



**FACULTAD DE CIENCIAS SOCIALES - UNIVERSIDAD DE BUENOS AIRES  
ARGENTINA FUTURA - JEFATURA DE GABINETE DE MINISTROS**

**INVESTIGACIÓN**

**Desafíos e impactos de la inteligencia artificial**

**Marcos normativos, riesgos y retos para la calidad democrática en la Argentina**

**EQUIPO**

**Dirección: Flavia Costa**

**Equipo: Pablo Rodríguez, Julián Mónaco y Ximena Zabala** (Historia de la IA, relevamiento bibliográfico y documental, entrevistas, glosario)

**Alejandro Covello y Iago Novidelsky** (Riesgo y seguridad)

**Mariano Zukerfeld, Celeste De Marco y Andrés Rabosto** (CamIA e Impacto en el mercado de Trabajo)

**Victoria Albornoz Saroff** (Perspectiva sindical)

**Jimena Durán Prieto** (Diseño)



## INFORME FINAL

<b>ÍNDICE</b>	2
Introducción	3
Resumen Ejecutivo	5
Recomendaciones	12
Investigación	15
Contexto	15
Breve historia de la IA	19
Perspectiva analítica	25
Marcos normativos de la IA	29
Análisis de los marcos normativos de EE.UU. y la Unión Europea	29
Dos perspectivas sobre la gestión de riesgos	32
Autoridades, gobernanza, responsabilidad	36
Accidentes, incidentes, notificaciones	38
Impacto de la IA generativa y los LLM en el mercado de trabajo	45
Relevamiento de bibliografía internacional	45
Aproximación a la problemática en el mercado de trabajo nacional	46
Estimación de impacto en la productividad y en los salarios en el sector de Software Argentino	48
Impacto de la IA generativa y los LLM en las industrias culturales y creativas: la perspectiva de los actores	50
Antecedentes del proyecto Centro Argentino Multidisciplinario de Inteligencia Artificial (CamIA)	55
Bibliografía	58
Anexos	
Presentación del Resumen Ejecutivo (PPT)	
Glosario	
Sitio web: <a href="http://tecnocenolab.ar/inteligencia-artificial/">tecnocenolab.ar/inteligencia-artificial/</a> (en línea)	



## INTRODUCCIÓN

Como campo de investigación, la Inteligencia Artificial (IA) tiene más de 70 años. Pero ha sido especialmente desde la última década que los avances en este ámbito han generado tecnologías o integrado grandes conjuntos tecnológicos que se comparan o superan a los humanos en tareas que requieren gran capacidad de cómputo, creatividad, razonamiento complejo, e involucran incluso la toma de decisiones. El creciente catálogo de aplicaciones y métodos de la IA, en particular la IA generativa, tiene el potencial de afectar profundamente las políticas públicas, así como a distintos sectores del trabajo y de la producción de conocimiento.

Hay diferentes proyecciones acerca del impacto económico positivo que puede tener para la Argentina y para la región la incorporación de IA en diferentes procesos productivos. Un estudio del Banco Interamericano de Desarrollo (BID) del año 2020 proyectaba que esta adopción podría representar para América Latina la oportunidad de elevar un 14% el Producto Bruto Interno.<sup>1</sup> Y un informe del Centro de Implementación de Políticas Públicas para la Equidad y el Crecimiento (CIPPEC) de 2018 señalaba que la adopción de IA podría duplicar la tasa de crecimiento inercial de la economía argentina.<sup>2</sup>

Asimismo, distintas fuentes señalan los desafíos y los efectos disruptivos que es necesario reconocer para mitigar en ese proceso de adopción. En marzo de 2023, un conjunto de expertos de nuestra región reunidos en Montevideo destacaron “el potencial productivo de los sistemas de inteligencia artificial, así como los riesgos que conlleva su crecimiento irreflexivo”.<sup>3</sup> Y señalaron la necesidad de “desarrollar criterios y estándares que permitan calificar estas tecnologías según sus riesgos de manera clara y transparente, para avanzar en políticas públicas que protejan el bien común sin obturar los beneficios del desarrollo tecnológico.”<sup>4</sup> Esto es así por el enorme poder de aceleración de procesos productivos y de toma de decisiones; porque la introducción de IA impacta en el mundo del trabajo y de la educación obligando a reescribir las reglas de industrias enteras; por su capacidad de crear instantáneamente contenidos y noticias que pueden ser falsas o erróneas; y debido a su capacidad de generar instancias en las que las regulaciones existentes ya no son adecuadas para enfrentar

---

<sup>1</sup> Gómez Mont C., Del Pozo C.M., Martínez Pinto C., Martín del Campo Alcocer A.V. (2020): “La Inteligencia Artificial al servicio del bien social en América Latina y el Caribe: panorámica regional e instantáneas de doce países”, Banco Interamericano de Desarrollo.

<sup>2</sup> Albrieu R., Rapetti M., Brest López C., Larroulet P., Sorrentino A. (2018): “Inteligencia artificial y crecimiento económico. Oportunidades y desafíos para Argentina”, CIPPEC.

<sup>3</sup> “Declaración de Montevideo sobre Inteligencia Artificial y su impacto en América Latina”, Montevideo, 10 de marzo de 2023. En: [fundacionsadosky.org.ar/declaracion-de-montevideo-fun/](http://fundacionsadosky.org.ar/declaracion-de-montevideo-fun/)

<sup>4</sup> ídem.



los problemas que aquejan a la sociedad, lo cual produce las llamadas brechas regulatorias.

En efecto, las características de las tecnologías de IA o que incluyen IA, como la opacidad (el efecto de caja negra), cierto grado de imprevisibilidad, la complejidad interactiva, su estructura de al menos ocho capas (Vercelli, 2023) y un comportamiento parcialmente autónomo pueden hacer difícil comprobar el cumplimiento de la normativa vigente –que protege derechos fundamentales y heterogéneos como la privacidad, los derechos de autor o los derechos laborales– o incluso impedir su cumplimiento.

A partir del desarrollo y la expansión de los modelos de lenguaje grandes (LLM, por sus siglas en inglés), como Chat GPT, LaMDA o PaLM, se espera que en los próximos años las inteligencias artificiales generativas tengan un impacto profundo en diversos aspectos de la sociedad, la economía, la política y la cultura, tanto a nivel global como regional y nacional.

En ese marco, este estudio es una exploración orientada a sistematizar la información existente sobre las iniciativas en materia de políticas y regulaciones de la Inteligencia Artificial que se vienen desarrollando en la Argentina, en relación con otras iniciativas internacionales de referencia. Busca delimitar una perspectiva teórico-analítica que se adecúe a las incumbencias definidas por el convenio entre Argentina Futura y la UBA, uno de cuyos horizontes es identificar los impactos de estas nuevas tecnologías en la calidad democrática y la participación ciudadana. Se propone presentar aspectos analíticos sobre las IA que consideramos útiles para generar un conocimiento más acabado acerca del objeto, así como para favorecer las discusiones sobre qué estrategias nacionales, qué políticas y qué regulaciones convienen a la Argentina y a la región. Y finalmente, ofrecer recomendaciones para una iniciativa integral orientada al desarrollo, la adopción, la implementación, el monitoreo y la mitigación de riesgos para una IA confiable y segura.



## RESUMEN EJECUTIVO

1. El desarrollo y la expansión de las inteligencias artificiales (IA) está generando profundos cambios en las sociedades. A partir sobre todo del desarrollo y la puesta en disponibilidad masiva de modelos de lenguaje grandes como ChatGPT o LaMDA, las inteligencias artificiales generativas se instalaron en el centro de la escena pública y se espera que en los próximos años tengan un impacto profundo en la economía, la política, la educación y la cultura, tanto a nivel nacional como a nivel regional y global.

