

Inteligencia Artificial (y Administración Pública)*

Artificial Intelligence (and Public Administration)

J. Ignacio Criado

Universidad Autónoma de Madrid

ORCID ID 0000-0002-9184-9696

ignacio.criado@uam.es

Cita recomendada:

Criado, J.I. (2021). Inteligencia Artificial (y Administración Pública). *Eunomía. Revista en Cultura de la Legalidad*, 20, pp. 348-372.

doi: <https://doi.org/10.20318/eunomia.2021.6097>

Recibido / received: 12/01/2021
Aceptado / accepted: 01/03/2021

Resumen

Los algoritmos y la Inteligencia Artificial (IA) en el sector público se están adoptando en diferentes contextos. Conceptualmente, todavía no están claras las implicaciones de la IA y los algoritmos en las administraciones públicas. Sin embargo, diferentes organizaciones internacionales han definido una serie de principios éticos y cuestiones legales a tener en cuenta. Las instituciones de la Unión Europea se han convertido en el actor clave para guiar la adopción de IA, estableciendo las primeras líneas para las políticas públicas a nivel nacional. El Gobierno de España ha comenzado a desarrollar una estrategia en esa línea, si bien con unos responsables de políticas tecnológicas encargados de su implementación, que cuentan con sus propias percepciones y prioridades. Este artículo se inicia con las siguientes preguntas de carácter descriptivo y exploratorio: ¿Cómo están enmarcando ética y jurídicamente la IA los principales organismos internacionales? ¿Cuál es el marco de prioridades de la política de IA promovida por el Gobierno de España en relación con la Unión Europea? ¿Cuál es la percepción sobre la IA de los responsables de gestión de tecnologías (CIOs) en los ministerios? Para responder a estas preguntas, este trabajo se nutre de una revisión documental y de parte de los resultados de un cuestionario sobre las percepciones de los responsables de gestión de las tecnologías en los ministerios españoles. Nuestro análisis permite responder exploratoriamente las preguntas de investigación del estudio. Además, los resultados señalan algunas claves para diseñar políticas y estrategias de uso de la IA en diferentes contextos político-administrativos.

Palabras clave

Inteligencia artificial, algoritmos, servicios públicos, innovación pública, administración pública, gestión pública, gobernanza pública, gestores públicos, CIOs, políticas públicas.

* Este trabajo se ha realizado con el apoyo del Programa On TRUST – CM H2019/HUM-5699 de la Comunidad de Madrid y el Fondo Social Europeo. Parte de una versión previa de este documento se presentó como ponencia en el XXV Congreso del CLAD.



Abstract

Algorithms and Artificial Intelligence (AI) in the public sector have been adopted in different contexts. Conceptually, the implications of AI and algorithms in public administration are still unclear. However, different international organizations have defined a series of ethical principles and legal issues to take into account. The European Union institutions have become a key player in guiding the adoption of AI, setting the first guidelines for public policies at the national level. The Spanish Government has developed a strategy according to these lines, albeit with those responsible for implementation having their own perceptions and priorities. Hence, this paper begins with the following descriptive and exploratory questions: How are the main international organizations framing AI ethically and legally? What is the framework of priorities of the AI policy endorsed by the Government of Spain in connection with the European Union? What is the perception about AI by IT managers (CIOs) in ministries? To answer these questions, this study is based on a documentary analysis and part of the results of a survey about the perceptions of public managers responsible for IT implementation in Spanish ministries. Our analysis encompasses an exploratory answer to the research questions guiding the study. In addition, the results indicate some key issues in designing policies and strategies to use AI in different politico-administrative contexts.

Keywords

Artificial intelligence, algorithms, public services, public innovation, public administration, public management, public governance, public managers, CIOs, public policy.

SUMARIO. 1. Introducción. 2. Inteligencia Artificial y Algoritmos. Una aproximación conceptual. 2.1. Definición y modelo de madurez de la Inteligencia Artificial en las administraciones públicas. 2.2. Algoritmos. Características e implicaciones para el sector público. 3. Gobernanza de la Inteligencia Artificial a nivel internacional. 4. Marco de la política de Inteligencia Artificial en la Unión Europea. 5. Hacia una política de Inteligencia Artificial en España. 6. La perspectiva de los CIOs ministeriales españoles. 7. Debates y conclusiones.

1. Introducción

El uso de algoritmos basados en sistemas de Inteligencia Artificial (IA) en el sector público es una realidad creciente, pero todavía poco explorada, a pesar de que el interés por el tema ha aumentado sin cesar en los últimos años. En las siguientes páginas nos acercamos a este fenómeno y a algunas de sus implicaciones desde una triple perspectiva. Primero, caracterizando lo que son IA y algoritmos, así como su relación con las administraciones públicas. Segundo, describiendo el marco internacional, europeo y nacional, así como algunos dilemas éticos, sobre la adopción de la IA en el ámbito público. Tercero, analizando las percepciones de responsables de tecnologías de información y comunicación en ministerios españoles sobre lo que es la IA, así como sus potenciales beneficios y desafíos dentro de las administraciones públicas. Aquí nos preguntamos por el alcance de la IA y los algoritmos en las administraciones públicas, en línea con lo que se ha planteado en diferentes contextos (Dutton, 2018; Margetts & Dorbantu, 2019; Meijer, 2017). En consecuencia, pretendemos enriquecer el debate en torno a uno de los temas más recurrentes en el estudio de las administraciones públicas durante los últimos años.

El estudio de algoritmos, IA y administración pública se encuentra en un proceso semejante respecto a otros ámbitos de las sociedades y economías contemporáneas. Gobiernos de todo el mundo han definido estrategias y políticas de IA frente a un futuro incierto (Criado, 2019; Lee, 2018; Valle-Cruz *et. al*, 2020). Muchos



de ellos han adoptado nuevas prácticas de IA en diferentes áreas del sector público, incluidas salud, educación, seguridad, emergencias, defensa, entre otras (Agarwal, 2018; Criado y Gil-García, 2019; Desouza, 2018; Janssen & Kuk, 2016). También se han desarrollado los primeros análisis de experiencias con impactos en diferentes dimensiones administrativas, tales como los procesos de adopción de decisiones, transparencia, organización del trabajo, privacidad, discrecionalidad administrativa o ética pública (Sun y Medaglia, 2019; Wirtz *et al.*, 2018). Hay quien habla de una nueva ola de innovaciones en el sector público que llevaría a hablar de Administración Pública 4.0 dentro del marco de una cuarta revolución industrial que tiene a las tecnologías disruptivas como la IA en el centro. Por ello, este trabajo parte de las siguientes preguntas de carácter descriptivo y exploratorio: ¿Cómo están enmarcando ética y jurídicamente la IA los principales organismos internacionales? ¿Cuál es el marco de prioridades de la política de IA promovida por el Gobierno de España en relación con la Unión Europea? ¿Cuál es la percepción sobre la IA de los responsables de gestión de tecnologías (CIOs) en los ministerios?

Este artículo analiza la realidad de los algoritmos, la IA y las administraciones públicas a través de una revisión de literatura y documental, así como datos primarios de un cuestionario. Si bien se recomienda algún trabajo previo para una revisión de literatura más exhaustiva (Valle-Cruz *et al.*, 2020), las siguientes páginas revisan algunas de las más recientes aportaciones académicas sobre el tema desde una perspectiva politológica asentada en los estudios de Administración Pública. Adicionalmente, el estudio también realiza una revisión documental referida a acciones y documentos procedentes de organismos internacionales (Naciones Unidas a través de la ITU o UNESCO, OCDE o Consejo de Europa), la Unión Europea y el Gobierno de España sobre IA en el sector público, tomando en cuenta otros casos (Gobierno de Colombia, 2019; Gobierno de EE. UU., 2020; Gobierno de Portugal, 2020; Gobierno de Uruguay, 2019). Además, la parte final del estudio ofrece datos primarios sobre las percepciones de los responsables de gestión de las tecnologías de la información y la comunicación en los ministerios españoles, en base a una encuesta administrada a todos ellos, siguiendo otros trabajos previos (Criado *et al.*, 2021; Ganapati & Reddick, 2012). Con esto se logra un contraste entre definición de políticas públicas y realidad en las percepciones de los responsables de su gestión en el sector público. También tiene el objetivo de mostrar algunas claves para diseñar políticas y estrategias de IA en diferentes contextos nacionales/federales, regionales/estatales y locales/municipales.

El documento se desarrolla de la siguiente forma. A continuación se elabora una aproximación conceptual de la IA y los algoritmos en el sector público. En la tercera sección se presenta la IA desde una perspectiva internacional, teniendo en cuenta el interés creciente en esta materia provocado en organismos internacionales. En la cuarta sección se analiza el marco de la política de IA en la Unión Europea, mientras en la siguiente (quinta) se muestran los principales hitos identificados en el ámbito español dentro del nivel nacional de gobierno. Esta sección sobre el caso español se contrasta con los datos de la perspectiva de los responsables de la política tecnológica en cada ministerio, sobre las políticas y adopción de la IA en el sector público. Finalmente, se completa el trabajo con una discusión de los resultados y unas conclusiones que abren la vía a futuros estudios.

2. Inteligencia Artificial y Algoritmos. Una aproximación conceptual

De cara a profundizar en el alcance de la Inteligencia Artificial (IA) y los algoritmos en el sector público, primero es necesario realizar una aproximación conceptual a ambos términos. A continuación no solo se presentan definiciones sobre IA y algoritmos, sino que desde el primer momento se tiene en cuenta su relación con las administraciones

públicas, con el objetivo de ir centrando el interés en sus implicaciones para lograr una gobernanza pública inteligente.

2.1. Definición y modelo de madurez de la Inteligencia Artificial en las administraciones públicas

La definición de lo que es Inteligencia Artificial (IA) se ha planteado de diferentes maneras. Se suele considerar que su origen se encuentra en *Dartmouth College* (1956) para caracterizar un ámbito emergente de investigación ligado a la fisiología del cerebro, análisis formal de la lógica proposicional y las ciencias de la computación (Sun y Medaglia, 2018). Por su parte, Valle-Cruz *et al.* (2020) indican que este ámbito de las ciencias se orienta a la creación de algoritmos que nutren máquinas inteligentes, simulando el comportamiento individual y colectivo, no sólo de los seres humanos, sino también de otros seres vivos, automatizando actividades, aprendiendo y evolucionando, así como mejorando el rendimiento y los resultados obtenidos con y sin ayuda de los seres humanos. Estas definiciones tienen en cuenta una dimensión científica del término y su identificación con un área de estudio.

