

Análisis de la Resolución de la Asamblea General de la ONU sobre inteligencias artificiales, desarrollo sostenible y derechos humanos¹

Ariel Vercelli

(INHUS, CONICET / FH - UNMdP)

arielvercelli@arielvercelli.org

Mar del Plata, Argentina

Resumen: Las inteligencias artificiales (IA) están generando cambios profundos en nuestras sociedades. Hace años que forman parte de nuestra vida cotidiana: son omnipresentes, poderosas, opacas y, muchas veces, invisibles al ojo humano. De allí que su regulación sea uno de los temas relevantes dentro de las agendas tecnológicas en el plano nacional e internacional. El 21 de marzo de 2024, en el contexto del 78° período de sesiones, la Asamblea General de Naciones Unidas adoptó una resolución sobre inteligencia artificial, desarrollo sostenible y protección de derechos humanos titulada “Aprovechar las oportunidades de sistemas seguros, protegidos y fiables de inteligencia artificial para el desarrollo sostenible”. Se trata de la primera resolución de la Asamblea General sobre IA que, habiendo nacido de un proyecto de los EE.UU., logró sumar el apoyo (co-patrocinio) de más de 120 países y fue adoptada por consenso. Entre otras preguntas que se buscan responder en la ponencia: ¿cuáles son los puntos más relevantes de esta nueva regulación sobre IA?, ¿qué recomienda sobre regulaciones nacionales e internacionales?, ¿qué propone para proteger los derechos humanos y evitar brechas tecnológicas? A continuación se describe la resolución y se analiza su relación con el desarrollo sostenible y los derechos humanos. La ponencia forma parte de una investigación mayor sobre las regulaciones de las IA tanto en el plano nacional como internacional.

Palabras clave: inteligencias artificiales, Asamblea General de Naciones Unidas, desarrollo sostenible, derechos humanos.

[1] Introducción: las Naciones Unidas, el desarrollo y su posición “táctica” sobre las IA

Las inteligencias artificiales (IA) están generando cambios profundos en nuestras sociedades [1]. Hace años que forman parte de nuestra vida cotidiana: se trata de diferentes tecnologías que se podrían definir como omnipresentes, poderosas,

[1] El artículo se desarrolló gracias al apoyo del Consejo Nacional de Investigaciones Científicas y Técnicas (CONICET), la Facultad de Humanidades (FH), de la Universidad Nacional de Mar del Plata (UNMdP) y Bienes Comunes A. C. La investigación se desarrolla dentro del Instituto de Humanidades y Ciencias Sociales (INHUS, CONICET / FH - UNMdP) y el Grupo de Investigación ‘Ciencia, Tecnología, Universidad y Sociedad’ (CITEUS), OCA 347/05, FH – UNMdP.

opacas y, muchas veces, invisibles al ojo humano [2]. Las IA se están transformando en un tema existencial: favorecen cambios radicales en relación al acceso a la información/datos, la creación de valor o la gestión del conocimiento. Hace años que desde la ONU se advierte sobre la aceleración del cambio tecnológico, una especie de crecimiento exponencial, y como éste podría afectar a la población mundial. Sin embargo, a pensar de las advertencias, la posición de la ONU y sus agencias sobre las IA no deja de ser ambigua. Por un lado, depositan grandes esperanzas en que las tecnologías digitales y, en particular, las IA puedan ser las herramientas tácticas que permitan alcanzar para 2030 los 17 Objetivos de Desarrollo Sostenible (ODS)². Por el otro, advierten sobre los riesgos que acarrearán estos cambios tecnológicos para los derechos humanos: entre otros, asimetrías en el acceso/disponibilidad, brechas digitales, monopolios, violación de la privacidad y la protección de datos personales, extractivismo de datos, sesgos algorítmicos, discriminación, noticias falsas, desinformación, armas autónomas, etc. [4] [5]. De allí que las regulaciones de las IA se hayan vuelto un tema recurrente en las principales agencias de la ONU.