2. La Argentina –como buena parte de los países de la región– se enfrenta con el desafío de diseñar estrategias de desarrollo, implementación, regulación y control de riesgos de sistemas de IA a contrarreloj. Las demandas en este sentido se entrecruzan: por un lado, se busca promover una tecnología que podría colaborar para que la Argentina retome el crecimiento sostenido de exportaciones de alto valor agregado que tuvo en la primera década del siglo XXI y los primeros años de la segunda.<sup>5</sup> Se procura generar empleos de calidad y acelerar el crecimiento económico, a la vez que afrontar y mitigar los impactos negativos de una tecnología que puede ser disruptiva, tanto en el empleo como en la calidad democrática, en particular por su potencial para facilitar la circulación de noticias falsas y desinformación. Alcanzar estos objetivos requiere fortalecer las capacidades científicas, productivas, tecnológicas y epistemológicas de los sectores público y privado, así como propiciar una indispensable actualización de las formaciones, tanto en las áreas de la informática, la ingeniería y la computación como en las comunicaciones, las ciencias sociales y políticas. Dirigir y acompañar la transformación requerirá de equipos “políglotas”, como los llama la Organización para la Cooperación y el Desarrollo Económico (OCDE), que profesen saberes sobre tecnología, normativa, seguridad, mediatización y sistemas sociotécnicos complejos (ver las recomendaciones de OCDE/CAF 2022).<sup>6</sup>

---

<sup>5</sup> Según Tacsir y Tacsir (2022), las exportaciones de bienes intensivos en conocimiento (BIC) pasaron de USD 1.200 millones aproximadamente en 2006, a más de USD 1.900 millones en 2011. Luego, se reducen gradualmente hasta US\$ 600 millones, aproximadamente, en 2021, reduciendo la participación argentina en el mercado global del 0,09% al 0,02%.

<sup>6</sup> En el documento “Estudios de la OCDE sobre Gobernanza Pública Uso estratégico y responsable de la inteligencia artificial en el sector público de América Latina y el Caribe”, de 2022, se señala la importancia de esta acción para los próximos años: “De cara al futuro, los gobiernos de América Latina y el Caribe necesitan garantizar que los servidores públicos de todos los niveles cuenten con las competencias y capacidades adecuadas en materia de IA, ya que los esfuerzos actuales tienden a poner el énfasis en el personal técnico. Es de vital importancia contar con un cuadro directivo superior que posea un alto nivel con conocimientos tecnológicos y una comprensión estratégica respecto de lo que la IA puede hacer, y del tipo de problemas que puede abordar, capaz de respaldar el despliegue de la IA en el Gobierno (...). Es fundamental que tanto los altos dirigentes como los responsables se encuentren preparados para gestionar el cambio” (OECD/CAF 2022: 170).



3. En consonancia con esa necesidad, este estudio se propuso brindar herramientas analíticas para promover en nuestro país el desarrollo de una iniciativa integral de IA, que se oriente al desarrollo, la adopción, la implementación y la gobernanza de sistemas de IA a la vez confiables, robustos y seguros.

4. A lo largo de la investigación se identificaron cinco rasgos o aspectos de las IA que, una vez asumidos, constituyen orientaciones epistemológicas y metodológicas. El primero es conocido pero no es ocioso recordarlo: la Inteligencia Artificial es una metatecnología, esto es, una tecnología de propósito general, aplicable a muy diversas actividades. El que muchas veces las IA estén indiferenciadas de los dispositivos y sistemas tecnológicos donde están incorporadas tiene efectos tanto en el nivel analítico como en el de la gobernanza. En breve: a los fines regulatorios, no alcanza con establecer *una* norma general para las IA, sino que es preciso identificar las capas y subsistemas que participan en las IA para analizar las diferentes legislaciones que las atraviesan, desde protección de datos y derechos de autor hasta legislación laboral y de protección del medioambiente.

5. El segundo rasgo proviene de una precaución teórico-metodológica, y consiste en que las inteligencias artificiales, en la medida en que integran y expanden el ecosistema digital, constituyen no una herramienta o dispositivo técnico, sino un *mundoambiente*. Esto significa que, en relación con los usuarios, no es suficiente un enfoque que las aborde desde la perspectiva de la instrumentalidad y de la relación sistema-usuario individual, sino que es necesario un enfoque sistémico, atento a las dinámicas multiescalares de la economía política del ecosistema digital.

6. El tercer rasgo –que se deriva de la literatura existente sobre IA generativa y LLM, pero no ha sido hasta el momento explorado en detalle– es particularmente significativo para esta investigación. Consiste en señalar que, a los efectos de un análisis epistemológico con epicentro en las ciencias sociales, las llamadas IA generativas y en particular los LLM son, no sólo *Inteligencia Artificial*, sino también *Sociedad Artificial*. Esto es así debido a que las IA generativas y en particular los LLM operan desde y sobre lo social a través del sistema Datos, Algoritmos, Plataformas (DAP). De allí que es deseable que en su análisis y monitoreo participen expertos en esos campos disciplinares, que deberán formarse también en competencias tecnológicas.

7. El cuarto rasgo consiste en que, en ciertos usos, las IA pueden ser tecnologías de alto riesgo, y por lo tanto requieren un tratamiento acorde a lo largo de todo su ciclo de vida. Es preciso que la iniciativa argentina de IA incorpore este enfoque, abordando



las inteligencias artificiales generativas desde una perspectiva de la seguridad y la gestión de riesgos.

8. El quinto elemento no es tanto un rasgo de las IA generativas sino una consecuencia analítica de tener en cuenta los aspectos anteriores. Para afrontar las IA generativas desde las ciencias sociales y humanas, para pensar eficazmente su gobernanza, no es suficiente con un enfoque desde la *ética profesional* de la IA sino que es preciso promover un enfoque desde la *ética organizacional* de la IA y desde el *pensamiento sistémico*, que busca establecer procedimientos de revisión transparente, estructuras de rendición de cuentas vinculantes, documentación de modelos y conjuntos de datos, auditoría independiente; en síntesis: defensas en profundidad a lo largo del sistema para que este sea más seguro y confiable.

9. Una vez identificados estos elementos, el enfoque de la investigación se orientó a combinar el necesario impulso del desarrollo de tecnología de IA en el país --que se corresponde con el propósito de promover la soberanía tecnológica-- con un enfoque de seguridad y gestión de los riesgos de la IA, poco desarrollado aun en la literatura local.

10. En cuanto a los marcos normativos, se recogieron como textos de base para el análisis la Recomendación sobre la ética de la inteligencia artificial de la UNESCO (2021), la Iniciativa Nacional de AI (NAII) de los Estados Unidos (2020), que remite a su vez al Marco de Gestión de Riesgos de Inteligencia Artificial del NIST (National Institute of Standards and Technology [Instituto Nacional de Estándares y Tecnologías], NIST, 2023); la Ley de IA de la Unión Europea (2023); los documentos del Observatorio de Políticas de IA (2020-2023) de la OCDE; y las Recomendaciones para una IA fiable emitidas por la Secretaría de Innovación Pública y publicadas en el Boletín Oficial argentino en junio de este año (2023). De esa primera selección, se decidió analizar en perspectiva comparada la Iniciativa Nacional de IA (NAII) de los Estados Unidos y la Ley de Inteligencia Artificial europea (2023), con especial énfasis en los marcos de referencia destinados a la gestión de los riesgos.

11. Del análisis surge que el principal objetivo de la Iniciativa Nacional de los Estados Unidos con respecto a la IA es liderar el desarrollo de IA en el mundo. Con respecto a los riesgos es una estrategia reactiva, en la medida en que no los aborda de manera directa, sino que encomienda al NIST, dependiente del Departamento de Comercio, la elaboración de un marco general y voluntario para la gestión de riesgos. Ese proceso dio como resultado el Artificial Intelligence Risk Management Framework (AI RMF 1.0), publicado el 26 de enero de 2023. En este documento se explicita que se trata de un documento vivo al menos hasta 2028 --para poder perfeccionarlo y adaptarlo a la



evolución de la tecnología emergente—, y el hecho de que sea un marco voluntario y no una norma obligatoria deja claro el énfasis.

12. La ley de la Unión Europea, en cambio, adopta desde su primera página una estrategia proactiva con relación a los riesgos, orientada a delimitar diferentes tipos de usos y prácticas en los que puedan participar sistemas de IA. Establece tres categorías de riesgo. Una es la de los riesgos inaceptables, lo que significa que hay usos de IA que están prohibidos: la manipulación maliciosa del comportamiento, la calificación social y la vigilancia masiva. Otros riesgos se consideran altos y deben ser obligatoriamente gestionados; esto implica un complejo circuito de gestión de calidad y riesgo, documentación, certificaciones, notificaciones. Ejemplos de esta segunda categoría son la identificación biométrica y la categorización de personas, la gestión y el funcionamiento de infraestructuras esenciales; la educación y la formación profesional; el empleo, la gestión de los trabajadores y el acceso al autoempleo; el acceso a servicios esenciales, o la aplicación de la ley. Una tercera categoría es la de los riesgos mínimos, en los que la UE exige transparencia para con el usuario: desarrolladores e implementadores deben informar al usuario que está interactuando con un sistema de IA.