De manera reciente, la aportación de la Unión Europea ofrece una visión muy concisa y aplicable a diferentes contextos. Se define la IA como «sistemas que muestran comportamiento inteligente mediante el análisis de su ambiente y el desarrollo de acciones –con cierto grado de autonomía– para alcanzar determinados objetivos (*systems that display intelligent behaviour by analysing their environment and taking actions –with some degree of autonomy– to achieve specific goals*)» (European Commission, 2019, p. 1). Esta noción de IA hace referencia explícita a la inteligencia, que se interpreta como la manera en que las máquinas son capaces de lograr algún tipo de racionalidad mediante la percepción del ambiente con el que interactúan, mediante sensores, adquiriendo y analizando datos, razonando sobre los mismos, de cara a adoptar algún tipo de decisión y acción, con el fin de alcanzar determinados objetivos definidos (o no) previamente (European Commission, 2019). Por consiguiente, los sistemas basados en IA pueden usar reglas simbólicas o aprender un modelo numérico. Además, pueden adaptar sus comportamientos a partir del análisis del entorno y de la interacción con el mismo. Esta definición acerca la IA a la robótica, uno de sus ámbitos, y se refiere a diferentes técnicas, incluyendo el *machine learning*, siempre teniendo como centro los datos masivos.

A partir de las definiciones presentadas, habría que preguntarse por qué ahora se ha situado la IA en el centro del debate público. Seguramente, el interés en la IA se refiere, precisamente, a la capacidad creciente del aprendizaje automático de los sistemas de *machine learning*, la disponibilidad de mayores cantidades de datos, así como el aumento de la capacidad de procesamiento de datos, a un coste cada vez menor (Joint Research Center, 2018). Unido a ello, el potencial para incrustar los algoritmos en diferentes ámbitos de la vida humana sobre los que percibir datos para predecir comportamientos, incluyendo el sector público, ha convertido a la IA, y todas sus aplicaciones, en protagonista de los procesos de digitalización de las sociedades actuales y también de las administraciones públicas. De hecho, incluso las relaciones de poder entre las principales potencias internacionales están determinadas por el control de los datos y el desarrollo de la IA del futuro.

Desde una perspectiva de madurez tecnológica sobre el desarrollo de la IA en las organizaciones, podemos plantear la cuestión conceptual desde otra vertiente. Aquí, la IA se puede definir tomando en cuenta dos de sus capacidades (Russell & Norvig, 2016): (i) máquinas desarrollando tareas mediante comportamientos inteligentes que reproducen los humanos, y (ii) máquinas comportándose racionalmente percibiendo lo que sucede en el ambiente y tomando decisiones para

alcanzar objetivos concretos. En los dos casos anteriores, se comienza a dar una mayor importancia a la cuestión de las capacidades que desarrolla la IA.

Esta perspectiva se amplía con otra idea de los mismos autores, quienes diferencian dos nociones de IA atendiendo a su complejidad. En efecto, Russell & Norvig (2016) mencionan la IA débil y la IA fuerte. En el primer caso (IA débil), se refiere a sistemas autónomos capaces de desarrollar tareas que requieren capacidades humanas sencillas (por ejemplo, percepción visual, comprensión del contexto, razonamiento probabilístico y tratamiento de la complejidad). En el segundo caso (IA fuerte), se centra en la idea de sistemas autónomos con capacidades e inteligencia super-humanas, que emulan la complejidad de la capacidad de los seres humanos para pensar, aprender y desarrollar tareas complejas (por ejemplo, juicio ético, razonamiento simbólico, gestión de situaciones sociales e ideación) (Brynjolfsson y McAfee, 2017). Las diferencias entre ambos tipos reflejan una evolución creciente de su complejidad a la hora de aplicarse dentro de entornos organizativos.

En términos de modelo madurez dentro del sector público también se puede diferenciar diferentes momentos de la IA. Lee (2018) ha sugerido cuatro estadios diferentes, ligados a la forma en que aprovechan el poder de la IA para generar disrupción en determinados sectores, así como para lograr entretenerse de manera más profunda en la vida diaria de la gobernanza de los sistemas en los que operan las organizaciones, servicios públicos y las personas. Esos estadios son los siguientes: IA Internet, IA Negocio, IA Percepción e IA Autónoma. Cada uno de ellos implica diferentes momentos en la evolución de la IA dentro de las organizaciones, en general, y las administraciones públicas, en particular.

Los dos primeros estadios de la IA (IA Internet e IA Negocio) se encuentran ya totalmente operativos en diferentes sectores de la vida humana. En concreto, estos dos estadios implican la reformulación de los mundos financiero y económico, la sustitución de determinados trabajos jurídicos (por ejemplo, decisiones que antes tomaban jueces), el diagnóstico médico, la intermediación financiera, el periodismo y las fuentes de información, etc. En todos esos ámbitos, la utilización masiva de algoritmos está generando transformaciones en su esencia, tanto en la forma en que interactúan los actores participantes en esos procesos, como en la propia naturaleza de la actividad.

Los dos estadios posteriores todavía se encuentran en una fase experimental o, incluso, de ideación. La categoría IA Percepción se refiere a expandir su poder a través del ambiente en diferentes situaciones de la vida, digitalizando el mundo alrededor de las personas a través de la proliferación de sensores y dispositivos inteligentes. Estos dispositivos irán transformando el mundo físico en datos digitales que podrían ser analizados y optimizados a través de algoritmos de aprendizaje profundo (Lee, 2018). Ejemplos concretos de este nivel incluyen la digitalización de los flujos de tráfico urbano a través de cámaras y el reconocimiento de objetos mediante IA, el apoyo a las experiencias de compra basándose en la monitorización de hábitos previos, la digitación del sonido ambiente de las casas y edificios de personas, así como la realización de ambientes de aprendizaje basados en la identificación visual, el reconocimiento del lenguaje y la creación de perfiles detallados pasados en comportamientos previos. Esta etapa implica lograr que los dos mundos, analógico (*off line*) y digital (*on line*), converjan.

Finalmente, la categoría IA Autónoma representa el último estadio en la evolución de la IA, integrando los tres momentos anteriores. Esta última fase supone la existencia de máquinas capaces de optimizar desde grandes cantidades de

conjuntos de datos hasta la generación de nuevas capacidades de sensorización de cualquier actividad humana. Aquí, la IA y las máquinas basadas en algoritmos complejos podrán trabajar de una manera totalmente autónoma, aprehender el mundo y darle forma con sus acciones independientes. Esta última etapa es la que en algunos contextos se denomina singularidad (*singularity*), el momento en el que las máquinas serían capaces de pensar y actuar de forma independiente y, por consiguiente, al margen de los seres humanos.

2.2. Algoritmos. Características e implicaciones para el sector público

La Inteligencia Artificial (IA) tiene una relación muy estrecha con los algoritmos que permiten su existencia, si bien la atención a los algoritmos ha sido mucho más escasa desde la perspectiva de sus implicaciones, sobre todo, en el ámbito de lo público. En términos amplios, el concepto de algoritmo se sostiene en el ideal de adquisición de datos de diferentes fuentes, así como su manipulación para lograr algún tipo de resultado (Janssen y Kuk, 2016). Tal y como estos autores plantean, los algoritmos disponen de una característica esencial para entender su potencial: los algoritmos co-evolucionan con datos, sistemas y humanos dentro de complejos sistemas socio-técnicos «*algorithms dynamically co-evolve with data, systems, and humans within a complex socio-technical system*» (Janssen & Kuk, 2016, p. 372). Por consiguiente, se puede esperar que las implicaciones de la IA y los algoritmos vayan más allá de otras tecnologías previas dado que refuerzan el poder de las interacciones mediadas por datos, además de mostrar una capacidad de aprendizaje que les permite evolucionar¹.

Así las cosas, los algoritmos nos interesan desde la perspectiva de las implicaciones derivadas de su materialización y su capacidad performativa. En este sentido, los algoritmos son importantes porque incorporan características cruciales para las transacciones económicas, la acción colectiva, las redes de gobernanza o las relaciones inter-personales. Introna (2015) sugiere en su estudio sobre escritura académica, por un lado, que los algoritmos adquieren relevancia porque parecen operar bajo la superficie o en el fondo, esto es, parecen ser inescrutables (*incrustability*). Además, los algoritmos son actores significativos porque resultan ejecutables (*executable*), lo que significa que operan de manera «automática» sin la necesidad de intervención humana. Ambas características son esenciales para entender su impacto en nuestras realidades sociales, políticas y administrativas.

Adicionalmente, Introna (2015) habla de la performatividad (*performativity*) de los algoritmos. Esta es la tercera característica que resulta clave cuando nos referimos a cuestiones relacionadas con la gobernanza de los algoritmos. La performatividad de los algoritmos se inspira en el trabajo de Foucault sobre «gobernanza» (*governmentality*), es decir, «*rooted in an ontology of becoming, often referred to as process thought*» (Introna, 2015, p. 7). En esta misma línea, el autor sugiere que «*in the relational flow of becoming, each action inherits from prior actions and imparts to subsequent actions*» (Introna, 2015, p. 8). En suma, la gobernanza de los algoritmos implica interacciones humanas y desde la dimensión de gobiernos y administraciones públicas, personas adoptando decisiones, resolviendo problemas o trabajando activamente, interaccionando con algoritmos. Todo ello se produce en un flujo constante de llegar a ser, donde cada acción recibe algo de las anteriores acciones e informa a las sucesivas.

¹ Para una revisión más sistemática sobre algoritmos se sugieren Brassard y Bratley (1998), Comen et al. (2009) o LaValle (2006).

Teniendo en cuenta lo anterior podemos concluir que los algoritmos tienen implicaciones en la acción de gobernar. Los gobiernos y administraciones públicas desarrollan sus actividades gracias a dispositivos y tecnologías que les permiten gestionar y controlar los datos, sobre todo, que emanan de la ciudadanía. En consecuencia, los algoritmos incorporan características que son diferenciales y suponen la necesidad de conocer su potencial, así como las dimensiones en las que tendrán un impacto directo en la sociedad y la gobernanza de lo público (Criado *et al.*, 2020). Las reglas incrustadas en los algoritmos tienen una directa consecuencia en aspectos como los procesos de adopción de decisiones del personal directivo, sobre qué empleados contratar, la elegibilidad para recibir ayudas públicas, la denegación de un permiso carcelario, la decisión de proporcionar un tratamiento médico, la elección de las rutas de patrullaje en una ciudad, etc.

La aproximación anterior a la noción de algoritmo tiene implicaciones de calado. Entre otros aspectos, el hecho de que los algoritmos tengan un papel cada vez más importante en las interacciones que se producen entre los individuos y las organizaciones supone que, por ejemplo, la individualización de los servicios que forma parte de las promesas de los algoritmos puede ser algo no benigno en determinados contextos o situaciones. En otras palabras, la generalización de los algoritmos estaría generando impactos directos en principios básicos para el buen funcionamiento de nuestras instituciones democráticas, tales como la equidad, la igualdad, la justicia y la propia democracia (Lazer, 2015). De hecho, hoy los algoritmos están incrustados en transacciones económicas (compra venta de acciones y clasificación de riesgos), activismo político (promoción de campañas electorales o difusión de noticias falsas), actividades organizativas (identificación de fraude fiscal o diagnóstico de enfermedades), así como interacciones humanas (búsqueda de amistades/relaciones o selección de un lugar donde realizar un viaje). Por tanto, esos ámbitos estarían ajustando su naturaleza y resultados con los de sus propios comportamientos en base a los algoritmos aplicados. No obstante, el alcance de la redefinición de comportamientos humanos y sociales en base a la aplicación de algoritmos de forma generalizada estaría por analizar.

