

La transparencia de las administraciones públicas en la era de la inteligencia artificial: conceptos y claves de actuación

Transparency in public administrations in the age of artificial intelligence: concepts and key actions

Antonia Ferrer-Sapena; Elena Sánchez-Arnau

Ferrer-Sapena, Antonia; Sánchez-Arnau, Elena (2025). "La transparencia de las administraciones públicas en la era de la inteligencia artificial: conceptos y claves de actuación". *Anuario ThinkEPI*, v. 19, e19a23.

<https://doi.org/10.3145/thinkepi.2025.e19a23>

Publicado en *IweTel* el 4 de noviembre de 2025

Antonia Ferrer-Sapena

<https://orcid.org/0000-0001-6432-917X>
<https://www.directorioexit.info/ficha706>
Universitat Politècnica de València
anfersa@har.upv.es



Elena Sánchez-Arnau

<https://orcid.org/0000-0002-2126-5329>
Universitat Politècnica de València
elenasanchez04@gmail.com

Resumen: La irrupción de la inteligencia artificial (IA) en todos los ámbitos de la actividad humana hace necesario redefinir el concepto de transparencia en las administraciones públicas. La opacidad de los algoritmos puede introducir y amplificar sesgos cognitivos, éticos y culturales, influyendo en la información, las fuentes y los resultados, pero también en las decisiones automáticas basadas en herramientas de IA que se toman desde las administraciones. Esto afecta a una gran diversidad de instrumentos utilizados por las instituciones, desde algoritmos de clasificación para la asignación de ayudas públicas hasta herramientas

para el diagnóstico médico. Para garantizar la transparencia de su funcionamiento, el Estado debe asumir la tarea de analizar, catalogar, clasificar y controlar los algoritmos que utiliza, y poner esa información a disposición de los ciudadanos, con la dificultad añadida de que, en muchos casos, la comprensión de estos sistemas requiere conocimientos técnicos con los que la población general no cuenta. En este trabajo se expone esta problemática, junto con algunas claves de actuación y casos de éxito.

Palabras clave: Inteligencia Artificial; Transparencia; Algoritmo; Datos abiertos; Sesgo.

Abstract: The emergence of artificial intelligence (AI) in all areas of human activity necessitates a redefinition of the concept of transparency in public administrations. The opacity of algorithms can introduce and amplify cognitive, ethical, and cultural biases, influencing information, sources, and results, as well as automated decisions based on AI tools made by public administrations. This affects a wide variety of instruments used by institutions, from classification algorithms for allocating public aid to tools for medical diagnosis. To guarantee the transparency of their operation, the State must undertake the task of analyzing, cataloging, classifying, and controlling the algorithms it uses, and making this information available.

lable to citizens. This is further complicated by the fact that, in many cases, understanding these systems requires technical knowledge that the general public lacks. This paper presents this problem, along with some key action points and success stories.

Keywords: Artificial Intelligence; Transparency; Algorithm; Open Data; Bias.

Introducción

Entre los muchos temas que se ven afectados actualmente por la inteligencia artificial (IA), los relacionados con la transparencia institucional, la representación ciudadana y el funcionamiento de la administración resultan de especial importancia, dada la trascendencia social que necesariamente tienen. Esto es así porque el acceso a la información que facilitan las herramientas de IA está profundamente marcado por dos características que las diferencian de los procedimientos tradicionales: la opacidad de sus fuentes (aunque aporten respuestas cuando se les requiere información sobre ellas) y la falta de mecanismos que garanticen su fiabilidad. La cuestión no es menor, ya que las consecuencias de esta carencia de información no se limitan al ámbito académico o científico, sino que afectan a todos los ciudadanos.

Aunque al hablar de IA solemos pensar en los instrumentos a los que accedemos a diario con nuestros ordenadores y que utilizamos tanto para recopilar información sobre los más diversos temas como para redactar correos o documentos administrativos, el uso de estas herramientas afecta también a numerosos ámbitos del funcionamiento de las administraciones. Por ejemplo, la asignación de ayudas públicas, la sanidad o ciertos procesos administrativos relacionados con la educación son gestionados cada vez más con el apoyo de estos instrumentos, e incluso algunas tareas son realizadas directamente por herramientas de IA. Podemos pensar, por ejemplo, en los algoritmos de tratamiento de imágenes que utilizan las máquinas para el diagnóstico médico o los sistemas de control de personas en aeropuertos, incluyendo el uso de cámaras y dispositivos de detección en fronteras.

