

POLÍTICAS PÚBLICAS E INTELIGENCIA ARTIFICIAL. ¿SOLUCIÓN BUSCA PROBLEMA?

Ismael Blanco*
Joan Subirats*

INTRODUCCIÓN

Atravesamos un periodo muy intenso de cambios de todo tipo, hasta el punto de que podemos hablar de cambio de época. Tal es la intensidad y profundidad de transformaciones que nos atraviesan en múltiples esferas. Sin posibilidad de profundizar en los aspectos más generales de estos cambios, queremos centrarnos aquí en la intensa revolución tecnológica que conlleva la emergencia de la Inteligencia Artificial (IA) y, más concretamente, en sus impactos en los procesos de elaboración de políticas públicas. La premisa de la que partimos es que la IA afecta a demasiados aspectos de nuestras vidas como para dejarlo todo en manos de los tecnólogos. Necesitamos miradas ‘inexpertas’ en tecnología que, desde enfoques multidisciplinares, adviertan sobre la forma en que la IA puede afectar distintos aspectos significativos de la vida y cómo hacer un buen uso de esta. El análisis de políticas públicas puede aportar una reflexión relevante sobre los impactos de esta tecnología en la calidad democrática, la eficiencia y la eficacia de la acción de gobierno.

Aunque imperativa y hasta cierto punto urgente, esta reflexión debería ser pausada, lejos de las exageraciones que a menudo se producen en el debate sobre el desarrollo de la IA. Innerarity habla en este sentido de “histeria digital” –visiones que tanto nos presentan la IA como una tecnología que puede “salvar a la humanidad” (alucinaciones, según Naomi Klein) o incluso suponer su extinción. Probablemente, este momento “hype” que vivimos con la IA forma

* Ismael Blanco y Joan Subirats son catedráticos de Ciencia Política e investigadores del Instituto de Gobierno y Políticas Públicas de la Universidad Autónoma de Barcelona.

parte de una estrategia de las grandes empresas tecnológicas que, al atraer la atención desmesurada sobre esta tecnología, captan inversiones y acaparan protagonismo político. Sin necesidad de compartir este tipo de visiones exageradas, creemos sin embargo que el impacto de la IA en la acción de gobierno y las políticas públicas es profundo e incuestionable. La rapidez con que se ha incorporado la IA en la esfera pública, sobre todo desde la incorporación del lenguaje generativo, está implicando grandes desafíos tanto desde el punto de vista instrumental como desde el punto de vista político y regulatorio. Resulta, pues, absolutamente necesario que reflexionemos no sólo sobre las oportunidades que representa para las políticas públicas la mejora de los sistemas de información y conocimiento, sino también sobre los desafíos que la IA plantea para la calidad de la democracia.

Este artículo pretende contribuir a esta reflexión a partir de tres ideas principales. La primera constata la incidencia de los humanos en el diseño y la aplicación de esta tecnología, a la vez que sus impactos en la distribución de costos y beneficios sociales. Esta visión ‘humanizada’ y ‘social’ de la IA nos obliga a considerar la dimensión política de la tecnología y por lo tanto la necesidad de ejercer un control democrático sobre ella. La segunda idea resalta el impacto de la IA sobre las relaciones de poder en que se asientan los procesos de elaboración de las políticas públicas: apartados de una visión estrictamente técnica de las mismas, resaltamos cómo el *policymaking* se fundamenta en relaciones de poder entre actores sobre las que la IA puede tener un impacto profundo. De aquí que reclamemos una agenda de investigación en políticas públicas orientada a desvelar los desequilibrios de poder y las posibilidades de fortalecimiento del bien común en una era de digitalización avanzada. Por último, cerramos este artículo sosteniendo que los poderes públicos deberían adoptar una estrategia proactiva (y no meramente reactiva) frente a un mundo tecnológico en profunda transformación. Es por ello por lo que terminamos elaborando algunas propuestas sobre las estrategias que los gobiernos deberían adoptar para aprovechar las oportunidades que representa la IA y contrarrestar algunos de sus efectos más perversos.

GOBERNAR LA IA: IMPACTOS SOCIALES Y CONTROL DEMOCRÁTICO

El debate suscitado por la IA conecta directamente con el que ya se planteó hace tiempo con relación a los impactos de Internet y la digitalización en el *policymaking* y en la actuación de las administraciones públicas¹. Decíamos entonces, siguiendo una afortunada expresión de Mark Poster (2007), que

¹ Véase la adenda al respecto en el libro de Dente-Subirats (2013).

