



Audidores de algoritmos

Enrique Benítez Palma

Economista

Ex Consejero de la Cámara de Cuentas de Andalucía

Resumen: En octubre de 2020, cinco Instituciones de Control Externo europeas publicaron un documento sobre la auditoría de algoritmos basados en el autoaprendizaje (machine learning) con el doble propósito de llamar la atención sobre este asunto y establecer un marco de referencia común a la hora de abordar la fiscalización de las decisiones públicas automatizadas por parte del control externo. En ausencia de un marco normativo común europeo, y a la vista del creciente uso de algoritmos por parte del sector público para fundamentar y agilizar la toma de decisiones, en un proceso de construcción de una administración pública digital basada en los datos (data driven), los primeros ejemplos de sesgos y errores en estas decisiones algorítmicas ponen de manifiesto la necesidad de establecer controles sólidos para hacer un seguimiento del uso de estas soluciones. La agilidad y la objetividad que prometen las herramientas basadas en el uso de los abundantes datos disponibles (big data) choca con una realidad en la que aparecen sesgos y se toman decisiones automáticas sin la más mínima participación de las personas implicadas. Decisiones importantes que afectan a la vida y a los derechos de esas personas. El artículo propone una revisión a la documentación disponible, y alerta sobre el papel secundario que puede corresponder al control externo si no da un paso adelante en este territorio confuso y por construir. Otras agencias reguladoras, o incluso la galaxia existente de organismos fiscalizadores de la sociedad civil pueden ocupar el espacio abandonado.

Palabras Clave: Control externo, algoritmos, inteligencia artificial, sector público, auditoría.

Abstract: In October 2020, five European Supreme Audit Institutions published a paper on the audit of machine learning algorithms with the dual purpose of drawing attention to this issue and establishing a common frame of reference when dealing with the audit of automated public decisions by external audit. In the absence of a common European regulatory framework, and in view of the increasing use of algorithms by the public sector to inform and speed up decision-making, in a process of building a data-driven public administration, the first examples of biases and errors in these algorithmic decisions highlight the need to establish robust controls to monitor the use of these solutions. The agility and objectivity promised by tools based on the use of the abundant data available (big data) clashes with a reality in which biases appear and automatic decisions are made without the slightest participation of the people involved. Important decisions that affect the lives and rights of these people. The article proposes a review of the available documentation, and warns of the secondary role that may correspond to external control if it does not take a step forward in this confused and unconstructed territory. Other regulatory agencies, or even the existing galaxy of civil society watchdogs, can occupy the abandoned space.

Keywords: External audit, algorithms, artificial intelligence, public sector, audit.



Introducción

El 14 de octubre de 2020, cinco importantes Instituciones europeas de Control Externo -las de Finlandia, Alemania, Países Bajos, Noruega y el Reino Unido, por orden alfabético en inglés- publicaron un documento conjunto de enorme interés e importancia para el futuro (inmediato) de la fiscalización de las decisiones públicas basadas en el uso de algoritmos. *Auditing machine learning algorithms. A White paper for public auditors*, pretende ser una guía básica de referencia para ayudar tanto a las instituciones como a los auditores individuales a llevar a cabo auditorías de algoritmos basados en *machine learning* que hayan sido utilizados por las agencias gubernamentales.

Aunque en España el uso de algoritmos en el sector público sea en estos momentos un asunto incipiente y sometido a discusión jurídica, limitado según los expertos en derecho administrativo a colaborar o apoyar la toma de decisiones, pero nunca a sustituir las potestades públicas o eludir la

necesaria motivación que impide la arbitrariedad¹, los planes de modernización del sector público que serán financiados con fondos europeos y la publicación de documentos como la *Estrategia Nacional de Inteligencia Artificial o España puede. Plan de recuperación, transformación y resiliencia de la economía*², anticipan la aplicación en las administraciones públicas españolas de este tipo de herramientas. De esta manera, surge en paralelo la necesidad de modernizar también las capacidades de las distintas instituciones independientes que tienen como misión ejercer el control de la acción del Gobierno, sea estatal, autonómico o local, o velar por el buen funcionamiento de la economía, como es el caso de la Comisión Nacional de los Mercados y la Competencia (CNMC).

Este trabajo propone conocer los motivos por los que se deben auditar los algoritmos que se usan en el sector público, analizar las distintas experiencias de las instituciones de control externo en el ámbito de la fiscalización del uso de algoritmos, y profundizar en el caso concreto de los Países Bajos. Asimismo, sugiere algunas lecturas adicionales que pueden ser de interés para establecer una primera toma de contacto con una cuestión que, sin duda, debe preocupar y ocupar a las diferentes instituciones de control externo durante los próximos años.

1. ¿Por qué auditar los algoritmos?

El uso de algoritmos en el sector público no es una novedad, al menos en los países occidentales. Desde hace varios años las organizaciones del sector público se apoyan en este tipo de herramientas para agilizar y objetivar la toma de decisiones. El paradigma de partida es la disponibilidad y neutralidad de los datos: existe abundancia de registros en manos de las administraciones públicas, y a partir de esta información se pueden diseñar fórmulas que ayuden a tomar decisiones más eficientes, eficaces y apoyadas en la evidencia.

Pero la cosa no es tan sencilla. Hay que saludar con entusiasmo la traducción y publicación en castellano del libro seminal "Automatizando la desigualdad", de la politóloga estadounidense Virginia Eubanks, al que ya he hecho mención en otros artículos anteriores³. Eubanks demuestra

1. Para no incluir una bibliografía muy prolija, se recomienda consultar Benítez Palma y Teré Pérez (2021).

2. El componente 16 del Plan de Recuperación, Transformación y Resiliencia del Gobierno de España desarrolla la hoja de ruta del despliegue de la Estrategia Nacional de Inteligencia Artificial.