3. Gobernanza de la Inteligencia Artificial a nivel internacional

Más allá de las definiciones, diferentes organismos internacionales y países han desarrollado durante los últimos años aproximaciones sobre cómo regular, promover o implantar algoritmos e Inteligencia Artificial (IA) desde y en los gobiernos y administraciones públicas. Para comenzar este recorrido es necesario conocer las propuestas realizadas desde diferentes organismos internacionales y multilaterales. Entre ellos, destacamos la Organización para la Cooperación y el Desarrollo Económico (OCDE), Naciones Unidas (NU) y algunos de sus organismos del sistema o el Consejo de Europa (CdE). En todas estas aproximaciones se ofrecen conceptos y puntos clave que guían o establecen un marco de referencia para los países miembros de estos organismos, de ahí su papel clave desde una perspectiva del encuadre de las políticas de IA a promover desde gobiernos y administraciones públicas (desde el sector público hacia fuera), pero también en gobiernos y administraciones públicas (dentro del propio sector público).

Organización para la Cooperación y el Desarrollo Económico (OCDE). La OCDE ha desarrollado una aproximación fundamentada en una recomendación legal, adoptada en mayo de 2019 por los países miembros. En concreto, el documento se denomina *Recommendation of the Council on Artificial Intelligence* (Recomendación del Consejo sobre Inteligencia Artificial). Dentro del mismo, se establecen una serie de principios y recomendaciones en relación con la IA que deberían plantear los países. En este caso se pone también el acento en la importancia que tiene el

conocimiento sobre IA, es decir, «las competencias y los recursos, como datos, código, algoritmos, modelos, investigación, *know-how*, programas de entrenamiento, gobernanza, procesos y mejores prácticas que son requeridas para entender y participar en ciclo de vida de la IA» (OCDE, 2020, p. 7), así como los actores y partes (organizaciones y personas) interesadas o afectadas por los sistemas de IA.

En concreto, la OCDE ha planteado unos principios para una gobernanza fiable de la IA. Los principios son (OCDE, 2020): (1) crecimiento inclusivo, desarrollo sostenible y bienestar; (2) valores centrados en los seres humanos y justicia; (3) transparencia y explicabilidad; (4) robustez y seguridad; y (5) rendición de cuentas. Junto a lo anterior, también ha planteado una serie de recomendaciones para incorporar en las políticas nacionales y en la colaboración internacional para una IA de confianza: (i) invertir en investigación y desarrollo de IA; (ii) promover un ecosistema digital para la IA; (iii) facilitar y configurar un ambiente de política pública para la IA; (iv) construir capacidades humanas y prepararse para la transformación del mercado de trabajo; y (v) cooperación internacional para una IA fiable. Adicionalmente, la OCDE ha establecido el Observatorio OECD.AI (<https://oecd.ai>), donde no solo se recoge información sobre los principios y recomendaciones para consolidar una IA fiable, sino también información sobre publicaciones, áreas de implementación de proyectos, países involucrados, así como tendencias y datos. Lo anterior supone un importante espacio de reflexión sobre cómo promover una IA desde y en los gobiernos y administraciones públicas.

Naciones Unidas (UN). Unión Internacional de Telecomunicaciones (UIT). Desde una perspectiva diferente, la UIT ha promovido desde el año 2017 una cumbre global con el lema: «IA para el bien» (*IA for good*), con una clara orientación técnica dentro del propósito de esta organización, muy centrada en la colaboración con los ministerios nacionales de ciencia y telecomunicaciones. Uno de los elementos diferenciales de la UIT es la vinculación de la IA con los 17 Objetivos de Desarrollo Sostenible (ODS) de NU, de manera que pretende identificar proyectos, iniciativas y organizaciones relacionadas que los permitan acelerar.

En este caso, las actividades se han centrado en promover la colaboración internacional entre países desde una perspectiva fundamentalmente técnica. Esto se ha llevado a término a través de la celebración de tres conferencias globales sobre IA para el bien (2017, 2018 y 2019), el desarrollo de dos grupos focales sobre Inteligencia Artificial para la Salud (FG-AI4H) y *Machine Learning* para 5G (FG-ML5G) o la creación de un grupo de colaboración (con más de 5,000 miembros, dentro de la red social LinkedIn <https://www.linkedin.com/groups/8567748/>). Pero sobre todo cabe destacar el repositorio sobre IA (<https://www.itu.int/en/ITU-T/AI/Pages/ai-repository.aspx>), donde se incluye información sobre proyectos e iniciativas concretas, por países o cada uno de los 17 ODS, siempre destacando la dimensión de las tecnologías emergentes involucradas.

Naciones Unidas (UN). *Department of Economic and Social Affairs* (UNDESA). También cabe decir que dentro del *United Nations eReadiness Index* de 2020 (Naciones Unidas, 2020) elaborado por UNDESA, el concepto IA aparece ya incluido como uno de los ámbitos emergentes clave de la digitalización de los gobiernos y administraciones públicas. En este sentido, se plantea que la IA tiene una dependencia directa de las políticas nacionales de datos abiertos, tanto datos estructurados, como no estructurados. Así, se dibuja un escenario de futuros servicios públicos en los que

estas tecnologías innovadoras pueden ser usadas para conectar fragmentos de datos e información relacionados de manera rápida; pueden mejorar la provisión de servicios

públicos a través de la racionalización de las transacciones digitales, la reducción de los errores, y la eliminación de la necesidad de solicitar a las personas la misma información varias veces. (Naciones Unidas, 2020, p. 159).

También se menciona el papel controvertido de la IA en la promoción de esta nueva ola de la revolución digital.

Naciones Unidas (UN), UNESCO. La UNESCO es uno de los organismos internacionales que ha liderado la atención a los problemas éticos de la IA. Los trabajos de la UNESCO parten de noviembre de 2019, cuando en la 40 sesión de la Conferencia General se adoptó una resolución para preparar un instrumento sobre la ética de la IA en forma de recomendación. Este trabajo se ha concretado en un completo documento publicado en abril de 2020 denominado «Texto para una Recomendación sobre la Ética de la Inteligencia Artificial» (*Toward a Draft Text of a Recommendation on the Ethics of Artificial Intelligence*). Este documento se compone de varios apartados, sobre la base de una visión ética de la IA, que se sostiene en el respeto por los derechos humanos y las libertades fundamentales, no dejar atrás a nadie, así como la protección del medio ambiente.

Seguramente, la parte de mayor interés se encuentra en la parte doctrinal y el debate sobre principios éticos, tal y como ya lo han planteado varios países en sus estrategias nacionales, pero también en su aproximación práctica para facilitar a los países su despliegue. En concreto, hay una serie de 24 potenciales principios relevantes que deberían tenerse en cuenta, entre los que se encuentran: derechos humanos, inclusión, bienestar, autonomía, explicabilidad, transparencia, responsabilidad, rendición de cuentas, sostenibilidad, buen gobierno, etc.

Junto a lo anterior, la UNESCO también sugiere una serie de acciones de política pública destacados tanto desde como en los gobiernos y administraciones públicas. En el primer caso, se indica la necesidad de promover la investigación, poner los derechos humanos en primer lugar, educar sobre costes-beneficios y desigualdades, practicar una gobernanza multi-actor, promover un ecosistema para la IA, generar capacidades humanas y preparar el mercado de trabajo o promover la cooperación internacional para lograr una IA fiable.

En el caso de la segunda vertiente (acciones de política pública en los gobiernos y administraciones públicas), también se subrayan otras cuestiones. Concretamente, se destaca la necesidad de (1) asegurar cambios sistémicos y en procesos; (2) proporcionar medidas técnicas y estrategia de gobernanza; (3) asegurar la responsabilidad y rendición de cuentas; (4) proporcionar soluciones técnicas, más investigación y enfoques regulatorios; (5) asegurar el beneficio, la confianza, la sostenibilidad y un enfoque de género adecuado; (6) así como usar la IA en la prestación de servicios públicos. En este último caso, se habla de, por un lado, «asegurar que el uso del sector público de la IA en los procesos de adopción de decisiones es transparente y consistente con las obligaciones de los derechos humanos». Por otro lado, se considera necesario «establecer guías y políticas de acceso, transparencia y rendición de cuentas, en el uso y desarrollo de sistemas de adopción de decisiones automatizadas, incluyendo el uso por el sector público» (UNESCO, 2020, p. 48). Lo anterior se completa con propuestas para reducir la brecha digital, establecer políticas de datos, asegurar mercados abiertos y consolidar las infraestructuras tecnológicas necesarias.

Consejo de Europa (CdE). El CdE ha desarrollado también una perspectiva propia, en este caso, centrada en diez recomendaciones concretas y en torno al impacto del uso de la IA desde la perspectiva de los derechos humanos. En concreto,

los ámbitos planteados son los siguientes (Consejo de Europa, 2019): (i) evaluación del impacto; (ii) consultas públicas para que los actores concernidos sean invitados a proporcionar *inputs*; (iii) obligación de los estados miembros a facilitar la implementación de estándares de derechos humanos en el sector privado; (iv) información y transparencia sobre cómo se usan los sistemas de IA, además de sobre cómo se adoptan las decisiones; (v) evaluación independiente de órganos autónomos respecto de las autoridades públicas y privadas; (vi) no discriminación e igualdad para prevenir y/o mitigar los riesgos para determinados grupos; (vii) protección de datos y privacidad; (viii) libertad de expresión, de reunión y asociación, así como el derecho al trabajo; (ix) remedios ligados al establecimiento de mecanismos de responsabilidad y rendiciones de cuentas por las posibles violaciones de derechos humanos que puedan suceder al implantar sistemas de IA; y (x) promoción de la alfabetización en IA. En suma, se trata de una colección de principios éticos y jurídicos generales en línea con los planteados anteriormente.

4. Marco de la política de Inteligencia Artificial en la Unión Europea

La agenda europea de Inteligencia Artificial (IA) se ha ido consolidando con los años en torno a varios aspectos. Primero, desde la Comisión Europea se ha tratado de insistir en el relato sobre la búsqueda de un espacio propio dentro de un contexto con dos líderes internacionales (China y EE. UU.), respecto de cuyas estrategias de IA se pretende marcar distancia. En concreto, en un estudio de la Comisión Europea se ha planteado que

existe una fuerte competición global sobre la IA entre Estados Unidos, China y Unión Europea (UE). Los Estados Unidos lideran por ahora pero China se acerca rápidamente y se espera que lidere en 2030. Para la UE, no es una carrera en la que ganar o perder, sino una vía de alcanzar las oportunidades que ofrece la IA de manera que sea humano-céntrica, ética, segura, y confiable en nuestros valores constitutivos” (Comisión Europea, 2018a, pp. 12-13).