Internet no era simplemente un “martillo” nuevo que servía para clavar más de prisa o con mayor comodidad los “clavos” de siempre, sino que el cambio digital modificaba la forma de relacionarnos e interactuar, alterando profundamente los procesos y posiciones de relación e intermediación. Defendíamos así que estábamos ante un cambio en profundidad de nuestras sociedades: un nuevo “país”, no forzosamente mejor, pero sí distinto. Aquellos que, hablando de “e-democracy” o “e-government”, defendían la neutralidad política y social de la transformación digital, y por tanto que de lo que se trataba era de mejorar la eficacia y la eficiencia del funcionamiento de los poderes públicos, aceptarán ahora que los cambios han sido mucho más intensos de lo entonces imaginable, dando lugar a fenómenos tanto de concentración de poder en grandes compañías tecnológicas (GAFAM) como a alteraciones significativas en la esfera pública y los procesos de generación y contraste de noticias y evidencias, tan importantes en el debate sobre problemas y alternativas de solución.

Sea como sea, no estamos ya en el periodo en que simplemente podamos celebrar la mejora de las relaciones entre la ciudadanía y las administraciones, facilitando la resolución a distancia de lo que hasta entonces eran complejos y costosos procedimientos de obtención de información, de permisos, de renovación de documentos, de liquidaciones fiscales o de obtención de información. Los avances en la seguridad de esos procesos a través de la aceptación de la firma electrónica, o la creciente coordinación entre distintos niveles de administración son un buen ejemplo de ello. Aquellos procesos, entonces novedosos, eran hasta cierto punto despolitizados, ya que no se cuestionaba el porqué de los servicios o a quién iban dirigidos, sino la mejor manera de prestarlos. Sin duda, existía y existe una cierta preocupación por los efectos perversos que podrían llegar a tener los grandes volúmenes de información que, sobre las personas, sus conductas, sus preferencias y sus hábitos irían acumulando las administraciones a través del uso de las TIC. Al lado del manejo de gran cantidad de datos administrativos, el creciente uso de videovigilancia, los nuevos programas de detección de personas sospechosas, etc., han ido suscitando preocupación y denuncias por la violación potencial de la privacidad. En todo caso, con la IA el salto de escala es evidente.

En los últimos años, han coincidido tres grandes procesos que explican la explosión que significa la IA: el enorme incremento del poder computacional que hace posible la instalación por todas partes de chips miniaturizados; el desarrollo de algoritmos cada vez más sofisticados; y, por último, pero no por ello menos importante, un gran salto adelante en el acceso a volúmenes de datos inimaginables hace algunos años, procedentes de cualquier sector y área de intervención (Nowotny, 2024). Por lo tanto, la consecuencia de todo ello

es que nuestro presente está cada vez más condicionado por sistemas tecnológicos que deciden sobre temas muy significativos de la vida cotidiana de las personas. Como se ha sido señalado reiteradamente en los años recientes, la IA y los sistemas automatizados de decisión pueden llegar a condicionar aspectos tan relevantes como la determinación de quién debe ser beneficiario de las ayudas públicas, la identificación de riesgos en la concesión de créditos, el control del rendimiento en el trabajo, la planificación de la movilidad, o la prevención del crimen. Cada una de esas decisiones, fundamentadas en algoritmos, acaban beneficiando y perjudicando a distintos individuos y grupos sociales. No son, por lo tanto, decisiones estrictamente técnicas, las que están en juego, sino decisiones que generan costes y beneficios, y que tienen por tanto un contenido político (AlgoRace, 2022). Debería preocuparnos, en definitiva, de quién decide los parámetros de la decisión, desde qué legitimidad, con qué criterios y con qué consecuencias.

Los expertos reunidos por la Unión Europea para definir la IA propusieron esta formulación:

Los sistemas de inteligencia artificial (IA) son sistemas de software (y posiblemente también de hardware) diseñados por humanos que, dado un objetivo complejo, actúan en la dimensión física o digital percibiendo su entorno mediante la adquisición de datos, interpretando los datos estructurados o no estructurados recogidos, razonando sobre el conocimiento, o procesando la información, derivada de estos datos y decidiendo la mejor acción o acciones a realizar para lograr el objetivo dado. Los sistemas de IA pueden utilizar reglas simbólicas o aprender un modelo numérico, y también pueden adaptar su comportamiento analizando cómo se ve afectado el entorno por sus acciones anteriores.²

Al margen de algunas afirmaciones del concepto que aún no son del todo operativas, lo cierto es que el estadio de desarrollo actual de la IA ha alcanzado niveles muy significativos en la evolución de los sistemas de decisión automática, incorporando mecanismos de aprendizaje automático o “machine learning”. Para poder desplegar sus capacidades, esta tecnología necesita disponer de un conjunto de datos etiquetados con los que pueda analizar sus distintos componentes y definir el tipo de problema con el que se enfrenta, proponer decisiones al respecto y también predecir su evolución futura. El establecimiento de agrupaciones o perfiles que permitan simplificar la heterogeneidad real existente permite así trabajar con grandes cantidades de datos de

² High-Level Expert Group on Artificial Intelligence (2019). A definition of AI: Main capabilities and disciplines. European Commission <https://www.aepd.es/sites/default/files/2019-12/ai-definition.pdf>

una manera manejable y analíticamente eficiente. En todo el proceso, resulta clave la labor humana.

