

# GOVERN OBERT

9

## **Transparencia algorítmica en el sector público**

---

**Pablo Jiménez Arandia**

---



## GOVERN OBERT

---

- 1** *La publicidad activa en la Ley de transparencia, acceso a la información y buen gobierno: posibilidades e insuficiencias.* Manuel Villoria
- 2** *Apertura y reutilización de datos públicos.* Martín Álvarez Espinar
- 3** *La justicia penal ante la corrupción.* Antonio del Moral García
- 4** *Contractació oberta.* Agustí Cerrillo i Martínez
- 5** *Nueve voces reflexionan en torno al gobierno abierto.* Jordi Rovira (coord.)
- 6** *Buen gobierno e integridad pública contra la corrupción.* Míriam Carles, Agustí Cerrillo, Laura Chaqués et al.
- 7** *Los datos abiertos y la inteligencia artificial, herramientas para la igualdad de género.* Storydata (coord.)
- 8** *El valor de los datos abiertos y casos de uso.* Alicia León Molina

# GOVERN OBERT

9

## Transparencia algorítmica en el sector público

---

**Pablo Jiménez Arandia**

---



**Jiménez Arandía, Pablo, autor**

Transparencia algorítmica en el sector público. – 1a edició. –

(Govern obert ; 9)

Bibliografia

ISBN 9788418601057

I. Catalunya. Generalitat. II. Títol III. Col·lecció: Govern obert. Castellà ; 9

1. Sistemes d'informació per a la gestió – Informàtica.

2. Transparència política

658:004

35.076

---



**Aviso legal**

Esta obra está bajo una licencia Reconocimiento 3.0 de Creative Commons. Se permite su reproducción, distribución, comunicación pública y transformación para generar una obra derivada, sin ninguna restricción siempre que se cite al titular de los derechos (Generalitat de Catalunya, Departamento de la Presidencia). La licencia completa se puede consultar en <http://creativecommons.org/licenses/by/3.0/es/legalcode.es>.

© Generalitat de Catalunya  
Departamento de la Presidencia

1.ª edición: enero de 2023

ISBN: 978-84-18601-05-7

DL B 23543-2022

Compaginación y producción:

Entitat Autònoma del Diari Oficial i de Publicacions

# Sumario

---

<b>Prólogo</b>	<b>7</b>
<b>Presentación</b>	<b>9</b>
<b>I. Introducción</b>	<b>13</b>
<b>II. Sistemas de decisión automatizada en el sector público</b>	<b>19</b>
<b>1. Automatizando las ayudas públicas</b>	<b>22</b>
<b>2. En busca del fraude al Estado</b>	<b>27</b>
<b>3. Algoritmos en la policía</b>	<b>30</b>
<b>4. ¿Se pueden predecir los delitos?</b>	<b>34</b>
<b>5. Otros casos de uso</b>	<b>38</b>
<b>III. Transparencia, explicabilidad y rendición de cuentas</b>	<b>41</b>
<b>IV. Herramientas e iniciativas para un uso transparente y ético de la inteligencia artificial (IA)</b>	<b>50</b>
<b>1. El modelo PIO del OEIAC: autoevaluación para un uso ético</b>	<b>54</b>

<b>2. El estándar algorítmico británico y su estrategia de datos</b>	<b>58</b>
<b>3. La estrategia de Barcelona para el uso de algoritmos y datos</b>	<b>60</b>
<b>4. ¿Cómo hacer cumplir la transparencia? La Ley Rider y la guía de información algorítmica</b>	<b>64</b>
<b>5. Los registros públicos de Ámsterdam y Helsinki</b>	<b>67</b>
<b>6. La agencia española para la supervisión de la IA</b>	<b>70</b>
<b>7. El Citizens' Biometrics Council y la participación ciudadana en el debate algorítmico</b>	<b>72</b>
<b>V. Reflexiones finales</b>	<b>76</b>
<b>VI. Glosario breve</b>	<b>78</b>
<b>Bibliografía</b>	<b>80</b>

## Prólogo

Imaginemos que un ordenador “decide” quiénes son los pacientes susceptibles de formar parte de una lista para ser intervenidos prioritariamente en un hospital público de acuerdo con un algoritmo informático, o “señala” las zonas agrícolas a las que habría que otorgar, en primer término, ayudas económicas de acuerdo con su potencial afectación por condiciones vinculadas al cambio climático. O imaginemos, del mismo modo, que un ordenador “determina” sobre qué población se han de centrar los esfuerzos para luchar contra el fraude fiscal.

Estos escenarios, y otros muchos parecidos que ya son una realidad hoy en día, tienen en común varias cosas. En primer lugar, la deseabilidad del objetivo que se plantea (ayudar, facilitar); en segundo lugar, la necesidad de tomar decisiones, distinguir y priorizar entre casos semejantes para optimizar unos recursos escasos. Finalmente, en ambos escenarios el “decisor”, o cuando menos el primer decisor, no es humano, sino un algoritmo debidamente programado y “entrenado”.

Las dos primeras características son típicas de toda decisión política, en general, y de las acciones de las administraciones públicas, en particular. Y se podría decir que no han cambiado durante siglos. A saber, priorizar en un marco de recursos finitos y optimizarlos para conseguir el máximo bien. Sin embargo, la tercera característica resulta toda una novedad histórica y tan solo somos capaces de intuir hacia dónde nos llevará: la aplicación masiva de algoritmos y sistemas de decisión automatizados, en un contexto de *big data*, para desarrollar las tareas definidoras de la buena política.

Pese a la novedad, existe suficiente bagaje, teórico y empírico, para saber que esta ayuda, que sin duda puede multiplicar la operatividad de las Administraciones, no debería en ningún caso resultar contraproducente. Como mínimo en dos sentidos: jamás debería substituir completamente el criterio humano y no debería introducir un elemento de arbitrariedad, aunque no sea (humanamente) interesado, en el resultado. En otras palabras, debería estar sujeta a la misma capacidad

de rendir cuentas (*accountability*) que toda la política “humana”. Y en ningún caso debería convertirse en un factor generador de desconfianza en la ciudadanía hacia las instituciones públicas.

Este es, precisamente, el objetivo de uno de los pilares del gobierno abierto: la transparencia (incluyendo, en este ámbito electrónico, la explicabilidad y la replicabilidad). Acabar con las zonas de oscuridad, incluso de penumbra, asociadas a la toma de decisiones debe ser una prioridad de cualquier gobierno democrático y es una de las funciones primordiales de la Dirección General de Buen Gobierno, Innovación y Calidad Democráticas, la unidad que vela por el buen gobierno y el gobierno abierto en la Generalitat de Catalunya.

El potencial de las nuevas tecnologías con la aplicación de algoritmos automáticos para la gestión pública, la multiplicación de capacidades que puede implicar, debe ser analizado con optimismo, pero sin ingenuidad. En un momento inicial de ciertas prácticas y herramientas, antes de una expansión que se antoja acelerada, conviene reflexionar sobre posibles límites e inconvenientes, pero, sobre todo, conviene que esa inmensa ayuda que pueden suponer los algoritmos esté gobernada bajo los principios del gobierno abierto.

Este volumen sobre transparencia algorítmica en el sector público supone una contribución de gran interés en este sentido. Agradezco al anterior director general de Transparencia, Datos Abiertos y Colaboración, el Sr. Albert Cañigueral, que firma la presentación, haberla promovido.

**Jaume López Hernández**

Director general de Buen Gobierno, Innovación y Calidad Democráticas



## Presentación

El volumen 9 de la colección Govern Obert está dedicado a un tema realmente innovador: la transparencia, la explicabilidad y la replicabilidad de los sistemas de toma de decisiones automatizadas que existen en las administraciones.

En un formato más breve suele hablarse también de la “transparencia algorítmica en el sector público”, pero consideramos que, en estas etapas iniciales de la temática, es importante que seamos claros en los conceptos y la terminología a utilizar. Es decir, en ponernos de acuerdo de qué hablamos y cómo nos referimos a ello.

De entrada, el sujeto de la frase son los “sistemas de decisión automatizada” (conocidos por la sigla ADMS, del inglés *automated decision-making systems*). Como se explica a lo largo de este volumen, estos sistemas de decisión pueden ir desde algoritmos bastante simples, como un árbol de decisión, hasta redes neuronales que se han configurado a partir de un conjunto de datos de entrenamiento y que siguen aprendiendo a medida que procesan nuevos datos. Es importante ser conscientes de esta diversidad dentro de los ADMS y de que, desde el ámbito jurídico, habrá que determinar qué se considera y qué no se considera un ADMS en el sector público.

Una vez sabemos de lo que hablamos —los ADMS—, debemos aclarar la definición de su transparencia, explicabilidad y replicabilidad, así como conocer el valor que estos elementos aportan a estas herramientas.

Desde un punto de vista de gobierno abierto, lo que queremos es promover la publicidad activa, la “transparencia” de estos sistemas de decisión automatizada dentro de las administraciones. Como se explica al inicio de este volumen, en los últimos años se han conocido varios casos de uso de ADMS desde el sector público sin el nivel de transparencia deseable. La datificación y automatización que implica la transformación digital de las administraciones debería ir acompañada

de mayor transparencia de esta parte de la acción de gobierno, con el fin de generar la confianza necesaria en el conjunto de la ciudadanía. Este aspecto, uno de los que queremos incluir en una próxima revisión de la Ley 19/2014, de 29 de diciembre, de transparencia, acceso a la información pública y buen gobierno, ha motivado también la publicación de este volumen. Cabe destacar, asimismo, que en el Plan del Gobierno de la XIV legislatura se explicita la voluntad de la Generalitat de avanzar hacia la creación de un inventario de los sistemas de decisión automatizada.

También hay que diferenciar la “transparencia” de la “explicabilidad” en relación con los ADMS. Si hablamos de transparencia de un sistema, pensaremos en tener acceso al código fuente del programa y a los esquemas de diseño. El hecho de tener acceso al código fuente de un algoritmo, para la mayoría de la población no será la mejor aproximación para garantizar el derecho de acceso a la información, ya que resultará muy complicado entender su funcionamiento interno y qué factores actúan, y de qué manera, en la toma de decisiones. Por lo tanto, preferimos el uso de la palabra “explicabilidad”, con la intención de encontrar los mecanismos que aporten una descripción, lo más comprensible posible, del diseño y el funcionamiento de estos instrumentos, así como del nivel de riesgo que implican. A lo largo del volumen se expone cómo varios gobiernos han abordado la explicabilidad de los sistemas. Cabe destacar el caso del Reino Unido, que hizo un ejercicio de participación cívica deliberativa para recoger e incorporar la opinión y las expectativas de un grupo diverso de personas.

Finalmente está la “replicabilidad” (o a veces también se habla de “reproducibilidad”). Siguiendo las prácticas habituales en el ámbito de la investigación académica, lo que queremos es que grupos independientes y expertos dispongan de toda la información necesaria (datos de entrenamiento, código fuente u otros parámetros de configuración) con el fin de replicar el ADMS que está en funcionamiento en la Administración. Poder replicar los resultados ayuda a verificarlos independientemente y a mejorar el algoritmo, tanto para posibles errores como para sesgos en su desarrollo. Es más en este ámbito de la “replicabilidad” y

no tanto desde el ámbito de la “explicabilidad” que el acceso al código fuente resulta de interés. A la vez, cuando el código es abierto, se hace reutilizable para otras administraciones, cosa que crea eficiencias compartidas. El acceso al código fuente es una cuestión que genera controversia, tal como queda patente en el caso de Civio y el sistema BOSCO para la gestión de ayudas públicas.

Esperamos que la publicación de este texto ayude a avanzar en la reflexión, desde la perspectiva del gobierno abierto, en un tema tan crucial como desconocido para el futuro del sector público. Son solo los primeros pasos de un camino que, en vista de su gran complejidad, hay que hacer acompañados. Acompañados del resto de departamentos con competencias en una materia tan transversal como esta. Y acompañados de los ciudadanos y ciudadanas de Cataluña, haciendo un buen uso de la inteligencia colectiva, para determinar la mejor forma de abordar la transparencia, explicabilidad y replicabilidad de los sistemas de toma de decisiones automatizadas en nuestras administraciones.

**Albert Cañigueral i Bagó**

Coordinador



## I. Introducción

En los Países Bajos, **Chermaine Leysner** recibió, en 2012, una carta de Hacienda solicitando la devolución del subsidio público que recibía por el cuidado de sus tres hijos. La factura era astronómica: más de 100.000 euros. La deuda se había generado durante los cuatro años en los que recibió una prestación que, según el Gobierno neerlandés, no le correspondía.

«Pensé “No te preocupes, esto es un gran error”. Pero no era un error. Era el inicio de algo gordo», decía diez años después Leysner al diario *Politico*.

Leysner, entonces una joven estudiante universitaria de origen surinamés, es una entre los miles de víctimas del **toeslagenaffaire** o escándalo de las prestaciones por el cuidado de los hijos.

Durante años, la autoridad fiscal del país centroeuropeo usó un **algoritmo** que trataba de identificar, sobre el papel, a los ciudadanos con más probabilidades de defraudar a Hacienda. En base a los resultados de este sistema, el Estado acusó erróneamente a familias como la de Leysner, que cayó en una depresión y acabó por separarse de su pareja. Miles de niños incluso fueron separados de sus padres, familias enteras se quedaron sin casa y otras muchas se hundieron en la ruina económica.

Varias investigaciones periodísticas revelaron los detalles del escándalo. El software usado por las autoridades asignaba una nota de riesgo a los ciudadanos a partir de una serie de variables. Pero lo hacía discriminándolos en función de su origen: los ciudadanos con doble nacionalidad —por ejemplo, los padres con orígenes en las antiguas colonias neerlandesas del Caribe— tenían más posibilidades de ser señalados.

Las familias acusadas de cometer fraude no podían saber los motivos que había detrás de esa acusación, y ni siquiera los funcionarios que

recibían la nota de riesgo conocían en base a qué criterios el sistema otorgaba un resultado u otro.

El uso de este algoritmo, como ocurre con cualquier tecnología implementada desde el sector público, se gestó en un contexto político y económico determinados. Los discursos contrarios a la inmigración y opuestos al gasto público dirigido a familias de bajos ingresos —cada vez más presentes en los Países Bajos— animaron al Ejecutivo neerlandés a poner en marcha este sistema.

Tal y como señala **Amnistía Internacional** en su informe *Máquinas xenófobas*, además de un diseño claramente discriminatorio, este software arrastraba muchas otras deficiencias. El algoritmo usaba técnicas de aprendizaje automático (o *machine learning*); es decir, era capaz de aprender de su propia experiencia a partir de los datos que se introducían en el sistema, pero ningún ser humano supervisaba cómo era este proceso.

Un informe parlamentario y las historias de las familias arruinadas tras ser acusadas erróneamente de fraude sacudieron a la opinión pública neerlandesa. El escándalo provocó incluso la dimisión del Ejecutivo en enero de 2021.

El *toeslagenaffaire* es uno de los casos más recientes y llamativos de mal uso de algoritmos e **inteligencia artificial (IA)** por parte de un gobierno, y ejemplifica los riesgos que estos sistemas conllevan si no se utilizan con las debidas garantías y precauciones.

La digitalización del sector público tiene un gran potencial para mejorar la vida de las personas. Pero, como veremos en las siguientes páginas, también puede provocar decisiones injustas y perjudiciales para muchos colectivos, especialmente si este proceso no se realiza de un modo transparente.

La evidencia recabada en los últimos años está llevando a cada vez más gobiernos a buscar mecanismos que eviten un mal uso de los sistemas algorítmicos.

En esta tarea — compleja y llena de obstáculos —, la transparencia y la rendición de cuentas juegan un papel clave. En los últimos años, hemos observado cómo la opacidad en el uso de herramientas tecnológicas facilita la implementación de errores y decisiones injustas por parte de la Administración. Y, además, dificulta a las personas afectadas la defensa de sus derechos y la exigencia de responsabilidades.

¿Pero cómo se están utilizando hoy en día los algoritmos y la IA desde el sector público? ¿Qué precauciones han de tomar los gobiernos para evitar su mal uso? ¿Y qué iniciativas y herramientas se están poniendo en marcha en nuestro entorno para lograr un uso más ético de estos sistemas?

Las páginas de este volumen de la colección Gobierno Abierto tratan de dar respuesta a algunas de estas preguntas.

### **Aclarando ideas y conceptos**

Antes de tratar estas cuestiones, queremos aclarar algunas ideas y conceptos que irán saliendo a lo largo de la publicación. También queremos explicar por qué usaremos unos términos y no otros para hablar de la digitalización de la acción de los gobiernos.

Para empezar, ¿de qué hablamos exactamente al referirnos a un algoritmo? Aunque existen muchas definiciones, una metáfora habitual es compararlos con una receta de cocina: para elaborar un plato, tenemos que seguir una serie de pasos concretos y finitos, que acaban resultando en un plato de comida específico. Las operaciones matemáticas básicas que hacíamos en el colegio cuando aprendíamos a sumar o a multiplicar también son sencillos algoritmos.

Si trasladamos este concepto al actual mundo digital, un **algoritmo** puede definirse como “una secuencia de instrucciones automáticamente ejecutadas por un ordenador”, propone la **UNESCO**. Pero estas definiciones de algoritmo se quedan cortas para entender los sistemas que nos afectan como ciudadanos.

Por ello, a lo largo de las siguientes páginas usaremos otro término que define mejor los sistemas basados en algoritmos que usan las administraciones públicas.

“Los **sistemas de decisión automatizada** o **ADMS** —la sigla en inglés de *automated decision-making systems*— son procesos pensados para optimizar la toma de una decisión mediante el uso de algoritmos y grandes cantidades de datos”, apunta la organización sin ánimo de lucro **AlgorithmWatch**.<sup>1</sup>

En estos sistemas, las decisiones son “inicialmente —parcial o completamente— delegadas a otra persona o entidad corporativa, que luego, a su vez, utiliza modelos de toma de decisiones automatizados para desarrollar una acción”.

En el caso del *toeslagenaffaire*, por ejemplo, el Gobierno neerlandés desarrolló un modelo automatizado en el que delegó la responsabilidad de analizar a los contribuyentes según el “riesgo” de que estuviesen cometiendo un fraude. El análisis servía luego a los funcionarios para abrir una investigación, cuya consecuencia inmediata era suspender la prestación que hasta entonces recibían las familias “defraudadoras”.

Para recuperar la ayuda y evitar tener que devolver el dinero ya asignado, estos hogares debían demostrar que cumplían con los requisitos exigidos para recibir el subsidio. La falta de colaboración de las autori-

---

1 AlgorithmWatch (<https://algorithmwatch.org/en/>) es una organización sin ánimo de lucro con sede en Berlín dedicada a “vigilar, desentrañar y analizar” los sistemas de decisión automatizada y su impacto sobre la sociedad en Europa. Entre otros trabajos, han publicado varios informes describiendo los ADMS usados por gobiernos y empresas en el continente.



dades acabó atrapando a miles de familias en un laberinto burocrático complejo y frustrante, con graves consecuencias para su bienestar.

Como demuestra este caso, el uso de sistemas algorítmicos y de IA se inscribe dentro de procesos complejos, que han de ser entendidos con un “enfoque holístico”, argumentan desde AlgorithmWatch.