Como ya se ha apuntado más arriba, los modelos de IA y gobernanza de los algoritmos ponen en cuestión muchos de los equilibrios éticos de nuestras sociedades y, de facto, plantean problemas relacionados con sesgos y discriminación, así como libertades fundamentales y privacidad de datos personales, que resulta esencial entender *desde* gobiernos y administraciones públicas (Criado, 2021). Y desde luego, la UE lo pretende hacer desde un enfoque propio, encuadrando las políticas de IA a desarrollar desde gobiernos y administraciones públicas (desde el sector público hacia fuera), pero también *en* gobiernos y administraciones públicas (dentro del propio sector público).

Desde un plano discursivo, las instituciones de la UE han puesto el acento en la necesidad de crear un modelo propio de IA basado en principios éticos, democráticos, así como el control de los datos personales y la privacidad de las personas. Desde un primer momento, lo anterior se ha destacado a través de la necesidad de atender la Carta de los Derechos Fundamentales de la UE o el Reglamento General de Protección de Datos (*General Data Protection Regulation*, GDPR, por sus siglas en inglés) para garantizar ese enfoque, fundamentado en principios éticos, seguridad y responsabilidad, así como en el empoderamiento de las personas y consumidores para lograr lo máximo de la IA (European Commission, 2018b). De hecho, el artículo 22 y siguientes del GDRP establecen la noción y limitaciones, por ejemplo, de las decisiones individuales automatizadas, incluida la elaboración de perfiles o todo aquello relacionado con la protección de los datos personales en un contexto de IA. En definitiva, parecería una apuesta por transitar

hacia un modelo propio que tenga la protección de los datos personales como eje central.

Las acciones más recientes sobre IA dentro de la UE están consolidando ese enfoque singular. Por un lado, el Libro Blanco de la IA (a partir de aquí Libro Blanco IA) es la propuesta de política que han aprobado las instituciones europeas (febrero de 2020) para establecer un verdadero marco de política pública en la materia (Comisión Europea, 2020). Aquí también tenemos que diferenciar entre la promoción de la IA *desde* gobiernos y administraciones respecto de la IA *en* los gobiernos y administraciones públicas. Adicionalmente, la propuesta europea también se fundamenta en los principios de excelencia y confianza.

En lo que se refiere a la excelencia, la UE pretende situarse al nivel de los ecosistemas de IA más importantes del mundo. Con ello se asume la necesidad de adoptar y desarrollar la IA en todos los sectores económicos y las administraciones públicas, a través de (1) la colaboración con los Estados miembros; (2) centrando los esfuerzos de la comunidad de investigación e innovación; (3) el desarrollo de habilidades para superar la escasez de competencias; (4) la preocupación por el acceso de las PYMES a la IA; (5) el aseguramiento de asociaciones público-privadas amplias para lograr compromisos sobre IA, datos y robótica; (6) la garantía de acceso a los datos y las infraestructuras informáticas; (7) asunción del liderazgo mundial en IA a través de la colaboración en foros internacionales con países que asuman las normas y valores de la UE en derechos humanos y privacidad; y (8) la promoción de la adopción de la IA por parte del sector público.

Desde el punto de vista del sector público, el Libro Blanco IA se plantea lo siguiente:

Acción 6: La Comisión iniciará conversaciones por sector abiertas y transparentes, en las que dará prioridad a la atención sanitaria, las administraciones rurales y los operadores de servicios públicos, para presentar un plan de acción que facilite el desarrollo, la experimentación y la adopción de la inteligencia artificial. Las conversaciones por sector se emplearán para preparar un “Programa de adopción de la IA” específico que respaldará la contratación pública de sistemas de inteligencia artificial, y ayudará a transformar los propios procesos de esta contratación. (Comisión Europea, 2020)

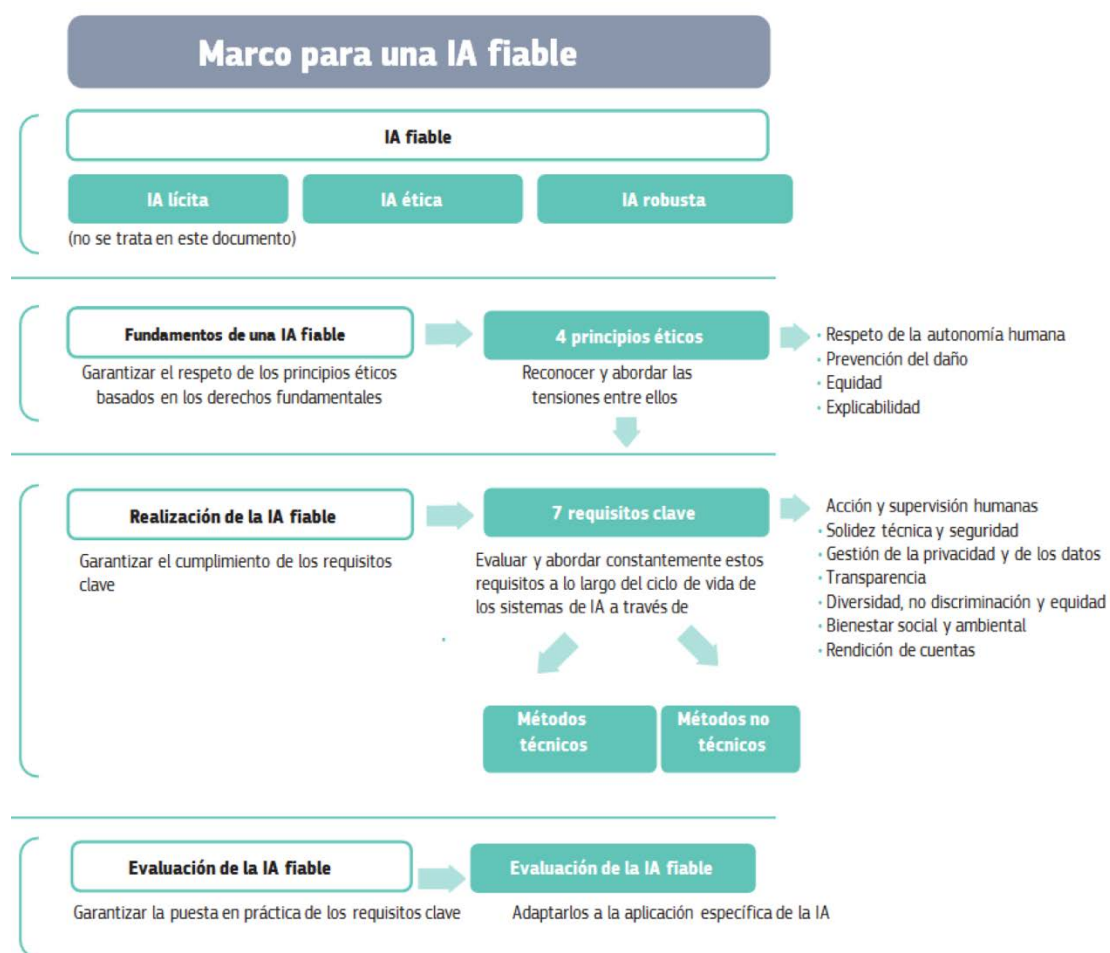
En este sentido se plantea que es necesario promover que los principales servicios públicos, transporte, supervisión financiera y otras áreas de interés adopten rápidamente productos y servicios basados en IA. En concreto, el marco de política que propone la UE sugiere la necesidad de focalizarse en los sectores públicos de salud y transporte, aquellos en los que se estima que las tecnologías de IA están lo suficientemente maduras para su adopción e implementación a gran escala. En todo caso, en este enfoque quedan fuera, de nuevo, cuestiones políticas de calado, especialmente, cómo la IA puede ayudar a mejorar los procesos de adopción de decisiones en el sector público, cuál es la aportación de la IA al diseño de las políticas públicas o qué oportunidades ofrece la IA para desarrollar modelos de gobernanza colaborativa e innovación abierta con la ciudadanía de cara a generar valor público.

Desde la perspectiva del ecosistema de confianza, la UE incide en la importancia de un marco regulador de la IA que, al mismo tiempo, facilite las oportunidades y se enfrente a las amenazas. Esta otra dimensión de la IA desde y en gobiernos y administraciones públicas se aproxima a la cuestión teniendo presentes los principales problemas que la IA puede generar, sobre todo, en lo que respecta a la protección de derechos fundamentales (como la protección de datos, privacidad o la no discriminación) como consecuencia de su opacidad. Así, siguiendo el trabajo del

Comité de Expertos de Alto Nivel creado por la propia Comisión Europea, se han definido una serie de requisitos para las aplicaciones y sistemas de IA (Comisión Europea, 2020): (a) datos y entrenamiento; (b) datos y registros de datos; (c) información que debe facilitarse; (d) solidez y exactitud; (e) supervisión humana; y (f) requisitos específicos en el caso de determinadas aplicaciones de IA, como las empleadas en la identificación biométrica remota.

También ha concretado una aproximación a las directrices éticas para una IA fiable. En concreto, aquí se consideran varios aspectos, comenzando por la necesidad de una IA lícita, ética y robusta, que se sostenga sobre cuatro principios éticos (Comisión Europea, 2019): respeto de la autonomía humana, prevención del daño, equidad y explicabilidad. Junto a ellos, operan siete requisitos clave que deben cumplir los sistemas basados en IA: (1) acción y supervisión humana; (2) solidez técnica y seguridad; (3) gestión de la privacidad de los datos; (4) transparencia; (5) diversidad, no discriminación y equidad; (6) bienestar social y ambiental; (7) rendición de cuentas. En este sentido, se pone el acento en la necesidad de integrar estos requisitos, que inciden en algunas de las cuestiones clave de la IA dentro de las administraciones públicas (ver Figura 1).

Figura 1. Directrices como marco para una IA fiable

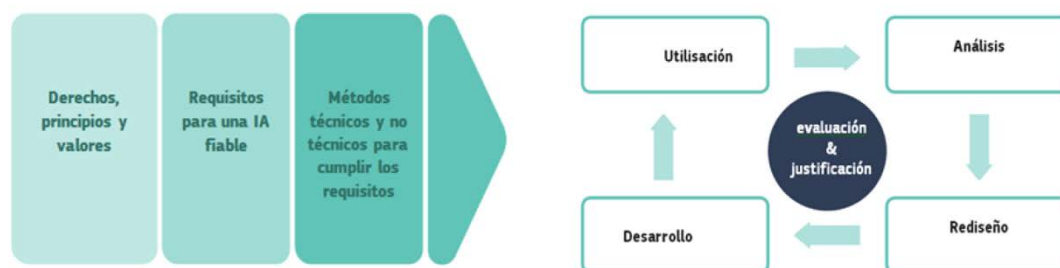


Fuente: Comisión Europea (2019, p.10).

