

See discussions, stats, and author profiles for this publication at: <https://www.researchgate.net/publication/337049607>

Estándares y seguridad en el uso humano de la IA | Standards and safety in the human use of AI

Preprint · November 2019

DOI: 10.13140/RG.2.2.26418.15045

CITATIONS

0

READS

2

1 author:



[Juan Antonio Lloret Egea](#)

FORMAEMPLEO

3 PUBLICATIONS 0 CITATIONS

SEE PROFILE

Some of the authors of this publication are also working on these related projects:



Los estados de la Inteligencia Artificial | The states of Artificial Intelligence [View project](#)

Capítulo 9-1º | Chapter 9-1º

Estándares y seguridad en el uso humano de la IA | [Standards and safety in the human use of AI](#)

Fig. I A. 9.1.1- Atrévete a soñar en GRANDE: los estándares empoderan a los innovadores | [Dare to dream BIG: Standards empower innovators](#). Crédito imag. (ISO). URL: https://youtu.be/N6ZLzzAZ_nQ

⇐ Ir al índice principal del libro	⇐ Go to the main index of the book
Contenidos 	Contents 
9.1-1.- Introducción	9.1-1.- Introducction
9.1-2.- La estandarización como beneficio de la IA en las industrias	9.1-2.- Standardization as a benefit of AI in industries
9.1-3.- La certificación de la inteligencia artificial	9.1-3.- The certification of artificial intelligence

Autor / [Author](#): [Juan Antonio Lloret Egea](#) | [Miembro de la Alianza Europea para la IA](#) / [Member to the European AI Alliance](#) | <https://orcid.org/0000-0002-6634-3351> | © 2019. Licencia de uso y distribución / [License for use and distribution](#): [[Los estados de la inteligencia artificial \(IA\)](#) | [The states of artificial intelligence \(AI\)](#)] [creative commons CC BY-NC-ND](#) | [ISSN 2695-3803](#) | Escrito / [Writed](#): 20/10/2019. Actualizado / [Updated](#): 20/10/2019 |

9.1-1.- Introducción | [Introducction](#)

El campo de la ética y la legalidad en la inteligencia artificial ha surgido en gran medida como una

respuesta a la gama de daños individuales y sociales que el mal uso, el abuso, el mal diseño o las consecuencias negativas no intencionadas de los sistemas de IA pueden causar, entre ellos:

Daños potenciales causados por sistemas de IA, denegación de autonomía, recurso y derechos individuales, resultados no transparentes, inexplicables o injustificables, invasiones de la privacidad, aislamiento y desintegración de la conexión social y en resultados poco confiables, inseguros o de baja calidad. Estos resultados pueden dañar directamente el bienestar de las personas y el bienestar público. También pueden socavar la confianza en el uso responsable de las tecnologías de IA beneficiosas para la sociedad. Asimismo hemos de tener en consideración aspectos como el conjunto de regulaciones y leyes que sean de aplicación de acuerdo al ordenamiento jurídico y a las normas técnicas para el diseño y construcción y uso de los servicios y procesos resultantes. ^{C9.1-1}

[English]

The field of ethics and legality in artificial intelligence has largely emerged as a response to the range of individual and social damages that misuse, abuse, bad design or unintended negative consequences of AI systems they can cause, and among them:

Potential damages caused by AI systems, denial of autonomy, recourse and individual rights, non-transparent, unexplained or unjustifiable results, invasions of privacy, isolation and disintegration of the social connection and in unreliable, insecure or poor quality results . These results can directly harm people's well-being and public welfare. They can also undermine confidence in the responsible use of AI technologies beneficial to society. We must also take into consideration aspects such as the set of regulations and laws that are applicable according to the legal system and technical standards for the design and construction and use of the resulting services and processes ^{C9.1-1}.

En la actualidad casi todo lo relacionado con la Inteligencia Artificial o está en vías de desarrollo o en proyectos que tratan de buscar las soluciones más adecuadas canalizando y focalizando distintos ángulos y puntos de vista que son aportados por diferentes estamentos de la política, la empresa y la sociedad. Algunos de los argumentos más consolidados son las **directrices marcadas por la UE**; o por empresas más notables relacionadas con este campo, como **IBM**, **Google** o **Microsoft**. Al respecto en este libro hemos recopilado los siguientes contenidos en el **capítulo 4**:

1. **La ética de la inteligencia artificial.**
2. **La ética de la IA en Europa.**
3. **Economía y antirreglas de la ética.**
4. **Sesgo racial en el procesamiento del lenguaje natural.**

5. [Confianza en la IA empresarial.](#)

[English]

Currently, almost everything related to Artificial Intelligence or is in the process of development or in projects that seek to find the most appropriate solutions by channeling and focusing different angles and points of view that are contributed by different levels of politics, business and the society. Some of the most consolidated arguments are the [guidelines set by the EU](#); or by more notable companies related to this field, such as [IBM](#), [Google](#) or [Microsoft](#). In this regard, we have compiled the following contents in [chapter 4](#):

1. [Ethics of Artificial Intelligence.](#)
 2. [The ethics of AI in Europe.](#)
 3. [Economy and anti-rules of ethics.](#)
 4. [Racial bias in natural language processing.](#)
 5. [Trust in business AI.](#)
-

A lo que se añaden **estándares técnicos** para que el arco de posibilidades que propone la IA tenga un adecuado sustrato futuro de desarrollo. Algunos de ellos, los más notables, son:

Los creados o en desarrollo por el [IEEE](#): de la familia de estándares P7000, grupo de trabajo de metodologías de ingeniería para asuntos éticos del ciclo de vida.

