

**Gobernanza ética de la IA y la robótica.
Construyendo posibilismo tecnológico desde
interacciones éticas**

*Ethical governance of AI and robotics. Building
technological possibilism from ethical interactions*

Jaime Rodríguez Alba¹

Universidad Autónoma del Estado de México, Toluca, México

Universidad Internacional de la Rioja, La Rioja, España

 <https://orcid.org/0000-0003-4612-3498>

albajaime@hotmail.com

Recibido: 07/10/2023

Aceptado: 29/10/2023

DOI: 10.5281/zenodo.10451864

RESUMEN

Las tecnologías exponenciales generan problemáticas éticas que han sido abordadas desde las tradiciones dominantes (principalismo, deontologismo, utilitarismo), con cierto olvido de las éticas del cuidado. Recuperar estas éticas permitiría pensar otras posibilidades de la IA. Entre ellas: un diseño que involucre elementos de vulnerabilidad o el horizonte de una robótica inclusiva. Pensar la gobernanza de la IA y la robótica no sólo desde estas éticas exige complementar la gobernanza *de* las mismas con la consideración de los horizontes éticos en dicha gobernanza implícitos. Si las éticas dominantes abonan alternativas de solución para los impactos de las tecnologías exponenciales, las éticas del cuidado —por cuanto ponen énfasis en la formación del carácter— son especialmente significativas para pensar la interacción entre tecnologías y la formación de la persona moral.

Palabras clave: Posibilismo tecnológico, gobernanza, ética del cuidado, gobernanza ética.

ABSTRACT

Exponential technologies generate ethical problems that have been approached from the dominant traditions (principalism, deontologism, utilitarianism), with a certain neglect of the ethics of care. Recovering these ethics would allow us to think of other possibilities for

¹ Doctor en Filosofía por la UNED (Universidad Nacional de Educación a Distancia) de España, Magister en Sociología por la UCM (Universidad Complutense de Madrid). Actualmente trabaja en el desarrollo de sistemas éticos integrales para la implementación de modelos de aplicación de la ética en organizaciones. También

AI. Among them: a design that involves elements of vulnerability or the horizon of an inclusive robotics. Reflecting on the governance of AI and robotics not only from the perspective of these ethics requires complementing their governance with consideration of the ethical horizons implicit in such governance. If the dominant ethics provide alternative solutions to the impacts of exponential technologies, the ethics of care — insofar as they emphasize character formation — are especially significant for thinking about the interaction between technologies and the formation of the moral person.

Keywords: Technological possibilism, governance, ethics of care, ethical governance.

Introducción

El desarrollo tecnológico exponencial que vivimos —caracterizado por una imbricación de innovaciones tecnológicas en segmentos o interacciones mínimas e imprevisibles pero con un crecimiento exponencial, más que lineal (López Portillo Romero, 2018)— es una oportunidad fabulosa para consolidar sociedades más humanizadas — justas, equitativas, que atienden al libre juego de las capacidades humanas, etc. Pero la tecnología por sí sola no genera la mismas. La racionalidad tecnológica, o tecno—científica, no se traduce en racionalidad social, política o moral de modo automático. El hecho tecnológico es ambivalente: tanto puede generar sociedades deseables como distópicas. Todo artefacto tecnológico supone una carga de valores en su propio diseño, así como impacta en los valores de la sociedad que lo alumbró, modificándola de modos diversos.

Frente al cambio tecnológico hay que tomar postura política desde una sociedad que articulada y conscientemente contempla planes, programas y políticas públicas orientadas a generar transformaciones tecnológicas que permitan crecientes dosis de libertad, igualdad y humanidad. Ni catastrofismo tecnológico, ni optimismo tecnológico, sino posibilismo tecnológico (Broncano, 2000): analizar las suturas valorativas

desarrolla temas de Filosofía Moral y Política orientados a fundamentar dichos modelos. Es coordinador de la RIEP (Red Iberoamericana de Ética Pública).

de los artefactos tecnológicos para orientarlos con fines sociales asumibles. No olvidando que toda transformación tecnológica, como bien supiera ver la sociología durkheimiana, altera las pautas de socialización, al generar disritmias en los tiempos de evolución de los sistemas y desacoples en las funciones sociales.

Marx señaló en sus apuntes (*Grundrisse*) que en cada momento histórico los seres humanos se plantean aquellos problemas que pueden resolver. En nuestros días las transformaciones tecnológicas son de tal magnitud que pudiera suceder que se planteen problemas que la humanidad, *quae tale*, no podrá resolver. Problemas que pueden mutar la naturaleza misma de tal humanidad. No en vano emergen con fuerza los discursos sobre Cyborgs, transhumanismo, posthumanismo, etc. Acá no entraremos en esto, más bien apelamos a un horizonte hermenéutico centrado en categorías de gestión pública. En concreto buscamos iluminar la noción de una gobernanza ética, como diferente de una gobernanza de la ética, en el sentido que posteriormente dilucidaremos.