Al mismo tiempo, también hay que insistir en que estos requisitos necesitan de una serie de métodos técnicos y de otro tipo para su puesta en marcha. Estos

métodos deberían abarcar todas las fases del ciclo de vida de un sistema de IA (ver Figura 2) (Comisión Europea, 2019). En el caso de los métodos técnicos, se consideran las arquitecturas para una IA fiable, la incorporación de la ética desde el diseño, los métodos de explicación, la realización de ensayos y validación de los algoritmos y sistemas de IA o los indicadores de calidad de servicio. Aquí se tiene la intención definir una lista de métodos de utilidad para lograr una IA fiable.

Figura 2. Esquema sobre métodos a aplicar en todo el ciclo de vida de los sistemas de IA



Fuente: Comisión Europea (2019, p. 26).

En lo que respecta a los métodos no técnicos, se estima que deberían evaluarse de una forma continua teniendo en cuenta la velocidad de desarrollo de los sistemas de IA. Particularmente, los principales métodos reseñados son (Comisión Europea, 2019): (1) la normativa de aplicación a la IA; (2) los códigos de conducta o documentos internos de política sobre cómo construir una IA fiable; (3) la certificación para acreditar ante el público que un sistema de IA es transparente, responsable y equitativo; (4) la rendición de cuentas a través de marcos de gobernanza tanto internos como externos a las organizaciones que se centren en los procesos de rendición de cuentas sobre las dimensiones éticas de las decisiones asociadas a los sistemas de IA; (5) educación y concienciación para fomentar una mentalidad ética entre la ciudadanía, incluyendo también a diseñadores, empresas y administraciones públicas y grupos afectados; (6) la participación de las partes interesadas y diálogo social; y (7) la diversidad y equipos de diseño inclusivos. Así, estos ámbitos tienen claramente un valor clave sobre el encuadre de las políticas y acciones de IA dentro de los gobiernos y administraciones públicas, dado que inciden directamente en su actuación, tanto en lo que respecta a las políticas y servicios públicos, como en la relación con la ciudadanía, así como en la gestión interna de las organizaciones.

En resumen, se puede decir que la UE ha establecido una ambiciosa agenda de IA, definiendo un marco de política pública, que a buen seguro va a tener una incidencia en los países miembros, aunque todavía está por ver cómo, cuándo y con qué alcance. Como se ha señalado, desde la retórica de las instituciones europeas parece que se ha entendido el desafío que supone la IA en el futuro, desde la noción de soberanía tecnológica y autonomía estratégica, hasta evitar los errores de otros países en contextos diferentes, cuyos modelos de desarrollo de IA y gobernanza de algoritmos parecieran contrarios a principios y derechos humanos básicos. Asimismo, también está por ver el alcance en los países miembros del marco de política pública sobre IA que se ha establecido desde la UE.

5. Hacia una política de Inteligencia Artificial en España

En el caso de España, también se han dado los primeros pasos orientados a la creación de una política de IA. Desde una perspectiva normativa, ya se vislumbró la incorporación de la actuación administrativa automatizada, con una lógica que se

acerca a la idea de adopción de la IA, en el conocido artículo 41 de la Ley 40/2015, de 1 de octubre, de Régimen Jurídico del Sector Público. En este artículo se habla de decisiones automatizadas ante «cualquier acto o actuación realizada íntegramente a través de medios electrónicos por una Administración Pública en el marco de un procedimiento administrativo y en la que no haya intervenido de forma directa un empleado público». Esta línea es la que después se ha desarrollado en el Reglamento General de Protección de Datos, tal y como han apuntado algunos autores, quienes han sostenido la necesidad de evitar cajas negras administrativas (Cerrillo, 2019), la promoción de sistemas de control de los algoritmos (Valero, 2019), o la protección de los derechos fundamentales a partir de un efectivo cumplimiento de los principios éticos (Cotino, 2019). En nuestro caso, nos centramos en conocer qué tipo de encuadre tiene la emergente política de IA del Gobierno de España, así como el alcance de la agenda desplegada hasta este momento.

Desde el Gobierno de España se concretó una primera aproximación a una política de IA en base a varios aspectos. El primer hito a reseñar es la Estrategia Española de I+D+I en Inteligencia Artificial (Ministerio de Ciencia, Innovación y Universidades, 2019) que se definió a partir de un Grupo de Trabajo en IA. Dentro de este documento se puso de manifiesto la existencia de diferentes estrategias nacionales de IA en el contexto europeo, así como la necesidad de que España se sumara a ese proceso. A partir de ello, se reconoció que el documento se adhería a los avances mencionados en el ámbito europeo. Lo anterior podría considerarse ya una muestra de europeización de las políticas de IA en España, en línea con lo acontecido en relación a otras olas pasadas de adopción de Gobierno Digital desde y en el ámbito público (Criado, 2009; 2010; 2012; 2021).

Además de este elemento, entre las prioridades se incluyó la necesidad de consolidar a las administraciones públicas como catalizadoras de la IA. Para ello, se indicaron áreas clave como la gestión de los datos de los individuos y sus análisis a través de aprendizaje automático o el desarrollo de *chatbots* basados en procesamiento de lenguaje natural como primera interfaz de interacción entre ciudadanos y administraciones públicas, focalizando a los empleados públicos en tareas de mayor valor (Ministerio de Ciencia, Innovación y Universidades, 2019). Además, se insistió en que dentro de los gobiernos locales hay tres áreas clave de interacción con el ciudadano (Ministerio de Ciencia, Innovación y Universidades, 2019:29): (a) transacciones de alto volumen y baja complejidad, como formularios; (b) alta complejidad y bajos volúmenes, como la asistencia social, y en el medio, (c) aquellos que tienen una complejidad media y volúmenes medios como la planificación de permisos y la recaudación de impuestos. Dentro de las prioridades, también se observó la preocupación por la ética de la IA y la mención a los ODS de Naciones Unidas como un paso para lograr un marco nacional de IA que se convierta en palanca transformadora de las políticas públicas.

Un segundo paso importante se produce a partir de la puesta en marcha del nuevo gobierno de coalición a partir de enero de 2020. Entonces se crea la Secretaría de Estado de Digitalización e Inteligencia Artificial (SEDIA), dentro del Ministerio de Asuntos Económicos y Transformación Digital (Real Decreto 2/2020). Por primera vez en España existe un órgano político de este nivel con la competencia directa de promover la IA en el país. En concreto, dentro de su marco competencial se han incluido algunas de sus principales objetivos, tales como (Real Decreto 403/2020 de desarrollo de la estructura orgánica del Ministerio de Asuntos Económicos y Transformación Digital, que también incluye la Secretaría General de Administración Digital): elaborar una estrategia española de IA; desarrollar las dimensiones normativa, ética y de regulación de la IA y sus tecnologías digitales habilitadoras; participar en comisiones, grupos de trabajo, etc., de carácter internacional sobre IA;

definición de programas de I+D+I en IA; impulsar de programas sobre IA en administraciones públicas y sector privado; o atender cuestiones de seguridad y normalización (entre otras).

Aquí hay que señalar algunos aspectos desarrollados durante la pandemia del COVID-19 que tienen relación con la IA. La crisis del COVID-19 llevó a la SEDIA a centrarse en el combate frente a la pandemia (Criado, Guevara-Gómez y Villodre, 2020; Meijer *et al.*, 2020). Pero también hay que señalar el hito de la creación de la primera Oficina del Dato dentro de la Administración General del Estado. En concreto, se justificó su creación sobre la base de que es necesario «impulsar nuevas capacidades y habilidades, el emprendimiento tecnológico y una inteligencia artificial inclusiva, ética, transparente, que permitan la igualdad de oportunidades y la creación de empleos» (Orden ETD/803/2020, de 31 de julio, por la que se crea la División Oficina del Dato). En este sentido, es clave la atención a las siguientes cuestiones: (1) el diseño de estrategias y marcos de referencia de gestión de datos; (2) la creación de espacios de compartición de datos entre empresas, ciudadanos y administraciones públicas de manera segura y con gobernanza (*sandboxes*, *data spaces* nacionales y europeos, ecosistemas de datos para uso sectorial tanto público como privado, etc.); (3) el empleo masivo de los datos en los sectores productivos de la economía mediante tecnologías de big data e IA; y (4) el desarrollo de mecanismos de acceso seguros a estas plataformas de datos, para la toma de decisiones públicas basadas en datos o para uso empresarial, garantizando sus seguridad y gobernanza. En suma, dentro de esta Oficina del Dato, junto a la gemela División de planificación y ejecución de programas, se espera que el encuadre pase a los siguientes pasos en la política española de IA.

Hasta el momento, el gran hito de política pública de IA en España se refiere a la adopción de la Estrategia Nacional de Inteligencia Artificial (ENIA), presentada el 2 de diciembre de 2020 en la Presidencia del Gobierno. Además de haber utilizado el mencionado trabajo del Ministerio de Ciencia, Innovación y Universidades (2019), la ENIA es el documento donde aparece la visión sobre el impulso de la IA en España, en directo alineamiento con la estrategia europea y la Agenda Digital 2025 de España. Aquí se plantea responder al hecho de que se espera que más del 20% de los fondos del mecanismo europeo de recuperación se centrarán en la transición hacia la digitalización (ENIA, 2020, pp. 12-13). Adicionalmente, se explicitan diferentes menciones a documentos de la Comisión Europea, incluyendo su Libro Blanco sobre IA. Este aspecto es esencial, dado que la ENIA se ha integrado desde el primer momento dentro del marco europeo sobre IA, al menos, formalmente.

Por otro lado, a día de hoy, sólo podemos señalar algunos rasgos adicionales sobre la ENIA, tal y como aparecen explicitados en el documento. De entrada, se detallan siete objetivos estratégicos a lograr en relación con la política de IA en España (ENIA, 2020): (1) excelencia científica e innovación en IA; (2) proyección de la lengua española; (3) creación de empleo cualificado; (4) transformación del tejido productivo; (5) entorno de confianza en relación a la IA; (6) valores humanistas en la IA; y (7) IA inclusiva y sostenible. Los objetivos definidos se han planteado siguiendo el trabajo mencionado antes de la OCDE, a la vez que se combina la dimensión más económica y tecnológica de la IA, con su vertiente más social y ética.

Sin embargo, los aspectos más destacados en relación con la agenda pública en España sobre IA aparecen explicitados como ejes de acción. De entrada, se señala el período 2020-2025 como horizonte y se establecen seis ejes de acción con una serie de líneas de actuación e hitos temporales específicos. Los ejes estratégicos son los siguientes (ENIA, 2020): (a) impulsar la investigación científica, el desarrollo tecnológico y la innovación en IA; (b) promover el desarrollo de capacidades digitales,

potenciar el talento nacional y atraer talento global; (c) desarrollar plataformas de datos e infraestructuras tecnológicas que den soporte a la IA; (d) integrar la IA en las cadenas de valor para transformar el tejido económico; (e) potenciar el uso de la IA en la Administración Pública y en las misiones estratégicas nacionales; y (f) establecer un marco ético y normativo que refuerce la protección de los derechos individuales y colectivos, a efectos de garantizar la inclusión y el bienestar social. A primera vista, aquí también se trata de desarrollar una política de IA *desde* gobiernos y administraciones públicas y *en* gobiernos y administraciones públicas.