Alcance: el estándar establece un modelo de proceso mediante el cual los ingenieros y tecnólogos pueden abordar la consideración ética a lo largo de las diversas etapas de iniciación, análisis y diseño del sistema. Los requisitos esperados del proceso incluyen la visión de gestión e ingeniería del desarrollo de nuevos productos de TI, la ética informática y el diseño del sistema de TI, el diseño sensible al valor y la participación de los interesados en el diseño ético del sistema de TI.

[English]

To which technical standards are added so that the arc of possibilities proposed by AI has an adequate future development substrate. Some of them, the most notable, are:

Those created or under development by the IEEE: from the P7000 family of standards, a working group of engineering methodologies for ethical life cycle issues.

Scope: the standard establishes a process model through which engineers and technologists can address ethical consideration throughout the various stages of system initiation, analysis and design. The expected process requirements include the management and engineering vision of the development of new IT products, IT ethics and IT system design, value-sensitive design and stakeholder participation in the ethical design of the IT system.

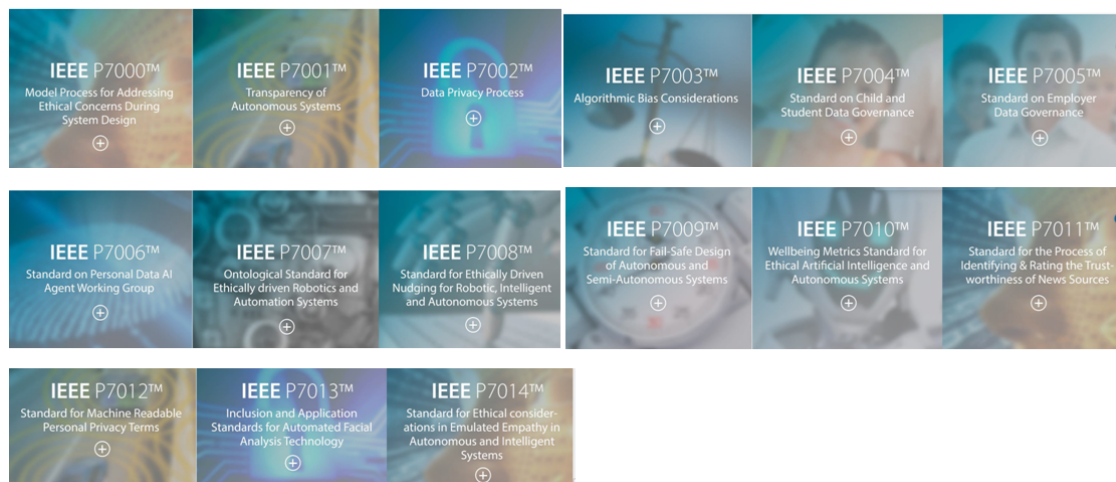


Fig. I A. 9.1.2- Familia de estándares P7000 | P7000 family of standards. Crédito imag. (Propia | Own). URL de origen: <https://ethicsinaction.ieee.org/>

P70XX: (01,02... 14)

1. Transparencia de sistemas autónomos | [Transparency of Autonomous Systems](#)
2. Proceso de privacidad de datos | [Data Privacy Process](#)
3. Consideraciones de sesgo algorítmico | [Algorithmic Bias Considerations](#)
4. Norma sobre gobierno de datos de niños y estudiantes | [Standard on Child and Student Data Governance](#)
5. Norma sobre gobierno de datos del empleador | [Standard on Employer Data Governance](#)
6. Grupo de trabajo sobre el agente de AI de datos personales (estándar) | [Standard on Personal Data AI Agent Working Group](#)
7. Norma ontológica para sistemas de automatización y robótica con ética | [Ontological](#)