Desarrollo

Comencemos por usos supuestamente benignos de artefactos tecnológicos —expresión que usamos en un sentido amplio— como las aplicaciones de Inteligencia Artificial (IA) en la gestión pública, para mostrar la necesidad de consideraciones éticas. Durante la gestión de la actual pandemia, en Reino Unido (Digital Future Society, 13 de Agosto de 2021) se decidió usar un sistema algorítmico de calificación (Ofqual) que promediando desempeños individualizados de los alumnos pudiera ofrecer una ponderación respecto a la calificación otorgada por los docentes en situaciones de virtualidad educativa. El resultado de tal software fue que los alumnos de colegios en contextos de vulnerabilidad

social salieron peor puntuados. Un alumno con buen desempeño en un centro educativo en contexto vulnerable resultaba así comparativamente peor valorado que uno con desempeño medio en otros centros. Además de las protestas ante tal información, por sus implicaciones discriminativas, se sigue investigando el algoritmo para indagar posibles sesgos en su construcción. Algo similar sucedió con el uso, desde 1998 en EEUU, del software COMPAS (Correctional Offender Management Profiling for Alternative Sanctions) que calcula el riesgo de incidencia de los delincuentes. El uso de este tipo de softwares en materia penal es controvertido doctrinalmente (Romeo Casabona, 2018), pues supone nociones del derecho penal cuestionadas desde visiones iusalternativistas o más garantistas. El uso de COMPAS se hizo célebre por el caso Loomis. Loomis es una persona de color que ante un robo fue apresado, mediante aplicación de COMPAS el juez decidió encarcelarlo. Ante el reclamo por el debido proceso, puesto que no se conocía el algoritmo del software, la Corte Suprema de Wisconsin desestimó el recurso alegando los derechos de propiedad de la empresa. Situación análoga, respecto a los derechos de acceso a la información pública, se produce en el recurso contencioso administrativo presentado en España por la negativa del Consejo de Transparencia a ofrecer el código del programa BOSCO con el que la administración española decide quién es beneficiado por el bono social eléctrico. Recurso presentado (CIVIO, 2 de julio de 2019) ante la petición de una ciudadana española a la que se le negó el bono social argumentando la valoración ofrecida por BOSCO.

Ante este tipo de problemáticas emerge la solución normativa vinculada a conceptos como transparencia algorítmica, trazabilidad algorítmica y afines, así como la inserción de estas temáticas en la infraestructura normativo—institucional relativa a transparencia, acceso a la información pública y buen gobierno.

Hay sin duda usos virtuosos de softwares vinculados de un modo u otro a la IA en la justicia. Tal es el caso del programa PROMETEA (Estévez, Fillotrani y Linares Lejarraga, 2020) de la Fiscalía de Buenos Aires. Dicho programa arroja sentencias posibles que siempre valida un juez —siendo su palabra definitiva—, pero además agiliza un montón de trámites judiciales y administrativos. PROMETEA trabaja sobre pilares éticos asociados a la transparencia y bajo el paradigma de la supervisión humana. Si bien no es una Inteligencia Artificial en el sentido más profundo del término.

La doctrina jurídica está respondiendo de modo profundo a los retos normativos, proponiendo por ejemplo que se consideren los algoritmos como reglamentos (Boix Palop, 2020). De este modo los mismos serán publicados y publicitados como normas jurídicas, siendo posible el recurso ante los mismos, lo que se traduce en las garantías respecto a su uso en la justicia, la gestión pública, etc. Además, también se debiera, según estas visiones, introducir requisitos vinculados a la programación y uso en la gestión pública en el procedimiento administrativo común.

Desde otros flancos jurídicos se está abordando la exigencia de tratamiento constitucional de los neuroderechos, como actualmente en Chile y su proceso constituyente, o se esgrimen herramientas diversas, como las Cartas de derechos digitales de España (ver tabla 1), destinadas a establecer derechos de las personas y organizaciones en los procesos de profunda transformación digital.

