Inteligencia Artificial y Administraciones Públicas

Juan Antonio Carrillo Donaire

donaire@us.es

ÍNDICE

1. CUESTIONES GENERALES, RIESGOS ASOCIADOS Y CLAVES DE FIABILIDAD
2. LA REGULACIÓN DE LA INTELIGENCIA ARTIFICIAL: PUNTO DE PARTIDA Y CUESTIONES A REGULAR. VIABILIDAD JURÍDICA Y MATERIAL DE LA TRANSFORMACIÓN DIGITAL DE LA ADMINISTRACIÓN PÚBLICA
3. LA DEFINICIÓN DE UN MODELO REGULATORIO EUROPEO BASADO EN LA GESTIÓN DE RIESGOS
4. EL PROBLEMA ÉTICO DE LA INTELIGENCIA ARTIFICIAL Y LA RAZÓN JURÍDICA: PROBLEMAS, IDEAS Y CONCLUSIONES DEL MODELO REGULATORIO

I. INTELIGENCIA ARTIFICIAL: CUESTIONES GENERALES

IA: "*sistemas capaces de procesar datos e información de una manera que se asemeja a un comportamiento inteligente, y abarca generalmente aspectos de (i) razonamiento, (ii) aprendizaje, (iii) percepción, (iv) predicción, (v) planificación o (vi) control*" (Res. UNESCO Nov. 2021 "La Ética en la IA").

Confusión conceptual: Inteligencia Artificial ≠ Digitalización ≠ Automatización ≠ Big Data ≠ software ≠ Machine Learning, etc…

Inteligencia artificial = Modelos y sistemas matemáticos algorítmicos + Datos digitalizados y etiquetados

Algoritmo: conjunto ordenado y finito de operaciones que permite hallar la solución a un problema. Más específicamente: **operaciones matemáticas sobre datos que integra y gestiona un software (modelo) al que se le asignan objetivos**. Se distingue:

1º) Por lo que hacen: priorizan (búsquedas en internet), clasifican (ofertan o retiran contenidos, etiquetan), asocian o correlacionan (encentran vínculos -reconocimiento facial-), o filtran (aíslan lo importante, -reconocimiento de voz). Lo normal es que hagan varias de estas funciones a la vez.

2º) Por cómo lo hacen: basados en reglas ≠ de aprendizaje automático –requieren factor tiempo-. Dentro de éstos últimos: Predictivos ≠ no predictivos.

Algoritmos predictivos: modelo matemático alimentado con **datos del pasado**, sobre los que se buscan correlaciones y se obtienen criterios y reglas.

3º) Para qué lo hacen:

Algoritmo aplicativo basado en el automatismos reglado: simplemente "ayuda a decidir" (p.ej. aplicaciones AEAT que liquidan tributos)

Algoritmo aplicativo complejo con elementos ponderativos y variables: el algoritmo proporciona el dato para la toma de decisión ("Euphemia": https://www.energynews.es/euphemia-algoritmo-calcula-precio-luz/) o "BOSCO"

Algoritmo predictivo. El algoritmo es capaz de decidir o de proporcionar el elemento valorativo de la decisión (p.ej. "RisCanvi" o "Viogen")

II. INTELIGENCIA ARTIFICIAL ADMINISTRACIÓN DE JUSTICIA

MINISTERIO
DE JUSTICIA

Anteproyecto de Ley de Medidas de Eficiencia Digital del Servicio Público de Justicia, por la que se transpone al ordenamiento jurídico español la Directiva (UE) 2019/1151 del Parlamento Europeo y del Consejo, de 20 de junio de 2019, por la que se modifica la Directiva (UE) 2017/1132 en lo que respecta a la utilización de herramientas y procesos digitales en el ámbito del Derecho de sociedades.

Las actuaciones automatizadas se regulan haciendo previsiones específicas para su uso para tareas repetitivas y automatizables (paginado de expedientes o declaración de firmeza, por ejemplo), pero también estableciendo límites. Estas actuaciones se ven especialmente facilitadas a través de la orientación al dato, que permite que algunas cuestiones que antes no se podían hacer de manera automática porque requerían que se leyera un documento, se procesara y se hicieran tareas en base a la información, se puedan ahora realizar de manera automática (por ejemplo, cálculo de plazos en base a fechas que aparecen como datos, comprobaciones automáticas de situación concursal de una empresa, en base a NIF y tipo de proceso judicial, etc.).

Igualmente se regulan, como subtipo de actuaciones automatizadas, las actuaciones proactivas, que aprovechan la información incorporada con un fin determinado, para generar efectos o avisos a otros fines distintos. Por ejemplo, notificaciones o avisos automáticos, sin necesidad de intervención manual.

Finalmente, como tipo diferenciado de las anteriores, pero con algunos requisitos en común, se definen las actuaciones asistidas, que generan un borrador total o parcial de texto, que puede servir de apoyo a la tarea del Juez o Jueza, Magistrado o Magistrada, Fiscal y Letrado o Letrada de la Administración de Justicia, manteniendo éstos siempre pleno control sobre el texto y sin que el borrador se constituya en resolución sin la intervención del operador. Para estas tareas, es fundamental la orientación al dato (para posibilitar la generación de borradores de calidad), así como la aplicación de nuevas tecnologías, como la Inteligencia Artificial.

I. LA DISRUPCIÓN DE LOS ALGORITMOS



REAL ACADEMIA ESPAÑOLA

disrupción

Del ingl. *disruption*, y este del lat. *disruptio, -ōnis*, var. de *diruptio, -ōnis* 'rotura, fractura'.

1. f. Rotura o interrupción brusca.

Quiebra brusca en **modelos de producción y mercado**: economía digital.

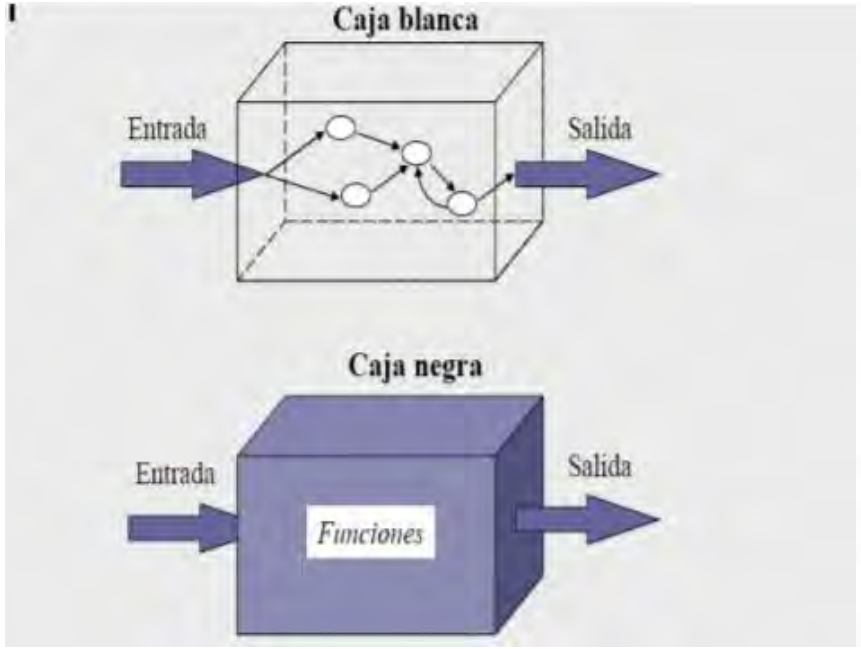
Quiebra en **modelos relacionales** (Persona-Persona – Administrado/Administración).