- [Standard for Ethically driven Robotics and Automation Systems](#)
8. Estándar para empujar éticamente sistemas robóticos, inteligentes y autónomos | [Standard for Ethically Driven Nudging for Robotic, Intelligent and Autonomous Systems](#)
 9. Norma para el diseño a prueba de fallos de sistemas autónomos y semiautónomos | [Standard for Fail-Safe Design of Autonomous and Semi-Autonomous Systems](#)
 10. Estándar de métricas de bienestar para inteligencia artificial ética y sistemas autónomos | [Wellbeing Metrics Standard for Ethical Artificial Intelligence and Autonomous Systems](#)
 11. Proceso de identificación y calificación de la confiabilidad de las fuentes de noticias | [Process of Identifying and Rating the Trustworthiness of News Sources](#)
 12. Términos de privacidad personal legibles por máquina | [Machine Readable Personal Privacy Terms](#)
 13. Estándares para la tecnología de análisis facial automatizado | [Standards for Automated Facial Analysis Technology](#)
 14. Norma para consideraciones éticas en la empatía emulada en sistemas autónomos e inteligentes | [Standard for Ethical considerations in Emulated Empathy in Autonomous and Intelligent Systems](#)

Los **creados o en desarrollo** por el subcomité **ISO/IEC JTC1/SC42** (sin incluir aquí los relacionados con el *Bigdata*).

Este subcomité se estableció en noviembre de 2017 para abordar los requisitos específicos de estandarización de AI. El alcance de JTC1 / SC42 es servir como entidad de enfoque y proponente del programa de estandarización de JTC1 sobre IA, y proporcionar orientación a los comités JTC1, IEC e ISO que desarrollan aplicaciones relacionadas con AI. Los temas que forman parte del trabajo de este subcomité son: Estándares fundamentales, enfoques computacionales y características de IA, confiabilidad, casos de uso y aplicaciones, *Big data* y Preocupaciones sociales ^{C9.1-2}.

[English]

Those created or under development by subcommittee **ISO/IEC JTC1/SC42** (not including those related to Bigdata here).

This subcommittee was established in November 2017 to address the specific requirements of AI standardization. The scope of JTC1 / SC42 is to serve as a focus entity and proponent of the JTC1 standardization program on AI, and provide guidance to JTC1, IEC and ISO committees that develop AI

related applications. The topics that are part of the work of this subcommittee are: Fundamental standards, computational approaches and characteristics of AI, reliability, use cases and applications, Big data and social concerns ^{C9.1-2}.

Inteligencia artificial — conceptos y terminología	Artificial intelligence — Concepts and terminology	ISO/IEC WD 22989
Marco para sistemas de inteligencia artificial (IA) que utilizan el aprendizaje automático (ML)	Framework for Artificial Intelligence (AI) Systems Using Machine Learning (ML)	ISO/IEC WD 23053
Tecnología de la información — Inteligencia artificial — Gestión de riesgos	Information Technology — Artificial Intelligence — Risk Management	ISO/IEC AWI 23894
Tecnología de la información — Inteligencia artificial (IA) — Sesgo en los sistemas de IA y la toma de decisiones asistida por AI	Information technology — Artificial Intelligence (AI) — Bias in AI systems and AI aided decision making	ISO/IEC NP TR 24027
Tecnología de la información — Inteligencia artificial (IA) — Descripción general de la confiabilidad en inteligencia artificial	Information technology — Artificial Intelligence (AI) — Overview of trustworthiness in Artificial Intelligence	ISO/IEC PDTR 24028
Inteligencia artificial (IA) — Evaluación de la robustez de las redes neuronales — Parte 1: Descripción general	Artificial Intelligence (AI) — Assessment of the robustness of neural networks — Part 1: Overview	ISO/IEC NP TR 24029-1
Tecnología de la información — Inteligencia artificial (IA) — Casos de uso	Information technology — Artificial Intelligence (AI) — Use cases	ISO/IEC NP TR 24030
Tecnología de la información — Inteligencia artificial — Visión general de las preocupaciones éticas y sociales	Information technology — Artificial intelligence — Overview of ethical and societal concerns	ISO/IEC AWI TR 24368
Tecnología de la información — Inteligencia artificial (IA) — Descripción general de los enfoques computacionales para sistemas de IA	Information technology — Artificial intelligence (AI) — Overview of computational approaches for AI systems	ISO/IEC AWI TR 24372
Tecnología de la información — Inteligencia artificial — Marco de gestión de procesos para análisis de Big Data	Information technology — Artificial intelligence — Process management framework for Big data analytics	ISO/IEC AWI 24668
Tecnología de la información — Gobernanza de TI - Implicaciones de gobernanza del uso de inteligencia artificial por parte de las organizaciones	Information technology — Governance of IT — Governance implications of the use of artificial intelligence by organizations	ISO/IEC AWI 38507

9.1-2.- La estandarización como beneficio de la IA en las industrias | Standardization as a benefit of AI in industries

Cómo la estandarización europea beneficia el despliegue de IA

La estandarización europea es un elemento esencial para evitar la fragmentación en el Mercado Único Digital y garantizar un enfoque centrado en las personas para la IA al garantizar que beneficie a la sociedad. Los estándares CEN y CENELEC crean un entorno de confianza y responsabilidad en torno al desarrollo y uso de la tecnología de inteligencia artificial, abordando problemas como la fiabilidad, la seguridad, la privacidad y la seguridad. Además, CEN y CENELEC tienen fuertes lazos con sus organismos internacionales de estandarización, la Organización Internacional de Normalización (ISO) y la Normalización Electrotécnica Internacional (IEC), donde el trabajo ya comenzó en IA en ISO / IEC JTC 1 SC 42 ^{C9.1-3}.