Tabla n°1
Categorías posibles de derechos digitales

Derechos de libertad	Derechos de igualdad	Derechos de participación y conformación del espacio público	Derechos del entorno laboral y empresarial	Derechos digitales en entornos específicos
Derecho a la identidad en el entorno digital.	Derecho a la igualdad y a la no discriminación	Derecho a la neutralidad de Internet.	Derechos en el ámbito laboral.	Derecho de acceso a datos con fines de

Derecho a la protección de datos.	en el entorno digital.	Libertad de expresión y libertad de información.	La empresa en el entorno digital	archivo en interés público, fines de investigación científica o histórica, fines estadísticos, y fines de innovación y desarrollo.
Derecho al pseudonimato.	Derecho de acceso a Internet.	Derecho a recibir libremente información veraz.		Derecho a un desarrollo tecnológico y un entorno digital sostenible.
Derecho de la persona a no ser localizada y perfilada.	Protección de las personas menores de edad en el entorno digital.	Derecho a la participación ciudadana por medios digitales.		Derecho a la protección de la salud en el entorno digital.
Derecho a la ciberseguridad.	Accesibilidad universal en el entorno digital.	Derecho a la educación digital.		Libertad de creación y derecho de acceso a la cultura en el entorno digital.
Derecho a la herencia digital.	Brechas de acceso al entorno digital	Derechos digitales de la ciudadanía en sus relaciones con las Administraciones Públicas		Derechos ante la inteligencia artificial. Derechos digitales en el empleo de las neurotecnologías.

Fuente: Elaboración propia en base a la *Carta de Derechos Digitales* (2020).

Iniciativas como PROMETEA (Berchi, 4 de marzo de 2020) son prometedoras para una justicia que ha de ser más eficaz y eficiente, razón por la que se siguen en otros países (Colombia aplica también este algoritmo abierto y trazable, y España lo está considerando). Pero PROMETEA no usa Deep Learning. Estamos así más ante una gobernanza de la ética que una gobernanza ética. Mientras la primera hace alusión al modo como podemos regular, comprender, apropiar socialmente, etc., el artefacto tecnológico en base a su uso (impacto, responsabilidad, etc.), la

segunda apela a un sentido más profundo, preguntando por los escenarios futuros y por el modo como se puede comprender la ética misma de las máquinas.

El modo como se piensa la ética de la IA viene signado por el lugar que la agencia humana tiene en la misma. Así (Marín García, 2019) para una IA débil —más automatización de procesos que inteligencia en un sentido cognitivo fuerte, esto es, una inteligencia artificial local— las cuestiones éticas se aproximan al paradigma de la responsabilidad humana en el diseño e implementación de los diseños tecnológicos. En cambio una IA fuerte —lo que algunos autores denominan IA general, aquella que puede sustituir y superar las capacidades humanas, al menos en procesos cognitivos vinculados al manejo de datos— involucra otro tipo de tematizaciones éticas, más ligadas al estatus del sujeto moral, a las temáticas de la vincularidad, las interacciones entre seres humanos y robots, la emergencia de un paradigma transhumano, etc. Tanto en unos casos como en otros las innovaciones precisan un abordaje complejo y multidisciplinar.

¿Por qué hablamos de gobernanza? ¿No sirven ya las usuales regulaciones por los mecanismos tradicionales? Por supuesto siguen siendo necesarias, pero a nuestro juicio la dimensión misma del fenómeno tecno-científico hace necesario superar paradigmas en exceso rígidos, partidarios de una concepción tecnicista y menos dialógica que el enfoque de una gobernanza multistakeholder. Gobernanza que tomamos como modelo para la comprensión del fenómeno tecnológico.

La noción de gobernanza surge durante los años setenta en el terreno de las ciencias del gobierno y la administración pública. Impulsada por la crisis de los modelos de bienestar ante los procesos de globalización. La capacidad de gobernar desbordaba las estructuras tradicionales del gobierno y la gubernamentalidad. A esto se añade que las sociedades exigían cada vez más presencia en los procesos de

autogobierno. Los procesos de gestión que definen los estados mutan así hacia nuevas problemáticas multidimensionales en las que lo ético resulta central, situándose junto al dominio técnico un dominio ético-político de corte deliberativo (Aguilar Villanueva, 2013). Gobernanza remite (Kooinan, 2003) a las interacciones de actores varios –públicos, privados, del tercer sector, del mundo académico, etc.– de cara a resolver problemas o crear oportunidades sociales, considerando las instituciones involucradas como contextos para las interacciones orientadas al proceso de gobernar y generando fundamentos normativos al efecto.