Por un lado, se pueden plantear algunas ideas básicas sobre la dimensión de la política de IA en relación a las cuestiones que afectan a la economía y la sociedad. Sin ánimo de ser exhaustivos, y sin disponer de evidencias más allá del análisis documental, la ENIA (2020) destila claramente inquietudes más allá de una mera aproximación tecnológica o económica. En concreto, se desarrolla un bloque completo sobre ética y normativa relacionada con la IA (eje estratégico 6), con medidas previstas como el desarrollo de un sello nacional de calidad IA, la puesta en marcha de observatorios para evaluar el impacto social de los algoritmos, el desarrollo de una carta de derechos digitales (cuya consulta pública se cierra el 20 de enero de 2021), la puesta en funcionamiento del consejo asesor de IA (creado mediante la Orden ETD/670/2020 de 8 de julio) o la promoción de diálogo, sensibilización y participación sobre IA. Sin embargo, a pesar de que se plantea al final de la ENIA un apartado sobre gobernanza y presupuesto, no se detallan cuestiones esenciales, incluyendo un mapa de actores e impulsores clave, indicadores de medida de avance de las acciones, impactos esperados de las intervenciones públicas, un modelo más específico sobre cómo se van a articular las relaciones intergubernamentales para integrar a gobiernos autonómicos y locales, mecanismos de colaboración público-privada, un cronograma de proyectos o qué financiación cuentan las acciones específicas. Es posible que estas dimensiones se concreten más adelante.

Por otro lado, aquí nos interesan las acciones previstas en la política de IA que tienen que ver con la transformación digital de las administraciones públicas. En este sentido, la ENIA indica cinco medidas (2020, p. 73): (1) incorporar la IA en la administración pública para mejorar la eficiencia y eliminar cuellos de botella administrativos; (2) poner en marcha un laboratorio de innovación para nuevos servicios y aplicaciones de la IA en la administración pública (*gobtechlab*); (3) fomentar las competencias IA en la administración pública; (4) programa IA para una gestión pública basada en datos; y (5) promover misiones estratégicas nacionales en el ámbito de la administración pública donde la IA puede tener impacto (foco en salud, justicia y empleo). Aquí tampoco podemos descender a un mayor detalle, si bien es cierto que en el documento se incide en la importancia de la IA para mejorar la transparencia administrativa, adoptar mejores decisiones, desarrollar una gestión pública basada en datos, incorporar nuevos mecanismos de compra pública innovadora, abrir mecanismos de regulación basados en *sandboxes* o prestar servicios digitales basados en sistemas de IA. También aquí se esperan ulteriores detalles sobre actores, gobernanza, indicadores de medida, presupuesto, cronogramas de implementación o mecanismos de revisión y evaluación.

En suma, la política española de IA, en el momento de completarse este análisis (mediados de enero de 2021) se encuentra todavía en la fase de formulación, sin embargo queda todavía por concretar detalles relacionados con la gobernanza del proceso de implementación y evaluación. En todo caso, y sin entrar en cuestiones que exceden el análisis de este documento, se podría decir que la IA se ha integrado en la agenda mediática, al menos en su apartado económico, no queda tan claro que suceda lo mismo con la dimensión político-administrativa. En este último ámbito los cambios asociados a la implantación de la IA y la gobernanza con algoritmos todavía

se perciben como algo que todavía está por llegar, si bien resultarán necesarios para lograr una verdadera transformación digital del sector público.

6. La perspectiva de los CIOs ministeriales españoles

Junto al encuadre que se ha planteado en las anteriores secciones desde una perspectiva conceptual y más general, nuestro trabajo también asume que es importante la visión que tienen los responsables de la política de tecnologías en los organismos públicos (*Chief Information Officers* o *CIOs*), como responsables de su definición y desarrollo. De cara a sumar evidencia más allá del análisis documental planteado, utilizamos parte de los resultados de un cuestionario desarrollado entre el 15 de junio y 15 de agosto de 2019 a los entonces responsables de las áreas TIC de todos los ministerios españoles (otros resultados más desarrollados han aparecido en Criado *et al.*, 2021). Particularmente, en esta sección de análisis nos centramos, por un lado, en su visión sobre el concepto IA, y por otro, en los beneficios y desventajas de su implantación en el sector público.

En primer lugar, nuestro estudio examina las percepciones de los CIOs sobre el concepto de IA que, como tecnología emergente, no se percibe de la misma manera, ni se adopta e implanta por igual. En este momento, no todos los miembros del mundo académico y personal directivo del sector público comparten el mismo enfoque sobre lo que es la IA. Además, este concepto se halla asociado a diferentes ideas y técnicas. En este sentido, las percepciones de los CIOs ministeriales ayudan a reconocer el enfoque y potencial de la IA en el sector público, ya que estos altos directivos públicos definen las políticas tecnológicas. Así, esta parte del estudio se centra en arrojar luz sobre los conceptos relacionados, así como las técnicas más comúnmente vinculadas a la IA por parte de los CIOs ministeriales.

Nuestro enfoque sobre la IA surge de una pregunta abierta para definir este concepto. La pregunta del cuestionario para los CIOs fue: «Desde su punto de vista, y brevemente, ¿cómo se puede definir Inteligencia Artificial?». Las respuestas a la pregunta del cuestionario fueron completamente abiertas. Los resultados se resumen en la Figura 3, con una nube de palabras (las palabras de tamaño superior se repitieron más en las definiciones dadas). Los datos confirman que los CIOs caracterizan principalmente la IA con la idea de tecnologías, sistemas o series de algoritmos inteligentes que ayudan en el proceso de toma de decisiones, pero también otras tareas que incluyen aprendizaje, razonamiento, etc. Estos directivos identifican la IA con nuevas capacidades y oportunidades para realizar tareas específicas.

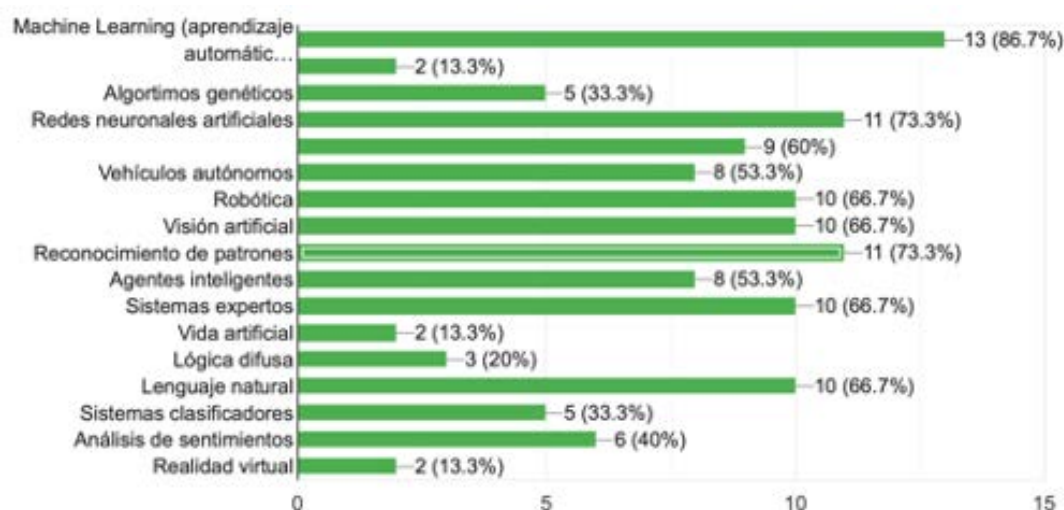
Figura 3. Definición abierta de Inteligencia Artificial



Fuente: Elaboración propia.

Otra categoría denota las técnicas más asociadas a la IA en el sector público. Aquí, las percepciones de los CIOs son críticas, ya que las políticas y estrategias nacionales tenderán a estar más o menos orientadas a determinadas técnicas, en la medida que sean más o menos familiares y se asocian más o menos con la IA por parte de estos actores directivos. Para comprender esta dimensión, la pregunta del cuestionario fue: «Entre los siguientes términos, señale cuáles tienen, a su juicio, mayor asociación con la Inteligencia Artificial». Las respuestas del cuestionario se plantearon a partir de una pregunta múltiple-respuesta (con un mínimo de tres opciones para responder del listado de opciones). Los resultados se resumen en la Figura 4, incluido el porcentaje de aparición de cada concepto/técnica. Las cinco técnicas más asociadas a la IA son: aprendizaje automático (86,7%), redes neuronales artificiales (73,3%), reconocimiento de patrones (73,3%), aprendizaje profundo (60,0%) y agentes inteligentes (53,3%). Este grupo de técnicas y sistemas de IA probablemente liderará la futura implementación de proyectos y aplicaciones en el sector público español, enmarcando la naturaleza y alcance de estas iniciativas.

Figura 4. Técnicas y conceptos asociados a la Inteligencia Artificial



Fuente: Elaboración propia.

Por otro lado, las expectativas sociales y políticas sobre la IA son altas en términos de logros y beneficios, pero también de posibles dificultades y problemas. Esta dimensión comprende las expectativas de los CIOs sobre las capacidades de transformación de la IA en el sector público. En particular, esta parte del estudio estuvo orientada a categorizar los beneficios y problemas esperados de la implementación de la IA en el sector público. Con ello se completa esta aproximación al encuadre que realizan los CIOs ministeriales en el caso de España.

Para examinar los beneficios esperados de la implementación de la IA en el sector público, el cuestionario abordó diferentes dimensiones de las organizaciones del sector público. La pregunta del cuestionario fue: «¿Qué beneficios considera que tienen los siguientes aspectos en relación con la implementación de la Inteligencia Artificial en los siguientes aspectos del sector público?». Las respuestas a la pregunta de la encuesta fueron en una escala *Likert*, desde «beneficio total» (7) a «ningún beneficio en absoluto» (1). En la Tabla 3, se identifican las ideas más destacadas de los CIOs españoles, en general, no demasiado optimistas, ya que identifican potenciales altos beneficios en aspectos como: eficiencia (6,7) y digitalización (5,7) (solo dos ítems por encima de 5 puntos en la escala). En el otro lado del espectro, los CIOs sugirieron menos beneficios esperados en temas centrados en los ciudadanos

y más políticos/democráticos, incluida la confianza (3.7), la equidad (4.0), la transparencia (4.1) o la seguridad de los datos personales (4.3).

Tabla 3. Beneficios de la aplicación de la Inteligencia Artificial

	CIOs ministeriales (n=15)	
	Media	Desviación típica
Eficiencia	6.7	0.5
Digitalización	5.7	1.3
Transparencia	4.1	2.0
Seguridad de los datos personales	4.3	1.9
Participación ciudadana	4.5	1.6
Interoperabilidad	4.1	2.1
Equidad	4.0	1.6
Rendición de cuentas	4.5	1.5
Confianza	3.7	1.8
Flexibilidad	4.6	1.6
Media	4.6	1.6

Fuente: Elaboración propia.