13. La Ley europea es exhaustiva al describir la red institucional a cargo de la gestión de esos riesgos. E incluye la obligación de informar accidentes o incidentes de IA en no más de 72 horas desde su ocurrencia a las autoridades nacionales y de la Unión. Dado que la IA acelera el análisis y la gestión de lo social, para gobernar su desempeño, para hacerla fiable es necesario crear nuevas mediaciones y/o fortalecer las existentes.

14. En cuanto a la perspectiva argentina, sobre la base de las Recomendaciones emitidas en junio por la Secretaría de Innovación Pública, entendemos que el texto acierta en identificar la necesidad de hacer un seguimiento de la IA a lo largo de todo el ciclo de vida, desde su concepción hasta su reciclado o descarte. Incorpora los valores de alineación o alineamiento sugeridos por las Recomendaciones de la UNESCO;<sup>7</sup> identifica los momentos de concientización, diseño, verificación, validación, implementación, operación y mantenimiento, así como la necesidad de establecer siempre un responsable humano en última instancia. Con todo, por un lado, al ser una recomendación voluntaria, su alcance es restringido. Y por otro, no establece instancias de monitoreo ni de investigación de accidentes e incidentes, tal como sí

---

<sup>7</sup> Proporcionalidad e inocuidad, seguridad y protección; equidad, sostenibilidad, derecho a la intimidad y protección de datos, supervisión y decisión humanas, transparencia y explicabilidad, responsabilidad y rendición de cuentas, sensibilización y educación; y gobernanza y colaboración adaptativa y de múltiples partes interesadas.



están comenzando a sugerir organizaciones supranacionales como la OCDE.<sup>8</sup> En nuestras recomendaciones sugerimos incorporar el análisis y la investigación de incidentes y accidentes de IA para robustecer el ecosistema de IA, volverlo más fiable para la sociedad y, además, disponer de equipos actualizados de monitoreo de estas tecnologías.

15. En cuanto a las autoridades a cargo de estudiar y dirigir el desarrollo de IA, en el inicio de la investigación advertimos, en confrontación con las iniciativas estadounidense y europea, una dispersión institucional,<sup>9</sup> algo que fue advertido también por las autoridades. En efecto, en septiembre de 2023 se creó la Mesa Interministerial sobre IA, por Decisión Administrativa 750/2023, con el objetivo de diseñar una estrategia integral “para el avance y aplicación de la IA en diversos sectores de la economía y sociedad, considerando un marco ético y de desarrollo sostenible”.

16. Con todo, es pensable que la iniciativa que más podrá contribuir a ordenar de manera estable la política pública argentina con relación a IA sea el Centro Argentino Multidisciplinario para la Inteligencia Artificial (CamIA), en incubación incipiente a partir del Programa de apoyo a las exportaciones de la Economía del Conocimiento (EDC) aprobado en junio de 2023. En la presentación se menciona el objetivo de crear “un centro de inteligencia artificial” que “tendrá entre sus objetivos generar capacidades de dirección y gestión de proyectos multidisciplinarios de desarrollo tecnológico basados en IA, articular las capacidades del sistema científico y tecnológico en IA y las necesidades del sector productivo, elaborar una agenda de política regulatoria en IA, desarrollar talentos en IA y contribuir a la internacionalización del ecosistema”. Al momento de cierre de este informe, la información disponible es que CamIA se emplazará en la Facultad de Ciencias Exactas de la UBA, y que será incubado durante un período máximo de 4 años.

17. A lo largo de la investigación enfrentamos la dificultad de participar de una conversación pública al mismo tiempo no del todo informada (por la novedad del fenómeno emergente) y con sobreabundancia de material no sistematizado (debido al interés que el tema suscitó en los medios de comunicación masivos y no masivos). Para ello, pensando en la instancia comitente de este informe, nos propusimos ordenar la conversación pública de dos maneras. Por un lado, identificando tres escalas de la IA: la escala macro, la meso y la micro, e identificando en la escala *meso* el ámbito más

---

<sup>8</sup> En efecto, el 20 de noviembre de 2023 se puso en línea la versión de prueba del primer Monitor de Incidentes de IA, que desarrolló a lo largo del año el grupo de expertos de la OCDE para IA.  
<https://oecd.ai/en/incidents>

<sup>9</sup> Solo en el ámbito del Poder Ejecutivo encontramos nueve áreas con iniciativas en torno a la IA, no coordinadas entre sí.



propio de intervención de las políticas públicas en el nivel de las políticas de Estado, en sus niveles nacional, provincial y municipal. La escala *micro* es la escala de los desarrollos e implementaciones específicos (por ejemplo la optimización de láseres para metrología, el monitoreo de pacientes en cuidados intensivos y muchísimas otras más), que se utilizan en muy diferentes industrias y disciplinas, y que en principio no parecen requerir un cuerpo de regulación específico. Por otro lado, la escala *macro* del desarrollo por parte de empresas como Google, Microsoft, Meta, Amazon o Baidu, quedan –en principio– por fuera del campo de la intervención estatal argentina. Es en el *nivel meso* donde se ubican las políticas públicas tanto para promover el desarrollo como para identificar riesgos y desafíos del desarrollo, así como impactos de la aplicación.

18. Los riesgos más habitualmente señalados por la literatura internacional en relación con la escala *meso* son: los sesgos, los datos inadecuados o malinterpretados, la suplantación de identidad, el *deep fake* o la desinformación (ya provenga de humanos o de máquinas), la vigilancia, la manipulación del comportamiento y la securitización del conocimiento (que significa que el conocimiento experto queda en manos de cada vez menos personas). Se trata en todos los casos de riesgos para la calidad democrática que están siendo investigados desde perspectivas, disciplinas y ámbitos poco comunicados entre sí. De allí que la recomendación principal en este sentido es unificar un equipo de trabajo experto para elaborar instrumentos analíticos y defensas en profundidad en el nivel tecnológico y normativo, así como sostener un equipo de investigación de incidentes de IA, que podría estar instalado en CamIA.

19. En cuanto a desafíos e los impactos sobre el trabajo y en particular sobre las industrias creativas, de las exploraciones realizadas en esos ámbitos se pueden obtener algunos datos nítidos. Por un lado, un 29% de las ocupaciones de la estructura ocupacional argentina están expuestas a los LLM. Por otro lado, en el corto lapso transcurrido desde que las IAG comenzaron a estar disponibles a nivel masivo, hay impactos observables en los trabajos intensivos en conocimiento. En tercer lugar, se registran diferentes niveles de preocupaciones entre los representantes de trabajadores de las industrias creativas que podrían ser abordados mediante políticas específicas. De allí que lo primero que surge es la necesidad de profundizar en estas indagaciones para comprender la magnitud de los impactos a lo largo del tiempo y desarrollar políticas orientadas a partir de datos. Posiblemente también CamIA podría ser una sede de estas iniciativas.

20. En cuanto a las ocupaciones con mayor exposición a los LLM, se identificó que entre ellas hay ocupaciones fundamentales para el desarrollo de tareas cotidianas: ocupaciones de gestión administrativa, de planificación y control de gestión, de



asistencia educativa, de gestión presupuestaria, contable y financiera, entre otras. Entre los desafíos que este escenario presenta subrayamos el de priorizar áreas clave tanto en el sector público como en el privado para emprender procesos controlados de transformación digital, de forma tal que puedan aprovecharse las potencialidades de las IA sin generar falsas expectativas ni efectos traumáticos contraproducentes.