3. Publicado a mediados de mayo de 2021 por la editorial Capitán Swing.

que, al menos en los Estados Unidos, es la gente más desfavorecida la que se relaciona con la administración pública, lo que introduce un llamativo sesgo en el manejo de la información disponible, que no abarca al conjunto de la población, ni incorpora todas las variables relevantes, sino que presenta carencias que además influyen en la decisión que se toma.

Para introducir la cuestión, podemos partir de cuatro ideas aportadas por los profesores Michael Kearns y Aaron Roth en su libro "El algoritmo ético". Se trata de una obra de lectura amena y recomendable que permite a las personas interesadas afrontar desde la curiosidad una aproximación a los problemas derivados del uso de algoritmos mal diseñados para tomar decisiones que afecten a nuestras vidas. Las cuatro ideas mencionadas son las siguientes:

1. "A medida que prolifera la captura y el análisis de datos, no sólo nos preocupa la privacidad, porque los algoritmos, esas pequeñas secuencias de código máquina que interfieren cada vez más en nuestros comportamientos a través de los teléfonos e internet, no se limitan a analizar los datos que generamos con cada uno de nuestros desplazamientos. También los emplean activamente para tomar decisiones que influyen en nuestras vidas" (Kearns y Roth, p. 13).
2. "¿Qué es un algoritmo? En su nivel más fundamental, no es más que una serie de instrucciones especificadas con gran precisión para llevar a cabo una tarea concreta" (Ibidem, p. 14).
3. "Sin embargo, los algoritmos, especialmente los modelos derivados directamente de datos a través de un aprendizaje automático, plantean retos diferentes. No solo permiten a un número importante de agentes tomar decisiones sin intervención humana, sino que, por su naturaleza tan compleja y opaca, ni siquiera sus diseñadores pueden prever cómo se comportarán en muchas situaciones" (Ibidem, p. 19). Esta afirmación enlaza con las advertencias realizadas hace unos años por Frank Pasquale en otro libro importante⁴, o más recientemente por Ryan Calo y Danielle Citron, cuando afirman que "los sistemas automatizados del Estado administrativo ponen de manifiesto hasta qué punto los funcionarios de los organismos han delegado sus responsabilidades en sistemas de terceros que son poco comprendidos incluso por sus creadores (*little understood even by their creators*)⁵.

4. "El resultado es que la toma de decisiones complejas automatizada que aporta el aprendizaje de las máquinas tiene un perfil propio, diferente del de su diseñador. Tal vez éste conozca bien el algoritmo que se utilizó para encontrar el modelo de toma de decisiones, pero no el modelo en sí" (Kearns y Roth, p. 23).

En definitiva, una cosa es el diseño de los algoritmos y otra muy diferente el control del funcionamiento de estos algoritmos una vez que se introduce en ellos un componente de aprendizaje autónomo. Esta divergencia es la que está en la base de la preocupación de las instituciones europeas de control externo cuando llaman la atención sobre la necesidad de auditar este tipo de sistemas.

Por lo tanto, hay dos cuestiones que fundamentan o explican la necesidad de controlar y vigilar el buen uso de algoritmos en el proceso de toma de decisiones en las administraciones públicas. Por un lado, están los sesgos derivados de la utilización de grandes bases de datos mal definidas, o incompletas, o que se basan en la experiencia histórica, lo que lleva a tomar decisiones que incorporan implícitamente mecanismos de discriminación (en los Estados Unidos si eres de raza negra un algoritmo puede tener la capacidad de clasificarte como una persona en situación de riesgo, o como un potencial criminal). Y por otro lado está la propia explicabilidad y fundamentación, es decir, la motivación administrativa, de las decisiones tomadas a partir del uso de un algoritmo concreto. Por utilizar de nuevo un ejemplo reciente y conocido, en los Países Bajos el gobierno tuvo que dimitir tras reconocer que un algoritmo de lucha contra el fraude había obligado a miles de familias de origen inmigrante a devolver ayudas sociales porque había categorizado erróneamente que esas familias estaban cometiendo fraude al acceder a las ayudas del Estado. Es decir, que en efecto se están tomando decisiones basadas en algoritmos que influyen en las vidas de los ciudadanos, en estos casos para perjudicarlas y causar un grave daño económico y reputacional.

En el Reino Unido, la utilización de un algoritmo para establecer las notas de los estudiantes de cara al acceso a la universidad, a partir de sus expedientes académicos y escuelas de procedencia, resultó ser un enorme fiasco que salpicó a la propia confianza de la sociedad en el sistema. El escándalo animó al gobierno a encomendar al *Centre for Data Ethics and Innovation* (CDEI) un informe sobre los sesgos algorítmicos, de obligada lectura para comprender la situación, publicado en noviembre de 2020: *Review into Bias in Algorithmic Decision-Making*.

4. *The Black Box Society: The Secret Algorithms That Control Money and Information*. Harvard University Press, 2015.

5. Calo, R. & D. K. Citron (2020): "The Automated Administrative State: A Crisis of Legitimacy." Boston University School of Law. He citado ya esta referencia en Benítez Palma (2021).

Uno de los párrafos del resumen ejecutivo es importante porque permite enmarcar la cuestión: “El uso de datos y algoritmos de forma innovadora puede permitir a las organizaciones comprender las desigualdades y reducir los prejuicios en algunos aspectos de la toma de decisiones. Pero también hay circunstancias en las que el uso de algoritmos para tomar decisiones que afectan a la vida puede considerarse injusto al no tener en cuenta las circunstancias de un individuo, o al privarle de su capacidad de acción personal. Este informe no se centra directamente en este tipo de injusticia, pero sí en la toma de decisiones, si la persona sujeta a la decisión no ha contribuido a la misma” (CDEI, 2020, p. 6).