“Un sistema de decisión automatizada o ADMS es, por tanto, un marco sociotecnológico que abarca un modelo de toma de decisión, un algoritmo que traduce este modelo a código, los datos que este código utiliza como *input* y el contexto político y económico que rodea a su uso.”

El uso de este término nos permite, además, establecer una diferencia clara con el concepto de **inteligencia artificial (IA)**. Como iremos viendo, algunos de los sistemas analizados usan la IA, pero muchos otros no.

Si acudimos a la definición propuesta por la Comisión Europea,<sup>2</sup> la IA se traduce en “sistemas que muestran un comportamiento inteligente, analizando su entorno y realizando acciones —con cierta autonomía— para alcanzar objetivos específicos”.

Estos pueden estar basados únicamente en un software: por ejemplo, los asistentes de voz, los programas para analizar imágenes, los motores de búsqueda o los sistemas de reconocimiento de voz y facial, o pueden estar integrados en un hardware: los robots avanzados, los coches autónomos o los drones son aplicaciones cada vez más comunes de esta tecnología.

Finalmente, otro concepto importante para entender los temas que vamos a abordar es el de **caja negra** (o *black box*). Así se conoce a los ADMS basados en algoritmos cuyos procedimientos y variables

---

2 Esta definición está extraída de la propuesta de regulación de la Comisión Europea para la inteligencia artificial (Artificial Intelligence Act) [<https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=COM%3A2018%3A237%3AFIN>]. A lo largo de este volumen nos referiremos en varias ocasiones a esta propuesta y la describiremos en detalle.

no son evidentes o conocidos. Es decir, al observar estos sistemas, generalmente podemos saber cuál es la información (o datos) que entra y sale, pero no la manera en la que se ha llegado a una decisión u otra.

Desde el punto de vista de los gobiernos, el uso de sistemas de caja negra deriva en graves problemas de opacidad. En el caso mencionado de los Países Bajos, la falta de transparencia del modelo de clasificación de riesgo fue una barrera para que miles de familias y cuidadores probasen su inocencia ante la acusación de fraude. Pero también supuso un obstáculo para que las autoridades no implicadas en su diseño —y para la sociedad civil en su conjunto, incluyendo a periodistas e investigadores— pudiesen llegar al fondo del asunto. Esto provocó que el carácter discriminatorio e injusto del sistema tardase varios años en salir a la luz.



### Algoritmo

Secuencia de instrucciones o pasos seguidos para lograr un resultado específico. En el actual mundo digital, los algoritmos generalmente son ejecutados por un ordenador de forma automatizada.



### Sistema de decisión automatizada o ADMS (*automated decision-making system*)

Sistema pensado para optimizar la toma de una decisión mediante el uso de algoritmos y grandes cantidades de datos. Este proceso se integra dentro de un marco social y tecnológico concreto, que determina sus efectos sobre las personas que interactúan con ellos.



### Inteligencia artificial (IA)

Sistemas que muestran un comportamiento considerado inteligente, analizando su entorno y realizando acciones concretas para un objetivo específico. Pueden basarse únicamente en un software (por ejemplo, los asistentes de voz, los programas de análisis de imágenes, etc.) o integrarse en dispositivos de hardware (robots avanzados, coches autónomos, etc.).



### Cajas negras

Sistemas automatizados y basados en algoritmos cuyos procedimientos y variables no son evidentes o conocidos. Al observarlos, generalmente podemos saber cuáles son los datos que entran y los que salen, pero no la manera en que el sistema llega a una decisión u otra.

## II. Sistemas de decisión automatizada en el sector público

El uso de sistemas de decisión automatizada por parte de los gobiernos ha crecido enormemente en los últimos años. Los legisladores y funcionarios públicos ven estos sistemas como una oportunidad para optimizar y agilizar procesos, y también como una vía para responder a la demanda ciudadana de servicios públicos más eficientes.

Esta combinación, unida al desarrollo de nuevas técnicas y herramientas digitales, hace pensar que el uso de ADMS seguirá creciendo en los próximos años. **La gran cuestión está en cómo se desplegarán estos sistemas y qué mecanismos de transparencia y rendición de cuentas tendremos los ciudadanos y ciudadanas a nuestra disposición.**

Pero antes de eso, es necesario conocer y poder ubicar dónde hay un algoritmo funcionando. El primer paso para saber cómo una tecnología nos afecta en nuestro día a día es saber que existe. A partir de ahí, si contamos con la información suficiente, podremos intentar entender cómo el sistema ha tomado una decisión u otra.

### ¿Dónde hay algoritmos?

Hoy en día es frecuente oír hablar de los algoritmos de Instagram, YouTube, Twitter o TikTok. Las grandes tecnológicas dueñas de las principales redes sociales usan complejos softwares que definen qué tipo de contenido vemos en nuestra pantalla. Entender su funcionamiento, o al menos ser conscientes de su existencia, es importante para tratar de sortear riesgos como la influencia de las noticias falsas o el filtro burbuja.<sup>3</sup>

---

<sup>3</sup> El *filtro burbuja* es el estado de aislamiento intelectual provocado por las redes sociales al filtrar y mostrarnos información basada únicamente en nuestros supuestos intereses y afinidades ideológicas y culturales.

Tal y como sostiene la investigadora **Anjana Susarla**, “la nueva alfabetización digital no pasa por usar un ordenador o estar en internet, sino por entender y evaluar las consecuencias de una vida siempre conectada”.

Esta reflexión es trasladable a otras áreas diferentes a las redes sociales. La automatización afecta a aspectos muy concretos de nuestras vidas. Desde si un banco nos concede un crédito hasta si el Gobierno nos investiga por haber cometido un supuesto fraude. Ámbitos en los que, sin embargo, resulta generalmente difícil detectar la presencia de un sistema algorítmico, especialmente si los actores que los implementan no hacen un esfuerzo por ser transparentes.

Más adelante en este volumen veremos como en los últimos años están creciendo las propuestas y los mecanismos para explicar a la ciudadanía qué procesos se están automatizando desde el sector público. Pero el camino por recorrer para los gobiernos todavía es largo.

La periodista **Karma Peiró**<sup>4</sup> señala en el pódcast “**Algoritmos y gobiernos**” que a día de hoy todavía es muy difícil ubicar estos sistemas, “porque todavía se conocen poco y porque la propia Administración hasta hace poco no ha visto necesario rendir cuentas a la ciudadanía”.

Antes de entrar a explicar las áreas donde el uso de ADMS es más frecuente, merece la pena mencionar varias iniciativas en pos de la transparencia surgidas desde la sociedad civil.

Una de ellas es el proyecto **Automating Society**, con el que AlgorithmWatch y la **Fundación Bertelsmann** pretenden mapear los sistemas automatizados usados por el sector público en Europa. En su último informe, publicado en 2021, sus autores destacan la velocidad con

---

4 Karma Peiró es una periodista especializada en nuevas tecnologías y en el impacto ético de la digitalización. Es también autora de un informe para la Autoridad Catalana de Protección de Datos (APDCAT) en el que se describen más de 50 ejemplos de sistemas automatizados usados en Cataluña, tanto en el ámbito público como privado [[https://apdcat.gencat.cat/ca/documentacio/intelligencia\\_artificial/](https://apdcat.gencat.cat/ca/documentacio/intelligencia_artificial/)].

la que estos sistemas se están extendiendo por el continente y lanzan una advertencia: “Necesitamos más transparencia sobre los objetivos y más datos sobre los ADMS que se están probando y desplegando”.

“Sin la capacidad de saber con precisión cómo, por qué y con qué fin se implementan los sistemas de decisión automatizada, todos los demás esfuerzos para la reconciliación entre dichos sistemas y los derechos fundamentales están condenados al fracaso”, apuntan.

En el capítulo dedicado al Estado español —el estudio recoge un análisis individualizado sobre dieciséis países del continente—, sus autores critican como los organismos públicos difunden poca información de los sistemas que utilizan y “son reacios a hacerlo cuando esta es solicitada por ciudadanos u organizaciones”.

Otro intento por mapear dónde se están usando ADMS, en este caso con un claro impacto social, es el **Observatorio de Algoritmos con Impacto Social** (OASI, por su sigla en inglés). Este proyecto de la Fundación Eticas, una entidad vinculada a la consultora Eticas Consulting, con sede en Barcelona, ha recopilado más de un centenar de sistemas implementados tanto por gobiernos como por el sector privado en todo el mundo. Las herramientas incluidas en este registro abarcan ámbitos como la salud, los servicios sociales, la policía, el comercio y el empleo, entre otras áreas.

### **Áreas de acción y casos de uso**

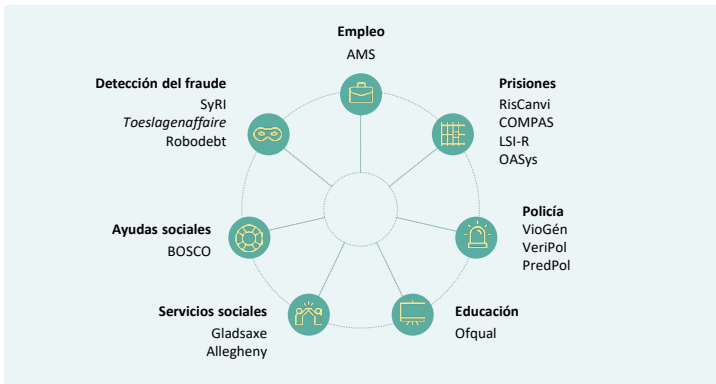
El uso de ADMS en el sector público se extiende a muchas áreas de la Administración. En las siguientes páginas describiremos varios casos que, por diferentes motivos, han sido documentados en los últimos años por medios de comunicación, organizaciones sociales e investigadores.

El que sigue no es ni mucho menos un listado exhaustivo de sistemas utilizados en nuestro entorno. Se trata, en cambio, de una aproxima-

ción para intentar entender cómo los ADMS se están introduciendo y desplegando desde los poderes públicos.

Geográficamente, algunos de estos sistemas se ubican en Cataluña o en el Estado español (BOSCO, VioGén, RisCanvi, VeriPol...), mientras que otros han aparecido en países europeos (SyRI, Gladsaxe, Ofqual, AMS...). También pondremos el foco en varios sistemas de Estados Unidos y el mundo anglosajón (COMPAS, PredPol, Allegheny...).

Como se observa en la **figura 2**, las áreas de acción donde se utilizan también son muy diversas. Aunque es cierto que en los últimos años los gobiernos han puesto un especial interés en automatizar procesos en el área de las **ayudas públicas y las políticas sociales**, también el ámbito policial y el sistema penal ofrecen numerosos casos de uso. En las siguientes páginas igualmente mencionaremos ADMS que se han usado o se usan en **educación** y en **empleo**.



## 1. Automatizando las ayudas públicas

Los algoritmos permiten a las administraciones públicas agilizar procesos que, hasta hace poco, recaían sobre funcionarios de carne y hueso. En esta labor, los gobiernos usan aplicaciones informáticas aparentemente sencillas, pero que en ocasiones pueden causar problemas. Así ha ocurrido en los últimos años con el **sistema BOSCO**.

Las recientes subidas del precio de la luz han agravado la pobreza energética de sectores cada vez más amplios de la población —en 2021, el 16 % de los hogares catalanes no pudo mantener su casa a una temperatura adecuada, según el Instituto de Estadística de Cataluña (Idescat)—.

Entre las ayudas sociales a las que pueden optar las familias que afrontan este problema está el **bono social eléctrico**, un descuento en el recibo de la luz que depende del Gobierno central.<sup>5</sup>

El bono social está en funcionamiento desde hace más de una década, un tiempo en el que ha sufrido diversos cambios, tanto en su contenido como en la forma de solicitarlo. En 2017, el Ejecutivo español cambió varios de los criterios para acceder a la ayuda e introdujo la aplicación informática BOSCO en el proceso de asignación.

Este programa funciona del siguiente modo. Las familias que quieren optar a la ayuda envían toda la documentación requerida a las distribuidoras eléctricas, que son las encargadas de gestionar la prestación. Estas introducen los datos en el sistema BOSCO y el programa realiza un cálculo que determina si cumplen o no con los requisitos marcados por el Gobierno.

BOSCO fue desarrollado internamente por el Gobierno español con el objetivo de agilizar y optimizar un proceso muy común en las administraciones: la revisión del papeleo presentado por un ciudadano que opta a algún tipo de ayuda o subvención.

## **Errores en la aplicación**

En 2018, la **Fundación Civio**, un medio periodístico que investiga temas relacionados con los poderes públicos, creó un simulador que

---

<sup>5</sup> La rebaja en la factura puede oscilar entre un 25 % y un 40 %, en función de si el consumidor es considerado “vulnerable” o “vulnerable severo”. Estas categorías las fija la Secretaría de Estado de Energía del Gobierno español a partir de unos criterios relacionados con la renta de las familias.

imitaba el funcionamiento de BOSCO. A través de su web, cualquier ciudadano puede todavía hoy calcular si cumple con los criterios fijados por el Gobierno y, por tanto, acceder a la prestación.

Desde Civio explican que al poco de publicar el simulador, muchos ciudadanos comenzaron a enviarles mensajes diciendo que algo no cuadraba. A pesar de que esta calculadora web sí les daba una respuesta positiva, las eléctricas les denegaban una y otra vez el bono.

“La teoría es que BOSCO se limita a implementar lo que dice la norma de quién tiene derecho y quién no. Lo que pasa es que, en el proceso de implementar en un software una especificación, se producen errores”, señala **David Cabo**, codirector de Civio.

Ante esta situación, en Civio decidieron solicitar, al amparo de la Ley de transparencia, toda la información sobre el funcionamiento interno de BOSCO, incluido su código fuente. Es decir, los comandos y baremos a través de los que funciona el algoritmo. Sin embargo, el Gobierno español accedió solo a entregarles las especificaciones técnicas del programa y los test de verificación, pero rechazó compartir el código fuente del software. Hacerlo, argumentó el Ejecutivo, implicaría problemas de seguridad pública para la defensa nacional y atentaría contra la propiedad intelectual.

El **Consejo de Transparencia y Buen Gobierno (CTBG)**, un órgano independiente del Ejecutivo creado para hacer cumplir la normativa sobre transparencia, también rechazó la apertura del código, aunque solo bajo el argumento de la propiedad intelectual.

### **Demanda para conocer el código fuente**

“El tema de la propiedad intelectual no puede ser un límite, porque por encima de ese límite está el derecho de la ciudadanía a conocer. Y la razón es la seguridad jurídica. ¿Nos van a regir entonces con un sistema de máquinas que nosotros no sabemos cómo funcionan?”, argumenta **Javier de la Cueva**, patrono y abogado colaborador de la Fundación Civio.



La Fundación decidió, en verano de 2019, presentar una demanda contra la resolución del CTBG, redactada por el propio De la Cueva. En el escrito se argumenta que el código fuente de BOSCO es equiparable a una ley, ya que “mediante su ejecución se generan derechos y obligaciones”.

“El ciudadano tiene tanto derecho a inspeccionar su funcionamiento como lo tiene con respecto a cualquier otra norma jurídica”, esgrime la demanda.

La pandemia ralentizó el proceso judicial hasta que, en febrero de 2022, un juez desestimó el recurso presentado por la Fundación, a partir de los mismos argumentos esgrimidos por el Gobierno. Civio decidió entonces presentar un recurso de apelación sobre el que tendrá que decidir la Audiencia Nacional.<sup>6</sup>

## ¿Abrir o no las tripas de los algoritmos?

**Liberar o no el código fuente de los algoritmos usados por la Administración es uno de los debates que están hoy sobre la mesa cuando hablamos de transparencia algorítmica.**

Hacerlo permitiría a investigadores, periodistas o auditores independientes inspeccionar cómo funciona el programa y detectar posibles problemas. En cambio, es frecuente que del otro lado se posicionen las empresas que desarrollan estas herramientas, aludiendo a que la propiedad intelectual de sus productos es una parte fundamental de su negocio.

Desde la Fundación Civio son contundentes sobre la necesidad de conocer las tripas de una aplicación informática que decide sobre un tema tan sensible como una ayuda pública a familias de bajos recursos.

---

<sup>6</sup> El hasta ahora último giro de guion de esta disputa judicial ocurrió en junio de 2022, cuando el CTBG cambió su postura inicial y decidió no oponerse al recurso de apelación presentado por la Fundación Civio.

“Si en lugar de ser un programa informático fuera un grupo de funcionarios con un protocolo para aplicar un decreto, ese protocolo sería público, porque lo que hace es decirles a los funcionarios cómo tienen que aplicar esa norma, que es pública. En este caso, no sabemos qué está haciendo la máquina dentro y si lo está aplicando bien o no”, señala **Eva Belmonte**, codirectora de Civio.

Hay que tener en cuenta, además, que en el caso del bono social, la aplicación fue íntegramente desarrollada por el **Ministerio de Transición Ecológica** y no por un tercero. Pero esto no es lo habitual cuando hablamos de sistemas algorítmicos, especialmente si tienen un cierto nivel de complejidad técnica. Lo más común es que las administraciones compren estos sistemas a consultoras tecnológicas u otras empresas especializadas que, como decíamos, no están dispuestas a dar a conocer los detalles de productos que requieren una gran inversión.

### Hacer público el código de los algoritmos

En 2020, con el objetivo de frenar la expansión de casos de COVID-19, el Gobierno español encargó a la empresa **Indra** el desarrollo de la aplicación móvil “**Radar COVID**”. El software pretendía rastrear los positivos e informar a los contactos estrechos.

En el contrato, el Ejecutivo puso como requisito que el código desarrollado fuese abierto, para permitir su supervisión externa y así evitar comprometer la privacidad de sus usuarios. La apertura del código de esta aplicación fue calificada como “hito” por el diario *El País*, y llegó después de que más de 200 académicos españoles firmaran un manifiesto a favor de la transparencia en el desarrollo de software público.<sup>7</sup>

Pocos meses antes, el **Departamento de Educación** de la Generalitat de Catalunya publicó el algoritmo que utiliza para asignar

---

7 El manifiesto a favor de la transparencia en el desarrollo de software público se puede leer íntegramente aquí: <https://transparenciagov2020.github.io/>.

el número de desempate en el proceso de preinscripción de los colegios públicos. **Carne Meléndez Cuffi**, de la Subdirección General de Gestión de Servicios a la Comunidad y Becas, explica que este paso se dio para hacer más transparente un proceso respecto del que se habían quejado varios ciudadanos.

“Fue la manera de demostrar la base científica del proceso y que este era el mejor modo de tomar una decisión equiprobable”, explica Meléndez Cuffi. Más allá del código fuente, en la web del Departamento de Educación se puede encontrar una explicación sobre el funcionamiento del algoritmo y los pasos que este sigue para calcular el número de desempate.<sup>8</sup>

## 2. En busca del fraude al Estado

La búsqueda del supuesto fraude al Estado en la recepción de ayudas públicas, como hemos visto en el caso explicado al inicio de este volumen, es un ámbito de mucho interés para algunos gobiernos, y en el que el uso de ADMS es cada vez más frecuente.

Además del *toeslagenaffaire*, los Países Bajos han sido escenario de otro caso de uso relevante de ADMS.

En 2014, el Gobierno neerlandés desarrolló un sistema bautizado como **SyRI (Systeem Risico Indicatie)** pensado para “prevenir y combatir el fraude” en áreas como los programas sociales, el pago de impuestos y las aportaciones a la Seguridad Social.