[2] La Resolución de la Asamblea General sobre IA y sus antecedentes inmediatos

El 21 de marzo de 2024, en el contexto del 78° período de sesiones, la Asamblea General de Naciones Unidas adoptó una resolución sobre inteligencia artificial, su regulación y la protección de derechos humanos. La resolución llevó por título “Aprovechar las oportunidades de sistemas seguros, protegidos y fiables de inteligencia artificial para el desarrollo sostenible” (A/78/L.49)³ [6]. La propuesta fue iniciada por EE.UU. y, tras varios meses de negociación, contó con el apoyo de 123 países y las relevantes no-objeciones de China y Rusia. Se trató de la primera resolución de la Asamblea General sobre temas de IA y fue adoptada por consenso⁴ [7]. Para la representante de los EE.UU., Linda Thomas-Greenfield, la principal virtud de la Resolución es abrir el diálogo sobre temas de IA a la gran comunidad mundial.

[2] En 2015 los países miembros de Naciones Unidas aprobaron la Agenda 2030 para el Desarrollo Sostenible y definieron 17 objetivos, conocidos como Objetivos del Desarrollo Sostenible (ODS). Los 17 Objetivos del Desarrollo Sostenible son globales y complementarios: 1) Fin de la pobreza; 2) Hambre cero; 3) Salud y bienestar; 4) Educación de calidad; 5) Igualdad de género; 6) Agua limpia y saneamiento; 7) Energía asequible y no contaminante; 8) Trabajo decente y crecimiento económico; 9) Industria, innovación e infraestructuras; 10) Reducción de las desigualdades; 11) Ciudades y comunidades sostenibles; 12) Producción y consumo responsables; 13) Acción por el clima; 14) Vida submarina; 15) Vida de ecosistemas terrestres; 16) Paz, justicia e instituciones sólidas; y 17) Alianzas para lograr los objetivos [3].

[3] Es parte de la aplicación y seguimiento integrados y coordinados de los resultados de las grandes conferencias y cumbres de las Naciones Unidas en las esferas económica y social y esferas conexas. Específicamente, el ítem 13 se presentó como “Facilitar el desarrollo sostenible aprovechando las oportunidades que ofrecen los sistemas de inteligencia artificial seguros y fiables - proyecto de resolución (A/78/L.49)”. En el orden del día apareció como “Adopción de medidas en relación con el proyecto de resolución” [7].

[4] La adopción de una resolución por consenso significa que el proyecto presentado no fue sometido a una votación formal (con votos afirmativos, negativos o abstenciones) sino que sobre el mismo no se han presentado objeciones por parte de los Estados Miembros [8].

Entre sus antecedentes inmediatos se destacan iniciativas nacionales y de documentos internacionales: la declaración de Bletchley, el Pacto Global sobre IA (GPAI), el Proceso de Hiroshima que firmó el G7, los principios del G20 para una IA Fiable y los principios de la OCDE sobre IA. Uno de los objetivos de la Resolución es amplificar la labor de las agencias de la ONU en temas de IA: entre otras, la Unión Internacional de Telecomunicaciones (UIT), la Organización de las Naciones Unidas para la Educación, la Ciencia y la Cultura (UNESCO), en particular, a partir de sus Recomendaciones sobre Ética de la Inteligencia Artificial [9], y el Consejo de Derechos Humanos. A su vez, se espera que la Resolución consolide las diferentes negociaciones para un pacto digital mundial y, en materia de gobernanza, ayude en la elaboración del informe final 2024 del AI Advisory Board. En rueda de prensa, la representante de EE.UU. junto a representantes de Bahamas, Japón, Países Bajos, Marruecos, Singapur y Reino Unido, consideró que la Resolución es un gran paso para gobernar las IA antes que estas tecnologías nos gobiernen a nosotros. Además, subrayó que la innovación y la regulación no son mutuamente excluyentes sino complementarias [10].

[3] Análisis de las consideraciones más relevantes sobre la Resolución de la Asamblea General

Aunque la Resolución es de lectura fácil, y está traducida en varios idiomas, por momentos su redacción se torna redundante, repetitiva y circular. A continuación, y de acuerdo a los objetivos de la investigación mayor, se describen y analizan algunos de los puntos que podrían considerarse de relevancia para el estudio de las regulaciones sobre las IA:

- **[3.1] Una definición (amplia) sobre sistemas de IA “seguros, protegidos y fiables”:** la Resolución define los sistemas de IA “seguros, protegidos y fiables” como aquellos que, perteneciendo al ámbito no militar y abarcando todo su ciclo vital⁵, se caracterizan por: [a] centrarse en las personas, [b] ser fiables, [c] explicables, [d] éticos e inclusivos, [e] plenamente respetuosos de los derechos humanos y el derecho internacional, [f] mantener la privacidad, [g] estar orientados al desarrollo sostenible y [h] ser responsables [6]. A pesar de la abundancia y diversidad de los elementos descritos, también presentes en varios documentos internacionales, sólo se trata de una definición aproximativa, amplia y ambigua sobre estos sistemas. Una definición estricta, tan necesaria en este campo, podría haber derivado en diferencias explícitas y demorado la adopción del proyecto por consenso. No obstante, la Resolución advierte sobre la urgencia de alcanzar un consenso mundial sobre qué son y cómo se desarrollarían estos “sistemas seguros, protegidos y fiables” de IA.

[5] El ciclo de vida de la IA incluye las etapas de prediseño, diseño, desarrollo, evaluación, puesta a prueba, despliegue, utilización, venta, adquisición, explotación y retirada de servicio.

- **[3.2] Las IA y su vínculo con los ODS de la agenda 2030:** La Resolución reconoce que los sistemas de IA que se consideran “seguros, protegidos y fiables” tienen el potencial de propiciar los avances hacia la consecución de los 17 Objetivos de Desarrollo Sostenible (ODS). Afirma que las IA pueden acelerar el desarrollo sostenible en sus tres dimensiones (económica, social y ambiental). A su vez, entiende que estos sistemas de IA podrían, de manera equilibrada e integrada, fomentar la transformación digital, promover la paz, salvar las brechas digitales (entre países y dentro de ellos) y, manteniendo la persona en el centro, promover y proteger el goce de los derechos humanos y las libertades fundamentales. Al respecto, la Resolución no cita casos de éxito sobre IA y los ODS, ni otras fuentes que permitan identificar como se dará esta articulación virtuosa entre las tecnologías digitales y la agenda 2030. Al respecto, es importante mencionar que en 2023 Naciones Unidas inició una evaluación de medio término sobre el grado de avance de dichos objetivos y los resultados fueron alarmantes [3]⁶.
- **[3.3.] El rechazo y no uso de sistemas de IA que violen los derechos humanos:** La resolución solicita a los Estados Miembros que se abstengan de utilizar (o dejen de utilizar) sistemas de IA que pongan en riesgo los derechos humanos. Específicamente, se invita a rechazar el diseño, el desarrollo y el uso (en realidad, todo el ciclo de vida) incorrecto o malicioso de los sistemas de IA que, entre otros elementos, puedan ser caracterizados por: no tener salvaguardas adecuadas, contrariar el derecho internacional, dificultar los avances hacia los ODS de la Agenda 2030, socavar el desarrollo sostenible (económico, social y ambiental), agrandar las brechas digitales (entre los países y dentro de ellos), reforzar las desigualdades y los sesgos estructurales, conducir a la discriminación, minar la integridad de la información y el acceso a ella, debilitar la protección, la promoción y el goce de los derechos humanos y las libertades fundamentales (incluida la vida privada) y aumentar el riesgo potencial de accidentes y amenazas complejas de actores malintencionados.
- **[3.4] Regulaciones nacionales, sector privado y otros actores socio-comunitarios:** La Resolución exhorta a los Estados Miembros a que elaboren marcos normativos y apoyen la gobernanza de “sistemas IA que sean seguros, protegidos y fiables”. Alienta a todos los Estados Miembros, e

[6] El informe, con datos a junio de 2023, analizó cada uno de los objetivos y mostró sin eufemismos como más de la mitad de la población mundial se está “quedando atrás”. Aunque algunos ODS registraron avances, otros muestran lentitud o, incluso, retrocesos preocupantes (pobreza, hambre, igualdad, clima). El informe convoca a actuar con urgencia, antes que la Agenda 2030 se convierta en una quimera o, peor aún, en el epitafio del mundo que podría haber sido [3]. Los escasos resultados de medio término y el mundo atravesando una policrisis (pandemia, cambio climático y guerras) obligaron a grupos de expertos de Naciones Unidas a replantear la estrategia y apostar por un uso intensivo de tecnologías digitales e IA para acelerar los resultados. En 2023 y 2024 se publicaron varios informes y planes estratégicos de diferentes agencias de Naciones Unidas donde se presentaron ideas y soluciones radicales para la Agenda 2030 [11]. Por ejemplo, el Plan Estratégico del Programa Información para Todos [12].