Esta aportación es clarificadora, en tanto en cuanto apunta a la problemática de sustituir la interacción directa de la administración con los ciudadanos por un sistema indirecto de toma de decisiones basado en los datos disponibles, y no en la aportación directa de información por parte de las personas sobre las que se decide. Es decir, que el riesgo de posible arbitrariedad basado en la subjetividad derivada de las siempre imperfectas relaciones entre seres humanos sería sustituido por otro riesgo más frío y difícil de detectar: el de prescindir de las personas sobre las que se deben tomar las decisiones, para confiar en datos históricos no siempre completos y a menudo portadores de sesgos de todo tipo.

El informe continúa: “Tomar decisiones justas e imparciales no sólo es bueno para las personas implicadas, sino también para las empresas y la sociedad. El éxito y la sostenibilidad de la innovación dependen de la construcción y el mantenimiento de la confianza del público. Las encuestas realizadas para esta revisión sugieren que, antes de la controversia de agosto [se refiere al fiasco descrito de las notas de acceso a la universidad], el 57% de las personas sabía que se utilizan sistemas algorítmicos para apoyar las decisiones sobre ellos, y sólo el 19% estaba en principio en desacuerdo con la sugerencia de que un algoritmo ‘justo y preciso’ ayude a ayudar a tomar este tipo de decisiones. En octubre, se constató que el conocimiento había aumentado ligeramente (hasta el 62%), al igual que el desacuerdo (al 23%). Esto no apunta hacia un cambio de actitud del público, pero está claro que aún queda mucho camino por recorrer para crear confianza en los sistemas algorítmicos. El punto de partida obvio para ello es garantizar que los algoritmos sean fiables” (CDEI, 2020, p.6). Las conclusiones y hallazgos del informe del CDEI están en sintonía con lo apuntado por el profesor Paul Daly, cuando afirma que “para que la

administración artificial sea eficaz, debe haber un ‘nivel suficiente de confianza social’: ‘no sólo es necesario que las aplicaciones gubernamentales de IA sean eficaces, sino que la población debe percibir las también como eficaces para que sean complementos legítimos o sustitutas de los funcionarios humanos”⁶.

Las recomendaciones del informe son las siguientes:

“Los responsables de la gobernanza de las organizaciones que despliegan o utilizan herramientas de toma de decisiones algorítmicas para apoyar decisiones significativas sobre las personas deben asegurarse de que haya líderes responsables de:

- Comprender las capacidades y los límites de esas herramientas.
- Considerar cuidadosamente si las personas serán tratadas de forma justa por el proceso de toma de decisiones del que forma parte la herramienta.
- Tomar una decisión consciente sobre los niveles adecuados de participación humana en el proceso de toma de decisiones.
- Poner en marcha estructuras para recopilar datos y supervisar los resultados para que sean justos.
- Comprender sus obligaciones legales y haber realizado las evaluaciones de impacto adecuadas.

Esto se aplica especialmente en el sector público, donde los ciudadanos a menudo no pueden elegir si quieren utilizar un servicio, y las decisiones que se toman sobre los individuos a menudo pueden afectar a su vida” (CDEI, 2020, p.10).

En relación con el uso de algoritmos y la necesidad de establecer un marco adecuado para vigilar la calidad de los datos y aprovechar el potencial de la Inteligencia Artificial para el sector público, es recomendable conocer y analizar las distintas iniciativas que se han llevado a cabo por el Reino Unido: la aprobación de un marco de referencia por la *Information Commissioner’s Office* (ICO), la definición de una Estrategia Nacional de los Datos (*National Data Strategy*), la reflexión sobre qué se debe entender por calidad de los datos (*data quality*) y la creación de una Oficina para la Inteligencia Artificial (*Artificial Intelligence Office*) como organismo específico dedicado a estos asuntos dentro del Ministerio con competencia en asuntos digitales.

6. Daly, P. (2019): “*Artificial Administration: Administrative Law in the Era of Machines*.” Working Paper, Ottawa University. He utilizado esta misma referencia en Benítez Palma (2021).



2. La situación en España

Aunque España no ha llegado aún al nivel de madurez de uso de algoritmos en la gestión pública que se observa, por ejemplo, en la Europa continental o el mundo anglosajón, en los primeros meses del año 2021 se han conocido dos iniciativas relevantes en torno a esta cuestión⁷.

En enero de 2021 se ha publicado la *Guía de auditoría algorítmica*, de Éticas Consulting, una organización con sede en Barcelona a cuyo frente se halla Gemma Galdón Clavell. Se trata de una guía muy accesible, disponible en castellano para quienes quieran acercarse a esta cuestión, y en la que se puede leer que “las auditorías algorítmicas se presentan como una forma necesaria de hacer que esta tecnología sea más explicable, más transparente, más predecible, y más controlable por la ciudadanía, las instituciones públicas y también las empresas, ya sea antes del desarrollo del sistema, durante su desarrollo o a posteriori” (Éticas, 2021, p. 9).

Esta Guía también recuerda que “de acuerdo con la normativa vigente en materia de protección de datos (RGPD y LOPDGDD), todo tratamiento automatizado que produzca efectos significativos sobre la vida de una persona debe

ser siempre supervisado por una persona” (Ibidem, p. 15), y plantea que todos los algoritmos utilizados por el sector público se sometan a algún tipo de auditoría. Su metodología se construye a partir de cuatro grandes principios rectores, complementarios entre sí (Éticas, 2021, pp. 23-27):

- a) Cumplimiento legal y ético: “todo algoritmo debe cumplir con lo dispuesto por las normas jurídicas y deontológicas vigentes”. Aquí se hace referencia expresa tanto a la protección de datos ya mencionada como a la Carta Europea de Derechos Humanos.
- b) Deseabilidad del sistema: “un algoritmo de impacto social debe ser siempre explicable, preciso, replicable, transparente y justo. Por este motivo es imprescindible prestar atención a cuál es el ‘problema’ al que pretende dar solución el sistema auditado, y examinar si la tecnología utilizada es, efectivamente, la mejor manera de abordarlo”.
- c) Aceptabilidad social: “un sistema que tiene efectos sobre la vida de las personas, ya sea de forma directa o indirecta, debe ser comprensible, controlable, sostenible y, en alguna medida, beneficioso para las partes afectadas por el mismo”.