El Ejecutivo central puso esta herramienta a disposición de cualquier administración del país, y varios ayuntamientos de ciudades relevantes como **Eindhoven, Rotterdam y Haarlem** lo adoptaron.

---

8 Tanto el algoritmo como la explicación de este sencillo algoritmo pueden consultarse en la web del Departamento de Educación de la Generalitat de Catalunya: <https://educacio.gencat.cat/ca/departament/transparencia/gestio-serveis-publics/algoritme-preinscripcio/>.

SyRI usaba datos provenientes de diferentes agencias públicas. La información almacenada por gobiernos locales, el sistema público de empleo, la Seguridad Social y las autoridades fiscales servía como base para analizar el riesgo de que un ciudadano estuviera defraudando a las arcas públicas, pero se desconoce exactamente qué tipo de variables utilizaba SyRI y los cálculos que realizaban sus algoritmos. En cualquier caso, este análisis servía para crear una especie de informe de riesgo para señalar a contribuyentes sospechosos y luego cada institución era responsable de investigar o no a los ciudadanos señalados.

### **Denuncia desde la sociedad civil**

Casi desde su puesta en marcha, SyRI fue objeto de críticas por parte de expertos y organizaciones de la sociedad civil. La principal era la opacidad mostrada por el Gobierno, que se negó a revelar qué elementos eran analizados por el sistema para supuestamente detectar el fraude, así como los criterios para realizar el cálculo.

Desde el Ejecutivo se escudaron en que, si revelaban estos detalles, los potenciales defraudadores podrían ajustar su comportamiento y acciones para escapar del análisis.

Otra crítica a SyRI tenía que ver con que el sistema se usaba solo para detectar el fraude en barrios pobres. En una carta dirigida al Ejecutivo neerlandés, el relator especial de la ONU sobre extrema pobreza y derechos humanos, **Philip Alston**, alertó de que el sistema ponía el foco exclusivamente en barrios de bajos ingresos, “victimizando y discriminando” a la población de esas zonas.

Tal y como escribió la periodista Sanne Blauw en un análisis sobre este caso en el periódico *The Correspondent*, “si solo buscas en ciertos lugares, solo encontrarás algo en esos lugares”. “Luego puedes referirte a los datos y decir: ves, hay fraude en este tipo de barrios, vamos a mirar más frecuentemente ahí; mientras otras violaciones no se comprueban”. Finalmente, varias organizaciones civiles y un sindicato llevaron a SyRI a juicio.

## **Sentencia histórica**

En febrero de 2020, un tribunal de la ciudad de La Haya declaró que SyRI era contrario al derecho a la protección de la vida privada y familiar, contemplado en la **Convención Europea de Derechos Humanos**, y suspendió su uso. La sentencia estableció que este derecho está por encima de la legítima lucha contra el fraude desde las administraciones públicas.

La ONG **Privacidad Internacional** habló entonces de una “sentencia histórica” y pionera. Nunca antes un juez había fallado en contra de un sistema de este tipo. “El tribunal consideró que el Gobierno no había logrado un equilibrio entre el derecho a la intimidad y el interés público por detectar el fraude a las arcas públicas, y que el uso de SyRI era desproporcionado en relación con el objetivo que pretendía alcanzar”, apuntaron desde la ONG.

Más allá de la legalidad o legitimidad de esta herramienta, varias investigaciones periodísticas revelaron también que el sistema resultaba poco efectivo y que no había servido para detectar casos de fraude entre los contribuyentes.

## **El esquema Robodebt en Australia**

De características similares, aunque con un recorrido de mayor uso, es el caso ocurrido en Australia mediante el esquema conocido como **Robodebt**.

En 2016, el Gobierno de la isla puso en marcha un sistema automatizado de reclamación de deudas a ciudadanos receptores de ayudas públicas o programas sociales. La herramienta cruzaba los registros de las prestaciones recibidas por los contribuyentes con sus datos sobre ingresos, en manos de las autoridades fiscales.

Varias investigaciones revelaron que el sistema caía con frecuencia en errores e inexactitudes, lo que no impedía el envío automático de no-

tificaciones oficiales a los ciudadanos señalados. Además, las familias que recibían estas cartas se veían bajo la presión de tener que probar su inocencia.

El funcionamiento de este esquema acabó provocando un reguero de víctimas y problemas de salud mental y física en muchas de ellas, además de varios suicidios.

En 2020, el Gobierno australiano canceló definitivamente el uso de este ADMS y anunció la devolución de cerca de medio millón de deudas inexistentes pero que ya habían sido pagadas.

Un año después, un tribunal condenó al Gobierno a una indemnización millonaria a las víctimas fruto de este “vergonzoso capítulo” en la Administración pública australiana, según la sentencia.

### **3. Algoritmos en la policía**

La policía es una de las áreas donde el uso de ADMS tiene un recorrido más consolidado. En el Estado español, el caso mejor documentado y analizado, además de ser el sistema más veterano en uso, es el de **VioGén**.

En 2004, el Congreso de los Diputados aprobó una ley contra la violencia de género pionera en Europa, que incluyó la creación de un protocolo de seguimiento policial para las mujeres que denuncian una situación de violencia, bautizado como VioGén (**sistema de seguimiento integral en los casos de violencia de género**).

El sistema incorpora un software que trata de predecir el peligro que corre la víctima en relación con su agresor. Desde el punto de vista técnico, VioGén analiza 37 indicadores de riesgo diferentes; por ejemplo, las características del presunto agresor, los episodios de violencia sufridos por la denunciante y la vulnerabilidad de la víctima y de sus hijos, entre otros factores.

Cuando una mujer va a la comisaría a denunciar una situación de violencia, un agente de policía le realiza una entrevista y rellena un formulario respondiendo a estos indicadores.

En base a la información introducida, el algoritmo de VioGén da como resultado cinco posibles niveles de riesgo: no apreciado, bajo, medio, alto y extremo.

### **Medidas de protección a las víctimas**

La Policía Nacional, la Guardia Civil y algunos cuerpos de policías locales en España usan VioGén. En Cataluña y Euskadi, los Mossos d'Esquadra y la Ertzaintza usan, respectivamente, protocolos propios. El agente que entrevista a la mujer puede aumentar de forma manual el riesgo asignado por el software si cree que este es demasiado bajo.

La estimación de riesgo calculada por este ADMS sirve a la policía y a los jueces para decidir qué medidas preventivas de protección ponen a la víctima. A fecha de octubre de 2022, desde su puesta en marcha en 2007, VioGén ha evaluado más de 700.000 casos de violencia de género en todo el Estado.

El **Ministerio del Interior** español ha ido mejorando y evolucionando el protocolo en estos años. Por ejemplo, en 2019 se incorporó una medición del riesgo que corren los hijos de las mujeres víctimas de violencia y el riesgo de letalidad para las denunciadas.

Este cambio llegó tras un caso muy polémico: en 2018, los dos hijos de una mujer que había denunciado varias veces a su expareja fueron asesinados por el hombre. En ese momento el sistema no incorporaba preguntas para medir el riesgo para los menores. El Estado reconoció *a posteriori* un fallo en cadena de sus protocolos.

## Deficiencias del sistema

Los expertos señalan que VioGén ha sido capaz de trasladar de forma efectiva el conocimiento científico sobre violencia machista —y el riesgo que corre una mujer que denuncia— a un cuestionario. También destacan como este protocolo ha permitido homogeneizar el tratamiento que reciben las víctimas.

Sin embargo, asociaciones de víctimas y supervivientes advierten que el sistema sigue teniendo deficiencias. Por ejemplo, el escaso número de agentes con formación específica en violencia de género, lo que en ocasiones se traduce en una entrevista deficiente a la denunciante. O la falta de apoyo psicológico en un momento tan delicado como la denuncia, aumentando así el riesgo de que las víctimas respondan de manera sesgada o condicionada por el miedo o el estrés.

Otra crítica habitual es que, a día de hoy, y a pesar de todos los años que lleva en funcionamiento, VioGén apenas ha sido analizado y auditado de forma independiente, lo que impide conocer sus posibles defectos y las potenciales oportunidades de mejora. En 2022, la Fundación Eticas presentó una auditoría algorítmica externa sobre VioGén.<sup>9</sup>

## VeriPol, el algoritmo para detectar denuncias falsas

En 2018, la Policía Nacional española comenzó a usar **VeriPol**, una herramienta de IA pensada para detectar denuncias falsas relacionadas con robos o hurtos.

El sistema analiza las palabras utilizadas por el denunciante y trata de identificar aquellas combinaciones que, según el registro de denuncias sobre el que está entrenado, son las más frecuentes cuando la persona está mintiendo. Este análisis se realiza a partir de la transcripción de la denuncia que los agentes hacen al lenguaje escrito.

---

9 La auditoría de la Fundación Eticas sobre VioGén alerta de la existencia de barreras que disuaden a las mujeres de denunciar a su agresor, así como de diversos problemas en la realización del cuestionario que nutre de información al software.



Los algoritmos de VeriPol se entrenaron a partir de algo más de mil denuncias, tanto falsas como reales, pero la información pública en torno a este ADMS es muy escasa. Varios expertos han señalado que la construcción de este software ha desatendido la diversidad de expresiones e idiomas que se utilizan en España, lo que puede desembocar en errores a la hora de identificar denuncias falsas. También se ha señalado que la muestra sobre la que está construida este ADMS no es suficientemente amplia.

Un estudio sobre la precisión de VeriPol publicado por uno de los impulsores de este sistema señala que tiene una precisión del 91%; es decir, que de cada 100 denuncias, el software acusa erróneamente a 9 personas.

### **Los riesgos de la actuación policial predictiva**

Otra área donde el uso de ADMS ha ganado fuerza en los últimos años es la conocida como **actuación policial predictiva** (*predictive policing*). Estos sistemas aseguran ser capaces de averiguar en qué zonas de una ciudad o región se cometerán más delitos, permitiendo así a los departamentos policiales enviar más o menos patrullas a cada zona.

Uno de los casos más controvertidos en este ámbito es el de **PredPol**, un software creado al inicio de la década de 2010 y que se ha aplicado en diversas jurisdicciones de Estados Unidos. PredPol funciona en base a los datos históricos de delitos de las áreas donde se implementa, aunque cada departamento puede elegir si incluye o no entre ellos las infracciones por “alteración del orden público”. Es decir, delitos menores como la venta o el consumo de pequeñas cantidades de droga y el vagabundeo, habituales en los barrios más pobres.

La autora e investigadora estadounidense **Cathy O’Neil**, en su libro *Armas de destrucción matemática. Cómo el big data aumenta la desigualdad y amenaza la democracia*, explica como PredPol acaba por generar un “bucle de retroalimentación” que perjudica a los barrios

más pobres: allí es donde ya hay más patrullas y, por lo tanto, donde se generan más datos para alimentar el sistema.

“Por desgracia, en la mayoría de las jurisdicciones, un mapa de la delincuencia generado de este modo lo que traza en realidad es el rastro de la pobreza”, apunta O’Neil.

En los últimos años, diversas investigaciones han desacreditado la precisión y utilidad de este tipo de sistemas. En el caso de PredPol, se ha probado que perpetúa la discriminación racial en las ciudades de Estados Unidos donde se ha usado, ignorando los delitos en las zonas de población mayoritariamente blanca y centrando su análisis en zonas más pobladas por personas de origen latino y afroamericano.

En 2020, el departamento de policía de Los Ángeles —uno de los clientes más antiguos de PredPol— anunció que cancelaba su uso. Otros departamentos en el país lo han hecho tras no poder demostrar la efectividad del sistema en la reducción de las tasas de criminalidad.

En el Estado español, el uso de sistemas para la predicción geográfica de delitos ha sido hasta ahora anecdótico. En 2015, la policía local del Ayuntamiento madrileño de Rivas Vaciamadrid comenzó una prueba piloto con Pred-Crime, un software desarrollado por la empresa Eurocop, pero pocos meses después el consistorio canceló su uso.

#### **4. ¿Se pueden predecir los delitos?**

La predicción del riesgo en el ámbito penal es una de las áreas donde el uso de algoritmos está más extendido. Hoy en día están en uso sistemas muy diferentes, pero en esencia hablamos de herramientas pensadas para estimar el riesgo de que una persona detenida —o ya en prisión— vuelva a cometer un delito.

El uso de la analítica predictiva o de la IA en este campo genera muchas dudas éticas. Entre ellas está el riesgo de discriminación hacia

ciertas comunidades. El caso probablemente más citado para ejemplificar este riesgo es el de **COMPAS** (Correctional Offender Management Profiling for Alternative Sanctions), un software de predicción de riesgo desarrollado por la empresa estadounidense Northpointe a finales de la década de 1990.

Como otros ADMS similares, COMPAS calcula, a partir del análisis de varios factores, el riesgo de que una persona vuelva a delinquir. En su caso, este cálculo sirve para que los jueces decidan si dejan en libertad o encarcelan preventivamente a un detenido, en base al riesgo de reincidencia estimado.

En 2016, la organización de periodismo de investigación **ProPublica** publicó un artículo en el que cuestionaba la precisión de las mediciones hechas por este algoritmo.<sup>10</sup>

El medio señaló que el software estaba generando más falsos positivos entre los detenidos negros que entre los blancos. Es decir, entre los detenidos que no volvieron a delinquir, los negros tenían el doble de probabilidades que los blancos de ser calificados como de “alto riesgo”.

La investigación de ProPublica reveló el carácter discriminatorio de COMPAS. El motivo era que este sistema se alimentaba de datos ya sesgados: en las cárceles estadounidenses la población afroamericana está sobrerrepresentada. Además, sus algoritmos no trataban de corregir este sesgo y acababan por funcionar bajo criterios racistas.

### **RisCanvi, el ADMS de las cárceles catalanas**

Las cárceles catalanas utilizan desde hace más de una década un ADMS para la gestión de su población penitenciaria. El **RisCanvi** (acrónimo de las palabras *risc* y *canvi*, ‘riesgo’ y ‘cambio’ en catalán) es un

---

10 La investigación de ProPublica, publicada en un artículo con el título *El sesgo de las máquinas* (Machine Bias) [<https://www.propublica.org/article/machine-bias-risk-assessments-in-criminal-sentencing>] ampliamente difundido, se basó en una pequeña filtración de datos sobre el uso de COMPAS en un condado de Florida.

protocolo de valoración de riesgo que se aplica a las personas que están en prisión y que sirve de apoyo a las decisiones de los funcionarios públicos.

Tal y como explica **Antonio Andrés Pueyo**, uno de los creadores de este protocolo y director del Grupo de Estudios Avanzados en Violencia (GEAV) de la Universidad de Barcelona, RisCarvi se desarrolló inicialmente para estimar el riesgo de reincidencia de delincuentes violentos, como homicidas o agresores sexuales, que estaban a punto de cumplir su condena.

Sin embargo, desde su puesta en marcha en 2010, el sistema ha evolucionado mucho. Hoy RisCarvi se aplica a cualquier persona que pase por una prisión en Cataluña, sea un interno preventivo o ya condenado. En el caso de los segundos, además de la posibilidad de cometer un delito violento, RisCarvi trata de medir cuatro tipos de comportamiento:

- La violencia autodirigida (suicidios, intentos de suicidio y autolesiones).
- La violencia intrainstitucional, es decir, conductas violentas o agresiones dentro de la cárcel a otros internos o a funcionarios.
- La reincidencia general o comisión de cualquier delito, violento o no, una vez en la calle.
- El posible quebrantamiento de condena, por ejemplo, al no regresar tras un permiso.

RisCarvi fundamenta su cálculo en el análisis de 43 variables relacionadas con la biografía criminal de la persona, su comportamiento en prisión, su biografía personal, factores de tipo clínico y variables de personalidad. Entre otras diferencias con el sistema COMPAS, RisCarvi por ejemplo no usa técnicas de aprendizaje automático para hacer su estimación de riesgo.

Su uso práctico también es diferente. Los funcionarios de prisiones —desde los trabajadores sociales hasta los psicólogos que están en contacto directo con el preso— se encargan de introducir toda la información relacionada con los factores de riesgo al sistema. Y el software

da como resultado tres niveles de riesgo, en forma de semáforo: rojo (*alto*), amarillo (*medio*) y verde (*bajo*). Cada seis meses el RisCanvi de cada preso se actualiza.

RisCanvi está pensado como un sistema de apoyo a las decisiones de los funcionarios públicos en diferentes situaciones. Por ejemplo, a la hora de elaborar un plan de mejora de conducta y reinserción laboral, pero también cuando un juez ha de decidir sobre una libertad condicional o un tercer grado.

Además, los funcionarios que interactúan con el sistema pueden modificar manualmente su resultado, aunque no hay datos públicos sobre con qué frecuencia ocurre esto.

### “Margen de objetividad” y críticas

Como decíamos, desde su creación, RisCanvi ha sufrido diversos cambios y actualizaciones. Si en sus comienzos el software se construyó en base a los datos de 600 casos, posteriormente se revisó usando información de hasta 30.000 presos de las cárceles catalanas.

Existen pocos estudios sobre la precisión de este sistema. Un informe del Departamento de Justicia señala que su capacidad predictiva es similar a la de otras herramientas parecidas usadas en el extranjero,<sup>11</sup> pero, como cualquier algoritmo, RisCanvi no es infalible.

De ahí que los resultados del software sean luego evaluados por el equipo de profesionales que trabaja con el preso, señala **Manuel Portela**, miembro del grupo Ciencia Web y Computación Social de la Universidad Pompeu Fabra (UPF). El equipo al que pertenece Portela ha tenido acceso al funcionamiento interno del algoritmo y ha entrevistado a profesionales que trabajan con él.

---

11 Otros sistemas predictivos usados en el ámbito penal ampliamente conocidos son el **LSI-R** (Level of Service Inventory-Revised), usado en Estados Unidos, y el **OASys** (Offender Assessment System), utilizado en las prisiones del Reino Unido.

“La mayor utilidad de RisCanvi tiene que ver con la objetividad que provee, ya que conoce históricamente muchos más datos que el propio profesional. Les da un margen de objetividad que les permite evaluarse a sí mismos y a la respuesta que dan”, apunta este investigador.

Con todo, el uso de RisCanvi no ha estado exento de críticas. El debate en torno a la utilización de herramientas automatizadas en las áreas penal y penitenciaria divide a los expertos. En el caso de este protocolo, una de las quejas habituales tiene que ver con el escaso número de estudios independientes que analicen la eficacia del sistema y sus efectos sobre los presos.

El científico de datos **Carlos Castillo Ocaranza**, que lidera el grupo de la UPF al que pertenece Portela, coincide con esta necesidad de aumentar la vigilancia sobre el sistema, pero también señala que, de algún modo, RisCanvi “está pagando el precio de ser pionero en esta área”.

“El RisCanvi fue implantado en 2010 y en ese momento había muy poca discusión sobre equidad algorítmica o procesos de gestión del riesgo algorítmico. Nada de eso existía”, reflexiona. La regulación sobre inteligencia artificial que está debatiendo la Comisión Europea incluye explícitamente a los predictores de riesgo en el ámbito penal dentro de los sistemas de alto riesgo y les requiere un esfuerzo especial de transparencia.