[English]

How European standardization benefits AI deployment

European standardization is an essential element to avoid fragmentation in the Digital Single Market and ensure a people-centered approach to AI by ensuring that it benefits society. The CEN and CENELEC standards create an environment of trust and responsibility around the development and use of artificial intelligence technology, addressing issues such as reliability, security, privacy and security. In addition, CEN and CENELEC have strong ties with their international standardization bodies, the International Organization for Standardization (ISO) and International Electrotechnical Standardization (IEC), where work has already begun in AI in ISO / IEC JTC 1 SC 42 ^{C9.1-3}.

Fig. I A. 9.1.3- Ficha informativa sobre inteligencia artificial | [Factsheet on artificial intelligence](#). Crédito imag. (Digital Single Market). URL: <https://ec.europa.eu/digital-single-market/news-redirect/624272>

A continuación extractamos y recogemos el *whitepaper*, *Artificial intelligence across industries* del IEC (International Electrotechnical Commission), en el que es notable su estudio sobre estándares en la sección 7, y sus conclusiones, sección 8.

El documento técnico fue desarrollado por la Junta de Estrategia de Mercado de IEC (MSB) con importantes contribuciones del Grupo Haier y el socio del proyecto del Centro de Investigación Alemán para la Inteligencia Artificial (DFKI). Los miembros del equipo de apoyo del proyecto incluyeron SAP, Huawei, NSW Data Analytics Center (DAC), China Electronic Standardization Institute (CESI), LG Electronics y Korea Electric Power Corporation (KEPCO)^{C9.1-2}.

[English]

Next we extract and collect the white paper, Artificial intelligence across industries of the IEC (International Electrotechnical Commission), in which its study on standards in section 7, and its conclusions, section 8 is remarkable.

The technical document was developed by the IEC Market Strategy Board (MSB) with significant contributions from the Haier Group and the project partner of the German Research Center for Artificial Intelligence (DFKI). Project support team members included SAP, Huawei, NSW Data Analytics Center (DAC), China Electronic Standardization Institute (CESI), LG Electronics and Korea Electric Power Corporation (KEPCO)^{C9.1-2}.

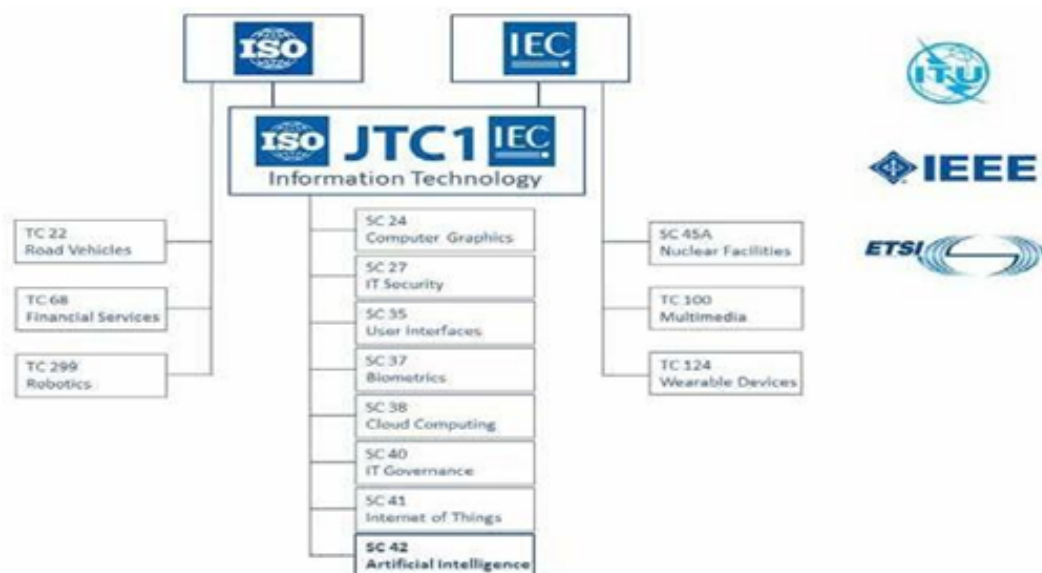


Fig. I A. 9.1.4- Paisaje de estandarización en IA | [Standardization landscape on AI](#). Crédito imag. (IEC).