De tal modo, la gobernanza de la ética en los diseños tecnológicos como la IA y la robótica ha de considerar los impactos, pero también elementos presentes en el diseño. Aunque no precise entrar en la cuestión de la naturaleza misma del sujeto moral que puede ser, en los casos de una Inteligencia Artificial General –aquella que pudiera reemplazar al ser humano en la toma y ejecución de decisiones–, de enorme importancia, habida cuenta del desplazamiento de la centralidad de la persona humana y sus características. En el terreno de la gobernanza de la ética encontramos enormes retos. El primero la presencia de usos tecnológicos que pueden desplazar al sujeto humano –pensemos en los autos eléctricos, la automatización de procesos industriales, etc. Pero también porque pueden ser usados con fines de eficientizar recursos públicos olvidando otros valores centrales en la gestión pública, como la inclusión, la equidad, etc. Piénsese en el caso de AMS (ArbeitsMarktService) de Austria que, usado para modelar perfiles laborales ha mostrado ser un algoritmo eficaz, pero que aleja principios de inclusión social de su funcionamiento, o en el caso del uso militar –drones, soldados aumentados, etc.– que pueden garantizar eficacia a costa de otros valores como el respeto por la vida. Sin duda esta gobernanza ética choca con el problema de las iniciativas de gobernanza por bloques geopolíticos (ver figura 2)

Tabla n°2

Iniciativas de gobernanza de la ética para IA

U.E.	OCDE y G20	EE.UU.	CHINA	RUSIA
Grupo de expertos de alto nivel. IA Confiable (legal, ética y robusta). Requisitos: Agencia y supervisión humana. Robustez y seguridad. Privacidad y gobernanza de datos y transparencia. Diversidad y no discriminación. Bienestar socio-ambiental. Responsabilidad y rendición de cuentas	Guía ética para la IA: ecosostenible, concorde con los DDHH, transparente y responsable, robusta y segura. Los gobiernos han de invertir en IA confiable, apostar por una transición justa Principios G20: Crecimiento incluso, sostenibilidad, bienestar, humanidad, equidad, seguridad, responsabilidad, robustez, etc.	IA con valores norteamericanos (Trump). Libertad, DDHH, estabilidad institucional, propiedad intelectual, “oportunidad para realizar sueños”, “devoción por ayudar a la gente”. Comprensible. Confiable. Segura. Responsable.	Principios de Pekin sobre IA: hacer el bien, humanidad, control de riesgos, apertura, responsabilidad. Gobernanza: optimizar el empleo, cooperación, moderación, planeamiento a futuro, adaptación, Armonía	Estrategia Nacional para el Desarrollo de la IA. No abordan directamente cuestiones éticas, salvo determinación de responsabilidades, propiedad intelectual. Considera que la regulación puede inhibir la innovación. Suscribe G20

Fuente: Elaboración propia en base a R. Carrillo (2020) y O. Klein (2020).

Desde el momento que la gobernanza dice interacción nos situamos en el terreno de ciertas características éticas que se han de satisfacer para contemplarla como buena gobernanza. En el caso de la gobernanza de los fenómenos tecnológicos, signados por la complejidad y el dinamismo, se hace especialmente relevante la consideración no sólo de la dimensión técnica, sino también ética. De un lado atender a los impactos –lo que remite a principios dialógicos de inclusión, simetría y reciprocidad entre interlocutores, así como veracidad–, por otro profundizar en la comprensión de la naturaleza del sujeto moral. De tal modo, la

gobernanza de la IA y la robótica antes que acogerse a un paradigma reductivo o sin tematización suficiente de la ética debería al tiempo problematizar la misma ética que está considerada en el proceso de gobernanza (gobernanza ética). Máxime si tenemos en cuenta que la ética está presente no sólo en el uso, sino también en los diseños mismos.

Cuando revisamos tematizaciones éticas de la IA se observa una cierta reducción a tópicos como los sesgos morales, prejuicios, afectación de derechos humanos, etc. Sin duda problemáticas que desde una ética de principios o una ética deontológica son más que relevantes. En algunos casos, como por el ejemplo los asociados a la valoración de la responsabilidad implícita en los automóviles autónomos, se aprecia una hegemonía de enfoques utilitaristas. Tal es el caso del experimento Moral Machine (Awad, et al, 2018), que a escala global ha rastreado las decisiones trágicas en manos de un automóvil automático –salvar una vida o varias, según sean vidas de niños, ancianos, mujeres, etc. Este experimento se apoya en la suposición de una suerte de aritmética moral según la cual el común denominador de las valoraciones decisionales tomadas –que el mayor número considere como lo menos malo determinada opción– se convertiría en una suerte de criterio moral a enseñar a un automóvil automático. El experimento ha mostrado también la presencia de diversos criterios valorativos: en oriente se prioriza más la vida del anciano que en occidente, por ejemplo.