Castillo apunta a la conveniencia de que las autoridades establezcan “unos parámetros de uso aceptables de la inteligencia artificial” y “hagan pasar a RisCanvi por esos parámetros”. “Hay que coger los sistemas que están funcionando y, retroactivamente, hacerlos pasar por la norma”, señala.

## 5. Otros casos de uso

En los párrafos anteriores hemos descrito cuatro áreas de la Administración pública en los que el uso de ADMS ya es frecuente. Pero

la automatización de procesos dentro de los gobiernos se extiende a muchos otros ámbitos.

A continuación mencionamos otros casos que han adquirido notoriedad en los últimos años para mostrar la variedad de usos, objetivos y áreas donde se utilizan.

- **La selectividad británica de 2020.** En verano de 2020, el Gobierno británico canceló los exámenes de acceso a la universidad por la pandemia y decidió usar un algoritmo que calculaba la nota de los alumnos en base a varios factores, entre ellos, los resultados históricos del centro al que pertenecía el estudiante. El software usado por el organismo público que regula los exámenes oficiales del Reino Unido (**Ofqual**) provocó una bajada generalizada de las notas de los estudiantes, siendo los más afectados los alumnos de escuelas públicas. Las protestas de los estudiantes hicieron que el Ejecutivo diese marcha atrás y retirase este algoritmo.
- **Detección de menores vulnerables.** El Gobierno local de la localidad danesa de **Gladsaxe** desarrolló un ADMS para detectar, en una fase temprana, menores que pudiesen necesitar atención de los servicios sociales, analizando datos sobre la realidad social y económica de los hogares. Organizaciones sociales advirtieron de que el sistema permitía imponer medidas muy duras sobre las familias, sin apenas control externo. Finalmente, las autoridades de protección de datos del país denegaron el permiso para almacenar información tan sensible y el sistema se quedó en un cajón.
- **El algoritmo de Allegheny.** La autora estadounidense **Virginia Eubanks** explica en su libro *La automatización de la desigualdad* cómo el condado de **Allegheny**, en Pensilvania (Estados Unidos), compró un algoritmo para estimar el riesgo de que un menor sufriera maltrato doméstico en sus primeros cinco años de vida. El programa analizaba más de un centenar de variables y, si la puntuación era alta, los servicios sociales abrían una investigación a la familia. Este ADMS señaló injustamente a familias de clase trabajadora que, en muchos casos, perdieron la custodia de sus hijos.

- **Puntuación a desempleados.** La agencia pública de empleo austriaca (**AMS**) inició en 2016 un programa para clasificar las posibilidades de inserción laboral de las personas desempleadas, asignándoles una puntuación que, en teoría, estimaba sus opciones de encontrar un trabajo. Varias investigaciones documentaron que el sistema discriminaba a varios colectivos, entre ellos, a las mujeres, a quienes asignaba una nota peor que a los hombres con la misma experiencia y habilidades. La autoridad de protección de datos austriaca declaró finalmente ilegal este sistema y suspendió su uso.



### III. Transparencia, explicabilidad y rendición de cuentas

La preocupación por la falta de transparencia, o la exigencia de la misma, viene, por tanto, porque los ADMS son sistemas complejos que, igual que los seres humanos, cometen errores y provocan decisiones injustas. Este es el motivo por el que los resultados de un sistema automatizado han de ser entendidos y explicados desde los poderes públicos.

En los últimos años, al mismo tiempo que la complejidad de la IA aumentaba, cada vez más investigadores han señalado la importancia de avanzar en la **inteligencia artificial explicable o XAI** (*explainable AI*),<sup>12</sup> en oposición al desarrollo de cajas negras. Como veremos más adelante, con modelos avanzados de IA esta es una tarea en muchas ocasiones compleja. Pero antes de eso vamos a acudir a otra definición concreta sobre qué es la *explicabilidad* de los algoritmos y de la IA, más centrada en la relación entre los sistemas y las personas que interactúan con ellos.

“Los impulsores de una IA han de comprometerse con la transparencia y la divulgación responsable de información sobre el sistema. Para ello deben proporcionar información relevante, adecuada al contexto y coherente con la tecnología utilizada”, establece la **Organización para la Cooperación y el Desarrollo Económicos (OCDE)** en sus recomendaciones para un buen uso de estas tecnologías (**figura 3**).

Este organismo internacional señala, además, los objetivos que esta transparencia activa busca. Entre ellos están que los usuarios o destinatarios de la IA sean conscientes de esta interacción —algo que, por lo general, no se cumple en los ADMS analizados— y que estas mismas personas entiendan los resultados del sistema y puedan impugnarlos si lo consideran necesario.

---

12 Este tipo de IA, y los procesos y métodos que la componen, “permiten a los usuarios humanos comprender y confiar en los resultados y productos creados por los algoritmos de aprendizaje automático”. La XAI “se utiliza para describir un modelo de IA, su impacto esperado y sus posibles sesgos”, según la definición propuesta por la compañía IBM.

### Transparencia y explicabilidad

“Los impulsores de una IA han de comprometerse con la transparencia y la divulgación responsable de información sobre el sistema. Para ello, deben proporcionar información relevante, adecuada al contexto y coherente con la tecnología utilizada.” OCDE, 2019

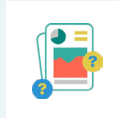
#### Objetivos:



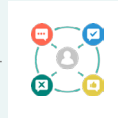
Fomentar una comprensión general sobre el sistema de IA.



Hacer que las personas sean conscientes de sus interacciones con el sistema.



Permitir que las personas afectadas puedan entender sus resultados.



Facilitar que puedan impugnar este resultado a partir de información sencilla y comprensible sobre por qué se llegó a una decisión.

Fuente: Recomendaciones del Consejo en Inteligencia Artificial. OCDE, 2019.

## Entender para poder rendir cuentas

**Albert Sabater**, director del **Observatorio de Ética en Inteligencia Artificial (OEIAC)** de la Universidad de Girona, sociólogo y experto en métodos de investigación social, establece dos grandes tipos de transparencia algorítmica en las administraciones.

La primera tiene que ver con la capacidad de los gobiernos para entender y explicar las cadenas de razonamiento que llevan a un sistema automatizado a dar un resultado u otro. Es decir, un entendimiento “técnico” del sistema que evite implementar algoritmos de caja negra.

La segunda, según Sabater, es una transparencia “más institucional”, que tiene que ver con cómo las administraciones encaran los procesos de automatización para que la ciudadanía pueda conocer estos sistemas y, llegado el caso, oponerse a ellos.

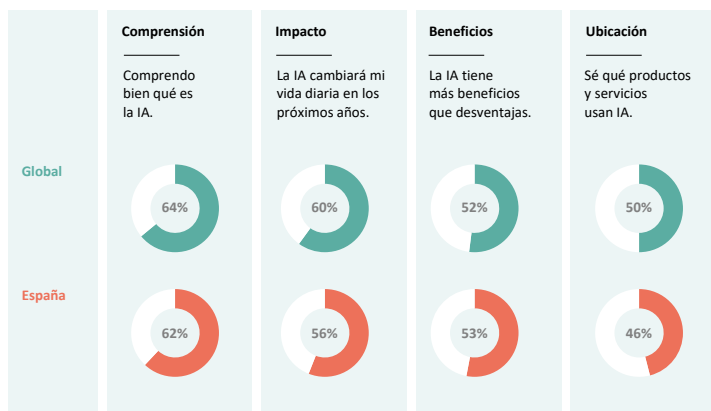
Para este experto, el primer tipo de transparencia “es mucho más difícil”, “porque tenemos poblaciones muy diferentes a las que asignamos recursos” y “porque la gran mayoría de los sistemas de inteligencia

artificial de última generación se basan en aprendizaje automático, usando muchísimos datos”, señala.

**Cristina Pombo**, asesora en transformación digital para el sector público en el **Banco Interamericano de Desarrollo (BID)**, cita otro valor crucial para la implementación de sistemas algorítmicos y de IA: generar confianza entre la ciudadanía.

**“En las conversaciones sobre transparencia hay algo muy importante y es ganarse la confianza de los ciudadanos siendo transparentes en el uso de la inteligencia artificial. Que los ciudadanos entiendan qué es lo que está pasando detrás de quien toma las decisiones”**, apunta. Esto implica proporcionar la información necesaria para que “exista una comprensión general” de la solución que se ofrece.

Una encuesta del instituto **Ipsos** sobre las opiniones y expectativas del uso de la IA en el mundo revela datos significativos sobre este punto (figura 4), como por ejemplo que en España apenas el 46% de la población es capaz de ubicar los productos y servicios que usan IA (frente al 50% global), o que poco más de seis de cada diez ciudadanos (tanto en España como a escala global) comprende bien qué es esta tecnología.



## **Criterios de responsabilidad**

La confianza en los sistemas de IA está íntimamente ligada con el principio de responsabilidad respecto a las decisiones que afectan a la ciudadanía.

**“Hay una tendencia a transferir la responsabilidad de la Administración al sistema. A decir: el sistema ha decidido. No, el sistema no decide nada. Nosotros decidimos lo que hace, y lo hace en unas determinadas circunstancias, con unos datos determinados y en un contexto específico”**, apunta Albert Sabater.

Establecer reglas claras sobre quién se hace responsable sobre el buen o mal funcionamiento de los sistemas, tal y como recomienda Pombo, no es sencillo.

Una opción para afrontar este reto es la creación de certificados específicos para los ADMS, señala Carlos Castillo. Una especie de “sello de calidad” entregado por una entidad independiente que certifique las garantías éticas y de buen uso de los sistemas. Cuando hablemos más adelante de las herramientas concretas a disposición de los gobiernos para mejorar la transparencia, veremos que muchas iniciativas van en esta línea.

Otro obstáculo para lograr una verdadera explicabilidad y transparencia tiene que ver con la propia manera en la que se desarrollan e implementan estos sistemas.

Como hemos mencionado, las administraciones públicas acostumbran a adquirir este tipo de herramientas a terceros. Pero en ocasiones los gobiernos no se aseguran de conocer el tipo de algoritmo que están comprando ni su funcionamiento técnico, tampoco la manera en la que llegan a una u otra decisión, una posición que dificulta o directamente imposibilita una rendición de cuentas real hacia la ciudadanía.

Una posible solución a este reto es establecer unos criterios claros en la compra pública. Es decir, unas cláusulas que garanticen unos

niveles mínimos de explicabilidad y transparencia para los proveedores tecnológicos de las administraciones.

## **Exigencias legales y derecho a la transparencia**

A día de hoy no existen todavía unos principios jurídicos claros que guíen el uso de los ADMS en el sector público,<sup>13</sup> tampoco en lo relativo a la transparencia y a la necesidad de explicar cómo estos sistemas llegan a una u otra decisión, pero ya existen varias normas o propuestas que van en esta dirección.

Para empezar, la **Artificial Intelligence Act europea** plantea unos requerimientos específicos de transparencia para los sistemas que considera de *riesgo alto*, es decir, los usados en áreas como la justicia, la educación, el empleo, las fronteras o en infraestructuras críticas, entre otras áreas. Aunque está por ver en qué se concretan estas exigencias.

En el Estado español, la **Estrategia nacional de inteligencia artificial (ENIA)** —un documento estratégico, no normativo— también menciona que los algoritmos involucrados en la toma de decisiones han de seguir criterios de transparencia y rendición de cuentas, así como de minimización de sesgos.

En la reciente **Ley integral para la igualdad de trato y la no discriminación** se incide en este punto al asegurar que las administraciones “favorecerán la puesta en marcha de mecanismos” para garantizar estos principios y se menciona la necesidad de “promover un sello de calidad de los algoritmos”.

En el ámbito catalán, por su parte, la Generalitat de Catalunya impulsó en 2019 una **Carta catalana por los derechos y las responsabilidades digitales**. Al hablar sobre la ética de la IA y la gobernanza algorítmica, el documento menciona el derecho de la ciudadanía a en-

---

13 <https://derechocal.es/opinion/control-judicial-de-los-algoritmos-robots-administracion-y-estado-de-derecho>.

tender estos sistemas, a que estos sean supervisados por un ser humano, y a que sean explicados para poder conocer los parámetros y datos utilizados (figura 5).

Carta catalana por los derechos y las responsabilidades digitales

 Derecho a entender	<ul style="list-style-type: none"><li>• IA auditable.</li><li>• Reconstruir decisiones.</li><li>• Entender sistemas.</li></ul>
 Derecho a la supervisión humana	<ul style="list-style-type: none"><li>• Impacto social supervisado.</li><li>• Medidas de contingencia.</li></ul>
 Derecho a la explicación	<ul style="list-style-type: none"><li>• Ser consciente de la IA.</li><li>• Conocer parámetros y datos.</li><li>• Definir responsabilidades.</li></ul>

Fuente: Carta catalana por los derechos y las responsabilidades digitales. Generalitat de Catalunya, 2019.

Finalmente, si recurrimos a normas ya consolidadas, hay que mencionar el **Reglamento General de Protección de Datos (RGPD)**, en aplicación en todos los países de la Unión Europea desde 2018.

Esta legislación comunitaria establece en varios de sus artículos (5, 12, 13 y 14) que el organismo responsable del tratamiento de los datos ha de ofrecer información sobre la lógica aplicada por el algoritmo, ya que el ciudadano tiene derecho a saber en qué se basa este algoritmo para tomar una decisión que le afecta directamente.

Igualmente, el RGPD habla, en su artículo 22, del derecho a no ser objeto de una decisión basada únicamente en un proceso automatizado, excepto en casos puntuales, e incluso en estos casos la persona tendrá derecho a la intervención de un ser humano si quiere oponerse a esta decisión.

## **Sentencias que marcan un camino**

Como comentábamos, el hecho de que las normativas sobre el uso de ADMS estén en debate, elaboración o recién aprobadas puede dificultar la labor de los gobiernos a la hora de saber qué se puede y qué no se puede hacer.

En los últimos años, sin embargo, se han dictado varias sentencias, relacionadas con casos de uso polémicos, que dan algunas pistas y que comienzan a sentar jurisprudencia.

**Ana Valdivia**, investigadora del King's College London sobre el uso de la IA y de los sistemas algorítmicos en las políticas migratorias, y el letrado Javier de la Cueva analizan varias de ellas en un artículo donde señalan las “fricciones” que a día de hoy existen entre la ley y los algoritmos.

Uno de los casos analizados es el software que el sistema público de educación italiano usó en 2016 para asignar plazas vacantes por todo el país a su profesorado. Varios enseñantes llevaron este sistema a los tribunales y uno de los demandantes advirtió que, al asignarle una plaza lejos de su lugar de residencia y de su familia, el sistema automatizado había infringido el derecho a la protección de la vida privada y familiar, recogido en el Convenio Europeo de Derechos Humanos.

En 2018, un juez italiano declaró ilegal el algoritmo al considerar que los principios de “participación, transparencia y acceso” en la relación de un ciudadano con los poderes públicos no podían ser “debilitados” al suplantar la actividad humana por un software.

La sentencia del tribunal de La Haya sobre el sistema SyRI que comentamos en el capítulo anterior es otra de las referencias señaladas por Valdivia y De la Cueva. Recordémoslo: este ADMS fue tumbado por la justicia neerlandesa porque, al usar datos provenientes de múltiples fuentes en manos del Estado, no respetaba la intimidad de la vida privada y familiar de las personas analizadas.

Además, esgrime el juez, la falta de transparencia suponía un riesgo de discriminación desconocido para las personas cuyos datos se estaban usando. Javier de la Cueva apunta el carácter “universal” de esta sentencia, ya que “lo que está en juego son los datos privados de las personas”.

**“Todos tenemos derecho a mirar qué está haciendo la Administración pública con nuestros datos. Y si quiere hacer sistemas predictivos, nos tiene que explicar el por qué”,** señala.

### **Definir el problema a resolver**

Antes incluso de tomar todas estas precauciones, los gobiernos han de valorar la conveniencia o no de aplicar un ADMS en un contexto determinado. “Hay que pensar realmente cuál es el problema que se quiere resolver y si ese problema va a poder ser resuelto con un sistema algorítmico”, señala Cristina Pombo, del BID.

En ciertos casos, los algoritmos pueden ayudar a agilizar procesos o a dotar de una mayor objetividad a ciertas decisiones. Pero en ocasiones se han usado para abordar cuestiones que no necesitaban un proceso de automatización. Por ejemplo, porque la información que se buscaba ya existía.

Carlos Castillo cita en esta línea las herramientas de actuación policial predictiva que describimos en el capítulo anterior. Para este científico de datos, dichos sistemas “están contruidos para resolver un problema que no existe, que es que la policía sepa cuáles son las áreas de la ciudad donde se cometen los delitos”. Uno de los argumentos esgrimidos por varios departamentos de policía de Estados Unidos que dejaron de usar el software PredPol fue precisamente que patrullando las calles ya recogían ese tipo de datos.

Además de la definición del problema a abordar, otro punto clave, como decíamos, es contar con los recursos suficientes para hacerlo del modo adecuado. Aquí se incluyen los recursos humanos y técnicos,



pero también las bases de datos necesarias, que han de ser de una calidad suficiente para evitar reproducir o perpetuar sesgos.

“Hay que evaluar y analizar en detalle la calidad de los datos que alimentan o que van a entrenar a los sistemas. Este es el primer paso para establecer su viabilidad”, señala Pombo, quien añade la importancia de que estas políticas se acompañen de leyes de protección y gestión de datos adecuadas.

Disponer de los datos adecuados para poder desarrollar y entrenar un sistema complejo de inteligencia artificial no es tarea fácil. Para la eticista e investigadora sobre la automatización y digitalización de las sociedades **Lorena Jaume-Palásí**, esto se da, además, en pocas ocasiones.

**“Una IA es algo que ayuda en pocas situaciones, porque se necesitan muchos datos. Hay pocos contextos en los que se consiguen los datos necesarios para cuidar del tipo de infraestructura necesaria para la IA”**, señala.

Jaume-Palásí defiende que se está “confundiendo el progreso con la digitalización o la *matematización*”. “Ese *hype* que tenemos con que el progreso está en que digitalicemos y convirtamos todo en inteligencia artificial es una confusión. Que algo sea digital no significa que sea moderno”, apunta.

## IV. Herramientas e iniciativas para un uso transparente y ético de la inteligencia artificial (IA)

La acción pública dirigida a una mayor transparencia en la automatización de procesos es todavía un área en construcción. Durante este capítulo, hablaremos de varias herramientas e iniciativas que los gobiernos y otros actores están desarrollando. El grado de madurez de estas acciones, como veremos, es muy diverso.

Clasificarlas no es una tarea sencilla. Así que para orientarnos mejor vamos a tomar prestada la tipología propuesta por el instituto de investigación británico **Ada Lovelace Institute** en un reciente informe (**figura 6**).<sup>14</sup>

Este informe recoge ocho categorías diferentes, en función de los mecanismos políticos que proponen, sus objetivos y la manera de conseguirlos. Una de ellas, por ejemplo, son los **principios y normas** que cada vez más gobiernos publican como guía para un uso ético de los ADMS. Otra son las **evaluaciones de impacto algorítmico**, que buscan estudiar el uso de un sistema en un contexto determinado y evitar potenciales efectos negativos.