21. A fines de brindar un ejemplo en un sector concreto, se relevó un estudio del CIECTI todavía en curso, que parte de una encuesta de 5500 casos entre programadores de software y que busca analizar si el uso de ChatGPT impacta en los niveles salariales y si varía en función de la experiencia laboral. Los resultados parciales arrojan que un 73% de la muestra declara haber utilizado al menos una vez herramientas de IA para la codificación (como ChatGPT o GitHub Copilot) y casi un 20% lo hace de manera muy frecuente. La frecuencia de uso y el nivel de experiencia están inversamente relacionados. Además, se identificaron brechas salariales positivas asociadas al uso de ChatGPT para todos los niveles de experiencia, sugiriendo que la adopción de esta herramienta podría estar vinculada a una mayor remuneración. Estos resultados parecen mostrar que las herramientas de IAG están redefiniendo los paradigmas de aprendizaje y desarrollo de habilidades en los trabajos intensivos en conocimiento no rutinarios, lo que conduce a repensar desde las estrategias y los currículos de la educación en diferentes niveles educativos, hasta los incentivos para los trabajos intensivos en conocimiento.

22. Finalmente, con respecto a los riesgos y la investigación de incidentes de la IA, se identificó la importancia de diferenciar dos enfoques transversales: el enfoque jurídico-normativo y el enfoque sistémico. Si bien el enfoque jurídico-normativo es fundamental para abordar el impacto de las IA, ya que estas metatecnologías pueden afectar derechos consagrados (a la privacidad, a la propiedad intelectual, a la protección del trabajo, entre otros), no es suficiente si de lo que se trata es de alcanzar una IA fiable, y segura (*safe*). Porque el enfoque jurídico señala el límite exterior de la Ley; allí donde, si un agente infringe la ley, puede ser legítimamente castigado. En caso de incidente, actúa *a posteriori* buscando causas y responsables, y emite sanciones, multas, penalidades. Mientras que el enfoque sistémico se enfoca en estudiar el incidente para prevenir y evitar que vuelva a repetirse. Busca comprender dónde estuvo la falla para fortalecer las defensas del sistema. Es un enfoque que, en caso de incidente, se interroga por los factores desencadenantes, identifica otros factores relacionados con el accidente, y emite recomendaciones.



## RECOMENDACIONES GENERALES

1. Impulsar una iniciativa nacional coordinada de desarrollo y gobernanza de la IA en relación con al menos seis objetivos: generar capacidades de dirección y gestión de proyectos multidisciplinarios de desarrollo y monitoreo de tecnologías basadas en IA; articular las capacidades del sistema científico y tecnológico en IA y las necesidades del sector productivo; elaborar una propuesta de política regulatoria de IA con perspectiva jurídica y enfoque sistémico; promover la formación de expertos “bilingües” capaces de comprender los desafíos sociales, políticos y educativos de la IA; promover la formación de expertos en gestión de riesgos e investigación de incidentes de IA; contribuir a la internacionalización del ecosistema de IA.
2. Promover que se revise, sistematice y eventualmente –de ser necesario– actualice la legislación existente referida a las distintas capas de sistemas de IA (ley de protección de Datos personales, ley de Acceso abierto, ley de protección al derecho de Autor, ley de protección del medioambiente, entre otras).
3. Potenciar la presencia de la Argentina en foros internacionales (como la Organización Mundial del Comercio) para fortalecer su situación relativa en el ecosistema digital global.

## RECOMENDACIONES GENERALES SOBRE SEGURIDAD Y RIESGOS DE LA IA

4. Establecer una escala de riesgos de sistemas y usos de IA que defina (1) prácticas prohibidas (riesgos inaceptables); (2) prácticas de IA de alto riesgo (deben cumplir requisitos obligatorios; los sistemas deben registrarse y se monitorean a lo largo de todo su ciclo de vida); y (3) prácticas de IA que se consideran de riesgo limitado (pueden tener requisitos obligatorios, como la transparencia, pero son menos estrictos que los del grupo 2).
5. Designar y/o impulsar autoridades normativas, fiscalizadoras, de certificación, monitoreo, notificación e investigación de incidentes y accidentes de IA.
6. Fomentar el estudio de incidentes de IA para los distintos campos de aplicación que involucren acciones sobre áreas críticas para la población y para la calidad democrática (como salud, ecosistema de la comunicación y la información, trabajo, justicia, toma de decisiones gubernamentales).

## RECOMENDACIONES DEL DESARROLLO HUMANO



7. Propiciar una iniciativa nacional de formación interdisciplinaria para el ecosistema digital que promueva el desarrollo de perfiles asociados a la planificación y el diseño digital; la gobernanza de datos; la gestión digital; la perspectiva del riesgo en sistemas de IA y el pensamiento sistémico de la seguridad (*safety*) y la protección (*security*) en IA.

8. Promover el desarrollo de laboratorios interdisciplinarios y federales de transformación digital (TD), que incuben proyectos de TD inclusivos y que fomenten la inserción de talentos en el sector productivo para evitar su “fuga”.

9. Propiciar un foro regional permanente de políticas para el ecosistema digital, que incluya el desarrollo de un mapa vivo del ecosistema digital y de la IA en América Latina. Dicho mapa puede incluir diferentes capas o capítulos: mapa de regulaciones, de infraestructura, de instituciones públicas y privadas de desarrollo, de expertos.

#### RECOMENDACIONES DEL ÁMBITO DE LA COMUNICACIÓN Y LA INFORMACIÓN

10. Promover el estudio de los sistemas y usos de la IA que puedan afectar al espacio de la información y la comunicación, y delimitar cuáles son los distintos tipos que existen (sistemas de curaduría de contenidos, sistemas vinculados a la publicidad y/u otros).

11. Desarrollar un enfoque basado en riesgos para analizar los sistemas de IA que afectan al espacio de la información y la comunicación. Esto implica delimitar diferentes usos y sistemas de IA, y hacer evaluaciones de riesgos para examinar si un sistema de IA es relevante para dicho espacio; y en caso afirmativo, definir requisitos.

12. Lo anterior implica que la evaluación de impacto de sistemas de IA debe incluir la pregunta por su impacto en los derechos relacionados con el ámbito de la comunicación y la información. Y que eventualmente los desarrolladores e implementadores de IA establezcan planes de gestión de riesgos para el espacio de información, que podrían incluir supervisión humana; adherencia a los hechos y moderación de contenido de los modelos de IA; marcadores de confianza [*trusted flaggers*] y/u otros procesos de marcado por parte de los usuarios; mecanismos de filtrado de contenido que los usuarios pueden aplicar para analizar y marcar contenido potencialmente problemático, y la notificación de incidentes relacionados con el espacio de la comunicación y la información.

#### RECOMENDACIONES EN REFERENCIA AL IMPACTO EN EL TRABAJO



13. Avanzar en la construcción de una base de datos de tareas laborales adaptada al Clasificador Nacional de Ocupaciones, con estructura similar a las bases de datos de tareas laborales de la Red de información Ocupacional (O\*net) y de la clasificación ISCO (*International Standard Classification of Occupations*), de acuerdo a la metodología propuesta por la Organización Internacional del Trabajo (2023).

14. Impulsar investigaciones que estudien el impacto de la implementación de las IA generativas en distintas ocupaciones, tareas e industrias, que evalúen los riesgos asociados a dichos impactos, y las percepciones de los actores en relación con esos u otros riesgos, para tomar decisiones de políticas públicas basadas en evidencia.

15. Una vez detectadas las ocupaciones más expuestas a las IA generativas, desarrollar una planificación e involucrar a los distintos sectores interesados para facilitar la transición de profesionales y trabajadores a nuevas ocupaciones, o nuevas formas de una misma ocupación.



## INVESTIGACIÓN

### CONTEXTO

Tal como mencionamos en la Introducción, a partir del desarrollo de modelos de lenguaje grandes (LLM), como ChatGPT, LaMDA, Megatron-Turing o PaLM, se espera que en los próximos años las inteligencias artificiales generativas tengan un impacto profundo en diversos aspectos de la sociedad, la economía, la política y la cultura, tanto a nivel global como regional y nacional.