7. Es necesario hacer referencia a los avances normativos que se han producido en Cataluña, con la aprobación del Decreto 76/2020 de Administración Digital, y la creación en enero de 2021 del Comité Ético de los Datos, un órgano consultivo y transversal, que tiene como objetivo “garantizar que el despliegue de los usos avanzados de los datos, propios de la administración digital, se hagan de forma equitativa y responsable.

- d) Protección y gestión adecuada de los datos: esta gestión “debe responder a los principios del tratamiento que establecen el RGPD y la LOPDGDD, como la exactitud, la limitación del plazo de conservación, la limitación de la finalidad o la integridad y confidencialidad de los datos”⁸.

Respecto a la auditoría en sí de los sistemas, propone un trabajo en cinco etapas (estudio preliminar, mapeo, plan de análisis, análisis e informe de auditoría), que comienza con una captación de información preliminar sobre el algoritmo para determinar “quién lo encarga, lo diseña, lo desarrolla, lo financia y lo implementa, y cuál es el problema que se pretende resolver usando ese algoritmo” (Éticas, 2021, p. 28) y termina con un informe de auditoría que “deje constancia del proceso realizado, así como del cumplimiento legal y ético, la precisión, la aceptabilidad y la deseabilidad del modelo en base a la interpretación de los resultados” (Ibidem, p. 61). Esta Guía permite entender, sin profundizar en conocimientos técnicos, de qué se habla cuando se habla de auditoría algorítmica.

Además de este interesante documento, en abril el Ayuntamiento de Barcelona ha hecho público su *Medida de Gobierno de la estrategia municipal de algoritmos y datos para el impulso de la ética artificial*. Las ciudades están tomando la delantera en la

toma de decisiones relacionadas con el uso de algoritmos, como lo ponen de manifiesto los casos de los registros públicos de algoritmos de Helsinki o Rotterdam (SAIDOT, 2021), o esta iniciativa de Barcelona. En esta *Medida de Gobierno* llama la atención, dentro de las recomendaciones técnicas, la parte dedicada al análisis de riesgos: “una aplicación de la IA debe considerarse de riesgo elevado tanto desde la perspectiva de la protección de la seguridad como desde la perspectiva de los derechos fundamentales de las personas. (...). Deben considerarse de riesgo elevado —y por lo tanto exigen un control especial— todas las aplicaciones de la IA con efectos jurídicos sobre los derechos de una persona, las que puedan causar lesiones, la muerte o daños materiales (o inmateriales significativos), así como las que puedan producir efectos imposibles de evitar por las personas” (Ayuntamiento de Barcelona, 2021, p. 37).

El documento del Ayuntamiento de Barcelona apuesta por la construcción de una Inteligencia Artificial fiable (lo que en el Libro Blanco de la IA de la Comisión Europea se define como *‘trustworthy AI’*; esto es, una IA digna de confianza, a partir del siguiente proceso, que también toma en consideración la reciente *Carta de derechos digitales*, impulsada por la Secretaría de Estado de Digitalización e Inteligencia Artificial (SEDIA) del Ministerio de Asuntos Económicos y Transformación Digital:



Fuente: Ayuntamiento de Barcelona (2021).

8. En este punto no me resisto a recomendar un artículo académico del máximo interés: Kaminski, M. E. (2019): *Binary Governance: Lessons from the GDPR's Approach to Algorithmic Accountability*. Southern California Law Review, vol. 92, pp. 1529-1616.

3. El control externo y el uso de algoritmos en la administración pública

La fiscalización de decisiones públicas basadas en herramientas algorítmicas no es nueva en el ámbito del control externo. Ya en 2017, la *Cour des Comptes* de Francia tuvo que auditar los resultados de la aplicación de estas técnicas en el acceso a la enseñanza superior. El informe, titulado (en inglés) *Admission post-bac and Access to Higher Education* concluía de manera rotunda que “la Admisión Post-Bachillerato (post-high school admission, o APB) es una plataforma digital pública utilizada por los bachilleres para presentar solicitudes con vistas a entrar en los cursos de primer año de la educación superior. Contiene dos algoritmos: un algoritmo de asignación, que cruza las preferencias de los centros educativos y de los solicitantes, y un algoritmo de clasificación, que atribuye un orden de

prioridad a las solicitudes de acceso a los cursos de grado. Los métodos de funcionamiento de la APB, que carecen de transparencia y no se ajustan al código educativo, ponen de manifiesto las incoherencias en las condiciones de acceso a los cursos de licenciatura en las universidades. La APB constituye un avance técnico útil, pero hoy en día es criticada con razón y necesita ser reformada” (*Web de la Cour des Comptes*)⁹.