El resto de categorías recogidas en el informe son **prohibiciones y moratorias; mecanismos de transparencia; auditorías e inspecciones algorítmicas; organismos de control externos o independientes; supervisión humana, y condiciones de contratación.**

---

14 El informe sobre responsabilidad algorítmica en el sector público del Ada Lovelace Institute (*Algorithmic Accountability for the Public Sector*) recoge iniciativas implementadas hasta 2021. Se puede consultar de forma íntegra aquí: <https://www.adalovelaceinstitute.org/project/algorithmic-accountability-public-sector/>.

### Principios y normas



- Guías de principios éticos para gobiernos.
- Identifican objetivos políticos de alto nivel y su implicación en el uso de ADMS.
- Orientan sobre cómo implementar estos principios en el diseño y uso de los sistemas.
- Ofrecen estándares normativos para evaluar su uso.

### Prohibiciones y moratorias



- Prohibiciones de uso de ciertos ADMS de alto riesgo.
- A veces son moratorias temporales.
- Hasta ahora se han aplicado sobre todo a tecnologías de reconocimiento facial en Estados Unidos.

### Mecanismos de transparencia



- Proveen información sobre los ADMS al público general: personas afectadas, periodistas y sociedad civil.
- Orientados a que la ciudadanía pueda entender los sistemas en uso y exigir respuestas y responsabilidades.
- Ejemplos: registros públicos de algoritmos, solicitudes de acceso al código fuente y explicaciones sobre la lógica algorítmica.

### Evaluaciones de impacto algorítmico



- Estas evaluaciones (*algorithmic impact assessments*, AIA) son cada vez más usadas por los gobiernos.
- Buscan estudiar el uso de un ADMS en un contexto determinado.
- Pretenden comprender, categorizar y responder mejor a los posibles daños y riesgos que plantean.

### Auditorías e inspecciones normativas



- Las auditorías inspeccionan un ADMS para comprenderlo y evaluarlo respecto a algún estándar normativo.
- Pueden ser auditorías técnicas que evalúen la fiabilidad, los posibles sesgos u otros aspectos.
- También pueden ser inspecciones normativas que evalúen un ADMS respecto a una legislación.

### Organismos de control externos/independientes



- Entes independientes del gobierno que supervisan el uso de ADMS en el sector público.
- Buscan garantizar la rendición de cuentas de las administraciones.
- También formulan recomendaciones, aplican sanciones o toman decisiones sobre el uso de ADMS.

### Supervisión humana



- Políticas que exigen que las decisiones apoyadas en ADMS garanticen la equidad y ofrezcan la posibilidad de reparación individual si esta es sesgada o errónea.
- Proporcionan espacios para que individuos o grupos afectados puedan debatir e impugnar decisiones.

### Condiciones de contratación



- Tratan de trasladar las normas de transparencia y rendición de cuentas a los ADMS adquiridos por las administraciones.
- Son un área de intervención creciente para lograr más transparencia y rendición de cuentas.

Algunos ejemplos de políticas e iniciativas que describiremos a continuación encajan en las categorías establecidas en esta lista, otras lo hacen de manera tangencial. Pero esta categorización es una buena referencia para conocer cómo los gobiernos están tratando de mejorar su transparencia en el campo de la automatización.

### **La Artificial Intelligence Act en el horizonte**

Antes de eso es relevante conocer un poco más en detalle en qué consiste la propuesta de regulación de la Comisión Europea sobre la IA presentada en abril de 2021.

La Artificial Intelligence Act (AI Act) es un documento clave para el futuro de la automatización de procesos en el continente, que influirá en los mecanismos para una mayor transparencia que se apliquen en los próximos años.

La norma busca guiar y controlar el desarrollo y uso de ADMS en los sectores público y privado europeos, pero también pretende influir en cómo otras regiones del mundo regulan estos sistemas en sus territorios. La Comisión, en palabras de la responsable de políticas digitales europeas, **Margrethe Vestager**, busca “hacer de Europa un referente mundial en el desarrollo y uso de una inteligencia artificial segura, confiable y centrada en el ser humano”.

Así, la AI Act está basada en el riesgo potencial que estas aplicaciones pueden tener sobre la sociedad y sobre los derechos de las personas. De forma resumida, el documento establece cuatro grandes niveles de riesgo.

En lo alto de la pirámide (**figura 7**) están los sistemas considerados de **riesgo inaceptable**, aquellos que pueden suponer “una amenaza para la seguridad, los medios de vida y los derechos de las personas”. Este tipo de herramientas quedarían, bajo esta norma, totalmente prohibidas.

El segundo escalón es el de **riesgo alto** e incluye sistemas que actúan en áreas especialmente sensibles, como la justicia, la salud y la educación. Para ellos la Comisión quiere implantar una supervisión más estricta y conocer de forma detallada cuál es su impacto sobre la ciudadanía.

El tercer nivel es el de los sistemas con un **riesgo limitado**, a los que Europa busca atar en corto con un mayor control. Finalmente estarían los de **riesgo mínimo**. A estos no se les impondría ningún control adicional.



Fuente: Propuesta de regulación de la IA, AI Act. Comisión Europea, 2021.

La AI Act, como decíamos, es por ahora una propuesta que aún tiene que ser refrendada por varias instituciones comunitarias. Si se cumple el calendario previsto, su aprobación definitiva podría llegar a lo largo de 2023.

Esta norma será una pieza más del andamio regulatorio sobre el mundo digital que la UE está tratando de levantar desde hace años, que incluye también textos como el RGPD y las más recientes leyes de servicios digitales y de mercados digitales (**DSA** y **DMA**, respectivamente, por sus siglas en inglés), como formas de contrapeso al poder adquirido en las últimas décadas por las grandes corporaciones tecnológicas.

## 1. El modelo PIO del OEIAC: autoevaluación para un uso ético

En febrero de 2020, la Generalitat de Catalunya presentó públicamente su estrategia de inteligencia artificial, bautizada como **Catalonia.AI**. Este documento analiza las capacidades existentes en el territorio y establece una serie de prioridades y líneas de actuación públicas con el objetivo de convertir a Cataluña “en un polo de innovación, liderazgo y atracción de talento y empresas en el terreno de la inteligencia artificial”.

Uno de los ejes de esta estrategia pasa por promover un desarrollo ético de estas tecnologías y centrado en las personas (*ética y sociedad*). Con este fin se creó el **Observatorio de Ética en Inteligencia Artificial de Cataluña (OEIAC)**, convertido en cátedra de la Universidad de Girona (UdG), para estudiar las oportunidades pero también las consecuencias éticas, sociales y legales del uso de la IA en el territorio.

En junio de 2022, el OEIAC presentó el **modelo PIO (principios, indicadores, observables)**, “una propuesta de autoevaluación organizativa sobre el uso ético de datos y sistemas de IA”, según lo definen sus creadores.<sup>15</sup>

“El modelo parte de que no hay todavía una legislación clara. Y mientras no exista, debemos hacer algo, debemos ser proponentes”, señala Albert Sabater, director del OEIAC y autor del modelo junto a la especialista Alicia de Manuel.

### Principios éticos y metodología

El modelo está basado en siete principios éticos fundamentales: transparencia y explicabilidad; justicia y equidad; seguridad y no maleficencia; responsabilidad y rendición de cuentas; privacidad; autonomía, y sostenibilidad.

---

15 Puede encontrarse un documento detallado sobre el modelo PIO y cómo aplicarlo, en catalán, castellano e inglés, en la web del OEIAC: <https://www.udg.edu/ca/catedres/oeiac/recursos-i-transferencia/informe-model-pio>.

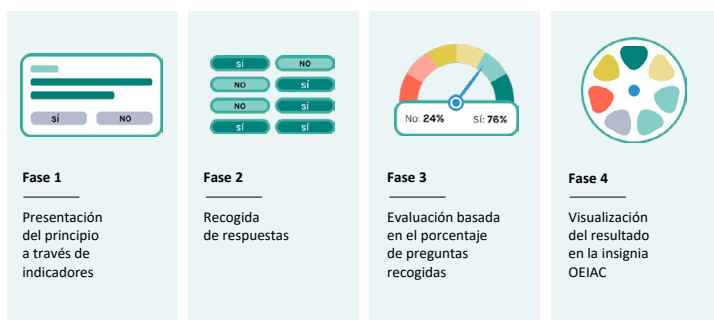
Su metodología consiste en cuatro fases diferenciadas (**figura 7**): presentación de los indicadores; recogida de las respuestas; evaluación, y visualización de los resultados. Las preguntas planteadas en el modelo, un total de 70, están relacionadas con cada uno de los principios mencionados. Por ejemplo:

- *¿Habéis documentado los tipos de datos, su procedencia y los cambios efectuados en cada fase del ciclo de vida de vuestro sistema de IA? ¿Habéis establecido mecanismos que permitan mitigar los posibles sesgos de clasificación en vuestro sistema de IA, sea por categorías cambiantes o inadecuadas según el contexto sociocultural?*
- *¿Habéis definido claramente los límites del nivel de control y participación humana dentro de vuestro sistema de IA?*

El formulario puede realizarse completamente en línea y las respuestas derivan en un resultado final, algo así como una nota al examen ético propuesto, que además se puede descargar en forma de insignia para darle visibilidad pública.

#### Fases del modelo PIO

(Principios, indicadores y observables)



Fuente: El modelo PIO (principios, indicadores y observables). OEIAC, 2022.

El modelo PIO, por tanto, funciona como un sistema de verificación (o *checklist*) aplicable por cualquier organización, tanto administraciones y entidades públicas como empresas privadas. La idea, señala Sabater, es ofrecer una herramienta práctica para aplicar una mayor transparencia interna a los ADMS y así poder detectar fallas en cada fase del proceso.

“Cuando documentas un modelo haces lo mismo que con cualquier investigación o proyecto: estás documentando cada uno de los pasos que llevas a cabo para obtener una solución. Si cometes un error, lo puedes identificar rápidamente. Si no lo cometes, puedes saber hasta qué punto puedes hacer ese sistema más eficiente”, explica.

### **Adaptación a la regulación**

La indefinición regulatoria que existe hoy en día en torno al uso de sistemas algorítmicos y de IA marca también las listas de verificación actualmente a disposición de los gobiernos. La AI Act y la evolución de las normas de cada Estado establecerán límites más claros sobre, por ejemplo, qué nivel de rendición de cuentas se les exige a sistemas aplicados en áreas muy sensibles.

De ahí que las actuales guías de uso —especialmente las promovidas desde observatorios y otros entes no implicados en esa labor regulatoria— puedan verse modificadas a corto plazo. “Queremos evolucionar de tal forma que nos podamos adaptar rápidamente cuando salga la legislación”, apunta Sabater. Esto implicaría, por ejemplo, que aquellos sistemas considerados de “uso inaceptable” no sean ni considerados para ser autoevaluados.

Pero también permitiría, según Sabater, insuflar una mayor confianza sobre otros sistemas menos problemáticos. “Uno de los objetivos principales es generar valor a través de estos principios éticos, pero también tener la capacidad de generar confianza en aquellos que no representan ningún tipo de riesgo, que no es algo menor”, apunta.

### **“Efecto contagio”**

Desde el OEIAC hablan también de intentar provocar un “efecto contagio” para que cada vez más entidades se adhieran a este tipo de filtro ético: “Cuando tienes la capacidad de convencer a diferentes organizaciones de que, al utilizar una serie de indicadores y principios, una de las cosas que puede ocurrir es que el propio principio adquiera



un valor económico o de optimización para la Administración”, señala Sabater, quien apunta la necesidad de convencer a cada vez más actores para que apliquen este tipo de autoevaluaciones.

“Cuando lo hacen unos pocos tienes que procurar que lo hagan muchos más, y convencer de que eso es suficientemente valioso para que exista un contagio”, reflexiona. Siguiendo esta línea, el Observatorio organizará en los próximos meses talleres de capacitación y seminarios para aplicar el modelo tanto en la Administración como en el sector privado y para diferentes sectores económicos. El objetivo, apuntan, es “llegar a un público que sea susceptible de cambiar las cosas”.

### **El proyecto EQUIA, un marco técnico de análisis**

En ocasiones, las iniciativas para aumentar la transparencia surgen de la cooperación entre el sector privado y público. En 2020, la consultora especializada Capgemini Engineering, en colaboración con la Dirección General de Administración Digital y de Organización (DGADO) de la Generalitat, impulsó el **proyecto “EQUIA. Inteligencia artificial fiable y fácil”**.<sup>16</sup>

EQUIA plantea un marco técnico de análisis, a través de un software específico, para que las administraciones públicas —u otras organizaciones— puedan evaluar las bases de datos utilizadas y los modelos algorítmicos asociados a cada sistema.

Esta “caja de herramientas técnica” analiza, entre otros puntos, la equidad del ADMS y los posibles sesgos existentes. A continuación muestra, a través de varios gráficos, el resultado de este análisis, permitiendo a los funcionarios detectar los puntos débiles del modelo.

Una vez desarrollado el marco técnico y práctico, EQUIA está actualmente testeándose sobre un sistema algorítmico centrado en el sector sanitario, explican desde la DGADO.

---

16 En este seminario web pueden conocerse más detalles sobre el desarrollo y los objetivos del proyecto EQUIA: <https://capgemini-engineering.com/es/es/webinar/technitive-xai-foederated-learning-sesg/>.

## 2. El estándar algorítmico británico y su estrategia de datos

En el Reino Unido, los potenciales efectos negativos de los ADMS permean desde hace años el debate público. Además, las administraciones usan estos sistemas con frecuencia —un tercio de los municipios del país los usaban ya en 2019 para la provisión de servicios públicos, reveló entonces *The Guardian*— y diversas investigaciones han revelado casos muy controvertidos.

Un ejemplo es el escándalo de los exámenes de acceso a la universidad en el verano de 2020 que mencionamos en el capítulo II. Este sistema automatizado fue finalmente retirado por el Gobierno de las islas tras las protestas de los estudiantes afectados y la sociedad civil.

Por otro lado, investigadores, activistas y entidades como el Ada Lovelace Institute y el Alan Turing Institute llevan años presionando a las administraciones para aumentar el control y la regulación sobre los ADMS.

Dentro de este contexto, en noviembre de 2021, el Ejecutivo británico lanzó su propio **estándar de transparencia algorítmica (Algorithmic Transparency Standard)**, una herramienta para ayudar al sector público a “proporcionar información clara sobre las herramientas algorítmicas que utiliza” y “por qué las está utilizando”.

Este estándar está dentro de la Estrategia nacional de datos (**National Data Strategy**) aprobada un año antes. Previo a eso, en 2018, el Reino Unido ya había fijado su propio marco ético para el uso de datos en el sector público, incidiendo en la necesidad de “la transparencia, la justicia y la rendición de cuentas” en el procesamiento algorítmico de datos.

### Orientación para las administraciones

La oficina digital y de datos del Gobierno británico (**Central Digital and Data Office, CDDO**) fue la encargada de desarrollar el actual estándar para la transparencia, que se puede consultar en línea.

Este modelo incluye un nombre, una breve descripción de la información solicitada y unas reglas de validación específicas para cada una de las 38 variables analizadas.

El primer nivel de la herramienta recoge información básica de los objetivos y el funcionamiento del ADMS. En el segundo, se entra en más detalle sobre cuestiones técnicas, como la lógica detrás de los algoritmos utilizados, los datos usados para entrenarlos y el nivel de supervisión humana garantizado en el proceso.

Además del estándar en sí, Downing Street también ha publicado una plantilla y una guía para orientar a las administraciones sobre cómo cumplimentar la información solicitada.

El buen funcionamiento de estas herramientas se basa en un equilibrio que en ocasiones es difícil de conseguir. Por un lado, se requiere un alto nivel de profundidad en la información solicitada, pero las herramientas también han de resultar accesibles y no ahuyentar a los funcionarios al mando, explica **Natalia Domagala**, jefa de datos y uso ético de la IA en el gabinete británico, en un artículo en *Open Government Partnership*.

“Queríamos facilitar todo el detalle posible, al mismo tiempo que nos asegurábamos de que el mecanismo de transparencia algorítmica que diseñábamos era práctico y mejoraba el entendimiento público sobre cómo se usan los algoritmos en el sector público.”

### **Construcción y testeo**

Un punto relevante dentro de estos mecanismos de rendición de cuentas es decidir qué tipo de preguntas hacerle al sistema. ¿Qué se quiere conocer exactamente del ADMS que se va a implementar?

Plantear esta cuestión a una representación de la sociedad civil, y no solo a un grupo de expertos en la materia, fue la opción elegida por las autoridades británicas. Como parte del diseño del estándar se hizo “un ejercicio de participación pública deliberativa” que reunió a 32 personas de diferentes características, contextos y orígenes.

Conjuntamente con el organismo para la innovación y un uso ético de los datos (**Centre for Data Ethics and Innovation, CDEI**), se trató de entender primero cuál era el nivel de conocimiento público sobre la materia, y después se crearon varios prototipos que “reflejaran las necesidades y expectativas” de los participantes en relación a la transparencia algorítmica.

Actualmente, el estándar británico está siendo testeado en diferentes departamentos del gobierno. También se está probando con ADMS que ya se usan en el sector público de las islas. Este trabajo de testeo servirá para actualizar el estándar y crear una versión final del mismo.

Miembros del Gobierno británico han manifestado su intención de que todos los ADMS del sector público pasen por este estándar y de que este filtro sirva, además, para construir un registro general de sistemas algorítmicos en las islas. Sin embargo, aún no está claro si su aplicación será un requisito previo imprescindible para la implementación de ADMS en la administración de las islas.

### **3. La estrategia de Barcelona para el uso de algoritmos y datos**

Las administraciones públicas locales juegan un papel relevante en el despliegue de ADMS. Los ayuntamientos son responsables de muchos de los servicios públicos con mayor incidencia sobre la ciudadanía, desde programas sociales hasta la gestión del espacio público.

El consistorio de Barcelona publicó, en abril de 2021, su propia **estrategia para el uso ético de los algoritmos y la inteligencia artificial**, una propuesta que puede ayudarnos a entender el rol de los municipios en la gestión de estos sistemas.<sup>17</sup>

---

17 El nombre completo de la estrategia aprobada por el Ayuntamiento de Barcelona es “Medida de gobierno de la estrategia municipal de algoritmos y datos para un impulso ético de la inteligencia artificial”. El documento se puede consultar en este enlace: [https://ajuntament.barcelona.cat/premsa/wp-content/uploads/2021/04/Mesura-de-Govern-Intel·ligencia-artificial\\_cat-v2.47-ca-ES\\_.pdf](https://ajuntament.barcelona.cat/premsa/wp-content/uploads/2021/04/Mesura-de-Govern-Intel·ligencia-artificial_cat-v2.47-ca-ES_.pdf).

**Michael Donaldson**, comisionado de Innovación Digital, Administración Electrónica y Buen Gobierno, explica la necesidad de este plan a partir del “potencial” que tienen los algoritmos y los datos para mejorar los servicios públicos, pero también por el riesgo de vulneración de derechos y discriminaciones que su uso puede implicar.

“Nos queremos blindar. Nos queremos dotar de unos criterios, de unos *checks and balance* para poder gestionar estos sistemas de manera ética y garantizar los derechos de privacidad, los derechos digitales y el derecho a la no discriminación cuando usamos esta tecnología”, argumenta Donaldson.