Estos casos, al igual que el uso de ciencia de datos en redes sociales, política, etc., permiten afirmar ciertos rasgos éticos imperantes (Calvo, 2020): positivismo digital –el dato moral como dato masivo sin problematizar la construcción social de los datos; consecuencialismo – enfoque utilitario centrado en el cálculo de las consecuencias de la acción; convencionalismo –consideración de la corrección en base a la mayor disposición a la corrección de la misma, por grupo humano; o

paternalismo –estimación según la cual la IA puede arrojar mejores decisiones y en ese modo orientar la acción humana.

Muchas de las problemáticas asociadas a la IA –sesgos, discriminación, usos de vigilancia, negación de derechos personalísimos, etc.– vienen por no desnaturalizar las aplicaciones y diseños tecnológicos. Considerar que una IA puede ser más justa, por ejemplo, que un juez, o gobernar mejor que un político, está en consonancia con los procesos de artificialización (Nurock, 2020), mediante los que los sesgos, discriminaciones, abusos, etc., que caracterizan las interacciones sociales pasan a ser ejecutados por las tecnologías, con el agravante que en las mismas se invisibilizan o naturalizan.

En lo político y social (Calvo, 2019), la IA implica retos como: el jaque a la autonomía moral y del sujeto político por la autonomización del complejo tecno–científico; la sustitución de la democracia por administración tecnocrática; o las perturbaciones en las interacciones sociales que producen sentido para la existencia de la persona.

La gobernanza ética, aún más que la gobernanza de la ética, exige innovación e investigación responsable, complementando los enfoques principalistas, deontológicos y utilitaristas hegemónicos con otros como el de las éticas del cuidado. Por ejemplo, una robótica interactiva e inclusiva (Pons, 2020) supone principios y objetivos orientados a mejorar justicia y bienestar de las poblaciones, democratizar las prácticas del cuidado, fomentar la autonomía y el desarrollo de las personas. Todo ello en un marco axiológico orientado por la seguridad de las innovaciones –la reflexividad crítica y el control sobre los riesgos que suponen– y la gradualidad de las mismas –para garantizar la no exclusión, así como la conformación de nuevos espacios de interacción entre entidades humanas y no humanas que sean apropiadas y apropiables por las comunidades.

Considerada la IA y la robótica desde la ética del cuidado (Nurock, 2020) hemos de redefinir en qué sentido una máquina cuida. Partir de la vulnerabilidad implícita a la dimensión del sujeto moral para así analizar cómo afecta a la identidad y desarrollo moral el uso de la IA y la robotización en los más variados ámbitos –robots en salud, ciencia de datos, internet de las cosas, etc. Por supuesto considerar la salvaguarda de la autonomía, por la vía de los neuroderechos y la vigilancia epistemológica sobre los procesos tecnológicos, pero también apuntalar todas aquellas tecnologías orientadas a equilibrar autonomía y dependencia.

El sujeto moral, considerado desde las éticas del cuidado, es un sujeto que se mueve en una tensión entre autonomía y dependencia, superando así visiones distorsivas que al reducir lo moral a la autonomía conducen a excluir una amplia gama de sujetos morales (niños, personas dependientes, con discapacidades, seres animales, etc.) Algunas propuestas en IA como la de Antonio Damasio caminan en esta dirección. En vez de reducir el lugar de lo humano en la IA, quizá lo más interesante sería dotar a la IA de vulnerabilidad, mediante diseños tecnológicos blandos que permitan que las máquinas, al enfrentarse a la precariedad constitutiva de su propio diseño, pudieran evolucionar hacia una inteligencia con sentimientos, al modo como ha hecho la inteligencia humana (ver Fundación Innovación Bankinter, s/f). Pero esto es algo problemático en sí, pues una evolución interna a las máquinas es algo que excede los modos como conocemos ha sido el proceso evolutivo humano (Broncano, 2000). Como consideran los transhumanistas, los seres artificiales no tienen por qué plantearse como análogos a los seres humanos ni a las máquinas actualmente existentes. En cualquier caso, el planteamiento de Damasio da cabida a escenarios de vulnerabilidad, superando así enfoques puramente principalistas, deontológicos y en exceso racionalistas.

El uso de IA en los procesos de gestión pública –big data en el ciclo de las políticas públicas, por ejemplo– sin duda generará impactos positivos, siempre que se establezcan marcos de gobernanza ética para considerar riesgos y oportunidades. La reducción deontológica del problema inducirá un exceso de organismos de control, la utilitarista conducirá a desestimar elementos fundamentales como el reconocimiento y actuación cuidadosa con el daño moral, y el principalista correrá el riesgo de ocultar todo un ámbito de problemas que sólo pueden emerger en la consideración prudencial y deliberativa de la situación concreta, contando con todos los actores (inclusividad) y permitiendo todas las narraciones de los mismos (simetría y reciprocidad) para garantizar la orientación hacia los bienes comunes.