En diciembre de 2022, el investigador en IA y profesor emérito de la Universidad de Nueva York Gary Marcus comentó:

Aún no está claro cuáles van a ser las aplicaciones de tecnologías como ChatGPT. La más inmediata es escribir trabajos escolares. Pero mi mayor preocupación es la desinformación, creo que va a acelerar de forma dramática la velocidad a la que se va a producir desinformación. Se va a usar para propaganda, para hacer páginas *web fake* y engañar a la gente.<sup>10</sup>

Poco después, en marzo de 2023, una carta abierta publicada por la organización no gubernamental *Future of Life* y firmada por investigadores, pensadores y empresarios de tecnología de la información como Steve Wozniak (co-fundador de Apple), Yuval Harari, Elon Musk y el inversor en IA Ian Hogart, autor desde 2018 del Informe sobre el estado de la IA, llamó a “poner en pausa” por “al menos 6 meses” el entrenamiento de aquellos sistemas de Inteligencia Artificial de potencia superior a la del sistema GPT-4.<sup>11</sup> La carta comienza advirtiéndolo:

Los sistemas de IA con inteligencia humana competitiva pueden plantear riesgos profundos para la sociedad y la humanidad, como lo demuestra una extensa investigación y lo reconocen los principales laboratorios de IA. Como se establece en los Principios de IA de Asilomar, ampliamente respaldados, la IA avanzada podría representar un cambio profundo en la historia de la vida en la Tierra y

---

<sup>10</sup> “Este veterano de la inteligencia artificial explica por qué ChatGPT es ‘peligrosamente estúpido’”. Gary Marcus entrevistado por Manuel Ángel Méndez para *El Confidencial*, 11 de diciembre de 2022. En Internet: [www.elconfidencial.com/tecnologia/2022-12-11/chatgpt-openai-gary-marcus-ia-ai-inteligencia-artificial\\_3537495](http://www.elconfidencial.com/tecnologia/2022-12-11/chatgpt-openai-gary-marcus-ia-ai-inteligencia-artificial_3537495)

<sup>11</sup> En Internet: <https://futureoflife.org/open-letter/pause-giant-ai-experiments/>



debe planificarse y gestionarse con el cuidado y los recursos correspondientes. Desafortunadamente, este nivel de planificación y gestión no está sucediendo, a pesar de que en los últimos meses los laboratorios de IA se han visto atrapados en una carrera fuera de control para desarrollar e implementar mentes digitales cada vez más poderosas que nadie –ni siquiera sus creadores– puede entender, predecir o controlar de manera confiable.

Los sistemas de IA contemporáneos se están volviendo competitivos desde el punto de vista humano en tareas generales y debemos preguntarnos: ¿Deberíamos dejar que las máquinas inunden nuestros canales de información con propaganda y falsedad? ¿Deberíamos automatizar todos los trabajos, incluidos los satisfactorios? ¿Deberíamos desarrollar mentes no humanas que eventualmente puedan superarnos en número, ser más astutas, quedar obsoletas y reemplazarnos? ¿Deberíamos arriesgarnos a perder el control de nuestra civilización? Estas decisiones no deben delegarse en líderes tecnológicos no electos. Sólo se deberían desarrollar potentes sistemas de IA una vez que estemos seguros de que sus efectos serán positivos y sus riesgos manejables.

Poco después, en la semana del 2 de junio de 2023, se publicaron en el Boletín Oficial de la Nación las “Recomendaciones para una inteligencia artificial fiable” en las que se señala que “la irrupción de la Inteligencia Artificial (IA) (...) empuja a los Estados a definir estrategias para encauzar el potencial transformador de esta tecnología en la resolución de problemas concretos y a favor del bien común”.<sup>12</sup> El documento, emitido por la Subsecretaría de Tecnologías de la Información, dependiente de la Secretaría de Innovación Pública, establece asimismo una serie de valores de alineamiento, a saber: Proporcionalidad e inocuidad, seguridad y protección; equidad, sostenibilidad, derecho a la intimidad y protección de datos, supervisión y decisión humanas, transparencia y explicabilidad, responsabilidad y rendición de cuentas, sensibilización y educación; y gobernanza y colaboración adaptativa y de múltiples partes interesadas.<sup>13</sup>

---

<sup>12</sup> *Recomendaciones para una inteligencia artificial fiable*, Subsecretaría de Tecnologías de la Información. Boletín Oficial: [www.boletinoficial.gob.ar/detalleAviso/primera/287679/20230602](http://www.boletinoficial.gob.ar/detalleAviso/primera/287679/20230602)

<sup>13</sup> Hubo una iniciativa nacional anterior, elaborada durante la presidencia de Mauricio Macri entre 2018 y 2019, el Plan Nacional de Inteligencia Artificial (ArgenIA, 2019), que señala la importancia de las IA, afirma la necesidad de formar recursos humanos, la relevancia de utilizar datos de calidad, pone el acento en la infraestructura computacional, y propone la creación de un Laboratorio de Innovación para “acelerar y canalizar el cumplimiento de objetivos propuestos en el Plan Nacional de IA”. Con el cambio de gobierno y la irrupción al año siguiente de la pandemia mundial del COVID-19, el Plan quedó sin efecto. Es posible consultarlo en Internet: <https://ia-latam.com/wp-content/uploads/2020/09/Plan-Nacional-de-Inteligencia-Artificial.pdf>



Esta misma dependencia —que entre sus competencias posee la de “Entender en la ciberseguridad y protección de infraestructuras críticas de información y comunicaciones asociadas del Sector Público Nacional y de los servicios de información y comunicaciones definidos en el artículo 1° de la Ley N° 27.078” – emitió en julio del mismo año una Guía de Notificación y Gestión de Incidentes de Ciberseguridad, que si bien no remite principalmente a la IA, contempla algunas de las potenciales intervenciones de sistemas de IA.

Poco antes, el 12 de junio de 2023, se había publicado en las páginas del portal oficial del Estado argentino<sup>14</sup> que el Banco Interamericano de Desarrollo (BID) había aprobado un préstamo a cinco años por un valor de 35 millones de dólares destinado al Programa de Apoyo a las Exportaciones de la Economía del Conocimiento con el fin de apoyar el desarrollo del sector y su inserción internacional. Fueron designados como organismos ejecutores el Ministerio de Economía y el Ministerio de Ciencia, Tecnología e Innovación a través de la Agencia Nacional de Promoción de la Investigación, el Desarrollo Tecnológico y la Innovación (Agencia I+D+i).

Por su parte, en la reciente Reforma Constitucional de la provincia de Jujuy, publicada en el BO provincial el 21 de junio de 2023, existe un artículo dedicado especialmente a la IA (número 76), donde se reconoce en dicha provincia “el derecho de toda persona a utilizar sistemas de inteligencia artificial o no humana, basados en métodos computarizados de algoritmos, datos y modelos que imitan el comportamiento humano y automatizan procesos complejos, así como otros futuros desarrollos que surjan en este campo”. Y se afirma que se sujetarán esos sistemas “a los principios de legalidad, transparencia, responsabilidad, privacidad y protección de datos, seguridad, no discriminación y rendición de cuentas, garantizando el acceso a la justicia en caso de vulneración de derechos y consagrando la acción de solicitud de revisión humana cuando sea necesario”.<sup>15</sup>

También en junio de 2023, la Unión Europea discutió y dio por terminado el proceso de enmiendas a la Ley de IA que regirá a la Unión por los próximos años. El 14 de ese mes, la plenaria del Parlamento Europeo aprobó un proyecto con enmiendas a la Ley de 2021 para regular el uso de la Inteligencia Artificial en la Unión Europea, dando inicio a una delicada negociación con los 27 países del bloque. La normativa sancionada (con

---

<sup>14</sup> “El BID aprueba destinar 35 millones de dólares al MINCYT en apoyo al desarrollo de la inteligencia artificial”, 12 de junio de 2023. En internet: [www.argentina.gob.ar/noticias/el-bid-aprueba-destinar-35-millones-de-dolares-al-mincyt-en-apoyo-al-desarrollo-de-la](http://www.argentina.gob.ar/noticias/el-bid-aprueba-destinar-35-millones-de-dolares-al-mincyt-en-apoyo-al-desarrollo-de-la)

<sup>15</sup> El texto puede consultarse en Internet: <http://www.saij.gob.ar/0-local-jujuy-constitucion-provincia-jujuy-lpy000000-1986-10-22/123456789-0abc-defg-000-0000yvorpyel>



499 votos a favor, 28 en contra y 93 abstenciones) ratifica la regulación de la Inteligencia Artificial según el nivel de riesgo: cuanto mayor sea para los derechos o la salud de las personas, mayores serán las obligaciones de los sistemas tecnológicos.<sup>16</sup>

En este marco vertiginoso, la Argentina –como buena parte de los países de la región– se enfrenta con el desafío de diseñar estrategias de desarrollo, implementación y regulación de sistemas de IA.