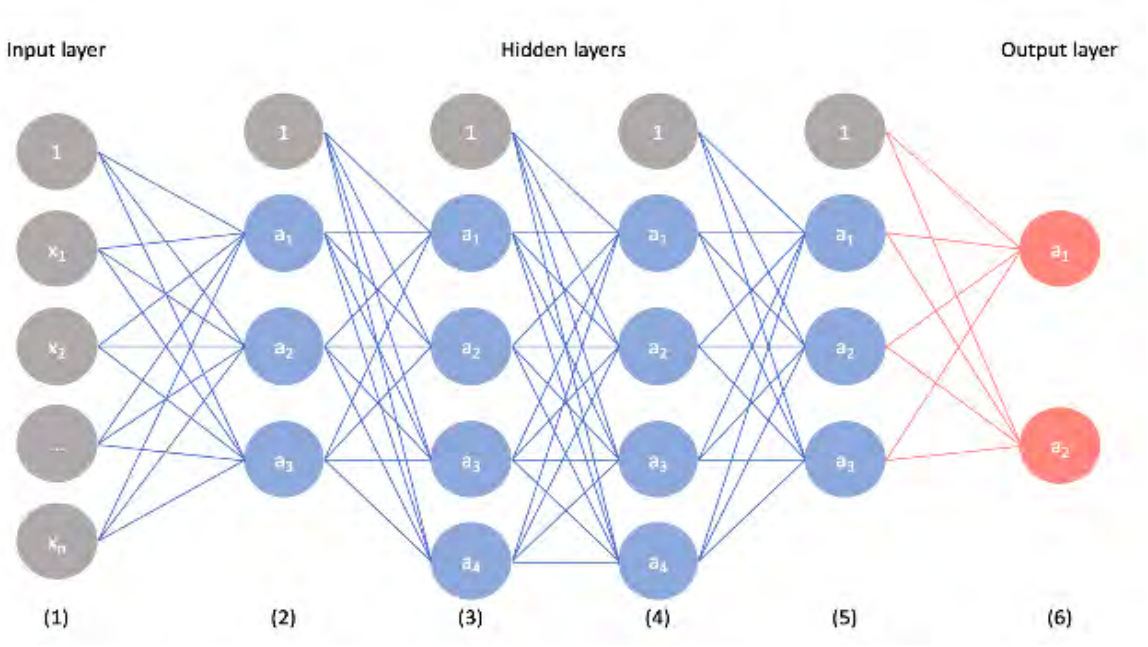
Quiebra brusca de **Derechos fundamentales**:

Derecho a la intimidad (protección de datos personales)
Derecho a la propia imagen (protección de datos personales).
Derecho a la tutela judicial efectiva (presunción de inocencia, inmediación judicial, defensa, acceso a la justicia…etc.).
Derecho al libre desarrollo de la personalidad (trabajo).
Derecho a la igualdad y no discriminación.
Derecho a la seguridad e integridad física.
etc.

I. INTELIGENCIA ARTIFICIAL: RIESGOS ASOCIADOS

1. Opacidad (Cajas Blancas vs Cajas Negras)





I. INTELIGENCIA ARTIFICIAL: RIESGOS ASOCIADOS

2. Sesgos (cognitivos y heurísticos)

Sesgo histórico: Influencia de elementos históricos (caso Loomis).

Sesgo representativo: Asunciones que no son representativas de la realidad. Fallo de lógica (circunstancias excepcionales: COVID 19).

Sesgo de medición: Se utilizan datos no relacionados con el objeto de la predicción (proxy o sustitutivos representativos).

Sesgo de agregación: Tendencia a considerar que las conclusiones de datos agregados pertenecientes a un grupo son extrapolables a otros individuos.

Sesgo evaluación: Los datos utilizados no representan a la población evaluada, sino solo una parte.

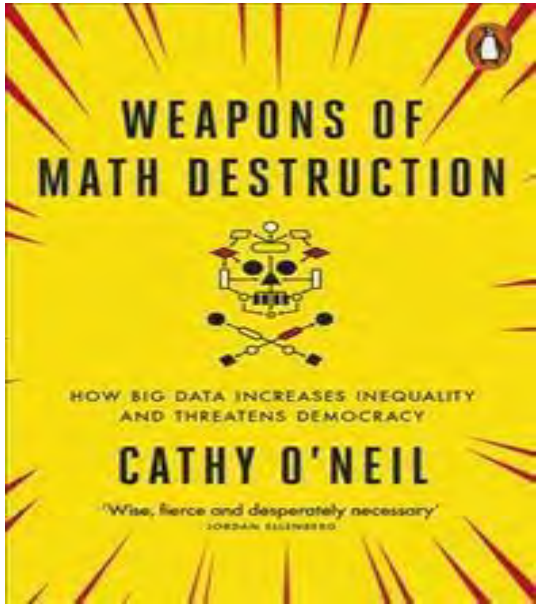
Etc… https://www.psicologia-online.com/que-son-los-sesgos-cognitivos-tipos-lista-y-ejemplos-5283.html

I. INTELIGENCIA ARTIFICIAL: RIESGOS ASOCIADOS

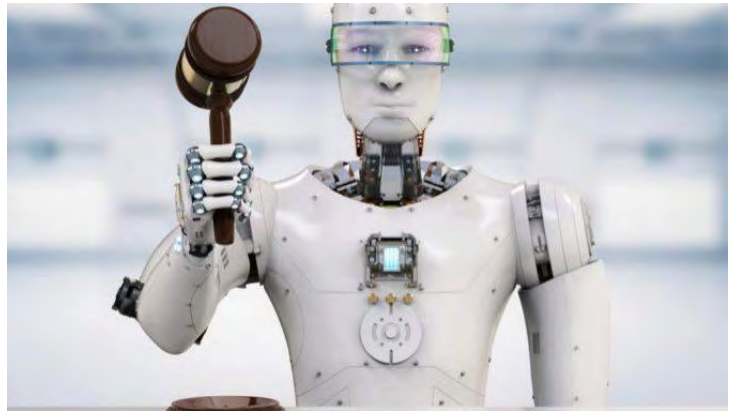
¿Por qué el sesgo artificial preocupa más que el sesgo humano?

1º) Inteligencia artificial e Inteligencia humana tienen **credibilidad y materialización diferente:** presunción de corrección –fe ciega-; uso a escala; menos coste. Rachel Thomas, Universidad de San Francisco