URL: <https://basecamp.iec.ch/download/iec-white-paper-artificial-intelligence-across-industries-en/>

La IA está generando serios desafíos en los campos sociales, económicos y regulatorios. Ya sea que esté relacionado con la confiabilidad, la privacidad, la seguridad de las interacciones hombre /

máquina, o el impacto en la fuerza laboral, una amplia base de interesados deberá abordar estos problemas con urgencia y de manera coordinada. Como estos desafíos serán comunes a todas las naciones del mundo, el Libro Blanco también ha insistido en la necesidad de un enfoque internacional para abordarlos. En particular, los gobiernos, los reguladores y los encargados de formular políticas deberán comprender y abordar cuidadosamente los principales problemas éticos que están surgiendo con el rápido despliegue de la IA. La industria y la comunidad de investigación también tendrán que abordar una serie de desafíos técnicos que pueden impedir el despliegue de IA en algunos dominios de aplicación.

Los problemas relacionados con los datos, los algoritmos, el *hardware* y las informáticas de infraestructuras se encuentran entre las limitaciones más difíciles de la actualidad para explotar todo el potencial y lograr las promesas más emocionantes de la IA. Desde esa perspectiva, se espera que la estandarización y la evaluación de la conformidad jueguen un papel instrumental para facilitar la adopción del mercado y la aceptación social de las tecnologías emergentes de IA ^{C9.1-2}.

[English]

AI is generating serious challenges in the social, economic and regulatory fields. Whether it is related to reliability, privacy, the safety of man / machine interactions, or the impact on the workforce, a broad base of stakeholders should address these issues urgently and in a coordinated manner. As these challenges will be common to all nations of the world, the White Paper has also insisted on the need for an international approach to address them. In particular, governments, regulators and policy makers should carefully understand and address the main ethical issues that are emerging with the rapid deployment of AI. The industry and the research community will also have to address a series of technical challenges that may prevent the deployment of AI in some application domains.

Problems related to data, algorithms, hardware and infrastructure infrastructures are among the most difficult limitations of today to exploit the full potential and achieve the most exciting promises of AI. From that perspective, standardization and conformity assessment is expected to play an instrumental role in facilitating market adoption and social acceptance of emerging AI technologies. ^{C9.1-2}.

Recomendaciones:

- Desarrollar pautas para la seguridad dentro de los sistemas de la IA

La IA tiene un profundo impacto en la seguridad de los sistemas de información. Por un lado, estos se pueden usar para crear nuevos tipos de hack bots o virus que se adaptan constantemente a las nuevas contramedidas. Por otro lado, también pueden crear agujeros de seguridad adicionales para las empresas al proporcionar nuevos puntos de acceso para ataques. Los datos de capacitación podrían usarse para tratar de influir en el sistema controlando específicamente un sesgo en los conjuntos de datos. Por lo tanto, es necesario que las empresas identifiquen los efectos de seguridad

que tendrá el uso de IA antes de aplicarlos, a fin de tomar las medidas de seguridad apropiadas.

- Regulatorias

Los despliegues de IA van acompañados de una incertidumbre regulatoria sustancial, que será vista por la industria y la comunidad de investigación como un obstáculo para invertir en IA. Además, es obvio que la IA contribuye a proyectos gubernamentales específicos, como el Mercado Único Digital en Europa. Por lo tanto, el despliegue exitoso de AI en el mercado requerirá una actividad reguladora coordinada entre múltiples partes interesadas (industria, gobiernos, comunidad de investigación, pymes, organizaciones de consumidores y otros). Se recomiendan las actividades reguladoras que se describen a continuación.

- Abordar la IA dentro de las políticas de privacidad

El GDPR europeo está abordando algunos aspectos del abuso de los derechos de privacidad individuales, como la toma de decisiones automatizada y la elaboración de perfiles sin el consentimiento del usuario. Sin embargo, sigue habiendo una serie de problemas de privacidad que deben abordarse.

Existe evidencia de que algunos piratas informáticos están realizando ingeniería inversa de los datos de capacitación para recuperar datos de privacidad de individuos y haciendo un mal uso de estos datos confidenciales. Además, los procesos de diseño y capacitación de algoritmos deben ser inclusivos y deben considerar plenamente los intereses de los grupos desfavorecidos. Deben establecer reglas especiales para situaciones éticas extremas. A ese respecto, las normas, leyes y políticas pertinentes sobre seguridad, ética y privacidad deberían mejorarse urgentemente. Cuando la IA se distribuye dentro de la nube, será fundamental que las personas tengan suficiente confianza en el proceso para enviar sus datos privados a la nube sin preocuparse por la violación de su privacidad. Además, los desarrolladores de IA deben adoptar los medios técnicos adecuados para proteger a las personas en el proceso de recopilación y uso de datos y para evitar que la información personal esté sujeta a fugas, falsificación y daños.

- Incluir la IA en la seguridad, la ética y otras políticas

Si bien la industria actual de IA se encuentra en un período de desarrollo vigoroso, debe considerar seriamente la formulación e implementación de futuras medidas de supervisión de la industria y debe considerar plenamente las responsabilidades y fallas en el proceso de desarrollo de IA y despliegue.