---

<sup>16</sup> Entre tanto, ya en septiembre de 2021 el Reino Unido había dado a conocer su Iniciativa Nacional sobre IA, basada en tres ideas-pilares con respecto al futuro: (1) los impulsores clave del progreso, el descubrimiento y la ventaja estratégica en la IA serán los accesos a personas, datos, computación y finanzas, todos los cuales enfrentan una enorme competencia global; (2) la IA se generalizará en gran parte de la economía y será necesario tomar medidas para garantizar que todos los sectores se beneficien de esta transición; y (3) los regímenes regulatorios y de gobernanza deberán seguir el ritmo de las demandas rápidamente cambiantes de la Inteligencia Artificial (IA), maximizando el crecimiento y la competencia, impulsando la excelencia en innovación y protegiendo la seguridad, las decisiones y los derechos de los ciudadanos.



## BREVE HISTORIA DE LA IA

La Inteligencia Artificial (IA) tiene su acta de nacimiento oficial en el verano de Estados Unidos de 1956, en una reunión de dos meses en el Dartmouth College, en Hanover, estado de New Hampshire. El director del Departamento de Matemáticas de esa institución, John McCarthy, convocó a un conjunto de especialistas, entre ellos Claude Shannon, el padre de la teoría matemática de la información, a crear un campo de investigación, la IA, que debía trabajar sobre siete ejes: computadoras automáticas; programación de computadoras para que usen un lenguaje; redes neuronales; teoría del tamaño de un cálculo; auto-mejoramiento [*self-improvement*] de una máquina; métodos maquínicos para formar abstracciones; y relación entre azar y creatividad.

Es importante detenerse brevemente en los antecedentes de esta propuesta para entender cuáles son los desafíos que atraviesa en la actualidad la IA en relación con aquel planteo original y sus sucesivas modificaciones a lo largo de medio siglo.

La IA es un campo de estudios que emerge en el cruce entre las ciencias de la computación, las ciencias cognitivas y la cibernética. En 1943 se publicaron dos artículos decisivos para el establecimiento del campo. En *A logical calculus of the ideas immanent in nervous activity*, Warren McCulloch y Walter Pitts elaboraron un modelo abstracto del funcionamiento de una neurona aplicando la lógica de Boole, una técnica algebraica para tratar expresiones de la lógica proposicional. Su idea era formalizar mediante un sistema de llaves (entrada-salida) el hecho de que las neuronas se comunican a través de impulsos eléctricos que se organizan de modo binario. Según McCulloch-Pitts, las neuronas realizan cálculos de ese modo, pero lo hacen de manera masiva e interconectada. Presentan para ello, por primera vez, la noción, tan extendida hoy en el campo de la IA, de las redes neuronales.

El segundo artículo de 1943 es *Behaviour, Purpose and Teleology*, fue escrito por Norbert Wiener, Arturo Rosenblueth y Julian Bigelow, y establece las bases de lo que será conocido como la “cibernética”, el campo de estudios interdisciplinario que se define como “ciencia que estudia la comunicación y el control en animales, seres humanos y máquinas”. Allí los autores plantean la noción de retroalimentación [*feedback*] y buscan formalizar los procesos de causalidad circular (en lugar de la causalidad lineal del esquema estímulo-respuesta del conductismo estadounidense). El objetivo era desencadenar acciones dirigidas por un propósito realizadas por máquinas cuya característica fundamental consistiera en ser capaces de ser sensibles al entorno y modificar su comportamiento en función de esa sensibilidad.

Ambas búsquedas encontrarían en la computadora la sede de su realización. Es John Von Neumann, dentro del Proyecto Manhattan –que tuvo su momento culmen en 1945, con el lanzamiento por parte de los Estados Unidos de dos bombas atómicas sobre Japón–, quien vinculó las redes neuronales de McCulloch-Pitts con la máquina sensible al entorno y dotada de un propósito de Wiener-Rosenblueth-Bigelow. El



resultado fue un dispositivo con sensores de entrada y mecanismos de salida, y entre ellos un estado interno, descrito por un programa, donde se realizan cálculos con base lógica a partir de circuitos eléctricos. Esa máquina, la computadora, fue planteada así como un sucedáneo artificial del cerebro.

A la salida de la Segunda Guerra Mundial se organizaron las conferencias Macy, de 1946 a 1952, donde estos y otros investigadores establecieron un suelo común de investigaciones que el propio Wiener reunió con el nombre de “cibernética”. Se destacaron dos elementos centrales para la historia de la IA: la reunión de las investigaciones sobre la relación entre corriente eléctrica y cálculos lógicos bajo el mote de información, definida por Claude Shannon, y la propuesta del investigador inglés Alan Turing –quien había trabajado junto con Shannon en el desciframiento de mensajes durante la Segunda Guerra– acerca de que se podían formalizar y mecanizar operaciones lógico-matemáticas al punto tal de abstraer las condiciones del pensamiento de su arraigo en una materialidad dada. En otras palabras, la conocida “máquina de Turing” permitía hacer una analogía entre mente (forma abstracta del pensamiento), cerebro (conjunto de neuronas) y computadora (mecanismo artificial que simula un cerebro).

Estas fueron las condiciones para que se constituyera, como desprendimiento de las reuniones cibernéticas, el campo de las ciencias cognitivas (Simposio de Hixon, 1948), que a partir de la analogía entre el cerebro y la computadora comenzó a trabajar en la hipótesis de que el acto de representarse el mundo para actuar sobre él podía ser realizado por una máquina y que, por ello mismo, esa máquina podía ser una herramienta metodológica para entender cómo funciona el cerebro. De ello surge el objeto de estudio de la *cognición*, definida como la relativa equivalencia entre representar, conocer y calcular.

La primera corriente de las ciencias cognitivas, denominada *cognitivismo*, sostenía que: (a) podía realizarse una teoría abstracta de la mente y de sus distintas realizaciones biológicas, sociales y/o artificiales; (b) es posible realizar un análisis interno de las representaciones que se producen entre los dispositivos de entrada y de salida de una mente cualquiera; (c) puede establecerse una lingüística “innatista” como método para conocer y reproducir esas representaciones interpretadas a través del análisis *sintáctico* (la gramática) de las proposiciones, sin preocuparse por el nivel *semántico* (la significación).

De las ciencias cognitivas se desprenderá el campo de la IA siguiendo tres premisas: (a) la computadora es un modelo eficaz para entender cómo funciona la mente (derivado de las ciencias cognitivas); (b) pueden realizarse programas que simulan funciones intelectuales (a través del *feedback* cibernético, el análisis del estado interno de la máquina y la sensibilidad de dicha máquina al entorno en el que funciona); (c) los procesos resultantes pueden ser mecanizados, automatizados y reproducidos por otras



máquinas. Por esa razón se sostiene que se trata de una “inteligencia” que es “artificial”, para la cual se definen las siete áreas de investigación y desarrollo que mencionamos al principio<sup>17</sup>.

Durante los años 60 del siglo pasado la IA fue una prioridad para las agencias de Defensa y Seguridad de los Estados Unidos y Gran Bretaña, especialmente, donde había surgido y se había desarrollado la cibernética. A fines de la década de 1950 el equipo de Frank Rosenblatt diseñó el Perceptrón en el Cornell Aeronautical Laboratory como una realización práctica de las redes neuronales de McCulloch-Pitts. El Perceptrón es uno de los antecedentes de los algoritmos diseñados para aprendizaje supervisado, que hoy se conoce como *machine learning*. Entre 1964 y 1966 el equipo de Joseph Weizenbaum en el Massachusetts Institute of Technology (MIT) puso a punto un *bot* conversacional llamado Eliza, que es el antecedente directo de los sistemas de procesamiento de lenguaje natural bajo modelo de diálogo como el actual chatGPT. Además, en esos años los fondos de estas agencias estaban dirigidos a la creación de sistemas de reconocimiento de patrones en imagen y en sonido. Finalmente, otra de las áreas de desarrollo de la IA fue la generación de sistemas expertos que simularan el razonamiento en un área específica para ayudar a la toma de decisiones en ámbitos como la medicina.