2º) Amplificación del sesgo (*feedback loop*): Cathy O´Neils, MIT y Berkeley



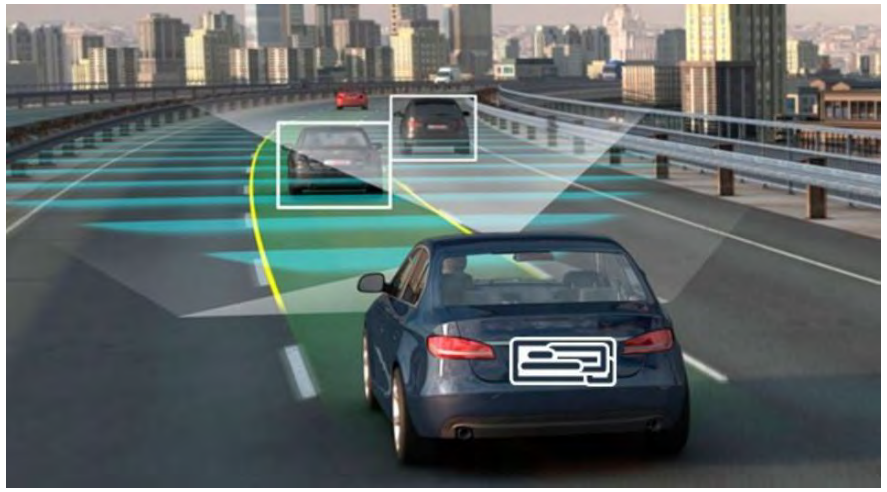


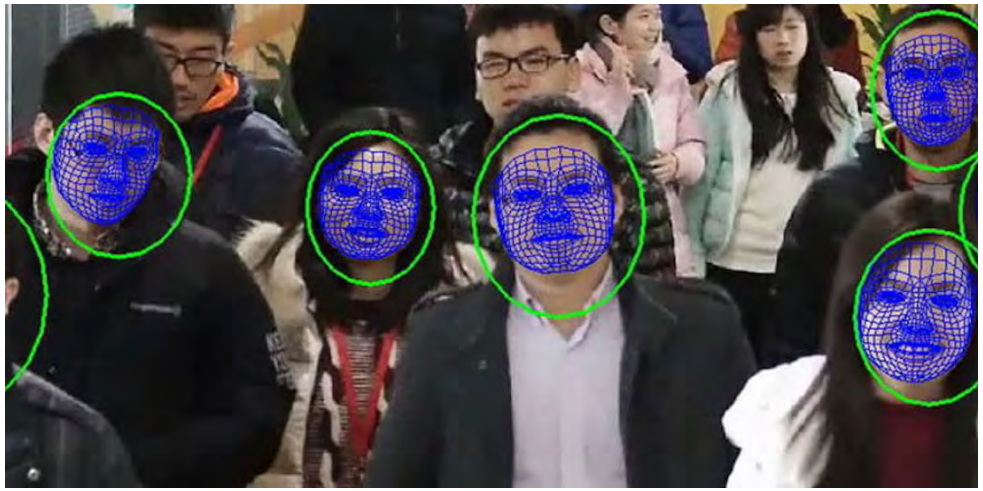


https://confilegal.com/20180402-justicia-robotica-el-caso-loomis/

I. INTELIGENCIA ARTIFICIAL: RIESGOS ASOCIADOS

3. PARADOJA: Minimización del riesgo y maximización del daño









I. INTELIGENCIA ARTIFICIAL: RIESGOS ASOCIADOS

Septiembre 2021: La Alta Comisionada DDHH de NNUU Bachelet pide una moratoria urgente en la venta y uso de la inteligencia artificial

El informe hace una serie de recomendaciones dirigidas a los Estados, entre ellas:

- Reconocer plenamente la necesidad de proteger y reforzar todos los derechos humanos en el desarrollo, el uso y la gobernanza de la inteligencia artificial como objetivo central, y garantizar el mismo respeto y aplicación de todos los derechos humanos en línea y fuera de ella
- Garantizar que el uso de la inteligencia artificial cumple con todos los derechos humanos y que cualquier interferencia con el derecho a la privacidad y otros derechos humanos está prevista por la ley, persigue un objetivo legítimo, cumple con los principios de necesidad y proporcionalidad y no perjudica la esencia de los derechos en cuestión
- Prohibir expresamente las aplicaciones de inteligencia artificial que no puedan funcionar de conformidad con el derecho internacional de los derechos humanos e imponer moratorias a la venta y el uso de sistemas que entrañen un alto riesgo para el disfrute de los derechos humanos, a menos y hasta que se establezcan salvaguardias adecuadas para proteger los derechos humanos
- Imponer una moratoria sobre el uso de tecnologías de reconocimiento biométrico a distancia en espacios públicos, al menos hasta que las autoridades responsables puedan demostrar que se cumplen las normas de protección de la intimidad y de los datos, y la ausencia de problemas significativos de exactitud y de efectos discriminatorios
- Adoptar y aplicar efectivamente, a través de autoridades independientes e imparciales, una legislación sobre la privacidad de los datos para los sectores público y privado como requisito previo esencial para la protección del derecho a la privacidad en el contexto de la inteligencia artificial
- Garantizar que las víctimas de violaciones y abusos de los derechos humanos relacionados con el uso de sistemas de inteligencia artificial tengan acceso a recursos efectivo

A/HRC/48/31

Advance Edited Version

Distr.: General
13 September 2021

Original: English

Naciones Unidas

Human Rights Council
Forty-eighth session
13 September–1 October 2021
Agenda items 2 and 3
Annual report of the United Nations High Commissioner for Human Rights and reports of the Office of the High Commissioner and the Secretary-General

Promotion and protection of all human rights, civil, political, economic, social and cultural rights, including the right to development

The right to privacy in the digital age*

Report of the United Nations High Commissioner for Human Rights

La Alta Comisionada también hace otras recomendaciones a las empresas, entre ellas:

- Hacer todo lo posible para cumplir con su responsabilidad de respetar todos los derechos humanos, incluso mediante la plena aplicación de los Principios Rectores sobre las Empresas y los Derechos Humanos
- Intensificar sus esfuerzos para luchar contra la discriminación vinculada a su desarrollo, venta o funcionamiento de sistemas de inteligencia artificial, entre otras cosas realizando evaluaciones sistemáticas y supervisando los resultados de los sistemas y los efectos de su despliegue
- Tomar medidas decisivas para garantizar la diversidad de la mano de obra responsable del desarrollo de la inteligencia artificial

I. INTELIGENCIA ARTIFICIAL: CLAVES DE FIABILIDAD

Recomendación Grupo de Expertos UNESCO sobre la ética en el uso de la IA (Noviembre 2021)

GRUPO EXPERTOS ALTO NIVEL UE

Directrices éticas para una AI fiable, abril 2019

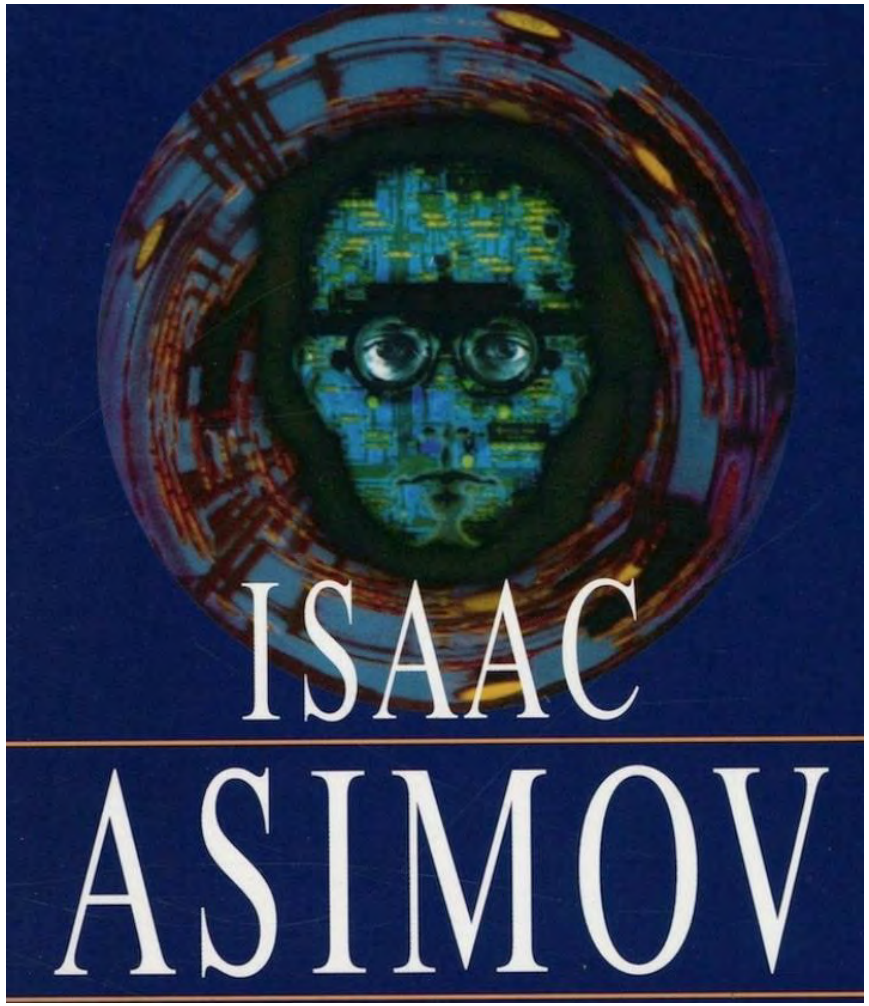
Policy and investment recommendations for trustworthy Artificial Intelligence, junio 2019

Recomendación del CONSEJO DE EUROPA, mayo de 2019 “Unboxing Artificial Intelligence: 10 steps to protect Human Rights”

Carta española de Derechos digitales de la ciudadanía 2021

1. **Lícita:** Ha de cumplir todas las leyes y reglamentos aplicables.
2. **Ética:** Ha de garantizar respeto de los principios y valores éticos (derechos fundamentales).
3. **Robusta:** Ha de ofrecer garantías de buen funcionamiento técnico, fortaleza frente a injerencias externas y ofrecer indemnidad a terceros en caso de daños accidentales.