Trabajos recientes intentan determinar hasta qué punto la IA está afectando ya a la vida diaria de los ciudadanos desde todos los puntos de vista, por ejemplo, en la educación o la investigación científica (**Lopezosa; Goyanes, 2024**), o en la gestión de negocios (**Font-Cot et al., 2025**). Es evidente que estas nuevas herramientas están siendo utilizadas a todos los niveles. El núcleo del problema, que suscita preocupación tanto en la administración como entre ciudadanos, asociaciones y plataformas sociales, es que no existe un control real sobre las fuentes que utilizan ni sobre su funcionamiento, porque en la mayoría de los casos son procesos no auditables que, por su magnitud o por restricciones legales (copyright, derechos patentados, etc.), quedan fuera del control humano.

Debe insistirse en la importancia de estos cambios, ya que, además de implicar modificaciones en hábitos y costumbres, sus consecuencias afectan profundamente a aspectos de gran relevancia social.

La clave de la explicabilidad algorítmica

En este contexto, el concepto clave es el de explicabilidad. Los resultados del uso de un algoritmo de IA son explicables cuando el usuario puede acceder al mecanismo por el cual la IA alcanza sus resultados. Es evidente por qué se trata de un elemento esencial para la transparencia en el uso de algoritmos. Sin embargo, aunque parece un concepto simple, su aplicación práctica puede resultar extremadamente compleja cuando se pretende que toda la ciudadanía entienda cómo funcionan los algoritmos. Se trata, en realidad, de un problema de acceso cualificado a la información, ya que la

falta de formación técnica de los usuarios puede hacer prácticamente imposible el deseado acceso generalizado.

Por ejemplo, es muy difícil explicar al público general por qué un sistema basado en redes neuronales (*deep learning*) es intrínsecamente menos explicable que un modelo lineal con variables claramente definidas. Parece, por lo tanto, necesario ofrecer esta información de forma simplificada, para que las personas puedan conocer los posibles riesgos de ciertos procesos sin necesidad de disponer de formación técnica específica.

En este sentido, las administraciones están comenzando a articular respuestas. Una iniciativa desarrollada durante 2024 y principios de 2025 fue el proyecto de creación de un *Registro de Algoritmos* por parte de la *Generalitat Valenciana*. Tras los trabajos previos de grupos multidisciplinares dedicados a analizar y facilitar el acceso a la información fundamental sobre los algoritmos a la ciudadanía —realizados desde el *Observatorio de Transparencia y Gestión de Datos UPV-GVA* y la *Cátedra PAGODA UV-GVA* (**Ferrer-Sapena; Sánchez-Pérez**, 2022; **Calabuig et al.**, 2023; 2024)— un grupo de académicos, científicos y técnicos, bajo el auspicio de la *Generalitat* y con la participación activa de su personal, diseñó un sistema de catalogación y registro de los algoritmos utilizados por la administración pública en la *Comunitat Valenciana*, que se está implementando actualmente (*GVA-Confía*, 2025). La idea es que el control sobre la transparencia y el uso de los algoritmos recaiga en la propia administración, en lugar de dejarlo en manos de los usuarios individuales.

Los sistemas basados en IA generativa son especialmente delicados, puesto que, por su diseño, sus resultados distan de ser fácilmente interpretables. El entorno digital de acceso a la información ha cambiado profundamente con la irrupción de la IA, pasando de un esquema basado en motores de búsqueda tradicionales a sistemas de IA generativa fundamentados en modelos de procesamiento del lenguaje natural que, por su estructura, son intrínsecamente poco explicables. Este cambio afecta de forma directa la manera en que los usuarios acceden y procesan la información.

La transparencia algorítmica

La transparencia algorítmica es, ante todo, una exigencia ética y un requisito indispensable en una sociedad democrática. Afecta directamente a la fiabilidad de la información, cuya ausencia puede ser perjudicial para el funcionamiento normal de la sociedad. La confianza en los resultados depende de la comprensión de los mecanismos subyacentes, de modo que la explicabilidad mencionada resulta esencial para que los procesos de decisión que afectan a los ciudadanos sigan siendo transparentes.

Un algoritmo transparente es aquel cuyo funcionamiento —desde la adquisición de datos hasta sus cálculos específicos— puede auditarse manejando elementos técnicos de dificultad razonable. Los criterios de ponderación de datos y gestión de sesgos deben ser auditables y comprensibles, al igual que el propio funcionamiento de los algoritmos. Por ejemplo, los procesos que intervienen en la IA generativa son inherentemente opacos, ya que se basan en una gran cantidad de parámetros y en el entrenamiento con bases de datos masivas que pueden ser parciales o sesgadas y, con frecuencia, no pueden analizarse en detalle. Esto dificulta la identificación y corrección de errores en los resultados y, por tanto, incide en la falta de transparencia.