El desarrollo de estas cuatro áreas –redes neuronales, sistemas expertos, modelos de diálogo “natural” y sistemas de reconocimiento de patrones–, que son las mismas que originaron la explosión actual de la IA, impulsó la delimitación de dos interpretaciones básicas de las perspectivas del campo. Una se denominó IA débil, que sostenía que la computadora *simulaba* actividades humanas pero no se confundía con ellas, de acuerdo a los aspectos nodales de los sistemas expertos y los modelos de diálogo “natural”. La segunda se denominó IA fuerte, y postulaba que, en tanto realización de la mente, la IA es *semejante* a la mente humana, básicamente a partir del funcionamiento masivo de las redes neuronales.

Esta distinción entre simulación y semejanza se establecía sobre la base de ciertos límites en el desarrollo de las computadoras, pues aún no existían los microprocesadores ni los circuitos integrados, ni la generalización del silicio como materia prima de la industria informática. Esto provocó que, por un lado, surgiera una impugnación epistemológica al cognitivismo, el llamado *conexionismo*, que enfatizaba la superioridad material del cerebro biológico y la plasticidad de sus interconexiones para “procesar más información” que cualquier computadora sobre la misma base del modelo de las redes neuronales. Y por el otro, que asomaran dudas sobre los resultados de las investigaciones en las áreas mencionadas. Hacia fines de los años 60,

---

<sup>17</sup> Computadoras automáticas; programación de computadoras para que usen un lenguaje; redes neuronales; teoría del tamaño de un cálculo; auto-mejoramiento [*self-improvement*] de una máquina; métodos maquímicos para formar abstracciones, y relación entre azar y creatividad.



para contrarrestar estas críticas, una de las figuras centrales de la IA, Marvin Minsky, sostenía que los límites de la IA no eran del orden de lo artificial, ni de su contrastación con la dotación biológica del cerebro, sino del orden de lo social, pues en el momento en el que las computadoras fueran situadas en un entorno tan complejo como la sociedad misma, en lugar de los laboratorios de computación, mostrarían todo su potencial.

La argumentación de Minsky sirvió de poco para convencer a las agencias de Defensa y Seguridad que sostenían a la IA, y así fue como, en la primera mitad de la década de 1970, este campo de investigación se congeló en el denominado “invierno de la IA”. En coincidencia con la crisis del petróleo en las economías occidentales, la Agencia estadounidense de Proyectos de Investigación Avanzados de Defensa (DARPA) suspendió el financiamiento de varios proyectos, algunos referidos a sistemas de reconocimiento del habla, mientras en Gran Bretaña el Informe Lighthill era lapidario respecto de las posibilidades futuras de la IA en ese país.

En la década de 1980, y hasta la constitución de internet en su variante comercial hacia mediados de los años 90, las computadoras evolucionaron en miniaturización de componentes, velocidad de procesamiento y capacidad de almacenamiento, pero eso no se tradujo en cambios significativos en el campo de la IA. Sin embargo, la arquitectura reticular de nodos de internet permitió que muchas computadoras pasaran a compartir datos y, a partir de las mejoras en los protocolos de comunicación entre ellas y en los sistemas comunes de codificación, a incluso procesar información de manera conjunta. Como consecuencia de esta transformación, y de la actividad comercial de internet, tanto la informática como la IA lograron atraer la inversión privada en los países más desarrollados, especialmente Estados Unidos, frente a la merma de la inversión directa de los agentes estatales.

A principios de los años 2000 comenzó el proceso que desemboca en los desafíos y preocupaciones actuales respecto de los dominios donde la IA “reemplaza” a los seres humanos y cuáles serían las consecuencias económicas, sociales, políticas y éticas. En el campo específico de la IA, la mayor velocidad de procesamiento y el acceso a grandes volúmenes de datos creó la posibilidad de elaborar modelos pre-entrenados [*pre-trained models*], como el actual chat GPT en sus diferentes variantes, esto es, un modelo o red de modelos que son automáticamente entrenados por grupos de datos para resolver determinados problemas; una suerte de auto-programación de los modelos computacionales, hasta entonces limitados, por un lado, a una provisión de datos más “artesanal” y, por el otro, a un tipo de programación ligada casi exclusivamente a secuencias lógicas equiparables a sistemas de lenguaje. Así fue como volvió a ganar importancia la tesis de las redes neuronales y del *conexionismo* como lógica de funcionamiento de red.



## DE LA IA AL SISTEMA DAP

Ahora bien, esta transformación de la IA vino de la mano de la reticulación de los nodos de internet, de manera que esos modelos pre-entrenados comenzaron a ser alimentados constantemente y de manera automática por grandes volúmenes de datos (*big data*). Y esto, a su vez, potenció a las corporaciones de *software*, que comenzaron a encontrar en los datos y la automatización de los procesos algorítmicos la clave para un nuevo modelo de negocios cuyo caso emblemático es Google, que ofrece una gran cantidad de datos y de productos digitalizados de la cultura (videos, música, imágenes) a cambio de controlar el entorno de los intercambios producidos en internet a través del sistema Android, para capturar cada vez más datos. En el seno de este modelo, que se conoce como modelo de plataformas, surgieron las redes sociales, desde Youtube hasta Instagram, pasando por Facebook, que no sólo aumentaron exponencialmente la comunicación y con ello la posibilidad de digitalizar grandes porciones de la vida social global, sino que también generaron incentivos para que las corporaciones lideraran la investigación y el desarrollo de la IA, en sentido contrario a lo que ocurría en los años 1960 y 1970.

Esta situación provocó transformaciones en diferentes planos. Ante todo, la IA encontró una salida a los dilemas de aquellos años: no se trata de que la inteligencia tenga mejores raíces biológicas que artificiales, sino de que las computadoras, como sede de la IA, se conecten directamente con lo social, confirmando en cierta manera las presunciones de Minsky sobre los niveles de complejidad que pueden alcanzar las máquinas “inteligentes” en entornos diferentes a los de un laboratorio. Luego, la opción más decidida por las tesis de las redes neuronales (sostenidas inclusive por algunos de los firmantes de la carta de abril de 2023 que aboga por una pausa de al menos seis meses y una reflexión sobre las consecuencias de la IA, como Geoffrey Hinton), basadas en mayor procesamiento y mayor volumen de datos, permitió justamente “desbloquear” las áreas que hacia 1970 ya habían experimentado límites en su desarrollo: los modelos de lenguaje natural y los sistemas de reconocimiento de patrones en sonidos, en voces especialmente, y en imágenes. Esto generó y genera un *feedback* “cibernético”, porque la automatización de los patrones de generación y reconocimiento de sonidos e imágenes acelera la capacidad de los modelos pre-entrenados para optimizar nuevos modelos y predecir patrones a una escala prácticamente no humana. Entre 2009 y 2012, por ejemplo, los sistemas de IA consolidaron el reconocimiento de fonemas y de una amplia variedad de objetos artificiales y naturales, lo que permitió aumentar exponencialmente la capacidad de generar sonidos e imágenes a partir de patrones dinámicos y generar diversas aplicaciones en las tecnologías de uso cotidiano, que tienden así a “reproducir” en entornos digitales casi cualquier aspecto de la vida social.



De ello se desprende que la división anterior entre una IA débil (simulación de la mente humana) y una IA fuerte (una mente artificial similar a la humana) se reconstituya en torno a una tripartición, que describió Raymond Kurzweil en su libro *La Singularidad está cerca* (2005): una IA estrecha [*narrow AI*], que se especializa en tareas limitadas según el modelo de los sistemas expertos (juegos, transacciones financieras, geolocalización, etc.), una IA general [*general AI*], que aspira a un desarrollo similar al de la mente humana en diferentes aspectos y actividades; y una Super IA [*Super AI* o *Singularity*] que se plantea como una inteligencia que ya no tiene como referencia a la inteligencia humana porque la supera tanto en velocidad de procesamiento como en cantidad de datos procesados. Se trataría de una inteligencia de la cual no conocemos sus rasgos fundamentales porque no tiene una escala antropométrica.