invita a múltiples interesados (sector privado, sociedad civil, organizaciones internacionales, medios de comunicación, academia, instituciones de investigación y personas del ámbito técnico), a crear ecosistemas propicios para la innovación, el emprendimiento y la difusión de conocimientos. Los invita a que adopten medidas para cooperar y prestar asistencia a los países en desarrollo a fin de lograr un uso inclusivo y equitativo de estas tecnologías. Alienta al sector privado a que se adhiera a las leyes internacionales y nacionales y actúe en consonancia con los Principios Rectores de las Naciones Unidas sobre las Empresas y los Derechos Humanos: entornos empresariales de libre competencia, ecosistemas competitivos, mercados con actividades económicas y comerciales justas, abiertas, inclusivas y no discriminatorias a lo largo de todo el ciclo vital de los sistemas de IA.

- **[3.5] Las necesidad de prevenir las brechas sobre IA:** La resolución resalta la necesidad de evitar y paliar una divisoria digital en materia de IA. Invita a los Estados Miembros a tomar todas la medidas que sean necesarias para salvar la brecha en materia de inteligencias artificiales y otras brechas digitales entre los países y dentro de ellos. Reconoce que existen diversos niveles de desarrollo tecnológico (entre países desarrollados y en desarrollo) y que los países en desarrollo se enfrentan a desafíos singulares para seguir la aceleración en materia de IA. La Resolución indica la necesidad de aprovechar la transferencia de tecnologías, aumentar la infraestructura de computación y de conectividad, mejorar la alfabetización digital y aumentar los fondos para investigación y desarrollo científico-tecnológico. En la Resolución no se aportan datos o indicadores al respecto de estas asimetrías. No obstante, algunos estudios indican que ALyC sólo aporta el 2,66% de los artículos sobre IA en revistas científicas [13]⁷. A simple vista, la brecha de IA, sumada a las ya conocidas brechas digitales y del desarrollo, podrían complejizar aún más la Agenda 2030 en ALyC.

[4] Reflexiones finales: sobre las necesidades y peligros en el uso intensivo de las IA

La Resolución de la Asamblea General de Naciones Unidas sobre IA, desarrollo sostenible y derechos humanos, tuvo la virtud de haberse adoptado por consenso en un momento de crecientes diferencias globales y fuertes tensiones geopolíticas por el desarrollo científico-tecnológico. La adopción de la Resolución muestra como el sistema de las Naciones Unidas, de conformidad con su mandato, pudo alcanzar un consenso (aunque precario y algo ambiguo) sobre qué podrían ser los “sistemas de IA seguros, protegidos y fiables”. Si bien la Resolución tiene como

[7] Porcentajes a nivel mundial: 47.14%, Asia del Este y Pacífico; 17.20%, Europa y Asia Central; 11.61%, América del Norte; 6.93%, Desconocido; 6.75%, Asia del Sur; 4.64%, Medio Oriente y África del Norte; 2.66%, América Latina y Caribe; 2.30%, Resto del Mundo; 0.77%, África Sub-Sahariana [13].

antecedentes documentos vinculados a EE.UU., Reino Unido y Japón, entre otros, también busca reafirmar el trabajo que la ONU y sus principales agencias vienen desarrollando sobre IA (ITU, UNESCO, CDH, AI Advisory Board). Es evidente que para la ONU y sus agencias las IA ya tienen implicancia existencial. Su posición, tal como se analizó en la ponencia, se puede caracterizar como dual: parece pendular entre las crecientes necesidades humanitarias de un mundo en polícrisis (pandemia, cambio climático y guerras) y la evidente violación de los derechos humanos detrás del uso masivo de las IA. Se trata de tecnologías con enorme potencial que, lejos de ser neutrales, transparentes o universales, se desarrollan actualmente con fines de lucro, fuertes intereses geopolíticos y, hasta el momento, no se caracterizan por proteger los derechos humanos. Uno de los mayores peligros, advertido ligeramente en la Resolución, es que las IA profundicen las desigualdades globales y se concentren sólo en una reducida y privilegiada lista de estados, corporaciones e individuos. El gran desafío con las IA, central para alcanzar los ODS y, en particular, el ODS N.º 10⁸, es si su potencial se podrá utilizar para servir al bien común y beneficiar a toda la humanidad.