La cuestión de la supervisión de la industria de IA no es un problema que enfrenta un solo grupo. Involucra una amplia gama de problemas sociales, sistémicos y complejos que requieren que empresas, gobiernos, usuarios, organizaciones tecnológicas y otros terceros participen y trabajen juntos para construir un ecosistema de IA sostenible.

Las normas de seguridad relevantes deben formularse y mejorarse. Los requisitos éticos de la tecnología de IA deben establecerse de acuerdo con un amplio consenso formado entre la sociedad

[English]

Recommendations:

- Develop guidelines for security within AI systems

AI has a profound impact on the security of information systems. On the one hand, these can be used to create new types of hack bots or viruses that constantly adapt to new countermeasures. On the other hand, they can also create additional security holes for businesses by providing new access points for attacks. Training data could be used to try to influence the system by specifically controlling a bias in the data sets. Therefore, it is necessary for companies to identify the security effects that the use of AI will have before applying them, in order to take appropriate security measures.

- Regulatory

AI deployments are accompanied by substantial regulatory uncertainty, which will be seen by industry and the research community as an obstacle to investing in AI. In addition, it is obvious that AI contributes to specific government projects, such as the Digital Single Market in Europe. Therefore, the successful deployment of AI in the market will require a coordinated regulatory activity among multiple stakeholders (industry, governments, research community, SMEs, consumer organizations and others). The regulatory activities described below are recommended.

- Address AI within privacy policies

The European GDPR is addressing some aspects of the abuse of individual privacy rights, such as automated decision making and profiling without user consent. However, there are still a number of privacy issues that must be addressed.

There is evidence that some hackers are reverse engineering training data to retrieve privacy data from individuals and misusing these confidential data. In addition, the algorithm design and training processes must be inclusive and fully consider the interests of disadvantaged groups. They must establish special rules for extreme ethical situations. In that regard, the relevant rules, laws and policies on security, ethics and privacy should be urgently improved. When AI is distributed within the cloud, it will be essential that people have sufficient confidence in the process to send their private data to the cloud without worrying about the violation of their privacy. In addition, AI developers must adopt appropriate technical means to protect people in the process of data collection and use and to prevent personal information from being subject to leakage, forgery and damage.

- Include AI in security, ethics and other policies

While the current AI industry is in a period of vigorous development, it must seriously consider the formulation and implementation of future industry monitoring measures and must fully consider the

responsibilities and failures in the AI development and deployment process.

The question of AI industry supervision is not a problem faced by a single group. It involves a wide range of social, systemic and complex problems that require companies, governments, users, technology organizations and other third parties to participate and work together to build a sustainable AI ecosystem.

The relevant safety standards must be formulated and improved. The ethical requirements of AI technology must be established in accordance with a broad consensus formed between society as a whole ^{C9.1-2}.

El *whitepaper* también se ha embebido para un mejor conocimiento del trabajo del IEC en el campo de la IA (a menudo conjunto y compartido con ISO) y por su naturaleza ejemplificante y educativa. (Este *whitepaper* está protegido por *copyright* y está alojado en su web de origen para su descarga libre y no está afectado por la licencia Creative Commons de este libro. Consulte *Notas legales*, por favor). | The whitepaper has also been embedded for a better understanding of the work of the IEC in the field of AI (often joint and shared with ISO) and for its exemplary and educational nature. (This whitepaper is copyrighted and is hosted on its source website for free download and is not affected by the Creative Commons license in this book. See Legal Notes, please).

Fig. I A. 9.1.5- Libro blanco inteligencia artificial en todas las industrias | [Whitepaper Artificial intelligence across industries](https://basecamp.iec.ch/download/iec-white-paper-artificial-intelligence-across-industries-en/). Crédito imag. (IEC). URL: <https://basecamp.iec.ch/download/iec-white-paper-artificial-intelligence-across-industries-en/>

9.1-3.- La certificación de la inteligencia artificial | [The certification of artificial intelligence](#)

Es de importante interés a juicio nuestro traer aquí la norma sobre la *Evaluación de la conformidad. Requisitos para organismos que certifican productos, procesos y servicios. ISO/IEC 17065:2012* (que citamos aquí mediante UNE).

Esta norma internacional contiene requisitos para la competencia, la consistencia de las actividades y la imparcialidad de los organismos de certificación de productos, procesos y servicios. No es necesario que los organismos de certificación que operan bajo esta norma internacional ofrezcan todos los tipos de certificación de productos, procesos y servicios. La certificación de productos, procesos y servicios es una actividad de evaluación de la conformidad de tercera parte. En esta norma el término 'producto' se puede entender como 'proceso' o 'servicio', excepto en aquellos casos en que se establecen disposiciones independientes para 'procesos' o 'servicios' ^{C9.1-4}.

[English]

It is of important interest in our judgment to bring the rule on Conformity assessment - Requirements for bodies certifying products, processes and services. [ISO/IEC 17065:2012](#) (which we cite here through UNE).