La IA reproduce los sesgos presentes en sus datos de entrenamiento. Esto no significa que exista algo deliberadamente diseñado para introducir sesgos ideológicos, económicos o sociales; simplemente, los algoritmos reproducen con más intensidad aquello que aparece con mayor frecuencia en los datos de entrenamiento. Así, si en la recopilación de información se prioriza la frecuencia de ciertos contenidos, el resultado no refleja necesariamente la veracidad de la información, sino los datos más repetidos o populares.

Conclusiones

Las actuaciones para garantizar la transparencia algorítmica son, y serán cada vez más, necesarias si se desea fomentar una administración pública que no oculte sus procesos de información. Es preciso emprender acciones que vayan más allá de la publicación de datos y la rendición de cuentas, para incluir programas de exposición sistemática de los procesos automáticos utilizados. Todo parece indicar que este será el horizonte de la transparencia administrativa en los próximos años: no solo los datos deben ser accesibles y auditables, sino también los instrumentos que los gestionan y procesan.

Las iniciativas necesarias son urgentes, y su objetivo debe ser la adopción de prácticas y recursos por parte de todos los agentes implicados para controlar un proceso que, en muchos aspectos, ya ha superado la capacidad de gestión de las administraciones. Desde las instituciones europeas existe una conciencia clara del problema, y se están realizando esfuerzos importantes para anticiparse a posibles situaciones de indefensión, tanto para ciertos colectivos como para la ciudadanía en general.

Las líneas de actuación deben perseguir objetivos básicos, como garantizar que los elementos de IA utilizados permitan auditorías externas sobre los criterios de entrenamiento y ponderación de los modelos, así como sobre su funcionamiento. Las instituciones deben también fomentar la transparencia y el control, por ejemplo, mediante la emisión de informes de impacto algorítmico que detallen posibles sesgos y sus efectos. Finalmente, la formación ciudadana en estos temas debe considerarse un objetivo prioritario, ya que sin unos conocimientos mínimos será difícil ejercer un control efectivo sobre estas herramientas.

Referencias

- Calabuig, José-Manuel; Cotino, Lorenzo; Ferrer-Sapena, Antonia; Sánchez-Pérez, Enrique-Alfonso** (2024). *Guía práctica: ¿Cómo abrir los algoritmos públicos? Recomendaciones para la implantación de registros de algoritmos públicos*. Universitat Politècnica de València; Universitat de València, versión 1.0, junio 2024. ISBN: 978 84 123363 5 1
<https://www.uv.es/cotino/publicaciones/Versioncotinoimagenes.pdf>
- Calabuig, José-Manuel; Ferrer-Sapena, Antonia; García-Raffi, Lluís-Miquel; Peset-Mancebo, Fernanda; Sánchez-Pérez, Enrique-Alfonso; Sánchez-del-Toro, Isabel** (2023). "Algoritmos matemáticos para una inteligencia artificial responsable, ética y transparente". *Revista Valenciana d'Estudis Autonòmics*, n. 68, pp. 283–305.
<https://dialnet.unirioja.es/servlet/articulo?codigo=9288120>
- Ferrer-Sapena, Antonia; Sánchez-Pérez, Enrique-Alfonso** (2022). "Análisis de las barreras para la explotación de los datos abiertos: estrategias y soluciones". En Cotino, Lorenzo; Todolí-Signes, Adrián (coords.). *Explotación y regulación del uso del big data e inteligencia artificial para los servicios públicos y la ciudad inteligente*, pp. 51–80. Tirant lo Blanch. ISBN: 978 8411133159
<https://www.uv.es/cotino/publicaciones/librobigindice.pdf>
- Font-Cot, Francesc; Lara-Navarra, Pablo; Sánchez-Arnau, Claudia; Serradell-López, Enric** (2025). "Scaling wheel framework: A holistic approach to startup scalability, governance models, and AI-driven innovation ecosystem competitiveness". *Journal of open innovation: Technology, market, and complexity*, v. 11, n. 3, 100619.
<https://doi.org/10.1016/j.joitmc.2025.100619>
- GVA-Confía** (2025).
<https://gvaoberta.gva.es/es/confia>
- Lopezosa, Carlos; Goyanes, Manuel** (2024). "Evaluación del uso ético de ChatGPT en investigación científica: principales tendencias, mejores prácticas y desafíos identificados". *Anuario ThinkEPI*, v. 18.
<https://doi.org/10.3145/thinkepi.2024.e18a19>