La estrategia recoge veinte acciones concretas, basadas en siete principios rectores:

- Acción y supervisión humana
- Robustez técnica y seguridad
- Privacidad y gobernanza de datos
- Transparencia
- Diversidad, inclusión y equidad
- Compromiso social y ambiental
- Responsabilidad, rendición de cuentas y control democrático

### **Clausulado ético y registro de algoritmos**

Entre las acciones previstas está la publicación de una guía técnica dirigida tanto al equipo directivo como al personal técnico encargado de los contratos y servicios del Ayuntamiento. Un complemento a esta guía será la creación de cláusulas específicas para la licitación de soluciones basadas en IA.

“Somos conscientes de que el desarrollo de tecnología algorítmica aplicada desde las administraciones es muy anecdótico. Y de que, por tanto, nos tenemos que dotar de soluciones que existen en el mercado. Teniendo en cuenta ese desequilibrio, tenemos que tener un marco y unos criterios para decidir qué tecnología sí y qué tecnología no”, razona Donaldson.

Esta especie de “clausulado ético” —todavía no desarrollado— aspira a parecerse a las cláusulas sociales o medioambientales que ya se aplican en muchas licitaciones del consistorio, explica el funcionario.

Otra de las medidas específicas de la estrategia es la creación de un registro municipal de algoritmos. Donaldson señala que están trabajando conjuntamente con ciudades como **Ámsterdam y Helsinki**, que ya cuentan con un registro similar, aunque reconoce sus dudas sobre el nivel de detalle que este registro podrá ofrecer. ¿Se incluirá, por ejemplo, en este registro el código fuente de los sistemas?

“Vamos a tener que buscar una fórmula que quizás no nos deje abrirlo en canal, por el derecho de propiedad intelectual. La palabra clave quizás sería la *explicabilidad*. ¿Qué decisiones se están tomando y en base a qué datos? Es decir, explicar bien no tanto el producto, sino más bien el servicio”, apunta Donaldson.

### **Órganos de gobernanza**

La estrategia municipal de Barcelona tiene tres órganos de gobernanza específicos. El primero es una comisión transversal dentro del Ayuntamiento que incluye a responsables de todas las áreas del consistorio. Está previsto que por ella pasen todas las soluciones y servicios basados en IA que ofrezca el Ayuntamiento. También será la encargada de aprobar iniciativas como el registro y el clausulado ético.

El segundo órgano es un consejo de inteligencia artificial integrado por investigadores, académicos y miembros de la sociedad civil, incluidas organizaciones en defensa de la justicia algorítmica, que asesorarán al Ayuntamiento en la materia, apunta Michael Donaldson.

Una tercera pata será “un pacto sobre humanismo tecnológico”<sup>18</sup> que impulse el buen uso de la tecnología, los algoritmos y la inteligencia artificial, y que apele también a las empresas, a las instituciones y al conjunto de la sociedad civil, señalan desde el consistorio.

### **Proactividad en los servicios públicos**

Uno de los puntos de la estrategia de Barcelona es la búsqueda de la “proactividad” en los servicios públicos a la ciudadanía. Según datos de la UE, alrededor del 30 % de los potenciales receptores de ayudas públicas en el continente no las solicitan porque desconocen su existencia o los procedimientos para hacerlo.

Desde Barcelona se aspira a utilizar los datos sobre la realidad ciudadana local para anticiparse y ofrecer un servicio “de forma personalizada y directa”, siendo sensibles a la diversidad local y garantizando una universalidad efectiva.

Uno de los proyectos piloto ya en marcha en la ciudad es la mejora de los **canales de comunicación públicos** mediante algoritmos de procesamiento del lenguaje natural, que extraen patrones de las incidencias trasladadas por los vecinos. Esta fotografía, apuntan desde el Ayuntamiento, servirá para mejorar su respuesta a las reclamaciones ciudadanas.

Esta búsqueda de servicios personalizados y proactivos a partir de los datos es también una de las prioridades de la DGADO de la Generalitat, explica Josep Maria Flores, jefe del Área de Impulso de la Mejora de los Servicios Digitales.

---

18 El humanismo tecnológico, una de las bases de la estrategia municipal de Barcelona, es “un modelo de desarrollo tecnológico basado en el ser humano, que reduce las desigualdades sociales y protege los derechos humanos, y en el que la tecnología está al servicio de las personas y de los intereses generales”, según la definición propuesta por el Ayuntamiento.

Tras un análisis interno, detectaron que la **tramitación de las becas universitarias Equidad** —ayudas para cursar estudios de grado o máster en universidades públicas catalanas— era uno de los casos de uso donde estas técnicas podrían ser más útiles.

En el curso académico 2021-2022 se comenzó un proyecto piloto que agiliza los trámites burocráticos asociados, enviando un correo electrónico de aviso y un formulario precompletado a los estudiantes que ya habían solicitado estas ayudas en años anteriores.

Flores señala que los datos del alumnado almacenados por la Administración sirven, previo consentimiento explícito, para hacer que esta experiencia sea más “plácida” para el estudiantado. En el curso 2022-2023 el proyecto continúa en marcha y se han enviado hasta 40.000 preavisos.

#### **4. ¿Cómo hacer cumplir la transparencia? La Ley Rider y la guía de información algorítmica**

Durante la pandemia de la COVID-19, los repartidores de las empresas de reparto de comida, o *riders*, adquirieron una visibilidad desconocida hasta entonces. Las imágenes de calles vacías durante el confinamiento solo transitadas por estos trabajadores aceleraron el debate sobre las condiciones laborales de un sector que ha vivido un crecimiento exponencial en los últimos años.

Empresas como Deliveroo, Glovo o Uber Eats funcionan a través de plataformas digitales donde los algoritmos juegan un papel clave. Cuestiones como la asignación de un pedido a un trabajador o el precio a pagar por reparto son parcialmente decididas por un software.

Con el objetivo de aportar una mayor transparencia a estos procesos, el Congreso español aprobó, el verano de 2021, la conocida como **Ley**



**Rider.**<sup>19</sup> La norma incluye por primera vez el derecho de los trabajadores, a través del comité de empresa, a ser informados por el empleador de los parámetros, reglas e instrucciones en los que están basados los algoritmos que les afectan.

### **Dificultades de acceso a la información**

Tras varios meses en funcionamiento, sin embargo, diferentes medios informaron de las dificultades que los comités de empresa estaban encontrando para acceder de forma efectiva a esta información algorítmica.

“Las empresas —especialmente las medianas y grandes— a las que estamos empezando a ir como comité de empresa a pedir esta información tienen una primera respuesta negativa. [...] Su primera reacción es ampararse en secretos comerciales o industriales, y la segunda es defender que no tenemos derecho a pedirlo porque no influyen en nuestras condiciones laborales”, señaló José Varela, responsable de digitalización del sindicato UGT, a *ElDiario.es*.

Otros sindicalistas alertaron de la invisibilidad de estos sistemas para muchos trabajadores y de la falta de conocimiento en las plantillas sobre su peso en las condiciones laborales, lo que acaba dificultando esa rendición de cuentas a la empresa.

Consciente de las dificultades para poner en práctica una norma de este tipo, el **Ministerio de Trabajo** designó un comité de expertos<sup>20</sup> para asesorar al Gobierno en la aplicación de esta Ley. Una de sus funciones principales hasta ahora ha sido la redacción de una guía

---

19 Esta Ley modificó el Estatuto de los Trabajadores en varios puntos “para garantizar los derechos laborales de las personas dedicadas al reparto en el ámbito de las plataformas digitales”. Un concepto que incluye no solo a las empresas de reparto.

20 El comité de expertos designado por el Ministerio de Trabajo está coordinado por la fundadora de Eticas Consulting, Gemma Galdón Clavell, y lo integran los profesores e investigadores universitarios Anna Ginès (ESADE), Ana Belén Muñoz (UC3M), Adrián Todolí (Universidad de Valencia) y Javier Sánchez-Monedero (Universidad de Córdoba).

sobre cómo aplicar de forma práctica la exigencia de transparencia algorítmica que ya recoge la legislación.

## Una guía práctica

En junio de 2020 se presentó públicamente un documento<sup>21</sup> que plantea 16 preguntas específicas que puede hacerse a sí misma cualquier empresa. Estas son algunas de ellas:

- *¿Qué tipo de tecnología utiliza el algoritmo? ¿Genera un modelo de caja negra? ¿Es un algoritmo de aprendizaje continuo? ¿Quién lo ha desarrollado?*
- *¿Cuáles son las variables que utiliza el algoritmo?*
- *¿Qué base de datos de entrenamiento y, en su caso, de validación, se ha utilizado?*
- *¿Se ha informado a las personas trabajadoras o candidatas respecto del uso de algoritmos para la toma de decisiones automatizadas?*

Además de esta guía práctica, el documento recoge también una explicación sencilla sobre qué es un algoritmo, cómo se diseña y se construye un sistema automatizado, y los ámbitos más frecuentes en los que se utilizan en el mundo laboral.

La publicación recoge igualmente cuestiones como qué tipo de derecho de información tienen los trabajadores: de tipo individual para los algoritmos que decidan de forma totalmente automatizada; o colectivo, a través de su representación legal, para decisiones semiautomatizadas con intervención humana, en base a la Ley Rider, pero también a partir del RGPD. O si existe la obligación de negociar el algoritmo con el comité de empresa — algo todavía no recogido por la legislación, pero que la guía sugiere como una posible mejora futura de la Ley —.

---

21 La guía *Información algorítmica en el ámbito laboral* se puede consultar íntegramente aquí: [https://www.lamoncloa.gob.es/serviciosdeprensa/notasprensa/trabajo14/Documents/2022/100622-Guia\\_algoritmos.pdf](https://www.lamoncloa.gob.es/serviciosdeprensa/notasprensa/trabajo14/Documents/2022/100622-Guia_algoritmos.pdf).

### **“Aterrizar” las garantías recogidas en la Ley**

“No queremos que las empresas hagan un ejercicio narrativo. Queremos que proporcionen información concreta de cómo toma decisiones el algoritmo”, explicó **Gemma Galdón**, coordinadora del equipo autor de la guía, en una entrevista tras su presentación. El objetivo de la guía es, por tanto, “aterrizar” en la práctica las garantías que ya recoge la Ley, pero también servir a las empresas para que verdaderamente entiendan los sistemas que están implementando.

Los miembros del comité señalan que muchas compañías adquieren a través de sus proveedores tecnológicos sistemas algorítmicos que, al ponerlos en práctica, no cumplen con lo prometido, o que simplemente no son explicados correctamente durante el proceso de compra.

Galdón apunta que la recepción de la guía por parte de patronales y asociaciones empresariales ha sido “muy buena”, en cierta medida porque esta herramienta no pide el código fuente de los algoritmos.

“Las empresas tienen miedo a que se les pida el código. Es un miedo que puede ser racional, porque a veces se cruzan temas de propiedad intelectual y de secreto empresarial que podrían generar problemas. **Lo que las empresas reciben muy bien es saber de qué estamos hablando cuando hablamos de transparencia**”, señala.

## **5. Los registros públicos de Ámsterdam y Helsinki**

Como hemos visto con Barcelona, los registros públicos de algoritmos son una de las iniciativas en pos de la transparencia algorítmica que más se están explorando en los últimos años. Los ayuntamientos de **Helsinki y Ámsterdam** fueron pioneros al lanzar sus propios registros en septiembre de 2020.

Las capitales finesa y neerlandesa definen, respectivamente, sus registros como “una ventana” y “un resumen” de los sistemas de IA y

algoritmos usados en la ciudad. Ambos han sido desarrollados por el mismo proveedor, la empresa **Saidot**, y comparten una misma estructura y diseño.

Primero hay una descripción general del sistema, que incluye sus objetivos y una explicación básica de cómo funciona, y en un nivel inferior se ofrece información detallada sobre las bases de datos utilizadas y cómo estos datos se procesan mediante comandos y algoritmos.

Además, se incluye también una descripción de cómo se han tratado de garantizar la equidad, la vigilancia humana y la mitigación de posibles riesgos en cada uno de los sistemas. Finalmente, se proporciona información de contacto sobre los equipos responsables de la implementación de cada sistema.

### **“Información entendible y actualizada”**

Esta especie de *padrón* de algoritmos está dirigido a la ciudadanía en general y busca generar confianza entre las personas que interactúan directamente con estos ADMS. “Todo el mundo debería tener acceso a información entendible y actualizada sobre cómo los algoritmos afectan a sus vidas”, argumentan desde el Ayuntamiento de Helsinki.

Actualmente, en el registro de la capital finesa aparecen siete sistemas, solo dos más que en el momento de su lanzamiento. Cinco de ellos son aplicaciones basadas en chatbots<sup>22</sup> sobre los canales de atención al ciudadano ofrecidos por el Ayuntamiento, el sistema de aparcamientos públicos de la ciudad, los servicios sanitarios municipales, el mercado del alquiler local y los servicios de maternidad y salud infantil. Además, se incluye un sistema de recomendaciones de la biblioteca central de Helsinki y otro de gestión inteligente de las colecciones de libros y otros materiales.

---

22 Los chatbots (o bots conversacionales) son aplicaciones que permiten simular una conversación humana al ofrecer respuestas automáticas ante preguntas o comentarios del usuario.

Por su parte, en el de Ámsterdam solo hay tres sistemas descritos. Los mismos que en el momento de su lanzamiento y lejos de la decena que aspiraba a aglutinar el registro cuando se hizo público el proyecto. Aquí aparecen un programa piloto sobre detección de fraudes en alquiler vacacional, un sistema de control de aparcamientos y una herramienta de procesamiento del lenguaje natural para tramitar quejas vecinales.

### **Herramienta de gobernanza**

Ya en el momento de desplegarse los registros, sus responsables reconocían las dificultades para ser transparentes con sistemas que, en muchas ocasiones, todavía están en fase de pruebas.

“El objetivo del proyecto es crear un algoritmo que sea justo, democrático y que no discrimine. Sin embargo, es complicado ser transparente sobre un algoritmo cuando todavía lo estás probando”, señalaba **Linda van de Fliert**, del Ayuntamiento de Ámsterdam, en declaraciones a *El País*.

La empresa Saidot explica que esta plataforma está dirigida tanto a administraciones públicas como a empresas privadas. En la plantilla desarrollada, se ha de especificar, por ejemplo, cómo cada ADMS evalúa la protección de la privacidad o su impacto sobre los derechos humanos de las personas afectadas.

**Meeri Haataja**, CEO de la compañía, explicó a *Equal Times* que más que una herramienta para la transparencia, la suya es “una plataforma de gobernanza de la IA”. “Les ofrecemos a las organizaciones públicas y privadas que usan IA a escala las herramientas para ayudarles a abordar los riesgos de estos sistemas”, apunta.

Preguntada sobre las limitaciones de este tipo de registros, Haataja sostiene que la “transparencia es un primer paso necesario” pero que por sí sola no es suficiente. “Una mejor gobernanza de la IA requiere de un enfoque integral”, apunta.

## Otros registros

Además de Helsinki y Ámsterdam, en los últimos años otras administraciones locales han desplegado registros similares: este es el caso de las localidades francesas de **Nantes**, **Lyon** y **Antibes**.

En Latinoamérica, el **Gobierno chileno**, en colaboración con la **Universidad Adolfo Ibáñez** (UAI), también ha creado su propio registro público: en él aparecen medio centenar de sistemas desplegados por diferentes departamentos del Gobierno. Sin embargo, hasta el momento, la información pública sobre cada uno de ellos es escasa. Además de una breve descripción de cada ADMS, este registro se limita a recopilar otras fuentes oficiales relacionadas con cada sistema.

## 6. La agencia española para la supervisión de la IA

En los presupuestos generales de 2022, el Gobierno español incluyó la creación de un supervisor estatal de los sistemas de inteligencia artificial (IA) usados tanto por empresas como por las administraciones. El objetivo de esta entidad, según el Ejecutivo, será “minimizar los riesgos significativos sobre la seguridad y salud de las personas, así como sobre sus derechos fundamentales, que puedan derivarse de su uso”.

Está previsto que la **Agencia Española de Supervisión de la Inteligencia Artificial (AESIA)** funcione como un organismo independiente y autónomo de otros poderes públicos. Dos ejemplos de autoridades similares, centradas en otros sectores, son la Agencia Española de Protección de Datos (AEPD) y la Comisión Nacional de los Mercados y la Competencia (CNMC).

La AESIA seguirá la estela de la propuesta regulatoria europea de la IA<sup>23</sup> y actuará, previsiblemente, como una entidad de auditoría algorítmica.

---

23 La AI Act recoge, en su artículo 56, la creación de un consejo a escala europea (European Artificial Intelligence Board) que garantice la aplicación de la nueva normativa en el continente en coordinación con los “supervisores nacionales”, es decir, con entidades como la AESIA.

mica y control sobre el funcionamiento de estos sistemas, sin importar si vienen del sector privado o público. En palabras del Gobierno, la agencia garantizará una “correcta interpretación” de la AI Act en territorio español.

### **Dudas sobre sus competencias**

En septiembre de 2022 se inició el procedimiento para elegir una ubicación física para la AESIA, que no será Madrid. Con todo, a fecha de cierre de esta publicación todavía seguía habiendo diversas preguntas en el aire sobre las funciones prácticas que tendrá.

¿Podrá, por ejemplo, este organismo multar a una administración que use sin la debida transparencia un ADMS considerado de alto riesgo por el estándar europeo? Entidades como las mencionadas (AEPD y CNMC) sí tienen esta capacidad para sancionar económicamente a un ente público o privado que incumpla la normativa en sus ámbitos, algo que todavía no está confirmado en el caso de esta agencia de nueva creación.

Más allá de una labor auditora y supervisora, la AESIA también asesorará en el desarrollo de sistemas eficientes y sostenibles, según la información difundida por el Gobierno, y creará guías y manuales de buenas prácticas para un uso ético y confiable de la IA, entre otras labores de divulgación.

Hay que tener en cuenta que la AESIA se enmarca dentro de la **Estrategia nacional de inteligencia artificial (ENIA)**, un documento presentado a finales de 2020 que marca las líneas estratégicas en este campo de la **Secretaría de Estado de Digitalización e Inteligencia Artificial (SEDIA)**. Los cinco millones de presupuesto inicial con los que contará la agencia proceden precisamente de esta estrategia. Además, según los documentos oficiales dados a conocer por *El País*, la AESIA contará inicialmente con un mínimo de 40 profesionales en su plantilla.

## Participación de colectivos sociales

Lograr una verdadera independencia respecto a las administraciones es precisamente uno de los principales retos de un organismo como la AESIA, todavía más si uno de sus objetivos es supervisar sistemas utilizados dentro del sector público español.

Varias organizaciones en defensa de la justicia algorítmica y el uso responsable de la IA han enviado a la SEDIA una carta abierta<sup>24</sup> para reclamar que este órgano supervisor funcione de manera autónoma, tanto en su estructura como en su cometido, y que promueva “la transparencia, auditabilidad y explicabilidad” de los sistemas algorítmicos “a través de herramientas e informaciones de fácil consulta y comprensión”. Las organizaciones firmantes, entre las que están entidades con sede en Cataluña como Algorights, Lafede.cat y la Fundación Éticas, reclaman también que se cuente con su voz y conocimiento en el diseño y creación de esta nueva entidad.

Finalmente, el documento también reclama que la agencia “ ejerza una supervisión prudencial que garantice la equidad de los procesos algorítmicos y la responsabilidad ante sus impactos negativos y/o discriminatorios”.

