

Inteligencia artificial y riesgos para la democracia

por Giuseppe Conte

estudioso de la epistemología de la complejidad

Resumen

La reciente difusión de las denominadas inteligencias artificiales generativas ha vuelto a poner en primer plano cuestiones relativas a la influencia de la tecnología en los sistemas sociales y políticos. Este artículo trata sobre dos clases de riesgos que surgen en la relación entre estas nuevas tecnologías y los sistemas democráticos.

Lo que en ciencia empírica se llama 'data' (datos),
siendo en realidad elegido arbitrariamente de la naturaleza de las hipótesis ya formuladas,
podría llamarse más honestamente 'tomado' (capta)
(Laing, 1985, pág. 33)

Premisa

Es necesario comenzar con definiciones, aunque concisas, de los dos protagonistas de este documento. En lo que respecta a la IA, sin entrar en los detalles técnicos, nos limitamos, para nuestro argumento, a una descripción resumida de sus principios de funcionamiento. Estos sistemas se basan esencialmente en el uso de "estadística aplicada". Una gran cantidad de datos, agregados de varias fuentes de información, es procesado por sofisticados algoritmos que identifican números, frecuencias, regularidades, correlaciones. Antes, durante y después de este paso hay una actividad humana continua encaminada a que los resultados finales se ajusten a los deseos iniciales.

Por democracia entendemos en cambio aquella forma de gobierno donde la soberanía es ejercida, directa o indirectamente, por el pueblo.

Diferentes axiomas

Los dos conceptos recién descritos, IA y democracia, ya chocan en sus respectivos supuestos básicos.

La IA es producto de una cultura técnica específica que asume implícitamente que existe una Verdad objetiva, que reside en datos que representarían información. Que esta información pueda ser procesada, a través de métodos cuantitativos y estadísticos, eliminando los errores de juicio y decisión humana, para destilar la Verdad misma. Precisamente estos mismos supuestos, a menudo

inconsciente, han permitido que el término "inteligencia" se utilice de manera despreocupada pero sugerente y casi supersticiosa en esta área. (Sini, 2023)

La democracia se basa estructuralmente en la construcción social de caminos 'viables', expresados mayoritariamente y orientados hacia el bienestar y la prosperidad de la comunidad. Verdades provisionales para ser continuamente sometidas a escrutinio a través de los diversos instrumentos de consulta popular.

"Quien considere que la verdad absoluta y los valores absolutos son inaccesibles al conocimiento humano no debe considerar posible solo su propia opinión, sino también la opinión de los demás [...] Por lo tanto, el procedimiento dialéctico adoptado por la Asamblea Popular o por el parlamento en el creación, procedimiento que se realiza a través de discursos y réplicas, ha sido propiamente reconocido como democrático".

(Kelsen, 2010, 146)

Esta apertura a las verdades posibles es el sine qua non para la existencia del pluralismo, el fundamento de la idea democrática.

"El pluralismo de la verdad nos abre los ojos, en primer lugar, a la contingencia: no tengo una visión de 360 grados; nadie lo tiene. En segundo lugar, y esta es la noción más audaz, la verdad es pluralista porque la realidad misma es pluralista, al no ser una entidad objetivable. Los sujetos somos igualmente parte de ella. No somos sólo espectadores de lo Real, sino también coactores e incluso coautores de él. Esta es precisamente nuestra dignidad humana".

(Panikkar, 1990)

Así, mientras la democracia se mantiene en el nivel del pluralismo complejo, el diálogo y el debate, incluso los largos y turbulentos, siempre con resultados provisionales, la IA se utiliza para reducir, simplificar, acelerar, optimizar, hacer más eficiente el proceso de toma de decisiones, guiar por un algoritmo no afectado por humanos, considerado defectuoso por naturaleza.

En este artículo abrazamos abiertamente los axiomas democráticos, rechazando aquellos que podemos llamar algocráticos (Pozza, 2020).

De los axiomas, o paradigmas, por usar un término muy querido por la filosofía de la ciencia, se derivan miradas y posturas específicas con las que nos relacionamos con el mundo y por tanto se derivan las consecuencias de nuestros actos.