En general las tecnologías exponenciales (López Portillo Romano, 2018) permiten posibilidades inmensas al respecto de los bienes comunes –por ejemplo los objetivos de desarrollo sostenible. Evaluar los impactos de las mismas será tarea ética a desarrollar colaborativamente por actores diversos. Supondrá también una nueva concepción del espacio público y su garante, el estado. Si han de mutar las capacidades y competencias laborales de las personas, prestando creciente atención a dimensiones como el pensamiento crítico o la flexibilidad (Kai-Fu Lee, 2019), no es menos cierto que la robotización y aplicación masiva de la IA supondrá un profundo cambio del mercado de trabajo. Para esto soluciones como la renta básica parecen ser insuficientes, puesto que no logran atinar en la médula misma de las mutaciones de la subjetividad moral y política de la persona, en tanto siguen de algún modo pensando al sujeto moral bajo el paradigma del individualismo liberal (Kai-Fu Lee, 2020). Una gobernanza ética sí permitiría al menos cristalizar horizontes de desarrollo estratégico para poner la IA y la robótica al servicio de las comunidades humanas.

El estado resultará en este proceso un actor central, y ejemplar, aunque no único. Pero se atisban falencias importantes, en especial en los países en desarrollo o subdesarrollados (Oszlak, 2020). Las propias dinámicas de la gestión pública inducen la imperiosa necesidad de superar los silos organizacionales (Oszlak, 2020) que hacen complicado desarrollar una visión estratégica y colaborativa del fenómeno tecnológico en sus usos e impactos públicos. La cuarta revolución industrial está ya en marcha situando una serie de problemáticas en la agenda de los diseños institucionales para el sector público, mostrando estos empero dinámicas y resistencias complejas (Pardo Beneyto & Abellán López, 2020).

Frente a las tecnologías exponenciales se precisa un estado exponencial (Oszlak, 2020) que oriente las innovaciones tecnológicas hacia los bienes comunes. Esto requiere, entre otras cosas, (Oszlak, 2020): desarrollar enfoques abiertos y flexibles; la experimentación, innovación, evaluación e interacción permanente; y, en especial, la producción de confianza (transparencia, inclusión, eficacia, ejemplaridad pública, estado amigable, etc.) Exige también nuevos valores públicos (Ramió, 2020). Valores como el fomento del aprendizaje continuo, big data para el bien común, la sostenibilidad intergeneracional, etc. Valores que se unirán a los tradicionales de equidad, igualdad de género, respeto a la privacidad, etc., sin duda también modificados por las tecnologías exponenciales.

Finalmente, y en el escenario geopolítico, que la IA –y en general las tecnologías exponenciales– no se convierta en la última amenaza para nuestra especie requiere superar los enfoques de competencia y beligerancia entre potencias de la IA (Kai-Fu Lee, 2020), apostando por construcciones interculturales que consideren los diversos escenarios de producción de IA conforme a valores específicos (Ortega Kein, 2020), así como reforzar el lugar de los organismos multilaterales y el derecho

internacional (Robles Carrillo, 2020). Desde las propias organizaciones técnicas ya se está apostando por enfoques dinámicos, abiertos y multisakeholder, como sucede con el IEEE (Instituto de Ingeniería Eléctrica y Electrónica) y su modelo FERE (Fundamental Ethical Responsibility of Engineers), el IGF (Internet Governance Forum) con su apuesta por la deliberación en torno a prácticas o la UIT (Unión Internacional de las Telecomunicaciones), que propone anclar gobernanza a ODS.

En todos estos espacios de deliberación debiera también sumarse a los enfoques éticos hegemónicos –deontológico, principalista o utilitarista– nuevos enfoques como el de las éticas del cuidado. El fundamento para pensar la gobernanza como gobernanza ética es autoevidente: los desarrollos tecnológicos inducen (Coeckelberg, 2020) una mutación de las nociones éticas fundamentales (autonomía, responsabilidad, etc.) afectando a los modos de toma de decisión, así como a las narrativas que conforman la subjetividad humana. Las tecnologías exponenciales pueden mejorar sustancialmente las condiciones de vida del ser humano si se controlan democrática y deliberativamente (Hughes, 2004). O pueden, por el contrario, dar pie a una suerte de escenario neofeudal (Moreno & Jiménez, 2018) marcado por la exclusión creciente y la pugna entre feudos geopolíticos de producción de IA.