Con respecto a las desventajas esperadas de la aplicación de la IA en el sector público, el cuestionario también abordó dimensiones clave de las organizaciones del sector público. Aquí, la pregunta del cuestionario fue: «Desde su punto de vista, ¿qué desventajas espera de la implantación de la Inteligencia Artificial en los siguientes aspectos del sector público?». Las respuestas a la pregunta de la encuesta fueron en una escala *Likert*, desde «desventaja total» (7) a «ninguna desventaja en absoluto» (1). Como se esperaba de la pregunta anterior (la Tabla 4 resume los resultados), los CIOs ministeriales asumen altas desventajas en aspectos como opacidad algorítmica (5.3), cuestiones éticas (5.1) y desconfianza (5.0), (en todos los casos por encima de 5 puntos en la escala). Sin duda, esta dimensión conlleva aspectos humanos de las tecnologías emergentes, en general, y de la IA, en particular.

Tabla 4. Desventajas de la aplicación de la Inteligencia Artificial

	CIOs ministeriales (n=15)	
	Media	Media
Reemplazo humano	4.9	1.8
Costos altos	4.1	1.9
Deshumanización	4.1	1.8
Desconfianza	5.0	1.8
Inflexibilidad	4.0	1.6

Vulnerabilidad de datos e información	3.9	1.8
Problemas éticos	5.1	1.8
Exclusión social	3.9	1.8
Desigualdad	3.5	1.6
Control invasivo	4.2	1.7
Opacidad algorítmica	5.3	1.5
Media	4.4	1.7

Fuente: Elaboración propia.

7. Debate y conclusiones

En esta sección completamos este trabajo resumiendo los principales hallazgos del estudio, estableciendo algunas conclusiones y detallando algunas vías de evolución futura en torno a los algoritmos y la Inteligencia Artificial (IA) en el sector público. En primer lugar, es necesario entender los aspectos básicos de las nociones de IA y gobernanza con algoritmos y cómo afectan al sector público. Si bien el concepto de IA no es unívoco, es clave la asociación entre inteligencia no humana, racionalidad, desarrollo de capacidades (para adoptar decisiones, acometer acciones, obtener resultados, etc.), así como diferentes niveles de madurez en su implantación en las organizaciones, también públicas. Los algoritmos, en la medida que actúan como reglas incrustadas en diferentes actividades humanas, determinan de una forma creciente muchos de nuestras actividades y comportamientos diarios (qué vemos en televisión, qué noticias leemos, con quiénes nos vinculamos en redes sociales, etc.), incluyendo cómo nos relacionamos con las administraciones públicas. Las experiencias algorítmicamente determinadas y ligadas a sistemas de IA son cada vez más en cantidad y más determinantes en términos de experiencias y comportamientos.

A partir de esa realidad, es claro que las organizaciones internacionales están marcando una serie de principios éticos y cuestiones legales para tener en cuenta. Por una parte, la Organización para la Cooperación y el Desarrollo Económico (OCDE) ha elaborado una recomendación legal que se une a un intento de generar conocimiento y liderar el diálogo sobre la dimensión ética de la IA, también desde gobiernos y administraciones, si bien su aplicación en el sector público todavía es emergente. Naciones Unidas ha hecho suyo el lema «IA para el bien» (*IA for good*), como una declaración de intenciones sobre el potencial de la IA, así como sus desafíos. También invita a desarrollar proyectos tecnológicos centrados en los Objetivos de Desarrollo Sostenible (ODS) y ha incluido la IA como uno de los retos de futuro del Gobierno Digital. En esa línea, la UNESCO ha establecido un marco sólido con los principales problemas y recomendaciones sobre la ética en la IA (incluyendo derechos humanos, inclusión, bienestar, autonomía, explicabilidad, transparencia, responsabilidad, rendición de cuentas, sostenibilidad, buen gobierno, etc.). La UNESCO propone claramente el uso de la IA para la prestación de servicios públicos, de manera que los gobiernos y administraciones públicas deben centrarse en las implicaciones de la IA en los procesos de adopción de decisiones, tanto en términos de transparencia como garantía de derechos humanos. Esta perspectiva se asemeja mucho a los planteamientos realizados por el Consejo de Europa (CdE) que, sobre todo, incide en la necesidad de que el uso de la IA se acerque a un bloque básico de derechos humanos. En definitiva, el componente ético y normativo es abundante dentro de las aproximaciones de estos organismos internacionales.

Por su lado, la Unión Europea (UE) ha consolidado un marco de referencia para el diseño e implementación de la IA en los países miembros, algo que ya se evidencia en casos como el español. Por un lado, hay que subrayar el papel catalizador de la UE y la intención de europeización desde las instituciones comunitarias de las políticas de IA dentro de los países miembros de cara a conformar un modelo propio. Quizá sea esta la principal conclusión en el caso europeo, la evidencia de que las instituciones comunitarias han pretendido desarrollar un modelo propio, en contraste con el chino y estadounidense, que permita a la UE presentarse como una tercera vía referente en IA. En concreto, ese modelo se sostendría, al menos, en los siguientes pilares: la necesidad de atender los derechos humanos y la protección de datos personales de una manera rigurosa (al menos, en línea con la Carta de los Derechos Fundamentales de la UE y el Reglamento General de Protección de Datos). A ello habría que añadir los principios de excelencia y confianza. Para todo lo anterior, ya se han concretado directrices centradas en la creación de una IA (con)fiable, incluyendo requisitos, métodos técnicos y no técnicos de aplicación (más arriba identificados). En definitiva, el Libro Blanco de la UE sugiere el enfoque que los países miembros deberían dar a sus políticas nacionales, y aunque no es imperativo su seguimiento, aquí operan los criterios de europeización que a buen seguro irán generando resultados de interés para el futuro.

El caso de España se convierte en reflejo de lo anterior. Primero hay que indicar que España es un país donde opera un alto grado de europeización de sus políticas públicas. Lo anterior viene a confirmar nuestro análisis documental, que ha evidenciado que la estrategia sobre IA en el caso español se adhiere plenamente a los planteamientos de la UE, en línea con lo acontecido y en relación con otras olas pasadas de adopción de tecnología desde y en el ámbito público (Criado, 2009; 2012; 2021). A nivel organizativo, destaca la creación de una Secretaría de Estado de Digitalización e Inteligencia Artificial, dentro del Ministerio de Asuntos Económicos y Transformación Digital. Este hecho es clave, no solo por ser la primera vez que en España existe un órgano político de este nivel con la competencia directa de desarrollar la IA. También lo es por el hecho de conectar la IA desde y en gobiernos y administraciones públicas. De hecho, será esencial la colaboración con la Secretaria General de Administración Digital, tal y como se está mostrando con algunas de las medidas adoptadas durante la crisis del COVID-19 (incluyendo la *app* de trazado y rastreo Radar COVID) o la creación de la Oficina del Dato y sus proyectos asociados. La publicación de la Estrategia Nacional de Inteligencia Artificial (ENIA) y su directo alineamiento con el marco europeo de IA, junto con algunas de las medidas ya puestas en marcha (carta de derechos digitales o la creación de un consejo asesor de IA) y otras que se mencionan (*sandboxes*, *gobtechlab*, ecosistemas de datos para uso sectorial tanto público como privado, etc.) van en buena dirección. Sin embargo, será necesario algo más de tiempo para caracterizar la política de IA en España, su alcance y, desde luego, los resultados de la misma.

Finalmente, también es crítico conocer la visión de los actores impulsores, sobre todo, los responsables TIC ministeriales (CIOs) españoles. Como responsables de la definición e implantación de las políticas tecnológicas dentro de sus organizaciones, disponer de evidencia sobre sus percepciones es esencial para comprender las potenciales limitaciones y oportunidades de la IA en las administraciones públicas españolas. En la primera dimensión sobre las percepciones acerca del concepto de IA en el sector público han mostrado varias ideas de interés, ya que comparten creencias comunes sobre cuál es la noción de IA (tecnologías, sistemas o series de algoritmos inteligentes que ayudan en el proceso de toma de decisiones, pero también otras tareas como aprendizaje, razonamiento, etc.). Asocian la IA con una serie de técnicas (sobre todo, aprendizaje automático, redes neuronales artificiales, aprendizaje profundo, reconocimiento de patrones y agentes inteligentes).

Por tanto, estos resultados confirman la existencia de visiones comunes sobre el concepto de IA. Además, respaldan nuestra hipótesis sobre los CIOs, que asignan una naturaleza tecno-céntrica a la IA en el sector público, dejando de lado los enfoques centrados en el ser humano/social, así como cuestiones éticas.

Nuestra segunda dimensión ha analizado las expectativas de los CIO sobre el desarrollo de la IA en el sector público en el futuro. Considerando los beneficios esperados, los CIOs españoles no son muy optimistas en relación con cuestiones como interoperabilidad, la rendición de cuentas, la transparencia, la digitalización, la seguridad de los datos personales o la flexibilidad, centrándose en beneficios como la eficiencia y la digitalización. Atendiendo las desventajas esperadas, están muy preocupados por la opacidad algorítmica, las cuestiones éticas y la desconfianza. En consecuencia, estos resultados corroboran diferentes expectativas sobre el potencial de la IA en el sector público. En concreto, esta dimensión también confirma nuestro supuesto de que los CIOs esperan resultados principalmente positivos (hiperpositivismo), más orientados a fomentar las operaciones de gestión pública (enfoque en servicios), que otras áreas probablemente más cruciales (de carácter político) en la actividad de nuestros sistemas de gobierno.

A partir de aquí, las vías de trabajo futuro son diversas. Primero, es necesario tener en cuenta el carácter multidisciplinar de los algoritmos y la IA, una vez que se generaliza su adopción y uso. Entonces, es crítico incorporar en su diseño y aplicación miradas diversas procedentes de las ciencias sociales, conductuales y económicas, pero también del mundo jurídico y ético, para construir instrumentos que beneficien al conjunto de la sociedad, y que interroguen sobre sus implicaciones reales en la vida de las personas y las organizaciones. Adicionalmente, cada vez es más importante disponer de evidencias sobre los efectos que están provocando los modelos algorítmicos y sistemas de IA en las organizaciones y personas que interaccionan con ellos. Más allá de la reflexión teórica se requieren evidencias que nos permitan encaminar más certeramente los siguientes pasos de la IA en nuestras comunidades políticas. Finalmente, las administraciones públicas se han convertido en actores clave en el desarrollo de la IA en nuestras sociedades, no solo porque se encargan de definir agendas nacionales y catalizar acciones y recursos, sino también como ámbitos de aplicación de tecnología. Por un lado, la generalización de estrategias y agendas nacionales evidencia el interés por liderar el proceso, así como definir modelos propios, tal y como sucede en el entorno de la UE. Por otro lado, casos como el español permitirán comprobar en la práctica hasta dónde llegan las promesas de la IA para transformar digitalmente el sector público. Esta próxima década será clave en el desarrollo de la IA y la gobernanza de los algoritmos en el futuro de las administraciones públicas.