I. INTELIGENCIA ARTIFICIAL: CLAVES DE FIABILIDAD



- Un robot no puede hacer daño a un ser humano o, por inacción, permitir que un ser humano sufra daño (**ética**).

- Un robot debe obedecer las órdenes dadas por los seres humanos, excepto si estas órdenes entrasen en conflicto con la 1ª Ley (**Legalidad/Intervención humana**)

- Un robot debe proteger su propia existencia en la medida en que esta protección no entre en conflicto con la 1ª o la 2ª Ley (**Robustez**).

ÍNDICE

1. CUESTIONES GENERALES, RIESGOS ASOCIADOS Y CLAVES DE FIABILIDAD
2. LA REGULACIÓN DE LA INTELIGENCIA ARTIFICIAL: PUNTO DE PARTIDA Y CUESTIONES A REGULAR. VIABILIDAD JURÍDICA Y MATERIAL DE LA TRANSFORMACIÓN DIGITAL DE LA ADMINISTRACIÓN PÚBLICA
3. LA DEFINICIÓN DE UN MODELO REGULATORIO EUROPEO BASADO EN LA GESTIÓN DE RIESGOS
4. EL PROBLEMA ÉTICO DE LA INTELIGENCIA ARTIFICIAL Y LA RAZÓN JURÍDICA: PROBLEMAS, IDEAS Y CONCLUSIONES DEL MODELO REGULATORIO

II. LA REGULACIÓN DE LA I.A.: PUNTO DE PARTIDA

REGLAMENTO GENERAL DE PROTECCIÓN DE DATOS UE 2016/679.

Artículo 22:

Todo **interesado** tendrá derecho a no ser objeto de una decisión **basada únicamente en el tratamiento automatizado**, incluida la elaboración de **perfiles**, **que produzca efectos jurídicos** en él o **le afecte significativamente** de modo similar (actos y actuaciones "ad extra")

Elaboración de perfiles: Tratamiento automatizado datos personales para evaluar aspectos (Rendimiento profesional, situación económica, salud, preferencias personales, rastros de geolocalización, rastro digital…etc.). **La CDD España lo permite si hay previsión normativa**

¿actos de inspección están incluidos?

Excepciones:

a) Que la decisión sea necesaria para la celebración o ejecución de un contrato.

a) Que esté autorizado con por el Derecho de la UE o EM (con medidas de garantía en favor del interesado).

a) Que exista consentimiento explícito del interesado.

II. LA REGULACIÓN DE LA I.A.: DERECHOS RECONOCIDOS

1.- Derecho a obtener **Intervención Humana.**

Human in the Loop

- No se refiere a intervención humana entre la inferencia de datos y el resultado (ello eliminaría la total automatización).
- No se refiere a intervención en la codificación o diseño del sistema algorítmico (fuera del proceso de toma de decisión).
- Se refiere a una intervención **significativa**, de **revisión** por persona con **autorización** y **capacidad** para **modificar** el resultado arrojado por el tratamiento automatizado de datos.

2.- Derecho a expresar su punto de vista (alegaciones).

3.- Derecho a impugnar la decisión.

II. REGULACIÓN DE LA I.A.: DERECHOS RECONOCIDOS

4.- Derecho a ser informado del tratamiento automatizado: Obligación general del responsable del tratamiento (art. 12.1 RGPD)

5.- Derecho a la información sobre la lógica aplicada.

6. **Derecho de acceso**: debate sobre la naturaleza jurídica del algoritmo (A. Huergo / A. Boix)

¿Existe el derecho a acceder al algoritmo y su código fuente?

DRECHLER, L. y BENITO SÁNCHEZ, J.C.: *"orientación significativa sobre el funcionamiento del algoritmo"*.

GIL GONZÁLEZ, E.: *"información genérica, más fácilmente comprensible para el usuario medio"*.

PALMA, A.: *"suficiente como para que el particular que se ve sometido a dichas decisiones, pueda llegar a hacerse una idea de cómo puede funcionar el programa informático"*.

GT29: *"formas sencillas de informar al interesado [...] No obstante, la información facilitada debe ser suficientemente exhaustiva para que el interesado entienda os motivos de la decisión"*.

Suficiencia para poder →

Formular oposición al tratamiento.
Solicitar intervención humana.
Formular alegaciones.
Impugnar la decisión.

II. CARTA ESPAÑOLA DE DERECHOS DIGITALES DE LA CIUDADANÍA



1º) Derechos de los ciudadanos ante el uso de la IA

- ✓ Derecho a ser informado cuando se interactúe con un sistema de IA
- ✓ Se prohíbe el uso de sistemas de IA dirigidos a manipular o perturbar la voluntad de las personas (perfilar, salvo consentimiento) o en cualesquiera aspectos que afecten a los derechos fundamentales
- ✓ A no ser discriminado por las decisiones basadas en algoritmos (sesgo)
- ✓ A que se asegure la transparencia, auditabilidad, explicabilidad, trazabilidad, accesibilidad, usabilidad y fiabilidad de los algoritmos y sistemas basados en IA
- ✓ A que se garantice la asistencia por un ser humano si la persona interesada lo solicita

2º) Derechos en relación con el uso y la toma de decisiones por los poderes públicos que afecten al ciudadano

- ✓ Derecho a las garantías procedimentales ordinarias. En particular, a que las decisiones en el entorno digital respeten los principios de buen gobierno y el derecho a una buena Administración digital.
- ✓ Derecho a no ser objeto de una decisión basada únicamente en procesos automatizados, lo que supone el derecho a:
  - A que la adopción de decisiones discrecionales quede reservada a personas, salvo que una norma con rango de ley permita la adopción de decisiones automatizadas
  - Derecho a solicitar una supervisión e intervención humana
  - Derecho a impugnar las decisiones automatizadas o algorítmicas
- ✓ Derecho a una motivación comprensible de las decisiones que se adopten en el entorno digital, con justificación de las normas jurídicas relevantes al caso y de los criterios de aplicación de las mismas
- ✓ En todo caso, será necesaria una evaluación de impacto en los derechos digitales en el diseño de los algoritmos en el caso de adopción de decisiones automatizadas o semiautomatizadas. Serán objeto de aprobación previa los sistemas algorítmicos que se vayan a usar para la toma de decisiones, con determinación de su ámbito concreto de aplicación y estructura de funcionamiento. Impulso a la certificación

II. INTELIGENCIA ARTIFICIAL AAPP: VIABILIDAD JURÍDICA

1. Sistemas de IA para actuación administrativa automatizada sin intervención humana.

"Artículo 41 LRJSP.

Actuación administrativa automatizada.

1. Se entiende por actuación administrativa automatizada, cualquier acto o actuación realizada ***íntegramente a través de medios electrónicos*** *por una Administración Pública en el marco de un procedimiento administrativo y* ***en la que no haya intervenido de forma directa un empleado público****.*

2. En caso de actuación administrativa automatizada deberá ***establecerse previamente el órgano u órganos competentes****, según los casos,* ***para la definición de las especificaciones, programación, mantenimiento, supervisión y control de calidad y, en su caso, auditoría del sistema*** *de información y de su código fuente. Asimismo, se indicará el órgano que debe ser considerado* ***responsable a efectos de impugnación****".*

Necesaria publicación previa de acto administrativo con el contenido informativo anterior

En caso de tratamiento de datos de personas físicas → Art. 22 RGPD (prohibición)

Desaconsejable uso de IA de caja negra

II. INTELIGENCIA ARTIFICIAL AAPP: VIABILIDAD JURÍDICA

Procedimientos reglados ¿ y discrecionales ?