En nuestra opinión, los paradigmas que subyacen en el uso actual de la IA producen o amplifican diversos riesgos, algunos de los cuales detallaremos a continuación. Seguirán ideas que potencialmente podrían mitigarlos.

No neutralidad de la técnica.

Somos conscientes de que el título de este párrafo puede despertar las primeras resistencias. De hecho, es una opinión común, extendida hasta el punto de convertirse en un mantra con propiedades ansiolíticas, que: "la tecnología es neutra, lo importante es cómo la usas".

Aquí afirmamos exactamente lo contrario: toda tecnología no es neutral. No es neutral porque toda tecnología se inserta en la relación mutuamente generativa entre los seres vivos y el medio ambiente, modificándolo. Actuar sobre esta relación significa influir en ambos polos, incluido, en el caso del hombre, el conocimiento de la relación misma. De ahí la naturaleza multifacética de la tecnología, donde el poder y la liberación siempre son también una pérdida. (Recuento, 2021)

Por tanto, es fundamental aceptar el hecho de que no 'utilizamos' la técnica, permaneciendo separados de ella, sino que cambiamos continuamente gracias a las herramientas técnicas que utilizamos en nuestras relaciones.

A continuación, enumeramos dos clases de riesgo que se dividen en intrínsecos a las tecnologías de IA, es decir, vinculados a su lógica operativa, y extrínsecos, es decir, vinculados a su posible usar.

riesgos inherentes

Pérdida de la biodiversidad cultural.

El proceso de 'destilación' de información operado por las IA, descrito sumariamente al principio, se traduce en el producto esperado por los diseñadores: una única salida. En esto, estas tecnologías difieren profundamente de otros algoritmos y programas, por ejemplo, los que se utilizan en los motores de búsqueda web, destinados a producir múltiples salidas correspondientes a una determinada entrada.

Las respuestas que brindan las IA son producto de una larguísima cadena de procesamiento de datos que comienza con las materias primas. Los llamados datos fuente, sometidos a largos y complejos procesos industriales, mínimamente transparentes, que producen el equivalente a un alimento ultraprocesado.

Este producto sintético se distribuye a nivel mundial y, a menudo, está disponible en el idioma del usuario. el final. Contribuye, día tras día, a la construcción de una cultura única y globalizada.

Cultura que incorpora de forma invisible los axiomas de la IA expresados al principio (existencia de una realidad objetiva, contenido informativo de los datos, máquinas como agentes racionales, etc.)

Lo que se acaba de describir puede representar, por lo tanto, una enorme aceleración del proceso de globalización y homologación que ya ha estado activo durante décadas. Una pérdida de biodiversidad cultural que inevitablemente tiene un fuerte impacto en el pluralismo y, por tanto, en la democracia.

Es razonable prever que esta homologación se articulará tanto verticalmente como históricamente – tendiendo a la producción de narrativas históricas homogéneas, tanto en sentido horizontal – geográficamente – tendiendo a la difusión de una sola cultura a nivel global.

“Hemos sufrido y sufrimos tanto por el fanatismo político, religioso y cultural que legítimamente tenemos sed de comprensión universal. Un ejemplo típico es el síndrome de la aldea global. Si bien es noble en sus intenciones, me parece otro digno sucesor de la mentalidad colonial. El colonialismo cree en el monomorfismo de la cultura, lo que significa que, en última instancia, solo hay una civilización: y ahora aquí está la unificación del mundo en una aldea global”.

(Panikkar, 1990)

De hecho, estos aspectos estarán muy influenciados por los datos de entrada de las IA, portadoras de pensamientos, valores, ideas y culturas que se encuentran en el centro de cada curva de distribución estadística de los datos suministrados como entrada a los algoritmos.

Riesgo de retroalimentación positiva, la segunda ola

Lo que se acaba de escribir es la base de las consideraciones posteriores.

Mientras escribimos, la web está literalmente inundada con los productos de las inteligencias artificiales generativas nacientes. En prácticamente todos los campos del conocimiento se producen cada vez más artefactos gracias a estas tecnologías.