Con motivo de la proliferación de uso de este tipo de herramientas, o soluciones, la revista ISACA ha prestado atención al tema, identificando los riesgos potenciales y las soluciones a adoptar para garantizar una buena auditoría de sistemas de inteligencia artificial:

Retos para el auditor de la IA	Claves para el éxito de la auditoría de la IA
– Marcos de auditoría o regulaciones específicas inmaduras para la IA	– Adoptar y adaptar los marcos y regulaciones existentes
– Precedentes limitados para los casos de uso de la IA	– Explicar y comunicar proactivamente sobre IA con las partes interesadas
– Definiciones y taxonomías inciertas de la IA	– Explicar y comunicar proactivamente sobre IA a las partes interesadas
– Gran variedad de sistemas y soluciones de IA	– Informarse sobre el diseño y la arquitectura de la IA para establecer un alcance adecuado
– Naturaleza emergente de la tecnología de IA	– Informarse sobre el diseño y la arquitectura de la IA para fijar el ámbito de aplicación
– Falta de orientación explícita sobre la auditoría de la IA	– Centrarse en la transparencia mediante un proceso iterativo. Centrarse en controles y gobernanza, no en los algoritmos
– Falta de puntos de partida estratégicos	– Implicar a todas las partes interesadas
– Posible curva de aprendizaje para el auditor de IA	– Informarse sobre el diseño de la IA y contratar a especialistas necesarios
– Riesgo para el proveedor creado por la subcontratación de IA a terceros	– Documentar las prácticas de arquitectura para la transparencia entre equipos

Fuente: ISACA (Auditing Artificial Intelligence, 2018)

9. <https://www.ccomptes.fr/en/publications/admission-post-bac-and-access-higher-education>

Como se puede ver, la novedad del tema, la confluencia con la metodología de las auditorías de sistemas de información y la indefinición de los propios marcos regulatorios son algunos de los principales problemas a la hora de afrontar una auditoría de sistemas de inteligencia artificial. En este sentido, es importante recordar la variedad de modelos y marcos existentes, tal y como pone de manifiesto un reciente documento de *Open Government Partnership* (BASU y PÉREZ, 2020). Según estos autores, hay tres modelos de gobernanza algorítmica en estos momentos:

1. **Obligado por la ley:** Francia tiene un marco legal que obliga a abrir los algoritmos (*Law for a Digital Republic*), por lo que el ámbito de aplicación de la política se establece en la ley. Las administraciones deben publicar en línea una lista de algoritmos cuando se utilizan para tomar decisiones que afectan a la vida de los ciudadanos.
2. **Obligado por la política (*Committed by policies*):** Canadá tiene una directiva emitida por el comité de supervisión de finanzas y operaciones del gobierno federal, que establece los requisitos sobre cómo se pueden utilizar los algoritmos para apoyar la prestación de servicios a los ciudadanos. El cumplimiento de esta directiva es obligatorio para los departamentos del gobierno federal canadiense, y de hecho se acaba de actualizar en abril de 2021 (GOVERNMENT OF CANADA, 2021).
3. **Voluntario:** En Nueva Zelanda, el enfoque de los algoritmos se basa en un compromiso público de los organismos gubernamentales, no en una norma legal. El país creó una carta de algoritmos (incluso como parte de un compromiso de la *Open Government Partnership*) que fue sometida a consulta pública y actualmente está trabajando en su aplicación (GOVERNMENT OF NEW ZEALAND, 2020). Como en cualquier enfoque voluntario, uno de los retos es que no todos los organismos se han comprometido a aplicar la carta. Es importante tener en cuenta que muchos de estos cambios se están produciendo antes de que se establezcan los marcos legales. Esto es una limitación, pero también una oportunidad que ha permitido un movimiento más rápido y ágil.

Para el control externo, la construcción de diferentes arquitecturas institucionales y la intersección de competencias con otras agencias reguladoras -por ejemplo, las de protección de datos, o las de protección de los derechos de los consumidores- dificultan la posibilidad de ejercer la misión de fiscalizar el uso de estas herramientas y el resultado de las decisiones implementadas a través de ellas. No obstante, el ejemplo holandés puede servir de marco de referencia en caso de que se plantee la voluntad de dar un paso adelante en este controvertido y apasionante terreno.

4. El caso del control externo en los Países Bajos

De las cinco Instituciones de Control Externo que promovieron el documento sobre la auditoría de algoritmos basados en *machine learning*, es el Tribunal de Cuentas de Holanda (*Netherlands Court of Auditors*) -y también la Cámara de Cuentas regional de Rotterdam (*Rekenkamer Rotterdam*)- el que más ha avanzado en la aplicación práctica de los principios desarrollados. Este comportamiento tiene que ver, de entrada, con su propia planificación estratégica, ya que en sus dos últimos planes se llama la atención sobre la evolución constante del sector público y el desafío que supone la utilización de nuevas tecnologías por parte de los gobiernos auditados. Holanda, en esta manifestación de interés, va por delante.

Así, en su documento estratégico del período 2016-2020, titulado "Confianza basada en la comprensión" (*Trust based on Understanding*), ya se señalan los riesgos derivados del uso de las tecnologías en el sector público, en un escenario de necesaria modernización pero en el que no siempre se tienen en cuenta las verdaderas necesidades y capacidades de la ciudadanía. En esta planificación estratégica, el Tribunal de Cuentas holandés señala que "estas constantes reorganizaciones han convertido al 'gobierno' en una compleja estructura de organizaciones y técnicas de financiación, sin una división clara de funciones y responsabilidades y con límites a menudo borrosos entre los flujos de dinero público y privado. La experiencia práctica nos ha enseñado que esta situación puede suponer un riesgo para la prestación eficiente de los servicios públicos y para la regularidad del gasto público. También dificulta el ejercicio del derecho del Parlamento a aprobar y modificar la política presupuestaria. El seguimiento del flujo de dinero del Tesoro a los ciudadanos no suele ser una tarea fácil. También hace más difícil saber exactamente quién es responsable de qué, y medir los resultados y el impacto de las políticas gubernamentales. Para el público en general, la situación es incluso aún más opaca. En otras palabras, hay deficiencias en la forma en que el gobierno informa sobre el gasto del dinero público. Esto repercute en nuestro trabajo" (Netherlands Court of Audit, 2016, p.8).