Conclusión

La gobernanza ética de la IA y la robótica habrá de trabajar en la construcción, para estas tecnologías exponenciales y disruptivas, de un “ecosistema cibernético” (Calvo, 2021, p. 38). Comenzando por el gobierno de datos –petróleo del que se nutren las IA–, creando infraestructuras éticas (códigos, comités, líneas éticas, procesos participativos, etc.)

orientadas a la gestión, monitorización y cumplimiento de todo lo involucrado en la producción y uso de datos masivos. Esta infraestructura se apoyará en un “entorno comunicativo y afectivo” (Calvo, 2021, p. 38)

Pero a su vez, y sin entrar en la problemática transhumanista relativa a la superación de la condición humana (mortalidad, precariedad, deficiencia cognitiva, etc.), se precisa avanzar, mediante interacción en redes cívicas, de expertos, empresarios, políticos, etc., en la consideración de la ética misma que está implícita en las máquinas. No tanto, aunque el tema es de la máxima relevancia, en lo relativo al posible diseño moral de una Superinteligencia (Bostrom, 2016) que desarrollara a partir de una voluntad coherente extrapolada –aquella que reduce la incoherencia con valores de partida– desde los propios algoritmos autocreados, consideraciones que pudiéramos estimar como morales desde una perspectiva humana. Ya en esta misma interpretación de una inteligencia superior a la humana –cuyas razones nos fueran oscuras– que sigue la lógica, existe la suposición ética de que la coherencia lógica es la que define el criterio de producción moral, lo que es sin duda estimar que el error no tiene valor ético. Esto chocaría sin duda con diversas fuentes de moralidad, como la cristiana que considera el error (pecado) como constitutivo de la condición misma de moralidad humana (salvación). Por esto, habida cuenta de la pluralidad valorativa humana, se ha de contemplar la gobernanza en sí misma ética, la gobernanza ética, desarrollando instancias de consideración metacognitiva en la que se expliciten los horizontes y experiencias morales mismos desde los que se está contemplando la problemática.

Una gobernanza ética así entendida permitiría además tematizar un problema central en los diseños normativo–institucionales de abordaje de la IA y sus aplicaciones: el disenso valorativo de partida. No es casual que la apuesta china por la IA, al partir de otros contextos valorativos, sea

diferente al occidental. Considerar perspectivas interculturales en la construcción de gobernanza, huyendo de la tendencia a imponer como criterio normativo la visión occidental –estimando la legitimidad de otras filosofías políticas presentes en los diseños y usos de la IA (Pak–Hang Wong, 2016)–, generaría, a nuestro entender, afrontar retos en la pugna bipolar en IA.

En síntesis, la consideración ética no ha de reducirse a enfoques normativos, mayormente hegemónicos en occidente. Ha de apuntalar espacios de deliberación, de intercambio de narrativas morales, construcción intercultural, abordaje de los conflictos éticos y morales derivados de los diseños tecnológicos. Por esto junto a la gobernanza de la ética ha de profundizarse en la gobernanza ética misma. Gobernanza como metodología de construcción de ecosistemas éticos, pero consideración también de la racionalidad de fines y el sustrato de experiencias morales que permiten comprender la compleja hermenéutica del fenómeno ético y sus vínculos con las tecnologías exponenciales.

Referencias

- Aguilar Villanueva, L. F. (2013). *El gobierno del gobierno*. INAP.
- Awad, E. et al. (2018). The Moral Machine experiment. *Nature* 563, 59–64.
- Berchi, M. (4 de marzo de 2020). La inteligencia artificial se asoma a la justicia pero despierta dudas éticas. *El País*. https://elpais.com/retina/2020/03/03/innovacion/1583236735_793682.html
- Boix Palop, A. (2020). Los algoritmos son reglamentos: la necesidad de extender las garantías propias de las normas reglamentarias a los