El etiquetado consiste en colocar el dato en categorías específicas. Ejemplos de ello pueden ser las variables personales, sociales, educativas y económicas que describen a una persona; los elementos de su análisis médico que permiten conocer su estado de salud; o el historial financiero o penal de una persona. Las etiquetas permitirán que el sistema analice esas variables y sus interrelaciones, pudiendo extraer de ello criterios de decisión o predecir comportamientos futuros.

La otra manera de gestionar el gran volumen de datos con la que opera la IA consiste en establecer algunas variables y darle un peso o valor a cada una de ellas, para así acabar obteniendo una valoración global a partir del cálculo final. Así como los sistemas de etiquetado permiten clasificar y procesar grandes cantidades de datos, los algoritmos son la base sobre la que se construye el entramado analítico que permitirá establecer relaciones entre informaciones y alcanzar resultados en forma de decisiones. Se trata de que máquinas entrenadas por humanos sean capaces de tomar decisiones de forma autónoma. Tales decisiones tendrán, en muchos casos, impactos potenciales enormes sobre los propios humanos. Algunos ejemplos extremos de esto último son aquellos sistemas de IA que permiten decidir quién es merecedor de una ayuda pública o de un crédito, qué presos presentan más riesgo de reincidencia en periodos de libertad condicional, quién debe ser considerado un sujeto de riesgo para la seguridad pública o, incluso, qué vidas deben priorizarse en caso de accidentes de vehículos autónomos o en las decisiones de las llamadas ‘armas inteligentes’.

Lo significativo, desde un punto de vista político, es entender que lo que acaba determinando en buena parte el resultado final es una labor humana que establece criterios, parámetros, etiquetas, con las que el sistema va a operar. Esa labor no está exenta de sesgos, conscientes o inconscientes, de prejuicios que forman parte del recorrido cultural y social de las personas implicadas, de subjetividades a la hora de dar distintos pesos a distintas variables, o a la hora de establecer la selección de grupos. Este tipo de cuestiones y de cautelas no son ninguna novedad en la evolución de los estudios de ciencia y tecnología, ni en la propia evolución de la IA³. Lo que estamos diciendo es que así como la intervención humana es necesaria, al mismo tiempo, exige incorporar un principio de precaución sobre los posibles efectos negativos no predecibles en este

³ Existe un Inventario Global de Guías de IA y ética que recoge más de 160 guías hasta el 2020 <https://inventory.algorithmwatch.org>. Más recientemente es destacable el informe de European Digital Rights, “Beyond Debiasing: Regulation AI and its inequalities” <https://edri.org/our-work/if-ai-is-the-problem-is-debiasing-the-solution/>. Ambos estudios referenciados en AlgoRace, 2022

momento, y ser conscientes que esa misma y necesaria intervención humana genera otros problemas asociados a los sesgos, prejuicios y discriminaciones en los procesos de automatización decisional que surjan de la utilización de la IA, como ha podido comprobarse en distintos casos (Ponce del Castillo, 2023).

El uso de los modelos de lenguaje (*Large Language Models*) en las nuevas y populares versiones de IA generativa han potenciado enormemente su uso y han generado la sensación de que sus capacidades son ilimitadas. Pero los límites siguen siendo significativos y en muchos casos no queda claro cuál es la lógica que hay detrás de los procesos decisionales generados (Ekbja, 2024). Algunos de los casos en los que se ha utilizado la IA para intervenir en procesos decisionales se han convertido en ejemplos de lo que no debería ocurrir. Así ocurre, por ejemplo, con el caso de las evaluaciones de alumnos en el Reino Unido en plena pandemia, con las ayudas sociales para la infancia en los Países Bajos, o las tarjetas sociales en Serbia (Ponce del Castillo, 2024). En todos estos casos, los más perjudicados fueron precisamente los más vulnerables, poniendo de relieve la evidente falta de neutralidad de lo que aparentemente se presenta como una mejora de la eficacia y eficiencia de las políticas públicas.