A este documento estratégico le ha sucedido uno nuevo referido al horizonte temporal 2021-2026, titulado "Confianza en la Rendición de Cuentas" (*Trust in Accountability*). De nuevo aparece la palabra confianza, lo que parece significativo, ya que en el ámbito de la aplicación de la inteligencia artificial a las decisiones públicas se utiliza el término "*trustworthy AI*"; es decir, inteligencia artificial confiable o digna de confianza. Se insiste en la idea de desentrañar cómo se toman las decisiones, para evitar su opacidad, la posible comisión de injusticias al basar la gestión en algoritmos, o incluso el rechazo social derivado de la incomprensión o la ausencia de motivación clara en las decisiones tomadas.



En este nuevo plan estratégico la mención a la tecnología y los sistemas de información es mucho más contundente: “la administración central apenas es capaz de seguir el ritmo de los avances en tecnología de la información, si es que lo hace. Sin embargo, la tecnología y los datos -el ‘oro del siglo XXI’- están cada vez más en el corazón del gobierno. La crisis de COVID-19 ha vuelto a poner de manifiesto que la continuidad de la prestación de servicios depende del buen funcionamiento de las tecnologías de la información, y los servicios públicos no son una excepción. Por eso damos tanta importancia a la gestión de la información en nuestra auditoría anual de responsabilidad gubernamental. Se ha convertido en el tercer componente de nuestra auditoría de la gestión operativa, junto con la gestión financiera y material. Cuando sea necesario, desarrollaremos normas adicionales sobre la gestión de la información. En este sentido, examinaremos el desarrollo y el control de los objetivos y el uso de algoritmos de autoaprendizaje en los próximos años” (Netherlands Court of Auditors, 2021a, p. 21).

Uno de los primeros resultados concretos de este reconocimiento estratégico de la importancia de las decisiones públicas basadas en datos (*data-driven*) y la necesidad de su vigilancia y control, es otro documento de 2021 titulado “Comprendiendo los algoritmos” (*Understanding algorithms*). El Tribunal de Cuentas de Holanda no sólo ha puesto el foco en la fiscalización de los algoritmos, sino que también está compartiendo con generosidad su conocimiento adquirido

con el resto de Instituciones de Control Externo interesadas en esta cuestión. Ha publicado también un marco para la auditoría de los algoritmos (*Algorithm Audit Framework*), de manera que la lectura conjunta de ambos documentos constituye, hoy por hoy, la mejor guía para las organizaciones y auditores preocupados por esta cuestión.

Sería muy prolijo, y desborda los objetivos de este artículo, describir la herramienta holandesa para la fiscalización de los algoritmos. Está a disposición del público interesado, en inglés, y basta con leerla y aplicarla, llegado el caso. Sí que es destacable mencionar que el marco de auditoría de algoritmos se basa en cinco pilares: la gobernanza de la herramienta y la rendición de cuentas sobre el uso de la misma; el modelo y los datos utilizados; la privacidad y el cumplimiento de las normas de protección de datos; los sistemas generales de tecnologías de la información; y finalmente la ética.

De lo que se trata en este momento es de llamar la atención sobre este inquietante asunto y de animar a tomar decisiones anticipativas para que el uso de algoritmos por parte de las administraciones públicas españolas, en cualquiera de sus niveles de gobierno -estatal, autonómico o local- no pille por sorpresa a las entidades de control externo. La experiencia holandesa aporta seis recomendaciones para el sector público, que se reproducen, resumidas, a continuación (Netherlands Court of Auditors, 2021b, pp. 38-42):

- a) “Publicar definiciones y requisitos de calidad claros y coherentes. Instamos al gabinete a que adopte un conjunto de términos claros y uniformes y unos requisitos de calidad específicos para los algoritmos. Las definiciones claras y coherentes y los requisitos de calidad fomentarán el intercambio de conocimientos, agilizarán los procesos y evitarán interpretaciones erróneas.”
- b) “Informar a los ciudadanos particulares sobre los algoritmos y explicarles cómo pueden obtener más información sobre ellos. Instamos al gabinete a que permita a los ciudadanos particulares acceder, de una manera lógica, a la información sobre qué datos se utilizan en los algoritmos, cómo funcionan básicamente estos algoritmos funcionan, y qué impacto tienen sus resultados.”
- c) “Documentar los acuerdos sobre el uso de los algoritmos y tomar medidas eficaces para supervisar el cumplimiento (*compliance*) de forma continua. Nuestra recomendación al gabinete es que garantice una documentación adecuada de los términos de referencia, la organización, el seguimiento (por ejemplo, en términos de gestión del ciclo de vida, mantenimiento y cumplimiento de la legislación vigente) y la evaluación del algoritmo, ya que esto dejaría claro si el algoritmo es y sigue siendo apto para el propósito inicial. Esto también permite ajustar el algoritmo, si es necesario.”
- d) “Garantizar que el marco de auditoría se traduzca en requisitos prácticos de calidad para los algoritmos. Recomendamos que el Consejo de Ministros dé instrucciones para que el *Chief Data Officer* de cada ministerio se encargue de convertir este marco de auditoría (diseñado para evaluar los algoritmos que ya se utilizan) en un conjunto práctico de normas de diseño o en requisitos de calidad para el desarrollo de algoritmos.”
- e) “Garantizar que todas las disciplinas pertinentes participen en el desarrollo de algoritmos. Nuestra recomendación al gabinete es que se involucren todas las disciplinas y tipos de conocimientos especializados pertinentes en el desarrollo de algoritmos. Esto significa que, junto a los especialistas técnicos, participen expertos jurídicos, especialistas en ética y asesores políticos.”



10. Agradezco la colaboración de Paul Hofstra en este punto del artículo.

- f) “Garantizar que se produzca una información clara, ahora y en el futuro, sobre el funcionamiento de Controles Generales de Tecnologías de la Información (CGTI). Recomendamos que el Consejo de Ministros dé instrucciones para que los ministros y secretarios de Estado competentes se encarguen de que los funcionarios que trabajan con algoritmos tengan y conserven el acceso a la información sobre la calidad de los CGTI en relación con los algoritmos en cuestión.”

