European Council, the Council, the European Economic and Social Committee and the Committee of the Regions. *Coordinated Plan on Artificial Intelligence* (Brussels, 12.7.2018. COM (2018) 795 final) and Communication of Commission to European Parliament, the European Council, the Council, the European Economic and Social Committee, and the Committee of the Regions. *Artificial intelligence for Europe.* {SWD (2018) 137 final})

[2] This, without considering the economic implications of AI.

[3] AI is increasingly present in the set of tools used as a defense and attack mechanism in cyberspace. See: "*Informe de Ciberamenazas y Tendencias – Edición 2019*" (Centro Criptológico Nacional; Centro Nacional de Inteligencia) y "*La inteligencia artificial aplicada a la defensa*" (Instituto Español de Estudios Estratégicos (IEEE), Documento de Trabajo 06/2018).

[4] European Economic and Social Committee. (526th Plenary Session of the EESC, May 31 to June 1, 2017). Opinion of the European Economic and Social Committee on ' *Artificial intelligence: the consequences of artificial intelligence for the single (digital) market, production, consumption, employment and society* '.

[5] https://www.csee.umbc.edu/courses/471/papers/turing.pdf

[6] https://www.muyinteresante.es/tecnologia/articulo/marvin-misnky

[7] I deprive the reader of the reproduction of the heated debates that are usually triggered after this statement, in which they inevitably end up appearing, philosophically, spiritual or religious concepts.

[8] There are several institutions that have tried to enumerate what such ethical principles should be. We point out those proposed by the European Group on Ethics of Science and New Technologies, of the European Commission, in its *Declaration on artificial intelligence, robotics and "autonomous" systems:* Human dignity; Autonomy; Responsibility, Justice, equity and solidarity; Democracy; Rule of law and accountability; Security, protection, and physical and mental integrity; Data protection and privacy and Sustainability.

[9] There is no consensus on a formal definition of AI. We like and have used Andreas Kaplan & Michael Haenlein's: "*Siri, Siri, in my hand: Who's the fairest in the land? On the interpretations, illustrations, and implications of artificial intelligence*".

[10] Especially, when it is the European Commission itself that intends to facilitate access to the latest technologies to all potential users: small and medium-sized enterprises, non-technology sector companies and public administrations, encouraging them to test them. (Communication from the Commission to the European Parliament, the European Council, the Council, the European Economic and Social Committee and the Committee of the Regions. *Artificial intelligence for Europe*)

[11] "Ethical principles" can be formulated from different points of view. The most common, however, is to make them coincide with respect for the rights formulated in the International Human Rights Letters, from the Universal Declaration of Human Rights of 1948, until our 1978 Constitution, through the EU Treaties, the Charter of Fundamental Rights of the EU, the European Convention on Human Rights or even the European Social Charter or the General Data Protection Regulation.
The European *Artificial Intelligence High-Level Expert Group* (AI HLEG) has pointed out the relationship between "fundamental principles-values-rights" as follows: "*Fundamental rights provide the basis for the formulation of ethical principles. Those principles are high-level abstract standards that developers, implementers, users and regulators must follow to defend the purpose of a man-centered and trusted AI. Values, in turn, provide more concrete guidance on how to defend ethical principles, while supporting fundamental rights*." (AI HLEG, *Ethics Guidelines for Trustworthy AI* (Draft - Dec, 2018). For the purpose of this paper, we will speak, generically, of "ethical principles".
Other bibliographical references about the principles that AI should conduct can be found in the following texts: *Asilomar AI Principles* , developed by the Future of Life Institute (2017); *Montreal Declaration for Responsible AI* , from the University of Montreal (2017) ; the second version of the general principles of the IEEE, *Ethically Aligned Design: A Vision for Prioritizing Human Well-being with Autonomous and Intelligent Systems* (2017), the *Ethical Principles* of the Group of Ethics of Science and New Technologies of the European Commission (2018); the *Five overarching principles for an AI code ,*

contained in §417 of the Report of the Committee for Artificial Intelligence of the British House of Lords (2018); and the *Tenets of the Partnership on AI* (2018).

[12] The most significant standards come from ISO (*International Standardization Organization*), IEC (*International Electrotechnical Commission*), IEEE (*Institute of Electrical and Electronics Engineers*), ITU (International Telecommunications Union) or, European scope, ETSI (*European Telecommunications Standards) Institute*).

[13] Although, in purity, the standards are part of what we have called "exogenous aspects", their generalized practical use has made them, de facto, elements consubstantial to the development of the algorithms, closer, on many occasions, to the "endogenous" components of the systems.

[14] An AI Technology , originally "harmless", inserted into an improperly configured or protected product or service could result in an undesirable product/service.

[15] Real Decreto 4/2010, of January 8, which regulates the National Interoperability Framework (ENI), already reflected the concern for extending the security guarantee to the entire life cycle of information systems, by expressing, in its article 1 .2: "*2. The National Interoperability Scheme will include the criteria and recommendations for security, standardization and conservation of information, formats and applications that must be taken into account by public administrations to ensure an adequate level of organizational, semantic and technical interoperability of the data, information and services that they manage in the exercise of their competences and to avoid discrimination against citizens because of their technological choice*."