Estos productos, día tras día, están poblando –o mejor dicho contaminando– las fuentes de datos que alimentarán a futuras versiones de las mismas Inteligencias Artificiales.

El resultado, como se comprende fácilmente, será la alteración de las premisas estadísticas que otorgan una fiabilidad, aunque relativa, a los resultados. Premisas que implican una correspondencia entre datos y fenómenos reales.

Cuando una salida de una IA entra en las entradas de otra IA, se desencadenará el conocido fenómeno de 'retroalimentación positiva', estudiado en cibernética. Un dato ultraprocesado se convertirá en entrada de un proceso que eventualmente producirá un dato ultraultraprocesado, y así sucesivamente.

El resultado práctico será una alteración de la base de información, con una 'compresión' de los datos hacia el centro de las curvas de distribución.

En otras palabras, las fuentes de datos originales de la IA serán cada vez más homogéneas y uniformes, por lo que los resultados se percibirán como cada vez más precisos e inequívocos. Y probablemente todo esto también sea bienvenido.

Estos resultados, a su vez, influirán en las ideas de los usuarios de IA, confirmando las premisas implícitas y alimentando el conocido fenómeno de la 'causalidad circular'.

En consecuencia, la alteración del contexto informativo en el que nos movemos cada uno de nosotros no puede dejar de repercutir en el pluralismo, pilar central de todo sistema democrático.

Riesgos externos a la tecnología

Riesgo de manipulación

La construcción de pensamientos, para cada uno de nosotros, depende en gran medida del entorno social en el que estamos inmersos y del nicho evolutivo e informacional en el que nos movemos.

Cada acción que tomamos proviene de nuestros valores, ideas, convicciones, creencias. Cada uno de nuestros productos está imbuido de estas premisas. Los productos de las tecnologías de IA no pueden ser una excepción a esto.

“Abandonemos, pues, la presunción de ser seres racionales, ya que todo dominio racional en el que cada uno de nosotros se mueve en cada instante se constituye como un dominio de coherencia operacional por la aceptación de las premisas fundamentales que lo definen en un acto emocional”.

(Maturana, 2006, p. 110).

Las IA están diseñadas en un contexto económico y de mercado preciso en el que son servicios en el sentido comercial del término.

“En inteligencia artificial, los conjuntos de datos nunca son materias primas que alimentan algoritmos: son actos políticos en sí mismos. Toda la práctica de recopilar datos, clasificarlos y etiquetarlos, y luego usarlos para entrenar esos sistemas es una forma de política, lo que nos ha llevado a lo que se llama "imágenes de trabajo", representaciones del mundo hechas exclusivamente para máquinas.

(Crawford, 2021, pág. 252)

Son servicios comerciales que se prestan a los clientes (Estados, organizaciones, empresas, particulares, etc.) y que se mueven en el marco del modelo de marketing global que claramente tiende a manipular al individuo.

“[...] todo el grupo social es parte integral del universo en red: psicólogos, economistas, mercadólogos, ergonomistas, terapeutas, diseñadores web... todos contribuyen a una ingeniería social en línea con la convicción expresada por Skinner según la cual la ciencia es mejor equipada que la política para 'manejar' al hombre”.

(Besnier, 2013, pág. 122)

Con base en las necesidades generales o específicas de estos clientes, los servicios son calibrados por sus productores. Como dicta la lógica del mercado, hay entonces un proceso continuo de mejora tratando de que el servicio se adhiera a las expectativas del público y, al mismo tiempo, tratando de cambiar el significado y las expectativas de los usuarios para que se adhieran al servicio ofrecido.

“La comunicación producida por el advenimiento de la industria del sentido habría generado un cortocircuito mental [...] precisamente en la construcción de una nueva percepción del yo, de la vida colectiva, de la sociedad y de la vida misma. La construcción de la idea de vida material, de la cotidianidad individual y social, salió del terreno de las “simples” relaciones humanas para convertirse en producto de una industria que produce ganancias en la construcción del “sentido de la vida”.