Aunque en este momento no parece probable una traslación inmediata de estas recomendaciones al sector público nacional, con pocas excepciones, marcan una línea de actuación de cara al futuro, que se puede concretar en diversas ideas fuerza, que van desde la presencia de un *Chief Data Officer* en cada ministerio -recordemos que los cimientos de todo esta modernización está en la disponibilidad y uso de los datos, y en la protección de la privacidad-, hasta la evaluación de los resultados, la transparencia, la observancia de claros principios éticos basados en el respeto a los derechos fundamentales, y el marco clásico de control de los sistemas de información.

Una última referencia a la experiencia holandesa debe hacerse, antes de finalizar este epígrafe, al trabajo realizado por la Cámara de Cuentas de Rotterdam¹⁰, cuyo informe concluye que “el municipio de Rotterdam utiliza algoritmos para apoyar su toma de decisiones. Aunque el municipio ha prestado atención al uso ético de los algoritmos, la concienciación sobre esta necesidad aún no es generalizada. Esto puede llevar a una falta de transparencia de los algoritmos y a resultados sesgados, como en el caso de un algoritmo destinado a combatir el fraude en las prestaciones”.

5. Conclusiones y sugerencias

En el año 2019, un informe de Nueva Zelanda proponía la creación ad-hoc de una agencia reguladora para el seguimiento y supervisión de algoritmos por parte del sector público. Las recomendaciones del informe son muy útiles, y aunque no hay constancia de que el Tribunal de Cuentas de Holanda las haya seguido, sí que parece claro y evidente que ha decidido asumir ese papel de controlador, en ausencia de un organismo específico que lo haga, en un entorno de uso creciente de algoritmos, no siempre con buenos resultados, como se ha visto en los casos de la lucha contra el fraude en la recepción de prestaciones básicas.

El informe de Nueva Zelanda recomienda lo siguiente: “Hay varios modelos posibles para una nueva agencia reguladora. Todos ellos tienen puntos fuertes y debilidades. En

la actualidad, hay muy pocos ejemplos internacionales de los que aprender, y los que existen están en sus primeras etapas.

Hemos propuesto una posible estructura para que la nueva agencia reguladora pueda trabajar con las agencias gubernamentales. El nuevo regulador podría desempeñar una serie de funciones, entre ellas:

- Elaborar directrices de buenas prácticas.
- Mantener un registro de los algoritmos utilizados en el gobierno.
- Elaborar un informe público anual sobre dichos usos.
- Realizar un seguimiento continuo de los efectos de estas herramientas.

Si la agencia reguladora va a recibir algún tipo de poder estricto, habrá que tener en cuenta su capacidad para supervisar y hacer cumplir estos poderes.

Si la agencia debe encargarse de controlar los algoritmos, hay que tener en cuenta que se trata de herramientas versátiles, capaces de ser reutilizadas para una variedad de usos. El escrutinio debe aplicarse a los nuevos usos/daños potenciales y no sólo a los nuevos algoritmos” (New Zealand Law Foundation, 2019, p.4).

Hay dos reflexiones fundamentales que debe hacerse el mundo del control externo en relación con el uso de algoritmos en el sector público, la fiscalización de estas herramientas y la supervisión de sus resultados.

La primera tiene que ver con la necesidad de estar al tanto de estos movimientos. Ya existen diversos grupos de trabajo internacionales -por ejemplo, el EUROSAI IT Working Group, que ha organizado una jornada específica sobre este tema- que están tratando el tema de los algoritmos. Pero una buena referencia en España puede ser la Comisión Nacional de los Mercados y la Competencia (CNMC), que ante la perspectiva del uso de algoritmos por parte de las grandes compañías globales para propiciar la colusión invisible de precios, ha decidido unirse al Computational Antitrust Project, de la Universidad de Stanford, junto a varias decenas de autoridades nacionales de la competencia, para aprender, compartir conocimientos y abordar de manera conjunta la estrategia de las diferentes autoridades nacionales en un mundo sin fronteras digitales. La creación de una red europea de instituciones de control externo interesadas en la supervisión del uso de algoritmos por parte del sector público puede suponer un buen paso adelante para asumir ese papel ahora inexistente que ya se reclama en Nueva Zelanda para evitar malos usos, enfoques erróneos o decisiones equivocadas.

La segunda cuestión tiene que ver con la primera, y también con la definición del papel que se desea jugar en un entorno cambiante, acelerado y cada vez más basado en la tecnología y los datos. Si los fondos públicos se van a gastar mediante decisiones automatizadas, el control externo debe asumir su inmediata actualización. Porque si no lo hacen estas instituciones, será la propia sociedad civil la que se convierta en la garante de la protección de los derechos colectivos e individuales de los ciudadanos. Ya lo están haciendo organizaciones como Algorithm Watch, Saidot, Transparent Internet, Montreal AI Ethics y un largo etcétera de “entidades fiscalizadoras sin ánimo de lucro”, más ágiles y preparadas para este cometido que muchas agencias gubernamentales. Es un dilema hamletiano, pero la respuesta parece más que evidente.