Se impone, pues, la necesidad de una regulación democrática del desarrollo de esta tecnología, una regulación que permita establecer los límites en su aplicación sin que ello signifique frenar su avance en direcciones que humanamente consideremos deseables. Ese ha sido el propósito, por ejemplo, de una regulación pionera como la Ley de Inteligencia Artificial de la Unión Europea, una ley centrada en la identificación y la categorización de los riesgos de la IA que prohíbe las aplicaciones que planteen ‘riesgos inaceptables’, establece criterios y obligaciones para los proveedores e implementadores de aplicaciones consideradas de ‘alto riesgo’, fija obligaciones de transparencia para aquellas que se consideran de ‘riesgo limitado’, y posibilita el uso gratuito de las aplicaciones de ‘riesgo mínimo o nulo’. Aunque celebrada por su carácter pionero, sin embargo, las organizaciones de la sociedad civil que promovieron el impulso de esta legislación, articuladas en la llamada Red Europea de Derechos Digitales (EDRi), advierten sobre la excesiva contención de la ley ante los intereses de la industria tecnológica y sus efectos limitados en materia de protección de derechos humanos. En palabras del Asesor de Amnistía Internacional para la regulación de la IA, Mher Hakobyan:

“La Ley de IA ofrece sólo unas protecciones limitadas a las personas afectadas y marginadas. No prohíbe el uso temerario ni la exportación de draconianas tecnologías de IA al mismo tiempo que tampoco garantiza una protección igual a las personas migrantes, refugiadas y solicitantes de asilo. Del mismo modo,

carece de disposiciones adecuadas sobre rendición de cuentas y transparencia, lo que probablemente exacerbará los abusos contra los derechos humanos”.⁴

CAMBIOS EN LOS PROCESOS ELABORACIÓN DE POLÍTICAS PÚBLICAS Y EN LAS RELACIONES DE PODER

A estas alturas, resulta evidente que las oportunidades de la IA para la hechura de políticas son múltiples y de gran calado. En Smart City Expo, por ejemplo, se presentan múltiples aplicaciones posibles de la IA en campos como el (re)diseño de las redes de transporte público, la gestión del agua, la gestión de los residuos y el cuidado de los ancianos. Desde su conocimiento acumulado sobre temas clave en el policymaking, el análisis de las políticas públicas está abriendo un importante campo de análisis y de reflexión orientado a la identificación de los desafíos y las estrategias que los poderes públicos deberían asumir ante la emergencia de esta tecnología. Huyendo de la fascinación acrítica que se manifiesta en algunos foros, el análisis de las políticas públicas no puede obviar los enormes desafíos y oportunidades que la IA representa para la acción de gobierno. Queremos aquí poner el acento en los potenciales impactos de esta tecnología en un aspecto clave de la elaboración de las políticas como es la distribución de recursos cognitivos entre los actores y la capacidad que se deriva de incidencia sobre los procesos de agenda-setting.

Resulta evidente que ni la incorporación de la lógica ni de los instrumentos derivados de la IA en el *policymaking* no son retos o problemas meramente técnicos. Tanto la generalización del cambio tecnológico y digital, como ahora la difusión de IA, tienen un impacto muy significativo en el proceso de formación de las políticas públicas. Los esquemas analíticos habituales de las políticas y decisiones públicas parten de la existencia de problemas colectivos sobre los que se quiere intervenir, bien por la presión de los propios afectados o interesados, bien por la voluntad de hacerlo por parte de las instituciones públicas implicadas. Todo ello genera escenarios de debate, conflicto y negociación con actores que canalizan y representan intereses, utilizando distintos recursos para ello. Cada tema en debate, cada política sectorial, tiende a generar su propio espacio en el que se debaten intereses e ideas, constituyendo una “comunidad” propia de esa política específica, con su propio lenguaje y circuitos de información (Stone, 1988).

Pues bien, la difusión y generalización de las TIC, su total integración en la cotidianidad, y ahora la generalización del uso de la IA, están abriendo y alte-

⁴ Declaraciones recogidas en la web de Amnistía Internacional. <https://www.amnesty.org/es/latest/news/2024/03/eu-artificial-intelligence-rulebook-fails-to-stop-proliferation-of-abusive-technologies/> Consultado el 10 de junio de 2024.

rando significativamente ese escenario. Las posibilidades de acción directa, de movilización *online*, de producción de contenidos, de búsqueda de información a escala internacional, de influencia en la propia producción de noticias, hace menos necesaria la articulación en entidades, asociaciones o grupos para poder actuar en los procesos vinculados a las políticas públicas. Se multiplican los actores potenciales, se diversifican sus intereses, se redistribuyen sus recursos. Y todo ello redunda en una mayor incertidumbre e impredecibilidad, tanto sobre los procesos de configuración de las políticas como de sus resultados decisionales.