This international standard contains requirements for the competence, consistency of activities and impartiality of certification bodies for products, processes and services. It is not necessary for certification bodies operating under this international standard to offer all types of product, process and service certification. The certification of products, processes and services is a third party conformity assessment activity. In this standard the term 'product' can be understood as 'process' or 'service', except in those cases where independent provisions are established for 'processes' or 'services' ^{C9.1-4}.

Asimismo por la ortodoxia, su carácter ilustrativo y educativo, y el análisis que propone, extractamos y embebemos el trabajo denominado *La certificación como mecanismo de control de la inteligencia artificial en Europa*, del doctor en informática Carlos Galán.

Sobre las conclusiones que efectúa el PhD. Carlos Galán entendemos de vital importancia que en España se incluya la certificación como parte de la Estrategia Española de I+D+i en Inteligencia Artificial, y que nosotros hemos recogido en el capítulo 3 en los apartados:

- [Estrategia española en I+D+i en IA](#)
 - [Posición española en tecnologías de la IA](#)
-

[English]

Also by the orthodoxy, its illustrative and educational character, and the analysis that it proposes, we extract and embed the work called Certification as a mechanism of control of artificial intelligence in Europe, by the computer doctor Carlos Galán.

On the conclusions made by the PhD. Carlos Galán We understand that it is vitally important that in Spain certification be included as part of the Spanish Strategy for R & D & I in Artificial Intelligence,

and that we have gathered in Chapter 3 in the sections:

- [Spanish strategy in R+D+i in AI](#)
 - [Spanish position in AI technologies](#)
-

Resumen:

«La utilización de la inteligencia artificial (IA) constituye una de las más significativas aportaciones tecnológicas que impregnará la vida de las sociedades occidentales de los próximos años, en muchas de sus actividades cotidianas y en sus sectores más representativos, desde la industria al sistema financiero, pasando por la educación, la salud, el transporte y, desde luego, la defensa y la seguridad, aportando significativos beneficios, pero evidenciando también riesgos que es necesario valorar y minimizar. Una realidad tan disruptiva como la IA exige que su tecnología y la de los productos y servicios sustentados en ella ofrezcan suficientes garantías de su adecuado funcionamiento. El presente trabajo analiza y hace una propuesta para la aplicación de los mecanismos de Certificación de la Conformidad al mantenimiento de las antedichas garantías»^{C9.1-5}.

[\[English\]](#)

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 aforementioned guarantees»^{C9.1-5}.

Conclusiones y acciones subsiguientes

- Siendo la IA un elemento indispensable para el desarrollo de la sociedad, es necesario garantizar, hasta donde sea posible, que la tecnología IA, los productos IA y los servicios IA se ajustan a lo dispuesto en la regulación legal que resulte de aplicación, manteniendo un escrupuloso respeto a los principios éticos universalmente aceptados.
- La garantía anterior sólo puede alcanzarse con una adecuada combinación de dos tipos de controles: *ex-ante* y *ex-post*, antes y después, respectivamente, del diseño, desarrollo, implementación o despliegue de la tecnología, productos o servicios implicados.
- El control *ex-ante*, aunque presenta indudables y lógicas ventajas, es complejo y está sometido, en cualquier caso, a la buena voluntad de los diseñadores, desarrolladores o implementadores.

- El control *ex-post*, materializado en forma de Auditorías independientes y Certificaciones de Conformidad, se configura como un mecanismo práctico y eficaz para alcanzar los objetivos.
- La implantación de las Auditorías de Conformidad IA exige disponer de un Esquema Europeo de Certificación de Tecnologías, Productos y Servicios IA; así como un elenco de Entidades de Evaluación de la Conformidad que dispongan de las capacidades técnicas requeridas, la independencia y la Imparcialidad debidas, y hayan sido debidamente habilitadas por una Entidad de Acreditación.
- Todo ello deberá estar perfectamente regulado a través de la normativa europea correspondiente.
- Las tecnologías, productos o servicios IA pueden construirse dentro de la UE o provenir de terceros países. La regulación anterior debe aplicarse a cualquier tecnología, producto o servicio de este tipo que pretenda desplegarse o comercializarse en Europa. Solo la tecnología, los productos o los servicios IA certificados deberían beneficiarse de ayudas públicas.
- Los esquemas de certificación de tecnologías, productos y servicios IA deben incorporarse a los planes derivados de la Estrategia Española de I+D+i en Inteligencia ^{C9.1-5}.

[English]

Conclusions and 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 Technologies, 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; 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
- All of this must be perfectly regulated, through the corresponding European regulations.
- 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 his type that is intended to be deployed or marketed in Europe. Only technology, certified AI products or services should benefit from public aid.
- 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 ^{C9.1-5}.