Bibliografía utilizada y recomendada

- **Ada Lovelace Institute (2020):** Transparency mechanisms for UK public-sector algorithmic decision-making systems.
- **Ada Lovelace Institute (2021):** Examining the Black Box. Tools for assessing algorithmic systems.
- **Ahamat, G., Chang, M. & Thomas, C. (2021):** The need for effective AI assurance. Centre for Data Ethics and Innovation blog.
- **Ahonen, P. & Erkkilä, T. (2020):** Transparency in algorithmic decision-making: Ideational tensions and conceptual shifts in Finland. *Information Polity*, 25, 419-432.
- **Ananny, M. & Crawford, K. (2018):** Seeing without knowing: Limitations of the transparency ideal and its application to algorithmic accountability. *New Media & Society*, Vol. 20(3), 973-989.
- **Ayuntamiento de Barcelona (2021):** Medida de Gobierno de la estrategia municipal de algoritmos y datos para el impulso de la ética artificial.
- **Bandy, J. (2021):** Problematic Machine Behavior: A Systematic Literature Review of Algorithm Audits. To Appear in the Proceedings of the ACM (PACM) Human-Computer Interaction, CSCW '21. 34 pages.
- **Bannister, F. & Connolly, R. (2020):** Administration by algorithm: A risk management framework. *Information Polity*, 25, 471-490.
- **Basu, T. y Pérez, P. (2020):** Reflexiones sobre la implementación de las políticas de algoritmos en el sector público. *Open Government Partnership*, 15 dic.
- **Benítez Palma, E. y Teré Pérez, A. (2021):** Entusiasmo robótico y externalización del derecho administrativo. Observatorio de transformación digital de la Universidad de Valencia.
- **Benítez Palma, E. (2021):** El control externo del gasto público en el Estado automatizado. *Revista Española de Control Externo*, nº 65, 98-111.
- **Brown, S., Davidovic, J. & Hasan, A. (2021):** The algorithm audit: Scoring the algorithms that score us. *Big Data & Society*, enero-junio, 1-18.
- **Busuioc, M. (2020):** Accountable Artificial Intelligence: Holding Algorithms to Account. *Public Administration Review*, vol. 00, pp. 1-12.
- **Canada Government (2020):** Algorithmic Impact Assessment Tool.
- **Canada Government (2021):** Directive on Automated Decision-Making.
- **Centre for Data Ethics and Innovation (2020):** Review into bias in algorithmic decision-making.
- **Cour des Comptes (2020):** Un premier bilan de l'accès à l'enseignement supérieur dans le cadre de la loi orientation et réussite des étudiants.
- **Criado, J. I., Valero, J. y Villodre, J. (2020):** Algorithmic transparency and bureaucratic discretion: The case of SALER early warning system. *Information Polity*, 25, 449-470.
- **Éticas Consulting (2021):** Guía de Auditoría Algorítmica.
- **Future of Privacy Forum (2017):** Unfairness by Algorithm: Distilling the Harms of Automated Decision-Making.
- **Giest, S. & Grimmelikhuijsen, S. (2020):** Introduction to special issue algorithmic transparency in government: Towards a multi-level perspective. *Information Polity*, 25, 409-417.
- **Information Commissioner's Office (ICO) (2020):** Guidance on the AI auditing framework.
- **Kaminski, M. E. (2020):** Understanding Transparency in Algorithmic Accountability. University of Colorado Law Legal Studies Research paper nº 20-34.
- **Kazim, E., Mendes Thame Denny, D. & Koshiyama, A. (2021):** AI auditing and impact assessment: according to the UK Information Commissioner's Office. *AI and Ethics*, feb. 2021.

- **Kearns, M. & Roth, A. (2020):** El algoritmo ético. La ciencia del diseño de algoritmos socialmente responsables. Wolters Kluwer, 305 pp.
- **König, P. D. & Wenzelburger, G. (2021):** Between technochauvinism and human-centrism: Can algorithms improve decision-making in democratic politics? *European Political Science*, enero 2021.
- **Koshiyama, A. et alia (2021):** Towards algorithm auditing. Technical report. Enero 2021.
- **Martin, K. (2019):** Ethical Implications and Accountability of Algorithms. *Journal of Business Ethics*, 160, 835-850.
- **Mitchell, S., Potash, E., Barocas, S., D'Amour, A. & Lum, K. (2021):** Algorithmic Fairness: Choices, Assumptions, and Definitions. *The Annual Review of Statistics and its Applications*.
- **Netherlands Court of Auditors (2016):** Trust based on Understanding.
- **Netherlands Court of Auditors (2021):** Trust in Accountability.
- **Netherlands Court of Auditors (2021b):** Understanding algorithms.
- **Netherlands Court of Auditors (2021c):** Audit framework for algorithms, v.1.0.
- **New, J. & Castro, D. (2018):** How Policymakers Can Foster Algorithmic Accountability. Center for Data Innovation.
- **New Zealand Government (2019):** Algorithm Charter.
- **New Zealand Law Foundation (2019):** Government use of Artificial Intelligence in New Zealand.
- **Peeters, R. (2020):** The agency of algorithms: Understanding human-algorithm interaction in administrative decision-making. *Information Polity*, 25, 1-16.
- **Reisman, D., Schultz, J., Crawford, K. & Whittaker, M. (2018):** Algorithmic Impact Assessments: a Practical Framework for Public Agency Accountability. AI Now Institute.
- **Rekenkamer Rotterdam (2021):** Gekleurde Technologie. Verkenning Ethisch Gebruik Algoritmes.
- **Richardson, R. (ed) (2019):** Confronting Black Boxes: A Shadow Report of the New York City Automated Decision System Task Force.
- **Soriano Arnanz, A. (2021):** Decisiones Automatizadas: Problemas y Soluciones Jurídicas más allá de la Protección de Datos. *Teoría y Método, Revista de Derecho Público*, vol. 3, 85-127.
- **USA Government (2019):** Algorithmic Accountability Act.
- **Varias Instituciones europeas de control externo (2020):** Auditing machine learning algorithms. A white paper for public auditors.
- **Vestri, G. (2021):** La inteligencia artificial ante el desafío de la transparencia algorítmica. Una aproximación desde la perspectiva jurídico-administrativa. *Revista Aragonesa de Administración Pública*, 56, 368-398.