La red, el acceso a las potencialidades de la IA, altera la distribución de costes para la acción colectiva y la distribución de recursos entre actores y, en consecuencia, las capacidades de incidencia en las políticas públicas. La ciudadanía en general, y por tanto cualquier grupo de afectados o implicados en procesos decisionales, cuenta con más recursos cognitivos (acceso a conocimiento en red), menos costes de organización y movilización (capacidad de identificar intereses comunes, de difundir mensajes, generar bulos o *fake news*, capacidad de comunicación), así como una menor necesidad o dependencia de recursos monetarios, de acceso a los medios de comunicación de masas y de grandes inversiones de capital para organizarse. Esto favorece lógicas organizativas menos rígidas, centralizadas y jerárquicas de la acción colectiva alterando la organización del *claim-making*, ganando capacidad de impacto en la conformación de la agenda pública, y, en consecuencia, de la agenda de los poderes públicos. En ciertos casos, pueden contar quizás menos los intereses y su nivel de formalización organizativa, y más la capacidad de establecer momentos relacionales potentes que marquen o determinen la agenda e influyan en las instituciones y sus actores.

Estaríamos pues ante una notable alteración de lo que sería el mapa de actores y de sus recursos, un tema clave, como ya indicó Lowi (1972), quien mostró cómo la distinta configuración de alianzas y conflictos entre actores terminaba caracterizando y diferenciando a unas políticas públicas de otras, haciéndolas, de alguna manera, más previsibles. Sin embargo, los nuevos sistemas informacionales en red y la nueva redistribución de recursos cognitivos e informacionales convierten a los procesos de conformación de las políticas en mucho más complejos e impredecibles. No es extraño pues que la sensación general es que ha aumentado la incertidumbre, y ello genera una mucha mayor complejidad tanto sobre los diagnósticos como en relación con las alternativas y su viabilidad técnica y social.

Una de las dimensiones más evidentes en las que se perciben las consecuencias del nuevo escenario configurado por Internet y la IA lo tenemos en cómo se han alterado las agendas y la conformación de los problemas que

acaban considerándose como “problemas públicos”. No es ajeno a ello el gran cambio que supone el manejo del “timing” o la mayor o menor capacidad de la sociedad civil de incidir en esos procesos al resultar menos dependiente de los *gatekeepers* mediáticos y de los intermediarios políticos. No todos los problemas sociales se convierten necesariamente en problemas públicos, es decir, en objetos de controversia política y de posible punto de arranque de una política pública (Kingdon, 1984). La definición de un problema público es esencialmente política. Para poder calificar un problema social como “problema público” es necesario que exista una demanda procedente de grupos sociales determinados, que ello de lugar al desarrollo de un debate público y que se genere un cierto grado de conflicto entre los grupos sociales organizados y las autoridades políticas. Se trataría, por así decirlo, de la conversión de una “demanda social” en “necesidad pública”, catalogada como tal. En todo ese proceso, los actores (promotores, actores institucionales, emprendedores políticos...) jugaban un papel esencial. Y, algunos de esos actores no sólo demostraban su fuerza logrando impulsar ciertos temas, sino también bloqueando que otros *issues* llegaran a formar parte de las “preocupaciones sociales”. En ese contexto de *agenda-setting*, se ha venido considerando que el debate sobre la definición de problema y la inclusión en la agenda se articulaba esencialmente en torno a los movimientos sociales, los medios de comunicación y los procesos institucionales de toma de decisiones.

En efecto, temas como la intensidad del problema (el grado de importancia que se da a las consecuencias del problema, tanto a nivel individual como colectivo); el perímetro o la audiencia del problema (es decir, el alcance de sus efectos negativos sobre los diferentes grupos sociales que se ven implicados en el mismo, la localización geográfica de tales efectos negativos y el desarrollo del problema en el tiempo); o la novedad del problema (es decir, su no cronicidad o su no reiteración); la urgencia del problema (que habitualmente facilita la apertura de una “ventana de oportunidad”), se han considerado extremadamente relevantes a la hora de evaluar las probabilidades de que un tema o conflicto social pudiera acabar incorporándose a la agenda pública y acabara desencadenando una política pública. Pues bien, la presencia de nuevos instrumentos tecnológicos que permitan presentar evidencias respaldadas por procesos de IA puede reforzar la capacidad de ciertos actores de penetrar en la agenda de configuración de problemas y de políticas de respuesta. Al ampliarse hasta difuminarse los espacios y arenas de intermediación, la interacción se produce de manera aparentemente caótica y agregativa, con flujos poco predecibles y con capacidades de impacto que no pueden, como antes, relacionarse con la fuerza del actor o emisor de la demanda, sino con su grado o capacidad para conseguir distribuir el mensaje, presentarlo con el formato

y la fuerza de las evidencias esgrimidas, y conseguir así alianzas que vayan mucho más allá de su “hábitat” ordinario.