Fig. I A. 9.1.4-**A** [ES]- Modelo de certificación propuesto en el documento *La certificación como mecanismo de control de la inteligencia artificial en Europa*^{C9.1-5}. URL:

http://www.ieee.es/Galerias/fichero/docs_opinion/2019/DIEEEO46_2019CARGAL-InteligenciaArtificial.pdf

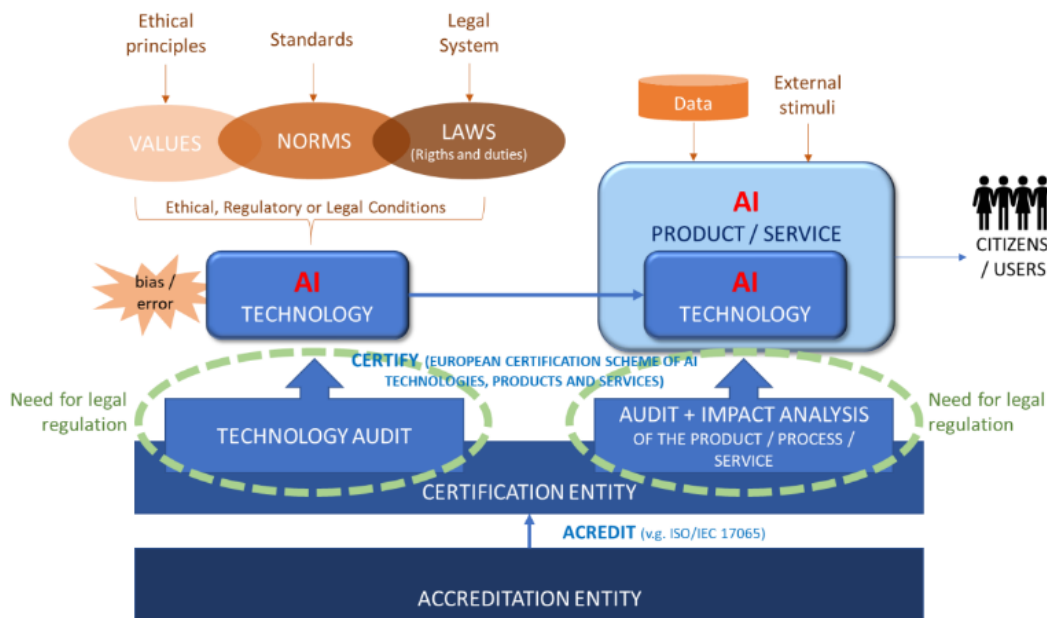


Fig. I A. 9.1.4-**B** [EN]- *Certification model proposed in the document Certification as a mechanism for controlling artificial intelligence in Europe*

Fig. I A. 9.1.5- *La certificación como mecanismo de control de la inteligencia artificial en Europa* | *The certification as a mechanism for control of artificial intelligence in Europe*. URL:

http://www.ieee.es/Galerias/fichero/docs_opinion/2019/DIEEEO46_2019CARGAL-InteligenciaArtificial.pdf

Bibliografía | Bibliography

[C9.1-1]. Leslie, D. (2019). *Understanding artificial intelligence ethics and safety: A guide for the responsible design and implementation of AI systems in the public sector*. The Alan Turing Institute. <https://doi.org/10.5281/zenodo.3240529>

[C9.1-2]. IEC. *Artificial intelligence across industries*. Whitepaper. [Recuperado (20/10/2019) de <https://basecamp.iec.ch/download/iec-white-paper-artificial-intelligence-across-industries-en/>]

[C9.1-3]. CEN / CENELEC. New initiative on Artificial Intelligence of the European Commission recognizes Standards as a competitive advantage for European industries. Brief News. [Recuperado (20/10/2019) de https://www.cencenelec.eu/news/brief_news/Pages/TN-2018-016.aspx]

[C9.1-4]. UNE. UNE-EN ISO/IEC 17065:2012. [Recuperado (20/10/2019) de <https://www.une.org/encuentra-tu-norma/busca-tu-norma/norma/?c=N0050466>]

[C9.1-5]. GALÁN, Carlos. *La certificación como mecanismo de control de la inteligencia artificial en Europa*. Documento de Opinión IEEE 46/2019.

http://www.ieee.es/Galerias/fichero/docs_opinion/2019/DIEEEO46_2019CARGAL-InteligenciaArtificial.pdf . [Consultado 18/10/2019].

© 2019. Licencia de uso y distribución / License for use and distribution: [[Los estados de la inteligencia artificial \(IA\)](#) | [The states of artificial intelligence \(AI\)](#)] [creative commons CC BY-NC-ND](#) | ISSN 2695-3803|

- [Notas legales](#) / [Legal notes](#)
 - [Página web de Formaempleo](#) / [Formaempleo website](#)
 - [Formulario de contacto](#) / [Contact Form](#)
-

Revisión #13

Creado el Sun, Oct 20, 2019 9:24 PM por [Antonio Seu Rueda](#)

Actualizado el Sun, Oct 27, 2019 2:37 PM por [Juan Antonio Lloret Egea](#)