Bibliografía

- Agarwal, P. K. (2018). Public Administration Challenges in the World of AI and Bots. *Public Administration Review* 78 (6), pp. 917-921. DOI: <https://doi.org/10.1111/puar.12979>.
- Brassard, G., & Bratley, P. (1997). *Fundamentos de algoritmia*. Madrid, España: Prentice Hall.
- Brynjolfsson, E., & McAfee, A. (2017). The business of artificial intelligence. *Harvard Business Review* 7, pp. 3-11.
- Cerillo, A. (2019). Com obrir les caixes negres de les Administracions públiques? Transparència i rendició de comptes en l'ús dels algoritmes (CA-EN). *Revista Catalana de Dret Públic*, (58), pp. 13-28. DOI: <http://dx.doi.org/10.2436/rcdp.i58.2019.3277>.

- Comisión Europea. (2020). Libro Blanco sobre la Inteligencia Artificial. Un Enfoque Europeo Orientado a la Excelencia y la Confianza. Recuperado de: https://ec.europa.eu/info/sites/info/files/commission-white-paper-artificial-intelligence-feb2020_es.pdf.
- Comisión Europea – Joint Research Center. (2020). AI Watch. Artificial Intelligence in Public Services. Overview of the Use and Impact of Artificial Intelligence in Public Services in the European Union. Recuperado de: https://publications.jrc.ec.europa.eu/repository/bitstream/JRC120399/jrc120399_misuraca-ai-watch_public-services_30062020_def.pdf.
- Comisión Europea. (2019). A Definition of AI. Main Capabilities and Disciplines. Recuperado de: <https://ec.europa.eu/digital-single-market/en/news/definition-artificial-intelligence-main-capabilities-and-scientific-disciplines>.
- Comisión Europea. (2018a). Artificial Intelligence. A European Perspective. Recuperado de: <https://ec.europa.eu/jrc/en/publication/artificial-intelligence-european-perspective>.
- Comisión Europea. (2018b). Artificial Intelligence for Europe. Recuperado de: <https://ec.europa.eu/digital-single-market/en/news/communication-artificial-intelligence-europe>.
- Consejo de Europa. (2019). Unboxing Artificial Intelligence. 10 steps to protect human rights. Recuperado de: <https://rm.coe.int/unboxing-artificial-intelligence-10-steps-to-protect-human-rights-reco/1680946e64>.
- Cormen, T. H., Leiserson, C. E., Rivest, R. L., & Stein, C. (2009). *Introduction to algorithms*. Cambridge, Estados Unidos: MIT Press.
- Cotino, L. (2019). Ética en el Diseño para el Desarrollo de una Inteligencia Artificial, Robótica y Big Data Confiable y su Utilidad desde el Derecho. *Revista Catalana de Dret Públic*, (58), pp. 29-48. DOI: <https://doi.org/10.2436/rcdp.i58.2019.3303>.
- Criado, J. I. (2021). La Política de Administración Digital en España. De los Servicios Públicos Digitales a la Gobernanza Inteligente y Administración Pública 4.0. En Ramió, C. (Ed.). *Repensando la Administración Digital y la Innovación Pública* (pp. 71-108). Madrid, España: Instituto Nacional de Administración Pública.
- Criado, J. I. (2019). Inteligencia Artificial y Administración Pública. Incorporando Nuevas Capacidades para la Innovación Pública Pública. *XXIV Congreso Internacional del CLAD*, Buenos Aires, 12-15 noviembre de 2019.
- Criado, J. I. (2012). Interoperability of e Government for Building Intergovernmental Integration in the European Union. *Social Science Computer Review*, 30 (1), pp. 37-60. DOI: <https://doi.org/10.1177/0894439310392189>.
- Criado, J. I. (2010). The Politics of E-government in Spain: Between Recent Innovations and Old Inertias. En Reddick, Ch. (Ed.) *Comparative E-government* (pp. 275-297). New York, Estados Unidos: Springer.
- Criado, J. I. (2009). *Entre Sueños Utópicos y Visiones Pesimistas. Internet y las TIC en la Modernización de las Administraciones Públicas*. Madrid, España: Instituto Nacional de Administración Pública.
- Criado, J. I., & Gil-García, J. R. (2019). Creating public value through smart technologies and strategies: from digital services to artificial intelligence and beyond. *International Journal of Public Sector Management*, 32 (5), pp. 438-450. DOI: <https://doi.org/10.1108/IJPSM-07-2019-0178>.
- Criado, J. I., Sandoval-Almazán, R., Valle-Cruz, D., & Ruvalcaba-Gomez, E. A. (2021). Chief Information Officers and Artificial Intelligence in Government. Implications and Challenges for the Public Sector. *First Monday*, 26(1). DOI: <https://doi.org/10.5210/fm.v26i1.10648>.
- Criado, J. I., Valero, J. & Villodre, J. (2020). Algorithmic transparency and bureaucratic discretion: The case of SALER early warning system. *Information Polity*, 25(4), pp. 449-470. DOI: 10.3233/IP-200260.

- Desouza, K. (2018). *Delivering Artificial Intelligence in Government*. IBM for the Business of Government Report. Recuperado de: <http://www.businessofgovernment.org/sites/default/files/Delivering%20Artificial%20Intelligence%20in%20Government.pdf>.
- Dutton, T. (2018). An Overview of National Artificial Intelligence Strategies. Recuperado de: <https://medium.com/politics-ai/an-overview-of-national-ai-strategies-2a70ec6edfd>.
- Ganapati, S., & Reddick, C. G. (2012). Open e-government in U.S. state governments: Survey evidence from Chief Information Officers. *Government Information Quarterly*, 29(2), pp. 115–122. DOI: <https://doi.org/10.1016/j.giq.2011.09.006>.
- Gobierno de Colombia. (2019). Política Nacional para la Transformación Digital e Inteligencia Artificial. Documento CONPES 3975. Recuperado de: <https://www.mintic.gov.co/portal/inicio/Sala-de-Prensa/Noticias/107147:CONPES-de-transformacion-digital-promovera-la-competitividad-del-pais-y-la-eficiencia-del-sector-publico>.
- Gobierno de Estados Unidos de América. (2020). American Artificial Intelligence Initiative. Year One Annual Report. Recuperado de: <https://www.whitehouse.gov/wp-content/uploads/2020/02/American-AI-Initiative-One-Year-Annual-Report.pdf>.
- Gobierno de Portugal. (2020). AI Portugal 2030. Recuperado de: <https://www.portugal.gov.pt/download-ficheiros/ficheiro.aspx?v=236848b1-fcb6-4c65-9773-292d1c5b9ad1>.
- Gobierno de Uruguay. (2019). Estrategia de Inteligencia Artificial para el Gobierno Digital. Recuperado de: <https://www.gub.uy/agencia-gobierno-electronico-sociedad-informacion-conocimiento/sites/agencia-gobierno-electronico-sociedad-informacionconocimiento/files/documentos/publicaciones/Estrategia%20IA%20-%20versión%20español.pdf>.
- Introna, L. D. (2015). Algorithms, governance, and governmentality: On governing academic writing. *Science, Technology & Human Values*, 41(1), pp. 17–49. DOI: <https://doi.org/10.1177/0162243915587360>.
- Janssen, M., & Kuk, G. (2016). The challenges and limits of big data algorithms in technocratic governance. *Government Information Quarterly*, 33(3), pp. 371–377. DOI: 10.1016/j.giq.2016.08.011.
- LaValle, S. M. (2006). *Planning algorithms*. Cambridge, Estados Unidos: University Press.
- Lazer, D. (2015). The rise of the social algorithm. Does content curation by Facebook introduce ideological bias? *Science*, 348(6239), pp. 1090–1091. DOI: 10.1126/science.aab1422.
- Lee, K.F. (2018). *AI Super-powers. China, Silicon Valley, and the New World Order*. New York, Estados Unidos: Houghton Mifflin Harcourt.
- Margetts, H. & Dorobantu, C. (2019). Rethink government with AI. *Nature* 568 (April), pp. 163-165.
- Meijer, A. (2017). Datapolis: A Public Governance Perspective on “Smart Cities” *Perspectives on Public Management and Governance*, 1(3), pp. 195-203. DOI: 10.1093/ppmgov/gvx017.
- Ministerio de Ciencia, Innovación y Universidades (2019). Estrategia Española de I+D+I en Inteligencia Artificial. Recuperado de: <http://www.ciencia.gob.es/portal/site/MICINN/menuitem.26172fcf4eb029fa6ec7da6901432ea0/?vgnnextoid=70fcdb77ec929610VgnVCM1000001d04140aRCRD>.
- Naciones Unidas. (2020). *United Nations eGovernment Index*. UNDESA. Recuperado de: <https://www.un.org/development/desa/publications/publication/2020-united-nations-e-government-survey>.
- OCDE. (2020). *Legal recommendation 0449*, 22 de mayo de 2019. Recuperado de: <https://legalinstruments.oecd.org/en/instruments/OECD-LEGAL-0449>.

- Russell, S., & Norvig, P. (2016). *Artificial Intelligence: A Modern Approach*. Englewood Cliffs, NJ, Estados Unidos: Pearson Higher Ed.
- Sun, T. Q., & Medaglia, R. (2019). Mapping the challenges of Artificial Intelligence in the public sector: Evidence from public healthcare. *Government Information Quarterly*, 36(2), pp. 368–383. DOI: 10.1016/j.giq.2018.09.008.
- UNESCO. (2020). Towards a Draft Text of Recommendation on the Ethics of Artificial Intelligence. Recuperado de: https://unesdoc.unesco.org/in/documentViewer.xhtml?v=2.1.196&id=p::usmarcd ef_0000373199&file=/in/rest/annotationSVC/DownloadWatermarkedAttachment/attach_import_5db2631f-9f2f-49a5-b3a5-5709fde182e8%3F_%3D373199eng.pdf&locale=es&multi=true&ark=/ark:/48223/pf0000373199/PDF/373199eng.pdf#%5B%7B%22num%22%3A35%2C%22gen%22%3A0%7D%2C%7B%22name%22%3A%22XYZ%22%7D%2C68%2C285%2C0%5D.
- Valero Torrijos, J. (2019). Las garantías jurídicas de la inteligencia artificial en la actividad administrativa desde la perspectiva de la buena administración. *Revista Catalana de Dret Públic*, 58, pp. 82-96. DOI: <https://doi.org/10.2436/rcdp.i58.2019.3307>.
- Valle-Cruz, D., Criado, J. I., Sandoval-Almazán, R., & Ruvalcaba-Gomez, E. A. (2020). Assessing the public policy-cycle framework in the age of artificial intelligence: From agenda-setting to policy evaluation. *Government Information Quarterly*, 37(4). DOI: <https://doi.org/10.1016/j.giq.2020.101509>.
- Wirtz, B. W., J. C. Weyerer, and C. Geyer. (2018). “Artificial Intelligence and the Public Sector. Applications and Challenges”. *International Journal of Public Administration*, 13 (7), pp. 1-20. DOI: <https://doi.org/10.1080/01900692.2018.1498103>.