Referencias bibliográficas

- [1] AI Advisory Body. (2023). Governing AI for Humanity: Interim Report. Naciones Unidas. Disponible en https://www.un.org/sites/un2.un.org/files/ai_advisory_body_interim_report.pdf
- [2] Vercelli, A. (2023). Las inteligencias artificiales y sus regulaciones: Pasos iniciales en Argentina, aspectos analíticos y defensa de los intereses nacionales. Revista de la Escuela del Cuerpo de Abogados y Abogadas del Estado, (9), 195–217. Disponible en <https://revistaecae.ptn.gob.ar/index.php/revistaecae/article/view/232>
- [3] ONU - Organización de Naciones Unidas. (2023). Informe de los Objetivos de Desarrollo Sostenible: Edición Especial. Disponible en https://unstats.un.org/sdgs/report/2023/The-Sustainable-Development-Goals-Report-2023_Spanish.pdf
- [4] Vercelli, A. (2021). El extractivismo de grandes datos (personales) y las tensiones jurídico-políticas y tecnológicas vinculadas al voto secreto. Revista Themis, N° 79, pps.: 111 – 125. Lima: Revista Themis. Disponible en <https://revistas.pucp.edu.pe/index.php/themis/article/view/24867>
- [5] Vercelli, A. (2024). Regulaciones e inteligencias artificiales en Argentina. InMediaciones de la Comunicación, 19(1), 105–135. Disponible en <https://doi.org/10.18861/ic.2024.19.1.3549>

[8] El Objetivo 10, “Reducir la desigualdad en y entre los países”, Inciso 2, plantea “*De aquí a 2030, potenciar y promover la inclusión social, económica y política de todas las personas, independientemente de su edad, sexo, discapacidad, raza, etnia, origen, religión o situación económica u otra condición*” [3].

- [6] ONU - Organización de Naciones Unidas. (2024, 21 de marzo). Aprovechar las oportunidades de sistemas seguros, protegidos y fiables de inteligencia artificial para el desarrollo sostenible. Disponible en https://digitallibrary.un.org/record/4040897/files/A_78_L.49-ES.pdf
- [7] UN WebTV. (2024, 21 de marzo). Asamblea General, 63ª sesión plenaria, 78º período de sesiones. [Video]. Disponible en <https://webtv.un.org/es/asset/k1d/k1dio1rvsf>
- [8] Library ONU. (2024, 31 de enero). What does it mean when a decision is taken "by consensus"? What is the difference between "by consensus" and "unanimous"?. Dag Hammarskjöld Library. Disponible en <https://ask.un.org/faq/260981>
- [9] UNESCO (2021) Recomendación sobre la Ética de la Inteligencia Artificial. Paris. UNESCO. Disponible en https://unesdoc.unesco.org/ark:/48223/pf0000376713_spa
- [10] UN WebTV. (2024, 21 de marzo). Joint stakeout by US, Bahamas, Japan, Netherlands... [Video]. Disponible en <https://webtv.un.org/en/asset/k1a/k1a3snry9c>
- [11] Vercelli, A. (2024). Libraries, access to information and artificial intelligences in Latin America and the Caribbean, en Garrido, M. & Wyber, S. (Editores) "*Development and Access to Information*", pps.: 53 – 61, International Federation of Library Associations and Institutions (IFLA): The Hague. Disponible en: https://repository.ifla.org/bitstream/123456789/3334/2/da2i-2024-full-report_rev-en.pdf
- [12] PIPT - Programa Información Para Todos. (2023). Plan estratégico del PIPT, 2023-2029. Paris: UNESCO. Disponible en https://unesdoc.unesco.org/ark:/48223/pf0000386173_spa
- [13] Maslej, N., Fattorini, L., Brynjolfsson, E., Etchemendy, J., Ligett, K., Lyons, T., Manyika, J., Ngo, H., Niebles, I., Parli, V., Shoham, Y., Wald, R., Clark, J. y Perrault, R. (2023). The AI Index 2023 Annual Report. Institute for Human-Centered AI. Stanford: Stanford University. Disponible en https://aiindex.stanford.edu/wp-content/uploads/2023/04/HAI_AI-Index-Report_2023.pdf