- La **Ley Alemana de procedimiento** prohíbe la adopción de actos automatizados en aquellos ámbitos en que resulte de aplicación la discrecionalidad administrativa y, en cualquier caso, siempre que exista margen de apreciación (pero sí semiautomatizados)

- **CDD España**: prohíbe en decisiones discrecionales "salvo que se prevean garantías adecuadas"

- **Sentencia Consejo de Estado Italiano** de 8 de abril de 2019. Concurso provisión de puestos de personal docente. La asignación de destinos la hizo un programa informático, un algoritmo, cuyo funcionamiento no se hizo público

  a) El algoritmo (como regla) posee pleno valor jurídico y administrativo. Está sometido a los principios generales de la actividad administrativa, como los de publicidad, transparencia, razonabilidad y proporcionalidad. Asimismo tiene que ser explicable, legible y comprensible

  b) No cabe dejar espacio a la discrecionalidad en la aplicación del algoritmo.

  c) Vigilancia: la Administración ha de velar por los intereses en presencia, realizando pruebas y actualizaciones de perfeccionamiento del algoritmo, en especial en el caso de aprendizaje progresivo y de *deep learning*.

  d) Desde el punto de vista del control judicial, el algoritmo o software se entiende a todos los efectos como un acto administrativo informático y el órgano jurisdiccional puede evaluar la corrección del proceso automatizado con plena cognición y en un doble plano: 1) por un lado, habría de determinarse la conformidad a Derecho del proceso informático en todas sus componentes: construcción, inserción de datos, validez y gestión de los mismos. De ahí la exigencia de transparencia en la fase administrativa; 2) en relación con lo anterior, el juez ha de poder evaluar la lógica y la razonabilidad de la decisión administrativa automatizada.

II. INTELIGENCIA ARTIFICIAL AAPP: TRANSPARENCIA y ACCESIBILIDAD

	Solución actual	Solución alternativa propugnada	Solución tendente propugnada
Solución	Art. 22 RGPD y Ley de Transparencia y Buen Gobierno.	Carácter Reglamentario de los algoritmos. Aplicación de garantías de los reglamentos	Control por auditorías de expertos designados por poderes públicos.
Crítica	Insuficiencia (personas jurídicas; régimen laxo y orientado a privados). Falta de seguridad jurídica.	Naturaleza no reglamentaria del algoritmo. Complejidad del algoritmo; coste.	Insuficiencia frente al ejercicio de poderes públicos.

Consejo de Transparencia y Buen Gobierno. Resolución 701/2018, de 18 de febrero (Civio vs BOSCO)

En el ámbito comparado, interesa mencionar aquí la pionera sentencia de 5 de febrero de 2020 del Tribunal de Distrito de la Haya (14). En este pronunciamiento se concluyó que se había vulnerado el artículo 8 del Convenio Europeo de Derechos Humanos -EDL 1979/3822- (derecho al respeto de la vida privada y familiar) por parte de las autoridades holandesas. En concreto, el pronunciamiento venía referido al uso de un instrumento para detectar fraude, instrumento de naturaleza algorítmica que realizaba predicciones sobre la base de datos insertados en el sistema. El órgano jurisdiccional holandés consideró que esta actuación administrativa contravenía el artículo 8.2 del Convenio, una vez realizada la correspondiente ponderación entre los derechos e intereses concurrentes, por cuanto el algoritmo no era transparente ni verificable.

II. INTELIGENCIA ARTIFICIAL AAPP: VIABILIDAD MATERIAL

2. Evaluación de suficiencia de datos y riesgos de error.

	IMPACTO BAJO DE UN ERROR DEL SISTEMA DE IA	IMPACTO ALTO DE UN ERROR DEL SISTEMA DE IA
DATOS REPRESENTATIVOS Y SUFICIENTES	Proyectos viables en producción. *[cuadrante II]*	Proyectos con potencial de impactos muy negativos de baja probabilidad[22] en producción. Para valorar su uso en producción se deben paliar o eliminar los impactos de los errores. *[cuadrante I]*
DATOS NO REPRESENTATIVOS Y/O INSUFICIENTES	Proyectos viables únicamente como pilotos o pruebas de concepto. Para valorar su uso en producción es esencial conseguir datos suficientes y representativos. *[cuadrante III]*	Proyectos con elevada probabilidad de consecuencias catastróficas en producción. No deberían emprenderse. *[cuadrante IV]*

*ZLOTNIK, A.: Inteligencia Artificial en las Administraciones Públicas: definiciones, **evaluación de viabilidad de proyectos y áreas de aplicación**. Boletic 84, julio 2019.

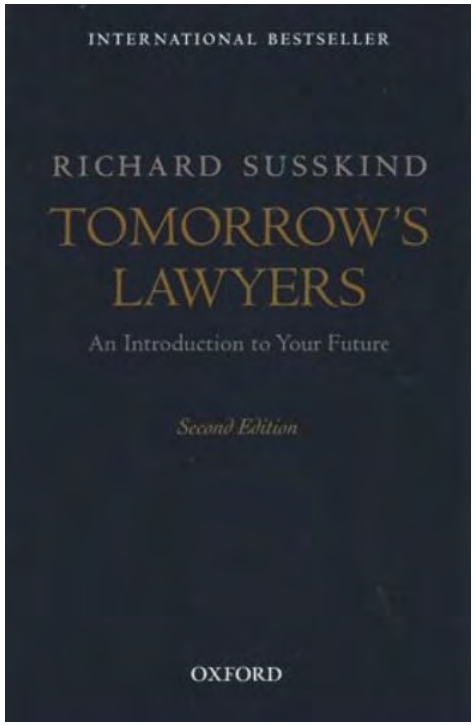
II. INTELIGENCIA ARTIFICIAL AAPP: VIABILIDAD MATERIAL

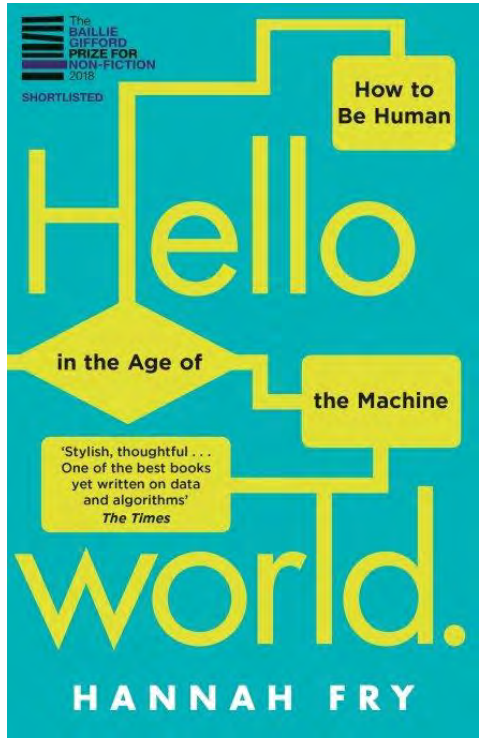
3. Suficiencia de recursos computacionales.

4. Equipo humano.

El "Científico de datos" (data scientist)







ÍNDICE

1. CUESTIONES GENERALES, RIESGOS ASOCIADOS Y CLAVES DE FIABILIDAD
2. LA REGULACIÓN DE LA INTELIGENCIA ARTIFICIAL: PUNTO DE PARTIDA Y CUESTIONES A REGULAR. VIABILIDAD JURÍDICA Y MATERIAL DE LA TRANSFORMACIÓN DIGITAL DE LA ADMINISTRACIÓN PÚBLICA
3. LA DEFINICIÓN DE UN MODELO REGULATORIO EUROPEO BASADO EN LA GESTIÓN DE RIESGOS
4. EL PROBLEMA ÉTICO DE LA INTELIGENCIA ARTIFICIAL Y LA RAZÓN JURÍDICA: PROBLEMAS, IDEAS Y CONCLUSIONES DEL MODELO REGULATORIO

III. INTELIGENCIA ARTIFICIAL: EL MARCO REGULATORIO EUROPEO

El posicionamiento Europeo y español sobre el papel de la IA

- ✓ **Reglamento del PARLAMENTO EUROPEO y del CONSEJO que establece el programa Europa Digital para el período 2021-2027 (julio 2018)**
- ✓ **COM (2018) 237 final "Artificial Intelligence for Europe", abril de 2018**
- ✓ **Conclusiones CONSEJO EUROPEO "Coordinated Plan on the Development and Use of Artificial Intelligence Made in Europe", febrero 2019**
- ✓ **COM (2020) 65 final. Libro Blanco UE sobre IA: un enfoque europeo orientado a la excelencia y a la confianza (febrero 2020)**
- ✓ **COM (2020) 66 final. Comunicación sobre una estrategia europea de datos (febrero 2020)**