¿Son más o menos equilibradas las relaciones de poder resultantes? El desarrollo tecnológico podría llevarnos a pensar en nuevos escenarios de democratización de acceso a la información y a los conocimientos técnicos, así como de influencia en la agenda de las políticas públicas. Sin embargo, no podemos obviar que la estrategia de la digitalización ha permitido a algunas grandes corporaciones hacerse con una cantidad enorme de datos sobre múltiples aspectos de la vida (por ejemplo, sobre los actos cotidianos de las personas, sus relaciones sociales, y sus preferencias y sus formas de consumo, entre muchos otros aspectos) ocupando así una posición preminente no sólo en el mercado sino en el conjunto de la vida pública y en una escala global. De hecho, podríamos decir que el problema más importante al que nos enfrentamos en relación a la IA y a sus potencialidades es el control casi absoluto que tienen las grandes corporaciones tecnológicas sobre su desarrollo y evolución. Y, por tanto, sobre los condicionantes que ello genera en su utilización masiva en todo tipo de procesos.

Un ejemplo de ello lo tenemos en el poder desmesurado que se confiere a aquellos encargados del etiquetado de los datos y del diseño de los algoritmos sobre los que las máquinas fundamentan sus procesos de aprendizaje y las decisiones que adoptan. Hablamos, pues, de un escenario clave para la conformación de políticas públicas, que no es solo más complejo e incierto, sino también, potencialmente mucho más desequilibrado como consecuencia de la concentración de las capacidades de manejo y análisis de información en las ya mencionadas grandes corporaciones privadas. Resulta necesario pues identificar patrones que desvelen las relaciones de poder en aspectos clave del *policymaking*. Los datos y los algoritmos en los que se basa la IA pueden reproducir importantes sesgos sociopolíticos, reforzando las posiciones de poder de aquellos que controlan la tecnología, y debilitando la de los grupos sociales más vulnerables y marginalizados, a menudo también más expuestos a la manipulación de la información (Innerarity, 2024).

LIDERAR EL CAMBIO TECNOLÓGICO: ESTRATEGIAS GUBERNAMENTALES ANTE LA IA

Las reflexiones que hemos desgranado en este artículo permiten identificar una serie de desafíos fundamentales derivados de la aplicación de la IA en los procesos de elaboración de políticas públicas y en la acción de gobierno. Queremos resaltar a modo de conclusión cuatro temas clave, ineludibles desde nuestro punto de vista en cualquier agenda pública en materia de IA.

El primero hace referencia al papel del conocimiento en la hechura de las políticas públicas y, concretamente, al riesgo del renacimiento de una visión tecnocrática de este proceso. La IA representa como hemos visto enormes oportunidades desde el punto de vista de la acumulación y la capacidad de interpretación de todo tipo de datos, ampliando de este modo las posibilidades de fundamentar las decisiones políticas en evidencia empírica. Sin desdeñar la importancia del conocimiento científico en la toma de decisiones, sin embargo, sabemos desde hace décadas que el desafío de las políticas públicas no es meramente técnico (la acumulación de datos, el diagnóstico de los problemas, la anticipación de efectos o su evaluación científica), sino que es, también, o fundamentalmente, un desafío político: qué tipo de valores e intereses se promueven en las decisiones públicas, quién gana y quién pierde con ellas. No es lo mismo incorporar conocimiento a partir del uso de los recursos de Internet y de la IA que un conocimiento surgido de la socialización y el debate público (Collins, 2024). La IA, en definitiva, puede acompañar el proceso de toma de decisiones aportando una capacidad inédita de diagnóstico y de anticipación de posibles escenarios, conectando datos y realizando análisis complejos probablemente impensables hasta el momento, pero no debería sustituir en ningún caso el debate público, basado en valores e intereses y orientado por una aspiración de transformación social.

El segundo tema que consideramos necesario abordar se refiere a los equilibrios entre valores públicos y privados, y concretamente al riesgo de profundización de las dinámicas de privatización de las políticas públicas. Lo hemos señalado anteriormente: el desarrollo de la IA, liderado por grandes corporaciones privadas que concentran cantidades ingentes de datos digitalizados y los dispositivos tecnológicos que los explotan, otorga a estas corporaciones un poder desproporcionado de influencia en los procesos políticos. El riesgo de ‘*agency capture*’ de las políticas públicas es evidente, lo que conlleva un desafío de fortalecimiento de las capacidades tecnológicas de las administraciones públicas, desde donde establecer una relación más equilibrada con respecto a los actores privados. Sin ello, corremos el riesgo que la IA promueva la colonización del sector público por parte del sector privado, un sector privado que, gracias a posibilidades inéditas de acumulación y análisis de datos, fortalece sus posibilidades de incidencia sobre las funciones de ‘timoneo’ de las políticas públicas, y ya no sólo de ‘remo’ (Ramíó, 2020).