Una industria que pone su "poder" a disposición de quien tiene los medios económicos para utilizarlo, ya sea una empresa que tiene que vender, ya sea un político o un directivo o un partido. Además, el advenimiento de la red y el poder de la comunicación social está ampliando enormemente las posibilidades de generar estallidos de consenso precisamente a partir de la estratificación del "sentido" de la vida y los acontecimientos.

(Bellucci, 2019, pág. 31)

Como se ha descrito, sostenemos que existe un abismo entre el enfoque de mercado descrito hasta ahora, con el que se desarrollan y gestionan estas tecnologías, y una visión democrática.

“Si las tecnologías robóticas, y no sólo ellas, fueran abandonadas a una pura lógica de poder, sustancialmente no muy diferente de la que ha «desregulado» la esfera de la economía, asistiríamos a un creciente divorcio entre humanidad y democracia, entendida como tal. en su carácter de régimen político donde el uso de cualquier medio no puede desligarse del respeto a los principios y derechos fundamentales”.

(Rodotà, 2012, p. 375)

Lo recién descrito no es un riesgo sin precedentes, sin embargo, creemos que la dinámica descrita podría crear las condiciones para su crecimiento significativo, tanto en términos de probabilidad como de impacto.

Combinación de riesgos

El resultado final, que combina el riesgo intrínseco -estructural a las IAs- con el riesgo extrínseco -el uso de la técnica en la lógica del mercado- creemos que puede conducir al crecimiento desmesurado de ese fenómeno que Alexis de Tocqueville ya denominaba "dictadura". de la mayoría" - esta vez 'guiados' por intereses privados e ideologías específicas.

Además, el potencial de la IA, sin una gobernanza adecuada, puede conducir a una exacerbación de ese aplanamiento del que ya había advertido Heidegger, utilizando la expresión “existencia inauténtica”:

“[...]. En este estado de irrelevancia e indistinción, el Si ejerce su típica dictadura. Nos lo pasamos bien y nos divertimos como nos divertimos; leemos, vemos y juzgamos de la literatura y el arte como uno ve y juzga. Nos alejamos de la masa como nos alejamos, encontramos escandaloso lo que encontramos escandaloso. El Sí, que no es allí un Ser específico sino todo (pero no como suma), decreta el modo de ser de la vida cotidiana. El Si tiene sus propias formas particulares de ser [...]. La promediabilidad es un carácter existencial del Si. En el Si, en cuanto a su ser, es esencialmente de él. Se mantiene, pues, en el medio de lo que se acuerda, de lo que se acepta y de lo que se rechaza, de lo que se acredita y de lo que se niega. Al determinar lo que es posible o lícito intentar, la medianía vela por toda excepción. Cada registro se nivela en silencio. Toda originalidad se disuelve en lo conocido, toda gran empresa se convierte en objeto de una transacción, todo secreto pierde su fuerza. La cura de la promedialidad revela una nueva y esencial tendencia del Dasein: la nivelación de todas las posibilidades del ser.

(Heidegger, 1976, pág. 163)

Propuestas para la mitigación de riesgos

Los riesgos estructurales y sociales como los descritos no pueden ser eliminados. Es razonablemente posible imaginar acciones de mitigación que intervengan sobre la probabilidad y sobre el impacto de su ocurrencia.

En este punto podemos vislumbrar la estructura común que está en el trasfondo de la dinámica destacada hasta ahora: la simplificación. 'Trivialización', término con un significado específico en el lenguaje de la epistemología de la complejidad.

La banalización como producto del proceso de selección que las IA están diseñadas para completar y que destilan la respuesta a la necesidad del usuario del vasto mar del conocimiento mundial.

De esta síntesis extrema surgen las siguientes ideas de mitigación de riesgos.

Mitigar el riesgo de pérdida de biodiversidad cultural

“Es más probable que un sistema interactúe constructivamente con el medio ambiente cuanto más diferenciada y diversificada sea su estructura interna”.