- programas empleados por la administración para la adopción de decisiones. *Revista de Derecho Público: Teoría y Método*, 1, 223–270
- Bostrom, N. (2016). *Superinteligencia. Caminos, peligros, estrategias*. Teell Editorial
- Broncano, F. (2000). *Mundos artificiales. Filosofía del cambio tecnológico*. Paidós
- Calvo, P. (2019). Democracia algorítmica: consideraciones éticas sobre la *datafización* de la esfera pública. *Revista CLAD Reforma y Democracia*, (70), 5–30.
- Calvo, P. (2020). Etificación. En Pizzi, J. y Cenci, M. (coord.), *Global Observatory for Social Pathologies* (36-94). Federal University of Pelotas.
- Calvo, P. (2021). El gobierno ético de los datos masivos. En Jon Rueda (ed.), *Tecnologías socialmente disruptivas. Dilemata, Revista Internacional de Éticas Aplicadas*, (34), 31–49
- CIVIO (2 de julio 2019). Que se nos regule mediante código fuente o algoritmos secretos es algo que jamás debe permitirse en un Estado social, democrático y de Derecho. *Civio: novedades*. <https://civio.es/novedades/2019/07/02/que-se-nos-regule-mediante-codigo-fuente-o-algoritmos-secretos-es-algo-que-jamas-debe-permitirse-en-un-estado-social-democratico-y-de-derecho/>
- Coeckelberg, M. (2020) *AI Ethics*. MIT Press
- Digital Future Society (13 agosto 2021) *Gobernanza algorítmica: proceda con precaución*. <https://digitalfuturesociety.com/es/experts-perspective-are-algorithms-deciding-for-us/>
- Estevez, E., Fillottrani, P., y Linares Lejarraga, S. (2020). *PROMETEA: Transformando la administración de justicia con herramientas de inteligencia artificial*. BID. <http://dx.doi.org/10.18235/0002378>

Fundación Innovación Bankinter 21 de mayo de 2020). ¿Es posible crear
máquinas que sientan?

[https://www.fundacionbankinter.org/noticias/es-posible-crear-
maquinas-que-sientan/?_adin=02021864894](https://www.fundacionbankinter.org/noticias/es-posible-crear-maquinas-que-sientan/?_adin=02021864894)

Gobierno de España (2021). *Carta de derechos digitales* [PDF] Plan de
recuperación, transformación y resiliencia

[https://www.lamoncloa.gob.es/presidente/actividades/Document
s/2021/140721-Carta_Derechos_Digitales_RedEs.pdf](https://www.lamoncloa.gob.es/presidente/actividades/Documentos/2021/140721-Carta_Derechos_Digitales_RedEs.pdf)

Hughes, J. (2004) *Citizen Cyborg. Why Democratic Societies Must Respond
to Redesigned Human of the Future*. Westview Press

INAÉTICA (s/f) *Ética en la era de la inteligencia artificial y el internet del
todo*. <https://ianetica.com/>

Kai-Fu Lee (2019). La inteligencia artificial y el futuro del trabajo: una
perspectiva china. En *El trabajo en la era de los datos*. BBVA

Kai-Fu Lee (2020). *Superpotencias de la inteligencia artificial*. Planeta

Kooiman, J. (2003). *Governing as Governance*. Sage Publications

López Portillo Romano, J. R. (2018). *La gran transición. Retos y
oportunidades del cambio tecnológico exponencial*. FCE

Marín García, S. (2019). Ética e inteligencia artificial. *Cuadernos de la
Cátedra CaixaBank de Responsabilidad social corporativa*, (42).

Moreno, L. y Jiménez, R. (2018). *Democracias robotizadas*. La Catarata

Nurock, V. (2020). ¿Puede prestar cuidados la Inteligencia Artificial?
Cuadernos de relaciones laborales, 38(2), 217-229

Ortega Klein, A. (2020). *Geopolítica de la ética en la Inteligencia Artificial*.
Real Instituto El Cano

Oszlak, O. (2020). *El estado en la era exponencial*. INAP

Pak-Hang Wong (2016). Responsible innovation for decent nonliberal
peoples: a dilemma?, *Journal of Responsible Innovation*, 3(2), 154-
168. DOI: 10.1080/23299460.2016.1216709

- Pardo Beneyto, G. y Abellán López, M. A. (2020). Tecnologías de la información y comunicación e innovación institucional. En B. Aldeguer Cerdá y G. Pastor Albaladejo, *Democracia, gobierno y administración pública* (288-300). Tecnos
- Pons, J. L. (ed.) (2020). *Inclusive Robotics for a Better Society*. Springer Natura Switzerland
- Ramió. C. (2020). Los valores públicos del futuro en el contexto de la Administración 4.0. En D, Vicher, *Códigos y preceptos éticos* (pp. 47-84). INAP
- Robles Carrillo, M. (2020). La gobernanza de la inteligencia artificial: contexto y parámetros generales. *REEI (Revista Electrónica de Estudios Internacionales)*, 24(1), 42-62. DOI: 10.17103/reei.39.07
- Romeo Casabona, C. M. (2018). Riesgo, procedimientos actuariales basados en inteligencia artificial y medidas de seguridad. *Revista de Derecho, Empresa y Sociedad (REDS)*, 13, 39-55. Disponible en: <https://dialnet.unirioja.es/descarga/articulo/6859383.pdf>