El tercer tema trata de las relaciones entre los poderes públicos y la ciudadanía y, concretamente de los riesgos de profundización de las desigualdades de acceso a bienes y servicios públicos a través de los sesgos que, como hemos visto, el etiquetado de datos y los algoritmos reproducen y acentúan. Los ejemplos de este tipo abundan, y señalan cómo la IA puede llegar

a representar retrocesos importantes en términos de discriminación racial, social y de género. Lo advertía Margaret Mitchell, antigua responsable del departamento de ética de la IA Google, para quien el problema principal es que “las personas a las que más perjudica la IA no deciden sobre su regulación”⁵; se trata, por el contrario, de un desarrollo tecnológico socio-demográficamente muy sesgado, protagonizado por hombres (muy) jóvenes, blancos y asiáticos, con una formación exclusivamente tecnológica y no humanista, que muestran a menudo escasa sensibilidad por las implicaciones sociales de los avances tecnológicos. Por ello, la proactividad de los poderes públicos en el terreno de la IA exige contrarrestar este tipo de sesgos a través del fortalecimiento de equipos multidisciplinares y socialmente diversos e inclusivos, donde se tengan en cuenta las distintas implicaciones del uso de estas aplicaciones tecnológicas.

El último tema que creemos necesario resaltar hace referencia a la cuestión crítica de la transparencia y de la rendición de cuentas (Hamon et al., 2024). Los algoritmos se convierten a menudo en cajas negras, basadas en datos masivos de procedencia difusa, cuyos ‘razonamientos’ son frecuentemente incomprensibles por parte de sus propios diseñadores, porque trabajan con una elevada autonomía. Sus propios diseñadores advierten de la posibilidad que produzcan alucinaciones y accidentes, con riesgos enormes cuando se aplican en campos altamente sensibles sobre la vida de las personas. ¿Cómo o por qué las máquinas deciden sobre nuestra privación de libertad o sobre si somos meritorios para la percepción de una vivienda pública, por ejemplo? Si la eclosión de Internet supuso importantes avances (aunque con límites y riesgos evidentes) en el campo de la transparencia administrativa, la IA, con el uso de algoritmos complejos diseñados por parte de corporaciones privadas, puede hacer de los procesos de toma de decisiones algo mucho más opaco e ininteligible para la mayoría de la población. La exigencia de transparencia y de rendición de cuentas sobre las máquinas es ineludible. La EU AI Act (EU Commission, 2021) establece una iniciativa pionera para obligar a la transparencia en la IA generativa, evitando que pueda haber suplantación de personas, objetos o situaciones sin explicitar su manipulación.

Se deriva de todo ello la necesidad de una actitud proactiva por parte de los poderes públicos, que deben asumir el desafío de promover un uso inteligente y democrático de la IA. Una actitud que debe comportar más regulación pública, el desarrollo de capacidades tecnológicas y competencias institucionales, y la creación de órganos especializados, con composición interdisciplinaria, que

⁵ <https://elpais.com/tecnologia/2023-11-27/margaret-mitchell-las-personas-a-las-que-mas-perjudica-la-inteligencia-artificial-no-deciden-sobre-su-regulacion.html> Consultado el 11 de junio de 2024.

elaboren códigos éticos, garanticen la contextualización de datos, examinen algoritmos y evalúen críticamente sus aplicaciones y resultados. La ya mencionada iniciativa reguladora de la Unión Europea, cuyo primer redactado se aprobó a finales del 2023 y que está pendiente de aprobación por parte del Parlamento y de la ratificación de los estados miembros, plantea por primera vez una estructura regulativa en la que se incorporan cautelas, se establecen distintas categorías de riesgos y se perfilan actuaciones explícitas de los poderes públicos al respecto (*El País*, 2023).