(Ceruti, 2009, p. 75)

- Ya en la actualidad, algunas grandes empresas compiten por dominar el lucrativo mercado de la IA. Es necesario establecer normas y reglamentos que limiten la difusión global de los sistemas individuales de IA y, al mismo tiempo, fomenten la proliferación de nuevos sistemas de IA.
- Hoy las Corporaciones productoras de sistemas de IA no declaran sus fuentes de datos y tampoco permiten la consulta. Se propone establecer organizaciones dedicadas a auditar y certificar estas fuentes. Se pueden emplear otros procesos de certificación para determinar la variedad en la composición de las bases de datos y para validar la cadena de transformación de datos. Como es fácil ver, se trata de propuestas que ya se han puesto en marcha en otros ámbitos, como el sector agroalimentario. Área que comparte la complejidad intrínseca, la multitud de actores y las lógicas aplicadas con los temas aquí tratados.
- La IA, para ser utilizada en áreas 'sensibles', y siempre exclusivamente en apoyo de decisiones humanas, debe ser de código abierto, y prever formas de auditoría de archivos y configuraciones.

Mitigar el riesgo de comentarios positivos

Cada producto de IA tiene cierta confiabilidad e importancia si se basa en el conocimiento humano, codificado digitalmente, no el resultado de otros algoritmos de IA.

“Al observar las capas de datos de entrenamiento que estructuran e informan los modelos y algoritmos de IA, podemos ver que la recopilación y el etiquetado de datos sobre el mundo es una intervención social y política, aunque disfrazada de acto puramente técnico.

La forma en que se interpretan, recopilan, categorizan y nombran los datos es fundamentalmente un acto de creación de mundos y perímetros, que tiene enormes ramificaciones sobre cómo la IA afecta al mundo y a las comunidades que se ven más afectadas por ella. el mito de

La recopilación de datos como una práctica benévola en la informática ha oscurecido sus operaciones de poder, protegiendo a quienes más se benefician y haciéndolos responsables de sus consecuencias”.

(Crawford, 2021, pág. 135)

Por estas razones es necesario alimentar estos sistemas exclusivamente con productos de la actividad humana. En ese sentido, se propone la creación de una 'Biocertificación', que distinga los artefactos producidos por el ser humano como una 'pegatina'. Estos objetos 'certificados' como productos humanos, junto con datos del procesamiento automático de información

(sistemas de información que no son de IA) deben ser los únicos artefactos que se pueden recopilar en los grandes datos que alimentan las IA.

Mitigar el riesgo de mercado

“Cada complejización organizacional se expresa en un aumento de variedad dentro de un sistema”.

(Edgar Morín, 2001, pág. 406)

Proponemos utilizar herramientas políticas y democráticas para mitigar este riesgo.

- Ampliación de la legislación antimonopolio en el ámbito de la IA.
- Fortalecer las herramientas y capacidades de los órganos de gobierno y supervisión existentes dentro de las organizaciones transnacionales, como las Naciones Unidas (ver apéndice).
- Definición y aplicación de estándares específicos, como ISO/IEC FDIS 42001, actualmente en desarrollo (ver apéndice).

Conclusiones

Queríamos, en estas líneas, sugerir ideas, ciertamente no soluciones listas para usar, precisamente porque la preparación de medidas adecuadas y operativas solo puede pasar por la mediación y la confrontación democrática.

El tema tratado en este artículo es extremadamente importante. En un futuro próximo, la política estará llamada a hacer frente al impacto de estas nuevas tecnologías. La apuesta es capital.

“Si el poder político ya no cumple su papel, si le falta, peor aún si se arrodilla ante poderes que le fascinan y de los que está convencido de que encierran la única verdad de nuestra época, entonces nos queda una sola opción : sustituir la política por la política, o con la reapropiación por parte de los ciudadanos, asociaciones, sindicatos, grupos institucionalizados o no, del derecho inalienable de ejercer individual y colectivamente su libertad de decisión y de

juicio. Y el alcance de nuestra participación depende nada menos que del futuro de nuestra civilización".