Singularidad de la Estrategia europea IA

- ✓ Otras **iniciativas** europeas recientes en el **campo de la justicia** "Use of innovative technologies in the justice field"

- ➢ **Plan ESPAÑA DIGITAL 2025 (julio 2020)**





III. MODELO REGULATORIO DE GESTIÓN DEL RIESGO en la UE (RISK-BASED REGULATION)

- **Conocimiento y reconocimiento del riesgo**: información disponible
- **Valoración del riesgo**: determinación del riesgo asumible/tolerable y prohibición del inasumible
- **Gestión del riesgo**: exigencias preventivas al riesgo tolerado y medidas de control. Principio de precaución

III. EL MARCO REGULATORIO EUROPEO: RESOLUCIÓN DEL PARLAMENT0 2020

Octubre de 2020: **Resolución del Parlamento Europeo con recomendaciones destinadas a la Comisión sobre un marco de los aspectos éticos de la inteligencia artificial, la robótica y las tecnologías conexas.**

Sectores de alto riesgo	— Empleo — Educación — Asistencia sanitaria — Transporte — Energía — Sector público (asilo, migración, controles fronterizos, sistema judicial y servicios de seguridad social) — Seguridad y defensa — Finanzas, bancos, seguros

III. EL MARCO REGULATORIO EUROPEO: RESOLUCIÓN DEL PARLAMENT0 2020

Usos o fines de alto riesgo	— Contratación — Clasificación y evaluación de estudiantes — Asignación de fondos públicos — Concesión de préstamos — Comercio, corretaje, fiscalidad, etc. — Tratamientos y procedimientos médicos — Procesos electorales y campañas políticas — Decisiones del sector público que tienen un impacto significativo y directo en los derechos y las obligaciones de las personas físicas o jurídicas — Conducción automatizada — Gestión del tráfico — Sistemas militares autónomos — Producción y distribución de energía — Gestión de residuos — Control de emisiones

III. EL MARCO REGULATORIO EUROPEO: RESOLUCIÓN DEL PARLAMENT0 OCTUBRE 2020

Principios éticos/jurídicos

- IA **antropocéntrica** y **antropogénica** (human in de loop)
- **Seguridad**, **transparencia** y **rendición de cuentas** (ciberseguridad - medidas de seguridad; explicabilidad y trazabilidad; información; auditabilidad)
- **Ausencia de sesgo** (será justificable si existe razón legítima y proporcional).
- **Responsabilidad social** (no interferencia en elecciones; no desinformación, derechos de los trabajadores; educación de calidad y no vulneración de propiedad intelectual) e **igualdad de género**.
- **Sostenibilidad medioambiental**.
- Respeto de la **intimidad** y protección de los **datos personales**.
- **Derecho de resarcimiento**.

PRINCIPIOS JURÍDICOS CLAVE: **(I)** principio de cognoscibilidad, **(II)** el principio de no exclusividad de la decisión algorítmica o derecho de intervención humana y (III) principio de no discriminación algorítmica (iv) principio de auditabilidad y certificación ética para su comercialización y distribución

III. EL MARCO REGULATORIO EUROPEO: PROPUESTA REGLAMENTO UE 2021

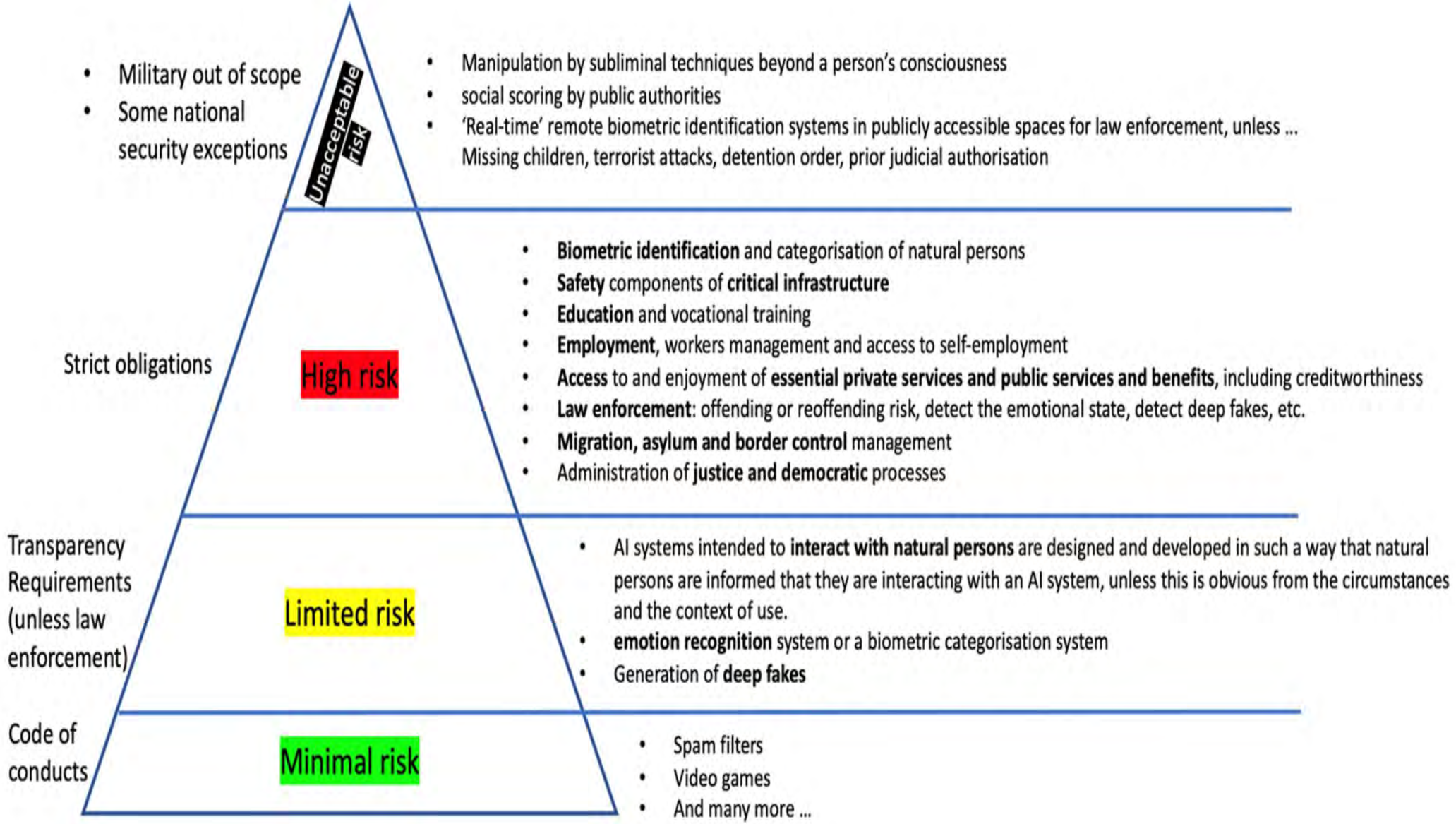
Abril de 2021: Propuesta de Reglamento del Parlamento Europeo y el Consejo estableciendo reglas armonizada sobre Inteligencia Artificial (Ley de Inteligencia Artificial)

- 1.200 contribuciones a su elaboración.
- Lectura previa a aprobación definitiva.
- Previsible ventaja competitiva frente a otros sistemas (EEUU, China).

4 niveles de regulación armonizada

- Prohibición total o parcial de usos de IA
- Requisitos para sistemas de alto riesgo (obligaciones en la cadena comercial)
- Obligaciones de transparencia para ciertos sistemas de IA
- Códigos de conducta voluntarios para sistemas sin alto riesgo.

III. EL MARCO REGULATORIO EUROPEO: PROPUESTA REGLAMENTO UE 2021



III. PROPUESTA REGLAMENTO UE 2021: PROHIBICIONES

- Sistemas que **manipulen subliminalmente** a las personas en su perjuicio.

- Sistemas que **exploten vulnerabilidades** de grupos de personas (edad, discapacidad…etc.) para manipular su comportamiento en su perjuicio.