Al final, la duda que podemos plantear es si, como cuestiona Nowotny, existe una *inteligencia artificial inteligente* (Nowotny, 2021). En algunos casos, los humanos pueden decidir que un conocimiento es mejor que quede dormido para conseguir así que el proceso que se está siguiendo sea más resiliente o consiga mejores resultados. Los sistemas tecnológicos de toma de decisiones están pensados para cumplir objetivos muy claramente determinados. No le resulta fácil a la IA funcionar en la ambigüedad. Es muy capaz de responder a preguntas o a instrucciones, pero le es muy difícil tener la cultura o el fondo de experiencia suficiente para dejar en suspenso alternativas o propuestas que, si bien son factibles, pueden acabar siendo contraproducentes. La sabiduría está en el espacio que hay entre el individuo y la comunidad, entre lo que podría hacerse, lo que conviene no hacer y lo que no es aún previsible. Aquella sabiduría que quiere evitar el control de todo lo que acontece para salvaguardar algo superior.

La difuminación de la responsabilidad nos debería asimismo preocupar, ya que en el *core* del sistema democrático se sitúa la importancia de la atribución de responsabilidades por las decisiones que se toman en nombre de todos. Necesitamos saber cuál es el camino que ha conducido a un algoritmo a tomar una decisión. La explicabilidad de una decisión que afecta a personas, que se toma en el marco de una política pública surgida de unos poderes democráticos, no puede ser meramente técnica (Cantens, 2023). Y precisamente la inexplicabilidad de las decisiones de la IA es una de las principales barreras para reforzar su rol de asistencia a los decisores públicos. Las decisiones políticas pueden ser menos precisas y acotadas que las que tomaría un proceso de IA, pero sabemos que en muchos casos esas decisiones se han tomado buscando aquellos equilibrios necesarios para que la coalición de actores e intereses que sustentan ese criterio mantenga su apoyo y legitimidad. Y en muchos casos ese recorrido se basa en lo que se llaman “ambigüedades constructivas” forjadas alrededor de objetivos compartidos (Nowotny, 2021).

REFERENCIAS

- AlgoRace (2022). *Una introducción a la IA y la discriminación algorítmica para movimientos sociales*, <https://www.algo-race.org/2023/10/27/informe-i-una-introduccion-a-la-ia-y-la-discriminacion-algoritmica-para-movimientos-sociales/>
- Cantens, T. (2024). “How Will the State Think With the Assistance of ChatGPT? The challenges of generative artificial intelligence for public administrations”, *AI & Society*. 10.1007/s00146-023-01840-9
- Collins, H. (2024). Why Artificial Intelligence needs sociology of knowledge, *AI & Society*, <https://doi.org/10.1007/s00146-024-01954-8>
- Crawford, K. (2023). *Atlas de la IA. Poder, política y costes planetarios de la Inteligencia Artificial*, Ned, Barcelona.
- Dente, B.-Subirats, J. (2013). *Decisiones Públicas*, Ariel, Barcelona.
- Ekbja, H.R. (2024). “In Humans we trust: rules, algorithms and judgement” en Ponce del Castillo, *Op. cit.*, pp. 47-57.
- El País* (2023). “Una ley pionera con muchos interrogantes” <https://elpais.com/tecnologia/2023-12-10/una-ley-pionera-para-una-tecnologia-con-muchos-interrogantes-las-claves-de-la-regulacion-de-la-ia-en-la-ue.html>
- European Commssion (2021). “Proposal for a Regulatory laying down harmonized rules on Artificial Intelligence”, <https://digital-strategy.ec.europa.eu/en/library/proposal-regulation-laying-down-harmonised-rules-artificial-intelligence>
- Hamon, R. et al. (2024). Generative AI Transparency: Identification of Machine Generated content, JRC Publications Repository, European Commision, <https://publications.jrc.ec.europa.eu/repository/handle/JRC137136>
- Innerarity, D. (2024). Inteligencia artificial y democracia, Informe Unesco, 2024 <https://www.unesco.org/es/articles/inteligencia-artificial-y-democracia>
- Nowotny, H. (2024). “In AI we trust: power, illusion and the control of predictive algorithms” en Ponce del Castillo, *Op. cit.*, pp. 59-63.
- Nowotny, H. (2022). *La fe en la Inteligencia Artificial*, Galaxia Gutenberg, Barcelona.
- Ponce del Castillo, A. (ed.). (2024). *Artificial Intelligence, Labour and Society*, European Trade Union Institute (ETUI), <https://www.etui.org/publications/artificial-intelligence-labour-and-society>
- Poster, M. (2007). “Internet Piracy as Radical Democracy”, en Dahlberg, L. -Siapera, E., (eds.), *Radical Democracy and the Internet. Interrogating Theory and Practice*, Palgrave-macmillan, Nueva York, pp. 207-225.
- Ramió, C. (2019). *Inteligencia Artificial y Administración Pública. Robots y humanos compartiendo el servicio público*. Barcelona: Catarata.

Stone, D. (1988). *Policy Paradox. The Art of Political Decision Making*, Norton, Nueva York.