[16] Based on the work of Stahl, Timmermans and Flick (*Ethics of Emerging Information and Communication Technologies*), we can distinguish problems that could have an impact on an **individual** level (such as autonomy, identity, dignity, privacy and data protection) and those with an impact at **social** (as the impartiality and the fairness, the collective identity and the welfare state, the responsibility, accountability and transparency, the privacy regarding surveillance , the democracy and the trust).

[17] No need for the purpose of this paper to distinguish between the concepts *hard-ethics* and *soft-ethics*, as they have been defined, for example, by Luciano Floridi in *Soft Ethics and the Governance of the Digital*.

[18] Term that we must interpret as the ability to explain the conclusions reached and their traceability.

[19] As of the writing of this work, the High-Level Expert Group on Artificial Intelligence has published a draft *Ethics Guidelines for Trustworthy AI* (December, 2018), prelude to the future final document.

[20] EESC opinion, *op. cit.*

[21] The European Parliament, in its *Civil Law Standards on Robotics* (European Parliament Resolution of February 16, 2017, with recommendations for the Commission on Civil Law Standards on Robotics (2015/2103 (INL)), it has included a *C ode ethical Conduct for robotics engineers* , whose principles could be in translatable part to a possible code of AI ethics. on the other hand, l to need for a code of conduct Ethics as guide instrument for AI developers it has also been revealed in " *Towards a Global Artificial Intelligence Charter* " (Thomas Metzinger, contained in *Should we fear artificial intelligence?* European Parliament, 2018).

[22] Vida, José; states: "... the initiatives that have been developed up now recommend increase and deepen knowledge about AI at all levels to recognize, define and control the disruptions in their development in order to regulate them properly and timely." (" *The challenges of the regulation of Artificial Intelligence: Some contributions from the European perspective* ", in *Digital Society and Law*. BOE, Nov., 2018).

[23] *UNE-EN ISO 19011: 2018 Guidelines for auditing management systems.*

[24] This also seems to be the claim of the European Economic and Social Committee, which, in the aforementioned work, states: "*The EESC advocates for an open source European AI infrastructure , which includes respectful learning environments of private life, real life test environments and high quality data sets for the development and training of AI systems. The EESC highlights the (competitive) advantage that the EU can gain in the world market by developing and promoting "European responsibility AI systems", provided with a European AI certification and labeling system.*" Adding,"*In this regard, propose using a standard system for verification, validation and control AI systems based on a wide range of standards of safety, transparency, intelligibility, surrender accounts and ethical values* ".

[25] Senior Officials Group - Information Systems Security (SOG-IS) Mutual Recognition Agreement (MRA) of Information Technology Security Evaluation Certificates.

[26] Regulation (EU) 910/2014 of the European Parliament and of the Council of 23 July 2014 on electronic identification and trust services for electronic transactions in the internal market and repealing the Directive 1999/93 / CE.

[27] Proposal for Regulation of the European Parliament and of the Council on ENISA, the "EU Cybersecurity Agency", and repealing Regulation (EU) 526/2013, and on Information and Communication Technology cybersecurity certification (''Cybersecurity Act'') .

[28] Royal Decree 3/2010, of January 8, which regulates the National Security Framework (ENS).

[29] Order PRE / 2740/2007, of September 19, which approves the Regulation of Evaluation and Certification of Information Technology Security.

[30] Despite the unquestionable benefits derived from the certification models, we must not forget, however, that a certification cannot fully guarantee that an ICT product or service is completely secure. So remember the *Cibersecurity Act* cited, in his defense of the scheme s certification technologies, products and services cybersecurity.

[31] The regulatory need is also in the Communication from the Commission to the European Parliament, the European Council, the Council, the European Economic and Social Committee and the Committee of the Regions. *Artificial intelligence for Europe ,* when it states: "*While self-regulation can provide a first set of benchmarks with respect to which it is possible to assess the applications and results that appear, public authorities must ensure that regulatory frameworks for development and the use of AI technologies are in line with those fundamental values and rights*."

[32] Activity that could very well be developed through the European Cybersecurity Certification Group.

[33] That it will be the national accreditation body designated in accordance with Regulation (EC) 765/2008 of the European Parliament and of the Council, of July 9, 2008, which establishes the requirements for accreditation and relative market surveillance to the marketing of products and repealing Regulation (EEC) 339/93 (OJ L 218 of 13.8.2008, p. 30) , for example in accordance with EN ISO / IEC 17065/2012 .

[34] "The translation of the values and the standards in the design and operation of AI systems should be part of the regulatory frameworks." (*Artificial Intelligence. A European Perspective.* Joint Research Center (JRC), the European Commission's science and knowledge service.)

[35] As the Communication from the Commission points to the European Parliament, the European Council, the Council, the European Economic and Social Committee and the Committee of the Regions. *Artificial intelligence for Europe : "*The guiding principle of all aid for AI research will be the development of a" responsible AI", centered on the human being; see the line of work of the Commission "Responsible Research and Innovation": https://ec.europa.eu/programmes/horizon2020/en/h2020-section/responsible-research-innovation