- Sistemas utilizados por las autoridades públicas para la clasificación y puntuación de personas (**perfiles**) basadas en datos sobre su comportamiento o características personales, que:

  - Puedan llevar a trato desfavorable en contextos no relacionados con los propios de los datos utilizados.

  - Puedan llevar a trato desfavorable que resulte desproporcionado a la gravedad de las conductas.

- Sistemas de **reconocimiento biométrico** en espacios públicos con objeto de control policial. Excepciones:

- Localización de posibles víctimas de delitos, incluidos menores desaparecidos. - Prevención de amenazas graves de ataques a la vida o actos terroristas.	**PROPORCIONALIDAD**
- Persecución de autores de ciertos delitos de relevancia comunitaria que conlleve penas de prisión de más de tres años.	**AUTORIZACIÓN JUDICIAL O DE AUTORIDAD ESPECÍFICA**

III. PROPUESTA REGLAMENTO UE 2021: OBLIGACIONES

Todas las aplicaciones de IA de alto riesgo habrán de cumplir con unos requisitos:

1. Disponer de un sistema de **análisis de riesgos continuo** durante su vida útil.

Los riesgos identificados deben comunicare al usuario.

2. Gobernanza de **datos**.

Entrenamiento, validación y prueba de datos que permita su calidad y ausencia de sesgo.

3. Documentación técnica demostrativa de cumplimiento de requisitos.

4. Trazabilidad de procesos (“logs”).

III. PROPUESTA REGLAMENTO UE 2021: OBLIGACIONES

5. **Transparencia**.

   Información al usuario sobre el propósito, grado de robustez, posibles riesgos, gobernanza de datos aplicada, medidas de intervención humana y vida útil aproximada de la aplicación.

6. **Intervención humana**.

   Posibilidad de monitorización, identificación de sesgos, interpretación, contradicción y apagado.

7. **Exactitud, robustez y ciberseguridad.**

   Resistencia a errores (mediante “backups “o “a prueba de fallos”)

III. PROPUESTA REGLAMENTO UE 2021: CONTROL

<table>
<tr><th colspan="3">Con intervención de tercero</th><th>Sin intervención de tercero</th></tr>
<tr><td colspan="3">Sistemas de identificación biométrica de personas</td><td>• Sistemas de control de infraestructuras críticas.
• Sistemas de IA predictiva aplicable a: Educación/Empleo/Servicios esenciales y públicos/control policial /Inmigración/Justicia y Democracia</td></tr>
<tr><td rowspan="2">Si se aplican estándares</td><td colspan="2">Opcional</td><td rowspan="3">Evaluación interna

*Aplicaciones de determinación de solvencia por entidades de crédito se evaluarán conforme a la Directiva 2013/36/EU.
**Aplicaciones bajo legislación armonizada de la EU se evaluarán conforme a su normativa particular.</td></tr>
<tr><td>Evaluación interna</td><td>Evaluación con intervención de tercero</td></tr>
<tr><td>Si no se aplican estándares</td><td colspan="2">Evaluación con intervención de tercero</td></tr>
</table>

III. PROPUESTA REGLAMENTO UE 2021: CONTROL

Tercero evaluador: Entidad designada por la autoridad nacional competente. A elección del interesado (salvo si es para uso policial, control de inmigración o asilo)

Resultado de la evaluación interna: Declaración de conformidad a disposición de autoridades por 10 años desde comercialización o puesta en funcionamiento.

Resultado de la evaluación por tercero: Certificado de conformidad ≤ 5 años prorrogables.

Resultado general: Marcado CE de conformidad. CE

III. PROPUESTA REGLAMENTO UE 2021: Informe del CESE septiembre 2021

Cuestiona la clasificación de riesgos del Proyecto de Reglamento. Particularmente con el **reconocimiento biométrico**. El reglamento prohíbe **la identificación biométrica a distancia "en tiempo real"** , que clasifica de alto riesgo cuando se utilice con fines no policiales. Por tanto:

- Permite implicitamente **el reconocimiento biométrico "posterior"**.
- Permite **el reconocimiento biométrico no destinado a identificar a una persona, sino a evaluar su comportamiento a partir de sus características biométricas** (microexpresiones, forma de andar, temperatura, frecuencia cardíaca, etc.): el «reconocimiento de emociones o afectos» o «reconocimiento de conductas», que califica **bajo riesgo.**

Así, **cualquiera puede utilizar la identificación biométrica** posterior y la "emocional" para cualquier fin y en cualquier lugar público, en particular en lugares de trabajo, comercios, acontecimientos, etc.; lo que deja la puerta abierta un "**Gran Hermano constante"**.

A juicio del CESE, estas prácticas de IA son **extremadamente invasivas** y pueden menoscabar **derechos fundamentales** contemplados en la Carta de los Derechos Fundamentales de la UE, como el derecho a la dignidad humana, el derecho a la integridad de la persona y el respeto de la vida privada.

Coincidiendo con la petición del **Supervisor Europeo de Protección de Datos** y el **Comité Europeo de Protección de Datos** (de junio de 2021), el **CESE** pide:

– La **prohibición del uso de la IA para el reconocimiento biométrico facial en espacios de acceso público y privado excepto con fines policiales o de autenticación**, para facilitar el acceso a espacios sensibles desde el punto de vista de la seguridad);

– La **prohibición de los sistemas de IA que utilicen la biometría para clasificar a las personas en grupos en función de su etnia, género, orientación política o sexual o cualquier otro motivo de discriminación**;

– La **prohibición de utilizar la IA para inferir las emociones, la conducta, la intención o los gestos/rasgos de una persona física**, excepto en casos muy específicos, como determinados fines relacionados con la salud, en los que sea importante reconocer las emociones del paciente.

ibangarciadb.eu

Introduce tu búsqueda

IBÁN GARCÍA PUBLICACIONES CONTACTO



Informe sobre principios éticos aplicables a la inteligencia artificial, la robótica y otras tecnologías.

Parlamento Europeo



Ibán García del Blanco
Eurodiputado y secretario federal de Cultura y Deportes del PSOE

ÍNDICE

1. CUESTIONES GENERALES, RIESGOS ASOCIADOS Y CLAVES DE FIABILIDAD
2. LA REGULACIÓN DE LA INTELIGENCIA ARTIFICIAL: PUNTO DE PARTIDA Y CUESTIONES A REGULAR. VIABILIDAD JURÍDICA Y MATERIAL DE LA TRANSFORMACIÓN DIGITAL DE LA ADMINISTRACIÓN PÚBLICA
3. LA DEFINICIÓN DE UN MODELO REGULATORIO EUROPEO BASADO EN LA GESTIÓN DE RIESGOS
4. EL PROBLEMA ÉTICO DE LA INTELIGENCIA ARTIFICIAL Y LA RAZÓN JURÍDICA: PROBLEMAS, IDEAS Y CONCLUSIONES DEL MODELO REGULATORIO

IV. El problema ético y la razón jurídica: Estrategia Nacional de Desarrollo IA



Plan de Recuperación, Transformación y Resiliencia

COMPONENTE 16

Estrategia Nacional de Inteligencia Artificial

Contenidos

1. Descripción general del componente.....
2. Principales retos y objetivos.....
3. Detalle sobre cada reforma/inversión del componente.....
4. Autonomía estratégica y seguridad.....
5. Proyectos transfronterizos y multi-país .....
6. Contribución del componente a la transición ecológica .....
7. Contribución del componente a la transición digital .....
8. Principio "Do not significant harm".....
9. Hitos, metas y cronograma.....
10. Financiación .....