(Sadin, 2018, pág. 159)

Concluimos compartiendo la lista de capacidades democráticas ideales expresadas por Martha Nussbaum, evidentemente no reducible o delegable a ningún algoritmo:

- La capacidad de razonar sobre los problemas políticos que afectan a la nación, de examinar, reflexionar, discutir y sacar conclusiones sin delegar en la tradición o la autoridad.
- La capacidad de reconocer en los conciudadanos a personas con iguales derechos, por muy diferentes que sean en cuanto a raza, religión, género y orientación sexual: mirarlos con respeto, como fines, no como instrumentos a manipular en beneficio propio. .
- La capacidad de preocuparse por la vida de los demás, de comprender qué políticas son significativas para las oportunidades y experiencias de los conciudadanos, de todo tipo, e incluso de personas fuera del propio país.
- La capacidad de visualizar la variedad de problemas de la vida humana a medida que se desarrolla: pensar en la infancia, la adolescencia, las relaciones familiares, la enfermedad, la muerte y mucho más teniendo en cuenta una amplia gama de historias personales, y no solo un conjunto estadístico.
- La capacidad de juzgar críticamente a los políticos, pero sobre la base de información precisa y consciente de las posibilidades reales que tienen a su disposición.
- La capacidad de pensar en el bien de toda la nación, no en el del propio grupo local.
- La capacidad de ver la propia nación como parte de un orden mundial complejo, en el que problemas de diversa índole requieren una discusión transnacional para su solución.

(Nussbaum, 2014, pág. 42)

Bibliografía y Sitografía

Bellucci Sergio, 2019, La industria de los sentidos, Ed. Harpo

Besnier Jean-Michel, 2013, El hombre simplificado, Vida y pensamiento

Ceruti Mauro, 2009, La coacción y la posibilidad, Raffaello Cortina pag. 75)

Conte Giuseppe, 2021, Esclavos de la libertad, en Bellucci Sergio. AI-WORK La digitalización del trabajo. Libro jaca.

Crawford Kate, 2021, Ni inteligente ni artificial, The Mill

Heidegger Martin, 1976, Ser y Tiempo, Longanesi.

ISO/IEC FDIS 42001 - Tecnología de la información - Inteligencia artificial - Sistema de gestión - <https://www.iso.org/standard/81230.html>

Kelsen Hans, 2010, Democracia, El Molino

Laing Ronald David, 1985, en Keeney Bradford P., La estética del cambio, Astrolabe

Maturana Humberto, 2006. Emociones y lenguaje en la educación y la política. Ed. Eleuthera

Morin Edgar, 2001, Método 1, Raffaello Cortina

Nussbaum Martha, 2014, Sin fines de lucro. Porque las democracias necesitan cultura humanista. El molino.

Panikkar Raimon, 1990, el pluralismo de la verdad

<http://www.dhdi.free.fr/recherches/horizonsinterculturels/articles/panikkarpluralism.doc>

Pozza Giuliano, 2020, "Algocracia": cómo la inteligencia artificial amenaza la libertad humana. <https://www.agendadigitale.eu/cultura-digitale/algocrazia-perche-senza-etica-con-lartigianato-artificiale-ci-giochiamo-la-liberta/>

Rodotà Stefano, 2012, El derecho a tener derechos, Ed. Laterza

Sadin Eric, 2018, La colonización siliconada del mundo, Ed. Giulio Einaudi

Sini Carlo, 2023, Astillas de pensamiento 10: la inteligencia artificial es una superstición. <https://www.youtube.com/watch?v=qvmwWp0SZml> (enlace verificado a partir de junio de 2023)

de Tocqueville Alexis, 2007, Democracia en América, Utet

Naciones Unidas 2023, Un pacto digital mundial: un futuro digital abierto, gratuito y seguro para todos <https://www.un.org/sites/un2.un.org/files/our-common-agenda-policy-brief-gobal-digi-compact-es.pdf>

Naciones Unidas, 2023, Proporcionando orientación global sobre inteligencia artificial. https://www.un.org/techenvoy/sites/www.un.org.techenvoy/files/general/Artificial_Intelligence_Summary_PDF.pdf