## 7. El Citizens' Biometrics Council y la participación ciudadana en el debate algorítmico

En febrero de 2020, medio centenar de ciudadanos británicos comenzaron a reunirse de forma periódica en las ciudades de Manchester y Bristol. Todos ellos formaban parte del **Consejo Ciudadano sobre Biometría** (Citizens' Biometrics Council), un proyecto del Ada Lovelace Institute.

---

24 El contenido íntegro de esta carta y las entidades que la impulsaron pueden consultarse en el siguiente enlace: <https://rightsinternationalspain.org/wp-content/uploads/2022/09/AESIA-cast.pdf>.



El estallido de la pandemia de la COVID-19, además de extender la duración de la iniciativa varios meses, llevó estos encuentros al formato en línea. En ellos, los 50 miembros del Consejo escucharon la voz de una amplia variedad de expertos en el uso de tecnologías biométricas.<sup>25</sup>

Desarrolladores informáticos, legisladores, activistas, académicos, juristas e incluso fuerzas del orden participaron en estos talleres. Además, los miembros del Consejo tuvieron la oportunidad de debatir entre ellos y decidir conjuntamente sobre la gran cuestión a la que querían dar respuesta: “¿En qué casos está bien el uso de estas tecnologías y en cuáles no?”.

### **Funcionamiento y recomendaciones**

“El Consejo no pidió un veto para ellas. Esto es algo interesante, teniendo en cuenta la presión que hay desde grupos por los derechos civiles para prohibirlas”, explicó **Aidan Peppin**, investigador sénior del Ada Lovelace Institute, durante su participación en las jornadas Democracia, Algoritmos y Resistencias.<sup>26</sup>

Las aplicaciones biométricas son uno de los usos más polémicos y discutidos del campo de la inteligencia artificial por las imprecisiones que estas tecnologías han mostrado hacia la población no blanca y por los riesgos para la privacidad de las personas analizadas. La red europea de entidades en defensa de los derechos digitales, **European Digital Rights (EDRI)**, ha pedido al Parlamento Europeo su prohibición en los espacios públicos.

“Entre las razones por las que no pidieron un veto está que consideraron que estas tecnologías tienen ciertos beneficios. Los miembros

---

25 El término *tecnologías biométricas* engloba desde las técnicas de reconocimiento facial que ya se usan en aeropuertos o pasos fronterizos hasta los registros digitales de huellas, utilizados, por ejemplo, por la policía, en oficinas o incluso en colegios.

26 Estas jornadas, organizadas por los colectivos Algorights, Algorace, Lafede.cat y el Espai Societat Oberta, se celebraron entre el 26 y el 28 de mayo en línea, en Barcelona y en Madrid. Las intervenciones de Peppin corresponden a la sesión “Políticas públicas transparentes y auditabilidad” [<https://www.espaisocietatoberta.org/es/blog/politicas-publicas-transparentes-auditabilidad>].

del Consejo con alguna discapacidad entendieron que el potencial de accesibilidad de ciertas aplicaciones biométricas es muy importante”, señaló Peppin.

Además de este posicionamiento, del trabajo del Consejo salieron un total de 30 recomendaciones dirigidas a los legisladores. Con tres conclusiones fundamentales: es necesario renovar, mejorar y fortalecer la actual regulación en torno a estas tecnologías; ha de existir un ente independiente que supervise su uso, y se han de fijar unos estándares mínimos de calidad para su desarrollo y uso. Estas conclusiones y recomendaciones ya han sido trasladadas a los legisladores del Reino Unido para que las tengan en cuenta, explicó Peppin.

### **Implicación ciudadana**

Más allá de sus conclusiones, este proyecto es un buen ejemplo práctico de cómo implicar a la ciudadanía en los procesos de gobernanza tecnológica. “Desde el Ada Lovelace Institute, entendemos la participación como una forma de empoderar a la gente, para que puedan moldear las leyes y decidir cómo estas tecnologías y sus datos han de usarse”, señalan.

Para la elección de los miembros del Consejo, trabajaron con una agencia de reclutamiento y a partir de una lista de criterios relacionados con la edad, el género, el contexto socioeconómico, el origen étnico, la religión, la ideología y el posicionamiento frente a la tecnología de los participantes. “Queríamos captar un reflejo lo más amplio posible de la población británica”, señaló Peppin.

Además, crearon varias sesiones especiales con miembros del colectivo LGTBIQ+, con personas con alguna discapacidad y con miembros de alguna minoría étnica, comunidades que, según su criterio, pueden ver amenazados sus derechos con el uso de estas tecnologías.

Sobre los pros y contras de esta forma de participación ciudadana, desde el Ada Lovelace Institute señalan la riqueza de perspectivas y

miradas que aportan estas iniciativas, pero reconocen que requieren tiempo y recursos económicos: “Es importante reconocer que la participación en la práctica es dura, pero se puede hacer”. Y apuntan a la necesidad de incorporarlas a otras áreas: “Uno de los retos es llevar estas formas de participación fuera de institutos de investigación como el nuestro e incorporarlas a los sistemas democráticos de gobernanza sobre estas tecnologías”, indicó Peppin.

## V. Reflexiones finales

A lo largo de esta publicación hemos visto cómo la automatización de procesos en la Administración pública afecta a todo tipo de áreas, y lo hace de formas muy diversas.

Un sistema como BOSCO es un algoritmo sencillo que automatiza la asignación del bono social eléctrico para familias de bajos recursos. Aunque en apariencia su diseño no reviste complejidad, un pequeño error puede dejar a muchos hogares sin acceso a esta ayuda.

En los Países Bajos, el Gobierno usó durante años un modelo de analítica predictiva discriminatorio. En un supuesto intento de acabar con el fraude en la recepción de una prestación para el cuidado de los hijos, cientos de ciudadanos fueron acusados erróneamente, con consecuencias nefastas en sus vidas.

El uso de ADMS en el sector público puede ayudar a mejorar los servicios públicos, pero, como hemos visto con los ejemplos descritos, también puede generar situaciones injustas, o amplificar problemas ya existentes en la sociedad, como el racismo o la discriminación hacia ciertos grupos.

Entre los casos de mal uso de sistemas algorítmicos conocidos en los últimos años, un rasgo común es su opacidad. Hasta hace poco, los gobiernos no sentían la necesidad de explicar cómo estaban usando modelos automatizados, lo que derivaba en el uso de cajas negras o de sistemas que carecían del control y las garantías necesarias. De alguna manera, esto parece estar cambiando.

En el capítulo III hemos visto cómo cada vez más legislaciones incluyen la transparencia —o la explicabilidad, un concepto más específico— entre los requisitos para implementar una política tecnológica. Igualmente, dichas normativas ya plantean y empiezan a poner coto a los escenarios donde el uso de ADMS puede traer más problemas que beneficios.

La propuesta de regulación europea y cómo esta finalmente se concrete definirá las líneas de la automatización del sector público en el continente para los próximos años.

Dentro de este contexto, a día de hoy ya existen diversas iniciativas concretas que tratan de garantizar —o al menos impulsar— un uso más ético y transparente de los modelos algorítmicos y de la inteligencia artificial. En el capítulo IV hemos querido dar algunas pinceladas sobre qué se está haciendo en este campo.

En ocasiones, dichas iniciativas provienen de las propias administraciones públicas; mientras que en otras son centros de investigación, colectivos organizados o empresas los que están creando herramientas prácticas para evitar un uso irresponsable de estas tecnologías.

Como hemos visto, su nivel de desarrollo y sus características son muy diversas, e incluyen la creación de estándares para orientar en el diseño de ADMS —como el estándar del Gobierno británico o, en un nivel diferente, el modelo PIO del OEIAC— o el impulso de registros públicos que ofrezcan información detallada sobre cómo se están utilizando estos modelos. También la aparición de entidades específicas para auditar y controlar la inteligencia artificial —véase la AESIA del Gobierno español— van en esta línea.

En los próximos meses y años veremos la aparición de iniciativas similares y la consolidación de algunas de ellas. El camino para un uso más responsable de estas tecnologías se está construyendo. Habrá que estar atentos para ver si todos estos esfuerzos concluyen en medidas efectivas y beneficiosas para el conjunto de la población, sin importar su origen, género o condición.

## VI. Glosario breve

### **AI Act**

La propuesta de regulación de la Comisión Europea sobre la inteligencia artificial (Artificial Intelligence Act) es un documento clave para el futuro de la automatización de procesos en el continente. Este borrador regulatorio, actualmente en discusión, propone un enfoque basado en el riesgo de los sistemas algorítmicos, así como exigencias específicas en materia de transparencia y rendición de cuentas.

### **Algoritmo**

Secuencia de instrucciones o pasos seguidos para lograr un resultado específico. En el actual mundo digitalizado, los algoritmos generalmente son ejecutados por un ordenador de forma automatizada.

### **Auditoría algorítmica**

Revisión del diseño, los resultados o el impacto que un sistema algorítmico genera sobre la población en general o sobre un grupo determinado de personas. El nivel de información al que tienen acceso los auditores determina cómo se realiza la auditoría.

### **Cajas negras**

Sistemas automatizados y basados en algoritmos cuyos procedimientos y variables no son evidentes o conocidos. Al observarlos, generalmente podemos saber cuáles son los datos que entran y los que salen, pero no la manera en la que el sistema llega a una decisión u otra.

### **Catalonia.AI**

Estrategia de la Generalitat de Catalunya para el desarrollo de la inteligencia artificial en el territorio y de soluciones tecnológicas asociadas. Incluye la creación de organismos específicos para la investigación y

el impulso de la IA, así como el fomento de un ecosistema público y privado puntero.

## **ENIA**

La Estrategia nacional de inteligencia artificial (ENIA) del Gobierno español recoge una serie de objetivos a largo plazo en el campo de la IA, además de actuaciones concretas que se inscriben dentro de la Estrategia España digital 2026.

## **Estándar algorítmico**

Mecanismo práctico que fija los criterios que ha de seguir un sistema algorítmico para garantizar un uso ético y responsable del mismo. Sirve como filtro previo a la puesta en marcha e implementación de los sistemas en un escenario real.

## **Inteligencia artificial (IA)**

Conjunto de sistemas que muestran un comportamiento considerado inteligente, analizando su entorno y realizando acciones concretas para un objetivo específico. Pueden basarse únicamente en un software (por ejemplo, los asistentes de voz, los programas de análisis de imágenes, etc.) o integrarse en dispositivos de hardware (robots avanzados, coches autónomos, etc.).

## **Conjunto de sistemas de decisión automatizada o ADMS (*automated decision-making system*)**

Sistema pensado para optimizar la toma de una decisión mediante el uso de algoritmos y grandes cantidades de datos. Este proceso se integra dentro de un marco social y tecnológico concreto, que determina sus efectos sobre las personas que interactúan con él.

## Bibliografía

A continuación se detallan las fuentes utilizadas para la elaboración de cada uno de los capítulos de esta publicación.

### I. Introducción

- ALGORITHM WATCH; BERTELSMANN STIFTUNG. *Automating Society. Taking Stock of Automated Decision-Making in the EU* [en línea]. Berlín: AW AlgorithmWatch gGmbH, enero de 2019. <[https://algorithmwatch.org/de/wp-content/uploads/2019/02/Automating\\_Society\\_Report\\_2019.pdf](https://algorithmwatch.org/de/wp-content/uploads/2019/02/Automating_Society_Report_2019.pdf)> [Consulta: 21/11/2022]
- AMNISTÍA INTERNACIONAL. “Dutch childcare benefit scandal an urgent wake-up call to ban racist algorithms” [en línea]. 25 de octubre de 2021. <<https://www.amnesty.org/en/latest/news/2021/10/xenophobic-machines-dutch-child-benefit-scandal/>> [Consulta: 21/11/2022]
- GEIGER, G. “How a discriminatory algorithm wrongly accused thousands of families of fraud” [en línea]. Londres: VICE News, 3 de marzo de 2021. <<https://www.vice.com/en/article/jgq35d/how-a-discriminatory-algorithm-wrongly-accused-thousands-of-families-of-fraud>> [Consulta: 21/11/2022]
- HEIKKILÄ, M. “Dutch scandal serves as a warning for Europe over risks of using algorithms” [en línea]. *Politico*, 29 de marzo de 2022. <<https://www.politico.eu/article/dutch-scandal-serves-as-a-warning-for-europe-over-risks-of-using-algorithms/>> [Consulta: 21/11/2022]
- UNESCO. “A lexicon for the digital age” [en línea]. *The UNESCO Courier*, julio-septiembre de 2017. <<https://en.unesco.org/courier/july-september-2017/lexicon-digital-age>> [Consulta: 21/11/2022]



## II. Sistemas automatizados en el sector público

- ALGORITHM WATCH; BERTELSMANN STIFTUNG. *Automating Society Report 2020* [en línea]. Berlín: AW AlgorithmWatch gmbH, octubre de 2020. <<https://automatingsociety.algorithmwatch.org/wp-content/uploads/2020/12/Automating-Society-Report-2020.pdf>> [Consulta: 21/11/2022]
- DIGITAL FUTURE SOCIETY. “Capítulo 6: epílogo. ¿Cómo construir algoritmos para el bien común?” [pódcast]. 20 de abril de 2022. <<https://digitalfuturesociety.com/es/podcasts/capitulo-6-epilogo-como-construir-algoritmos-para-el-bien-comun/>> [Consulta: 21/11/2022]
- SUSARLA, A. “The new digital divide is between people who opt out of algorithms and people who don’t” [en línea]. *The Conversation*. 17 de abril de 2019. <<https://theconversation.com/the-new-digital-divide-is-between-people-who-opt-out-of-algorithms-and-people-who-dont-114719>> [Consulta: 21/11/2022]

### 1. Automatizando las ayudas públicas

- ALSINA, F.; CAMPUZANO, M. “Més pobresa energètica a les llars catalanes” [en línea]. *Ara*. Barcelona, 21 de julio de 2022. <[https://www.ara.cat/opinio/mes-pobresa-energetica-llars-catalanes-francina-alsina-maria-campuzano\\_129\\_4438131.html](https://www.ara.cat/opinio/mes-pobresa-energetica-llars-catalanes-francina-alsina-maria-campuzano_129_4438131.html)> [Consulta: 21/11/2022]
- CIVIO. “El Gobierno se niega a explicar cómo funciona su aplicación para conceder o no el bono social” [en línea]. *Tu derecho a saber*. 28 de noviembre de 2018. <<https://civio.es/tu-derecho-a-saber/2018/11/28/el-gobierno-se-niega-a-explicar-como-funciona-su-aplicacion-para-conceder-o-no-el-bono-social/>> [Consulta: 21/11/2022]
- CIVIO. “Que se nos regule mediante código fuente o algoritmos secretos es algo que jamás debe permitirse en un Estado social,

democrático y de Derecho” [en línea]. *Civio*. 2 de julio de 2019. <<https://civio.es/novedades/2019/07/02/que-se-nos-regule-mediante-codigo-fuente-o-algoritmos-secretos-es-algo-que-jamas-debe-permitirse-en-un-estado-social-democratico-y-de-derecho/>> [Consulta: 21/11/2022]

- CIVIO. “La Justicia impide la apertura del código fuente de la aplicación que concede el bono social” [en línea]. *Tu derecho a saber*. 10 de febrero de 2022. <<https://civio.es/novedades/2022/02/10/la-justicia-impide-la-apertura-del-codigo-fuente-de-la-aplicacion-que-concede-el-bono-social/>> [Consulta: 21/11/2022]
- CIVIO. “El Consejo de Transparencia ya no se opone a nuestra petición de liberar el código fuente del bono social” [en línea]. *Civio*. 8 de junio de 2022. <<https://civio.es/novedades/2022/06/08/el-consejo-de-transparencia-entiende-ahora-que-la-sentencia-que-denego-el-codigo-fuente-es-contraria-a-derecho/>> [Consulta: 21/11/2022]
- CIVIO. “Te ayudamos a pedir el bono social de electricidad” [en línea]. *Civio*. <<https://civio.es/bono-social/>> [Consulta: 21/11/2022]
- DIGITAL FUTURE SOCIETY. “Capítulo 3. BOSCO y el bono para pagar la luz” [pódcast]. 7 de marzo de 2022. <<https://digitalfuturesociety.com/es/podcasts/capitulo-3-bosco-y-el-bono-para-pagar-la-luz/>> [Consulta: 21/11/2022]
- GARCÍA, L. “El coste de Radar COVID: 50 euros por código introducido en la aplicación” [en línea]. *Newtral*. 16 de septiembre de 2021. <<https://www.newtral.es/coste-radar-covid-codigo/20210916/>> [Consulta: 21/11/2022]
- PÉREZ COLOMÉ, J. “El Gobierno publica un código incompleto y confuso de la ‘app’ Radar Covid” [en línea]. *El País*. 9 de septiembre de 2020. <<https://elpais.com/tecnologia/2020-09-09/el-gobierno-publica-hoy-el-codigo-de-radar-covid-casi-medio-ano-despues-de-lanzar-el-proyecto.html>> [Consulta: 21/11/2022]

## 2. En busca del fraude al Estado

- ALSTON, P. *Brief by the United Nations Special Rapporteur on extreme poverty and human rights as Amicus Curiae in the case of NJCM c.s./De Staat der Nederlanden (SyRI) before the District Court of The Hague (case number: C/09/55982/ HA ZA 18/388)* [en línea]. 26 de septiembre de 2019. <<https://www.ohchr.org/sites/default/files/Documents/Issues/Poverty/Amicusfinalversionsigned.pdf>> [Consulta: 21/11/2022]
- BLAUW, S. “An algorithm was taken to court – and it lost (which is great news for the welfare state)” [en línea]. *The Correspondent*. 10 de febrero de 2020. <<https://thecorrespondent.com/276/an-algorithm-was-taken-to-court-and-it-lost-which-is-great-news-for-the-welfare-state/36504050352-a3002ff7>> [Consulta: 21/11/2022]
- HENRIQUES-GOMES, L. “‘Robodebt-related trauma’: the victims still paying for Australia’s unlawful welfare crackdown” [en línea]. *The Guardian*. 20 de noviembre de 2020. <<https://www.theguardian.com/australia-news/2020/nov/21/robodebt-related-trauma-the-victims-still-paying-for-australias-unlawful-welfare-crackdown>> [Consulta: 21/11/2022]
- HENRIQUES-GOMES, L. “‘Robodebt: court approves \$1.8bn settlement for victims of government’s ‘shameful’ failure” [en línea]. *The Guardian*. 11 de junio de 2021. <<https://www.theguardian.com/australia-news/2021/jun/11/robodebt-court-approves-18bn-settlement-for-victims-of-governments-shameful-failure>> [Consulta: 21/11/2022]
- PRIVACY INTERNATIONAL. “The SyRI case: a landmark ruling for benefits claimants around the world” [en línea]. 24 de febrero de 2020. <<https://privacyinternational.org/news-analysis/3363/syri-case-landmark-ruling-benefits-claimants-around-world>> [Consulta: 21/11/2022]

- VERVLOESEM, K. “How Dutch activists got an invasive fraud detection algorithm banned” [en línea]. *Algorithm Watch*. 6 de abril de 2020. <<https://algorithmwatch.org/en/syri-netherlands-algorithm/>> [Consulta: 21/11/2022]