IV. El problema ético y la razón jurídica: Robustez, fiabilidad y certificado ético digital

Certificación por tercera parte. Variables y condicionantes: tipo de algoritmo + alcance (función) = > < necesidad de control

- **RIESGO BAJO:** Algoritmo aplicativo basado en el automatismos reglado: simplemente "ayuda a decidir" (p.ej. aplicaciones AEAT que liquidan tributos)

- **RIESGO MEDIO:** Algoritmo aplicativo complejo con elementos ponderativos y variables no normativas: el algoritmo proporciona el dato para la toma de decisión ("Euphemia": https://www.energynews.es/euphemia-algoritmo-calcula-precio-luz/) o BOSCO

- **RIESGO ALTO:** Algoritmo predictivo. El algoritmo es capaz de decidir o de proporcionar el elemento valorativo de la decisión (p.ej. "RisCanvi" o "Viogen")

ESTRATEGIA ESPAÑOLA IA 2019: 6 PRIORIDADES

PRIORIDAD 3

Crear una plataforma computacional donde los agentes de I+D+I puedan experimentar sus desarrollos antes de lanzarlos al mercado.

Herramientas de capacitación dirigidas a habilidades y certificaciones en IA.

- La creación de una plataforma y un ecosistema computacional donde la universidad, los centros científicos y tecnológicos y las PYME nacionales puedan experimentar sus desarrollos innovadores antes de lanzarlos al mercado. Esta plataforma debería actuar como ventanilla única, brindando y mostrando servicios, experiencia, algoritmos certificados, modelos de programación, herramientas de desarrollo, componentes, módulos, bases de datos no sesgadas, recursos informáticos, funciones de creación de prototipos y acceso a financiamiento público y privado.
- Ofrecer herramientas de capacitación que permitan a diferentes comunidades de usuarios obtener habilidades y certificaciones en el área de la IA.

PRIORIDAD 5

1. Transparencia y certificación de la gobernanza de todo el ecosistema digital;
2. Procesos de certificación a lo largo de todo el ciclo de vida del dato, especialmente cuando. (i) el dato abierto se "entrega" a un tercero para su uso y explotación, o (ii) un ciudadano, empresa o sistema software desea verificar si un dato de la Administración proporcionado por un tercero es correcto y está actualizado. El objeto es evitar el *spam* en los datos o los *fake data* de datos del sector público[89].

GOBIERNO DE ESPAÑA · MINISTERIO DE CIENCIA, INNOVACIÓN Y UNIVERSIDADES · ESTRATEGIA ESPAÑOLA DE I+D+I EN INTELIGENCIA ARTIFICIAL

Recomendación 7: Velar por un uso ético de la IA en todos sus campos de aplicación

- **Certificado ético europeo**: Recomendación Parlamento UE 2020
- **Sistemas de evaluación de impacto ético**: Recomendación UNESCO 2021

IV. El problema ético y la razón jurídica: La Carta de Derechos Digitales julio 2021



1. DERECHOS DE LIBERTAD

1. Derechos y libertades en el entorno digital
2. Derecho a la identidad en el entorno digital
3. Derecho a la protección de datos
4. Derecho al pseudonimato
5. Derecho de la persona a no ser localizada y perfilada
6. Derecho a la ciberseguridad
7. Derecho a la herencia digital

2. DERECHOS DE IGUALDAD

8. Derecho a la igualdad y a la no discriminación en el entorno digital
9. Derecho de acceso a Internet
10. Protección de las personas menores de edad en el entorno digital
11. Accesibilidad universal en el entorno digital
12. Brechas de acceso al entorno digital

3. DERECHOS DE PARTICIPACIÓN Y DE CONFORMACIÓN DEL ESPACIO PÚBLICO

13. Derecho a la neutralidad de Internet
14. Libertad de expresión y libertad de información
15. Derecho a recibir libremente información veraz
16. Derecho a la participación ciudadana por medios digitales
17. Derecho a la educación digital
18. Derechos digitales de la ciudadanía en sus relaciones con las Administraciones Públicas

4. DERECHOS DEL ENTORNO LABORAL Y EMPRESARIAL

19. Derechos en el ámbito laboral
20. La empresa en el entorno digital

5. DERECHOS DIGITALES EN ENTORNOS ESPECÍFICOS

21. Derecho de acceso a datos con fines de archivo en interés público, fines de investigación científica o histórica, fines estadísticos, y fines de innovación y desarrollo
22. Derecho a un desarrollo tecnológico y un entorno digital sostenible
23. Derecho a la protección de la salud en el entorno digital
24. Libertad de creación y derecho de acceso a la cultura en el entorno digital
25. Derechos ante la inteligencia artificial
26. Derechos digitales en el empleo de las neurotecnologías

6. GARANTÍAS Y EFICACIA

27. Garantía de los derechos en los entornos digitales
28. Eficacia

IV. EL PROBLEMA ÉTICO Y LA RAZÓN JURÍDICA

Informe 2021
Artificial Intelligence Index

8 **La ética en el campo de la IA carece de puntos de referencia y de consenso: Aunque varios grupos están produciendo una serie de resultados cualitativos o normativos en el ámbito de la ética de la IA, el campo generalmente carece de puntos de referencia que puedan utilizarse para medir o evaluar la relación entre los debates sociales más amplios sobre el desarrollo de la tecnología y el desarrollo de la propia tecnología. Además, los investigadores y la sociedad civil consideran que la ética de la IA es más importante que las empresas privadas.**

United Nations
Educational, Scientific and
Cultural Organization

Organisation
des Nations Unies
pour l'éducation,
la science et la culture

Organización
de las Naciones Unidas
para la Educación,
la Ciencia y la Cultura

Организация
Объединенных Наций по
вопросам образования,
науки и культуры

منظمة الأمم المتحدة
للتربية والعلم والثقافة

联合国教育、
科学及文化组织

Conferencia General
41ª reunión - Paris, 2021

41 C

Principios generales muy vagos: proporcionalidad, inocuidad, equidad, no discriminación, sostenibilidad, respeto a la privacidad, supervisión humana, transparencia y explicabilidad, responsabilidad y rendición de cuentas

Ámbitos de actuación Estados: (i) modelos de implantación éticos, con medición del impacto ético durante todo el ciclo de vida (ii) implantación "blanda": certificado ético, reconocimiento mutuo y presunción de conformidad (iii) política de datos que garantice su calidad y fiabilidad… (hasta 11: medioambiente, género, cultura, salud, cooperación internacional, etc.)

41 C/23
14 de septiembre de 2021
Original: inglés

Punto **8.2** del orden del día provisional

PROYECTO DE RECOMENDACIÓN SOBRE LA ÉTICA DE LA INTELIGENCIA ARTIFICIAL


ELI
EUROPEAN
LAW
INSTITUTE
Model Rules on Impact Assessment of Algorithmic Decision-Making Systems Used by Public Administration
Report of the European Law Institute

IV. EL PROBLEMA ÉTICO Y LA RAZÓN JURÍDICA: ALGUNAS CONCLUSIONES PARA EL MODELO REGULATORIO

I. **Dilema:** Los algoritmos se desarrollan más fácilmente en los ámbitos de libertad que en los de potestad/autoridad, pero las AAPP y el poder judicial no puede renunciar a ellos.

II. Quedan **muchas cuestiones por definir**: ámbito aplicación ¿qué decisiones? Garantías aplicables ¿las mismas o hacen falta nuevas?

III. **Modelo regulatorio**: la gestión de los riesgos de la I.A. aboca a un planteamiento jurídico esencialmente precautorio y discrecional, en el que la evitación de un riesgo potencial, con independencia de su mayor o menor certidumbre:

1º) inclina a adoptar acciones preventivas formuladas en sentido negativo: lo que no habría de hacerse o hay que impedir para evitar la consumación de lo que hoy es simple incertidumbre (excluir decisiones plenamente automatizadas en decisiones discrecionales “ad extra”, garantizar presencia humana).

2º) Hace que las exigencias de cumplimientos sean finalistas y de apreciación discrecional: fiable, explicable, auditable, proporcional, etc (“la discrecionalidad que paramos en la puerta sale por la ventana” del aplicador)

3) Desmaterializa el derecho aplicable, que abandona la determinación sustantiva a la definición del riesgo tolerable, cediendo su validación al conocimiento experto (al margen de la Administración), lo que reduce la capacidad decisora de la Administración en el control de su cumplimiento

4) Apela al principio de precaución (arts. 15 y 16 TR Ley Consumidores y usuarios. Art. 191 TFUE ámbito ambiental “todas las medidas necesarias y proporcionadas”)

Gracias por su atención

Juan Antonio Carrillo
donaire@us.es