### 3. Algoritmos en la policía

- ÁLVAREZ, P. “El Estado reconoce ‘un fallo en cadena’ de Interior y Justicia en el asesinato de dos niñas a manos de su padre” [en línea]. *El País*. 24 de diciembre de 2021. <<https://elpais.com/sociedad/2021-12-24/el-estado-reconoceun-fallo-en-cadena-de-interior-y-justicia-en-el-asesinato-de-dos-ninas-a-manos-de-su-padre.html>> [Consulta: 21/11/2022]
- BELLIO, N. “Spanish police plan to extend use of its lie-detector while efficacy is unclear” [en línea]. Berlín: AW AlgorithmWatch gmbH. 27 de octubre de 2020. <<https://algorithmwatch.org/en/spain-police-veripol/>> [Consulta: 21/11/2022]
- DIGITAL FUTURE SOCIETY. “Capítulo 4. VioGén, el software contra la violencia machista” [pódcast]. 17 de marzo de 2022. <<https://digitalfuturesociety.com/es/podcasts/capitulo-4-viogen-el-software-contra-la-violencia-machista/>> [Consulta: 21/11/2022]
- FUNDACIÓN ETICAS. “La auditoría externa del sistema VioGén” [en línea] <<https://eticasfoundation.org/es/gender/the-external-audit-of-the-viogen-system/>> [Consulta: 21/11/2022]
- GARCÍA, J. “VeriPol, el polígrafo ‘inteligente’ de la policía, puesto en cuestión por expertos en ética de los algoritmos” [en línea]. *El País*. 9 de marzo de 2021. <<https://elpais.com/tecnologia/2021-03-08/veripol-el-poligrafo-inteligente-de-la-policia-puesto-en-cuestion-por-expertos-en-etica-de-los-algoritmos.html>> [Consulta: 21/11/2022]

- GARCÍA, T. “Minority Report en las policías europeas: llegan los sistemas de predicción de delitos” [en línea]. *El Salto Diario*. 12 de mayo de 2019. <<https://www.elsaltodiario.com/tecnologia/minority-report-policias-europa-sistemas-algoritmos-prediccion-delitos>> [Consulta: 21/11/2022]
- MILLER, L. “LAPD will end controversial program that aimed to predict where crimes would occur” [en línea]. *Los Angeles Times*. 21 de abril de 2020. <<https://www.latimes.com/california/story/2020-04-21/lapd-ends-predictive-policing-program>> [Consulta: 21/11/2022]
- PODER JUDICIAL. *Estadística del Sistema de Seguimiento Integral en los casos de Violencia de Género (Sistema VioGén)*. <<https://www.poderjudicial.es/cgpj/es/Temas/Estadistica-Judicial/Estadistica-por-temas/Datos-penales--civiles-y-laborales/Violencia-domestica-y-Violencia-de-genero/Sistema-de-Seguimiento-Integral-en-los-casos-de-Violencia-de-Genero--Sistema-VioGen-/>> [Consulta: 21/11/2022]
- SANKIN, A.; MEHROTRA, D.; MATTU, S.; GILBERTSON, A. “Crime Prediction Software Promised to Be Free of Biases. New Data Shows It Perpetuates Them” [en línea]. *The Markup*. 2 de diciembre de 2021. <<https://themarkup.org/prediction-bias/2021/12/02/crime-prediction-software-promised-to-be-free-of-biases-new-data-shows-it-perpetuates-them>> [Consulta: 21/11/2022]

#### 4. Predictores de riesgo en prisiones

- ARAGÓ, L.; SAURA, G. “El algoritmo de prisiones que no rinde cuentas a nadie” [en línea]. *La Vanguardia*. 7 de diciembre de 2021. <<https://www.lavanguardia.com/vida/20211207/7911428/algoritmo-prisiones-rinde-cuentas-nadie.html>> [Consulta: 21/11/2022]
- CENTRO DE ESTUDIOS JURÍDICOS Y FORMACIÓN ESPECIALIZADA. “La taxa de reincidència penitenciària 2014 a Catalunya” [en línea]. *Justidata*.

Núm. 59, marzo de 2015. <[https://cejfe.gencat.cat/web/.content/home/publicacions/butlleti\\_justidata/justidata59.pdf](https://cejfe.gencat.cat/web/.content/home/publicacions/butlleti_justidata/justidata59.pdf)> [Consulta: 21/11/2022]

- DIGITAL FUTURE SOCIETY. “Capítulo 2. RisCarvi (II): ¿Se puede predecir el próximo delito?” [pódcast]. 17 de febrero de 2022. <<https://digitalfuturesociety.com/es/podcasts/capitulo-2-riscarvi-ii-se-puede-predecir-el-proximo-delito/>> [Consulta: 21/11/2022]
- GENERALITAT DE CATALUNYA. *Manual d'aplicació del protocol de valoració RisCarvi* [en línea]. Mayo de 2019. <[https://justicia.gencat.cat/web/.content/home/ambits/reinsercio\\_i\\_serveis\\_peni/manual-aplicacio-protocol-avaluacio-riscarvi.pdf](https://justicia.gencat.cat/web/.content/home/ambits/reinsercio_i_serveis_peni/manual-aplicacio-protocol-avaluacio-riscarvi.pdf)> [Consulta: 21/11/2022]

## 5. Otros casos de uso

- ALFTER, B. “Automating Society Report 2020. Denmark” [en línea]. *Algorithm Watch*. Berlín: AW AlgorithmWatch gGmbH. <<https://automatingsociety.algorithmwatch.org/report2020/denmark/>> [Consulta: 21/11/2022]
- BBC NEWS. “A-levels and GCSEs: How did the exam algorithm work?” [en línea]. *BBC*. 20 de agosto de 2020. <<https://www.bbc.com/news/explainers-53807730>> [Consulta: 21/11/2022]
- KAYSER-BRIL, N. “Austria’s employment agency rolls out discriminatory algorithm, sees no problem” [en línea]. *Algorithm Watch*. Berlín: AW AlgorithmWatch gGmbH. 6 de octubre de 2019. <<https://algorithmwatch.org/en/austrias-employment-agency-ams-rolls-out-discriminatory-algorithm/>> [Consulta: 21/11/2022]
- SIMONITE, T. “Meet the Secret Algorithm That’s Keeping Students Out of College” [en línea]. *Wired*. 10 de julio de 2020. <<https://www.wired.com/story/algorithm-set-students-grades-altered-futures/>> [Consulta: 21/11/2022]

### III. Transparencia, explicabilidad y rendición de cuentas

- BOYON, N. “Opinions about AI vary depending on countries’ level of economic development” [en línea]. *Ipsos*. 5 de enero de 2022. <<https://www.ipsos.com/en/global-opinions-about-ai-january-2022>> [Consulta: 21/11/2022]
- DE LA CUEVA, J.; VALDIMA, A. “The Paradox of Efficiency: Frictions Between Law and Algorithms” [en línea]. *Verfassungsblog on Matters Constitutional*. 2 de abril de 2022. <<https://verfassungsblog.de/roa-the-paradox-of-efficiency/>> [Consulta: 21/11/2022]
- DE LA SIERRA MORÓN, S. “Control judicial de los algoritmos: robots, administración y estado de derecho” [en línea]. *Derechocal.es*. 12 de mayo de 2021. <<https://derechocal.es/opinion/control-judicial-de-los-algoritmos-robots-administracion-y-estado-de-derecho>> [Consulta: 21/11/2022]
- GENERALITAT DE CATALUNYA. *Carta catalana per als drets i les responsabilitats digitals* [en línea]. Dirección General de Sociedad Digital. Diciembre de 2019. <[https://politiquesdigitals.gencat.cat/web/content/00-arbre/ciutadania/drets-responsabilitats-digitals/Carta\\_v2.pdf](https://politiquesdigitals.gencat.cat/web/content/00-arbre/ciutadania/drets-responsabilitats-digitals/Carta_v2.pdf)> [Consulta: 21/11/2022]
- OCDE. *Recommendation of the Council on Artificial Intelligence. OECD/LEGAL/0449* [en línea]. 22 de mayo de 2019. <<https://legalinstruments.oecd.org/en/instruments/OECD-LEGAL-0449>> [Consulta: 21/11/2022]
- ZAPATA LOPERA, R. “The algorithm that decided the destiny of thousands of families” [en línea]. *Medium*. 7 de diciembre de 2019. <<https://rzapatal.medium.com/the-algorithm-that-decided-the-destiny-of-many-families-7b374fa2574b>> [Consulta: 21/11/2022]

## **IV. Herramientas e iniciativas para un uso ético y transparente de la inteligencia artificial (IA)**

### **1. El modelo PIO del OEIAC: autoevaluación para un uso ético**

- OEIAC. *El modelo PIO (principios, indicadores y observables): Una propuesta de autoevaluación organizativa sobre el uso ético de datos y sistemas de inteligencia artificial*. Junio de 2022. <[https://www.udg.edu/ca/Portals/57/OContent\\_Docs/modelpio-CAS-v5.pdf](https://www.udg.edu/ca/Portals/57/OContent_Docs/modelpio-CAS-v5.pdf)>

### **2. El estándar algorítmico británico y su estrategia de datos**

- DOMAGALA, N. “Strengthening Algorithmic Transparency in the United Kingdom” [en línea]. *Open Government Partnership*. 30 de noviembre de 2021. <<https://www.opengovpartnership.org/stories/strengthening-algorithmic-transparency-in-the-united-kingdom/>> [Consulta: 21/11/2022]
- GOBIERNO DE LA GRAN BRETAÑA. *Algorithmic Transparency Standard*. Central Digital and Data Office. 29 de noviembre de 2021. <<https://www.gov.uk/government/collections/algorithmic-transparency-standard>> [Consulta: 21/11/2022]
- GOBIERNO DE LA GRAN BRETAÑA. *Data Ethics Framework*. Central Digital and Data Office. 12 de junio de 2018. <<https://www.gov.uk/government/publications/data-ethics-framework>> [Consulta: 21/11/2022]
- GOBIERNO DE LA GRAN BRETAÑA. *Provide information on how you use algorithmic tools to support decisions (pilot version)*. Central Digital and Data Office. 29 de noviembre de 2021. <<https://www.gov.uk/guidance/provide-information-on-how-you-use-algorithmic-tools-to-support-decisions-pilot-version>> [Consulta: 21/11/2022]



### 3. La estrategia de Barcelona para el uso de algoritmos y datos

- AYUNTAMIENTO DE BARCELONA. *Mesura de Govern de l'estratègia municipal d'algoritmes i dades per a l'impuls ètic de la intel·ligència artificial*. Comisionado de Innovación Digital. Abril de 2021. <[https://ajuntament.barcelona.cat/premsa/wp-content/uploads/2021/04/Mesura-de-Govern-Intel·ligencia-artificial\\_cat-v2.47-ca-ES\\_.pdf](https://ajuntament.barcelona.cat/premsa/wp-content/uploads/2021/04/Mesura-de-Govern-Intel·ligencia-artificial_cat-v2.47-ca-ES_.pdf)> [Consulta: 21/11/2022]
- DUBOIS, H.; LUDWINEK, A. *Access to social benefits: Reducing non-take-up*. Eurofound. 21 de septiembre de 2015. <<https://www.eurofound.europa.eu/publications/report/2015/social-policies/access-to-social-benefits-reducing-non-take-up>> [Consulta: 21/11/2022]

### 4. ¿Cómo hacer cumplir la transparencia? La Ley Rider y la guía de información algorítmica

- DEL CASTILLO, C. “Las empresas se niegan a desvelar sus algoritmos laborales un año después de aprobarse la obligación legal” [en línea]. *EIDiario.es*. 18 de mayo de 2022. <[https://www.eldiario.es/tecnologia/empresas-niegan-desvelar-algoritmos-laborales-ano-despues-aprobarse-obligacion-legal\\_1\\_9002361.html](https://www.eldiario.es/tecnologia/empresas-niegan-desvelar-algoritmos-laborales-ano-despues-aprobarse-obligacion-legal_1_9002361.html)> [Consulta: 21/11/2022]
- DEL CASTILLO, C. «Gemma Galdón: “Las empresas van a tener que dar datos concretos sobre cómo toman decisiones sus algoritmos”» [en línea]. *EIDiario.es*. 10 de junio de 2022. <[https://www.eldiario.es/tecnologia/gemma-galdon-empresas-dar-datos-concretos-toman-decisiones-algoritmos\\_128\\_9073291.html](https://www.eldiario.es/tecnologia/gemma-galdon-empresas-dar-datos-concretos-toman-decisiones-algoritmos_128_9073291.html)> [Consulta: 21/11/2022]
- RUIZ, M. “La ‘ley Rider’ cumple un año en pleno pulso entre las empresas y Trabajo” [en línea]. *Newtral*. 12 de agosto de 2022. <<https://www.newtral.es/ley-rider-cumple-ano/20220812/>> [Consulta: 21/11/2022]

## 5. Los registros públicos de Ámsterdam y Helsinki

- CASSAWERS, T. “Can algorithmic registers solve automated bias?” [en línea]. *Equal Times*. 24 de marzo de 2021. <<https://www.equaltimes.org/can-algorithmic-registers-solve?lang=es#.Y05tjOxBxQK>> [Consulta: 21/11/2022]
- CITY OF HELSINKI AI REGISTER. <<https://ai.hel.fi/en/ai-register/>> [Consulta: 21/11/2022]
- GEMEENTE AMSTERDAM. City of Amsterdam Algorithm Register. <<https://algorithregister.amsterdam.nl/en/ai-register/>> [Consulta: 21/11/2022]
- HIDALGO PÉREZ, M. “Ámsterdam y Helsinki crean un registro de algoritmos para mejorar la confianza ciudadana en la inteligencia artificial” [en línea]. *El País*. 7 de octubre de 2020. <<https://elpais.com/tecnologia/2020-10-06/amsterdam-y-helsinki-crean-un-registro-de-algoritmos-para-mejorar-la-confianza-ciudadana-en-la-inteligencia-artificial.html>> [Consulta: 21/11/2022]
- UNIVERSIDAD ADOLFO IBÁÑEZ. *Repositorio algoritmos públicos*. <<https://www.algoritmospublicos.cl/>> [Consulta: 21/11/2022]

## 6. La agencia española para la supervisión de la IA

- GONZALO, M. “España tendrá una agencia nacional de supervisión de algoritmos en 2022” [en línea]. *Newtral*. 22 de diciembre de 2021. <<https://www.newtral.es/agencia-nacional-supervision-algoritmos/20211222/>> [Consulta: 21/11/2022]
- MARCOS, J. “El Gobierno inicia con la Agencia de Inteligencia Artificial el traslado de sedes fuera de Madrid” [en línea]. *El País*. 12 de septiembre de 2022. <<https://elpais.com/espana/2022-09-12/el-gobierno-inicia-con-la-agencia-de-inteligencia-artificial-el-traslado-de-sedes-fuera-de-madrid.html>> [Consulta: 21/11/2022]

- GOBIERNO DE ESPAÑA. “El Gobierno inicia el proceso para elegir la sede de la Agencia Española de Supervisión de la Inteligencia Artificial”. Ministerio de Asuntos Económicos y Transformación Digital. 13 de septiembre de 2022. <<https://portal.mineco.gob.es/es-es/comunicacion/Paginas/agencia-esp%C3%B1ola-de-supervisi%C3%B3n-de-inteligencia-artificial.aspx>> [Consulta: 21/11/2022]
- LA FEDE.CAT ET AL. “Organizaciones de la sociedad civil reclamamos nuestra participación en la Agencia Estatal de Supervisión de Inteligencia Artificial (AESIA)” [en línea]. Septiembre de 2022. <<https://rightsinternationalspain.org/wp-content/uploads/2022/09/AESIA-cast.pdf>> [Consulta: 21/11/2022]

## **7. El Citizens’ Biometrics Council y la participación ciudadana en el debate algorítmico**

- ADA LOVELACE INSTITUTE. *The Citizens’ Biometrics Council* [en línea]. <<https://www.adalovelaceinstitute.org/project/citizens-biometrics-council/>> [Consulta: 21/11/2022]
- EDRI. *Will the European Parliament stand up for our rights by prohibiting biometric mass surveillance in the AI Act?* [en línea]. 10 de mayo de 2022. <<https://edri.org/our-work/will-the-european-parliament-stand-up-for-our-rights-by-prohibiting-biometric-mass-surveillance-in-the-ai-act/>> [Consulta: 21/11/2022]

Para la redacción de este volumen también se han llevado a cabo entrevistas con los siguientes expertos:

**Albert Sabater Coll**, director del Observatorio de Ética en Inteligencia Artificial de Cataluña (OEIAC).

**Cristina Pombo**, asesora en transformación digital en gobiernos, ética e IA en el Banco Interamericano de Desarrollo (BID).

**Carlos Castillo Ocaranza**, científico de datos e investigador de la Universidad Pompeu Fabra.

**Carme Meléndez Cuffi**, subdirectora general de Servicios a la Comunidad y Becas del Departamento de Educación de la Generalitat de Catalunya.

**Javier de la Cueva**, abogado y patrono de la Fundación Civio.

**Josep Maria Flores**, jefe del Área de Impulso de la Mejora de los Servicios Digitales de la Dirección General de Administración Digital de la Generalitat de Catalunya.

**Lorena Jaume-Palasi**, eticista e investigadora. Fundadora de The Ethical Tech Society.

**Michael Donaldson**, comisionado de Innovación Digital, Administración Electrónica y Buen Gobierno del Ayuntamiento de Barcelona.

A continuación se indican las fuentes normativas y los reglamentos consultados y referenciados en esta publicación:

Reglamento del Parlamento Europeo y del Consejo por el que se establecen normas armonizadas en materia de inteligencia artificial (Ley de inteligencia artificial) y se modifican determinados actos legislativos de la Unión. Comisión Europea. 21 de abril de 2021. <<https://eur-lex.europa.eu/legal-content/ES/TXT/HTML/?uri=CELEX:52021PC0206&from=EN>> [Consulta: 21/11/2022]

Reglamento (UE) 2016/679 del Parlamento Europeo y del Consejo, de 27 de abril de 2016, relativo a la protección de las personas físicas en lo que respecta al tratamiento de datos personales y a la libre circulación de estos datos y por el que se deroga la Directiva 95/46/CE (Reglamento general de protección de datos). <<https://eur-lex.europa.eu/legal-content/ES/TXT/HTML/?uri=CELEX:32016R0679&from=EN>> [Consulta: 21/11/2022]

Real Decreto-ley 9/2021, de 11 de mayo, por el que se modifica el texto refundido de la Ley del Estatuto de los Trabajadores, aprobado por el Real Decreto Legislativo 2/2015, de 23 de octubre, para garantizar los derechos laborales de las personas dedicadas al reparto en el ámbito de plataformas digitales. *Boletín Oficial del Estado* núm. 113, de 12 de mayo de 2021, páginas 56733 a 56738. <[https://www.boe.es/diario\\_boe/txt.php?id=BOE-A-2021-7840](https://www.boe.es/diario_boe/txt.php?id=BOE-A-2021-7840)> [Consulta: 21/11/2022]

Ley 15/2022, de 12 de julio, integral para la igualdad de trato y la no discriminación. *Boletín Oficial del Estado* núm. 167, de 13 de julio de 2022. <<https://www.boe.es/eli/es/l/2022/07/12/15/con>> [Consulta: 21/11/2022]

ISBN 978-84-18601-05-7



9 788418 601057