En esta visión, de acuerdo a los aspectos que hemos delineado hasta aquí de la historia de la IA, los sistemas informáticos como base de la IA ya no buscan *simular* o *asemejarse* a una mente humana, sino que calculan algo para ella incalculable, y por ende, siguiendo a McCulloch y Pitts, también representan algo no representado ni representable por ella. La IA hoy puede *crear*, *inventar*, y sobre todo *operar* sobre el mundo humano sin tener ya como referencia a un ser humano aislado, como ocurría en los años 1960, sino estando inmersa en la vida social, cultural y política de millones de seres humanos. De acuerdo a las analogías con las ciencias del lenguaje que acompañan a la IA desde sus inicios, ya no hace falta concentrarse en el nivel *sintáctico* (las reglas gramaticales, más próximas a las reglas lógicas y a la idea tradicional de cálculos), ni discutir qué ocurre con el nivel *semántico* (si la IA tiene conciencia de ser una inteligencia a partir del seguimiento de esas reglas, si comprende lo que hace), sino que la IA opera en el nivel *pragmático*, en el uso concreto y social de la lengua y el habla, por la mera posibilidad de procesar automáticamente miles de millones de expresiones humanas de diversa índole en nanosegundos.

De este modo se “cumplen” los siete ejes que dieron inicio al campo de la IA: las computadoras funcionan automáticamente, hacen uso del lenguaje a través de redes neuronales que amplían sin cesar la capacidad de cálculo e introducen mejoras en ese funcionamiento, y no sólo logran formar abstracciones semejantes a un “pensamiento humano”, sino que también expresarían algún grado de creatividad. Sin embargo, difícilmente los impulsores iniciales de la IA hubieran imaginado el escenario actual.



## PERSPECTIVA ANALÍTICA

Desde el inicio, para abordar el análisis de los riesgos y desafíos de las IA generativas, entendimos que debíamos tomar una serie de decisiones que, una vez asumidas, fueron parte relevante de la delimitación epistemológica y metodológica de esta investigación. Identificamos así cinco rasgos o aspectos de las IA.

El primero es conocido, pero no es ocioso recordarlo: las Inteligencias Artificiales no son *una* tecnología, sino que son **metatecnologías**, esto es, tecnologías de propósito general, aplicables a muy diversas actividades. Y esto es así en al menos tres sentidos.

Por un lado, y principalmente, porque –como escribe Ariel Vercelli en su artículo “Las inteligencias artificiales y sus regulaciones: Pasos iniciales en la Argentina, aspectos analíticos y defensa de los intereses nacionales” (2023)--, las IA “pueden ser analizadas como redes heterogéneas, híbridos y ensamblajes tecnológicos. Su mera existencia evidencia la articulación e integración de éstas con otras redes, prácticas y procesos científico-tecnológicos más amplios. Al igual que ocurrió con el software (los programas de computación) en las etapas tempranas de la computación electrónica digital, en muchas ocasiones las IA también están indiferenciadas de los dispositivos y sistemas tecnológicos donde están incorporadas” (2023, 208).

El hecho de que muchas veces las IA estén indiferenciadas de los dispositivos y sistemas tecnológicos donde están incorporadas tiene efectos tanto en el nivel analítico como en el de la gobernanza. Para decirlo sintéticamente, a los fines regulatorios, no alcanza con establecer *una* norma general para las IA, sino que es preciso identificar las capas y subsistemas que participan en las IA para analizar las diferentes legislaciones que las atraviesan, desde protección de datos y derechos de autor hasta legislación laboral y de protección del medioambiente.

Por otro lado, tal como señala el filósofo Luciano Floridi, conocido por su trabajo en ética de la IA, se denominan metatecnologías aquellas tecnologías que “operan y regulan otras tecnologías” (Floridi 2011: 91. Y si bien no siempre es así, en ocasiones las IA son metatecnologías como las leyes o las tecnologías de seguridad, porque son “parte de las condiciones de operación de otras tecnologías” (idem).

Finalmente, porque tal como afirman Ajay Agrawal, John McHale y Alex Oettl en su artículo “Finding needles in haystacks: Artificial Intelligence and recombinant growth [Encontrar agujas en pajares. Inteligencia artificial y crecimiento recombinante]”, las IA también nos ayudan a producir conocimientos que jamás tendríamos sin ellas. En este sentido, la reciente explosión en la disponibilidad de datos y los avances informáticos en las capacidades para descubrir y procesar esos datos “pueden ser vistos como



‘metatecnologías’, esto es: tecnologías para la producción de nuevos conocimientos” (2018: 3).

“Por supuesto –agregan–, las metatecnologías que ayudan en el descubrimiento de nuevos conocimientos no son nada nuevo. [...] [Pero] la promesa de la IA como metatecnología para la producción de nuevas ideas es que facilita la búsqueda en espacios de conocimiento complejos, permitiendo tanto un mejor acceso al conocimiento relevante como a una mejor capacidad para predecir el valor de nuevas combinaciones”. Es en este mismo sentido que Alex Trollip (2021) escribe que “las metatecnologías son tecnologías o invenciones que tienen la capacidad de ayudar a nuevos descubrimientos o estimular la innovación en otras áreas”.

En cierta medida, las IA son nuevas modalidades de existencia de la “máquina universal”, en el sentido que utilizaba Alan Turing esa imagen para referirse a la máquina de computar como un instrumento capaz de realizar innumerables tareas y colaborar en su desarrollo y amplificación.

El segundo rasgo a considerar es que las inteligencias artificiales, en la medida en que integran y expanden el ecosistema digital, constituyen no una herramienta o dispositivo técnico, sino que –particularmente después del shock de virtualización que implicó la última pandemia (Costa 2021)-- han comenzado a ser para nosotros un **mundoambiente**. Las tecnologías del ecosistema digital están dejando de ser instrumentos que podemos elegir usar o no, y se han vuelto cada vez más indispensables para realizar actividades cotidianas como guiarnos en una ciudad o iniciar un trámite de documentación obligatoria (Zuboff, 2021), Esto significa que, en relación con los usuarios, no es suficiente un enfoque que las aborde desde la perspectiva de la instrumentalidad y de la relación sistema-usuario individual, como si cada usuario pudiera decidir qué hacer en cada caso con la IA, sino que es necesario un enfoque sistémico, atento a las dinámicas multiescalares de la economía política del ecosistema digital.

El tercer rasgo –que por momentos emerge pero no ha sido suficientemente explorado en la literatura sobre IA– es particularmente significativo para esta investigación, y tiene implicaciones importantes. Consiste en señalar que, a los efectos de un análisis epistemológico con epicentro en las ciencias sociales, las llamadas IA generativas y en particular los LLM son, no sólo Inteligencia Artificial, sino **Sociedad Artificial**. Esto es así debido a que los LLM operan desde y sobre el mundo social a través del sistema de Datos, Algoritmos y Plataformas (DAP).

Veamos esto: los elementos básicos de los LLM son tres: (1) enorme capacidad de cómputo (a grandes rasgos, la capa del hardware), (2) métodos de procesamiento de información (aprendizaje profundo, redes neuronales, etc.; las capas del software y de las aplicaciones de IA) y (3) grandes conjuntos de datos (materiales “de la vida social”, obtenidos en buena medida a través de plataformas; las capas de input y de usuarios).



Es decir: su alimento (input) y su producción específica (output) son los intercambios lingüísticos en diferentes idiomas, las figuras retóricas y las reacciones emocionales, las relaciones sociales de diversas culturas. Estos sistemas sociotécnicos complejos que son los LLM y las IA generativas aceleran el procesamiento, la gestión y la (re)producción de lo social. *Producen sociedad*. Y si las capas 1 y 2 son el producto de desarrollos históricos de las ciencias de la informática y la computación, el estudio y el trato con la capa 3 es el dominio de las ciencias del lenguaje, las ciencias de la comunicación, la sociología y la ciencia política. De allí que es deseable que en su desarrollo, su análisis y su monitoreo participen expertos en esos campos disciplinares.

El cuarto rasgo consiste en que, en ciertos usos, **las IA pueden ser tecnologías de alto riesgo**, y por lo tanto requieren un tratamiento acorde a lo largo de todo su ciclo de vida. Es preciso distinguir cuáles son esos casos o usos, y abordarlos con la perspectiva sistémica de la seguridad y la gestión de riesgos. Volveremos a esto en seguida.

El quinto elemento no es tanto un rasgo de las IA generativas sino una consecuencia analítica de tener en cuenta los rasgos anteriores. Para afrontar las IA generativas desde las ciencias sociales y humanas, para pensar eficazmente su gobernanza, **no es suficiente** la perspectiva de **la ética de la IA**, que es la forma más habitual en la que los saberes de las ciencias humanas se presentan en la discusión (por ejemplo, en la Recomendación sobre la ética de la inteligencia artificial, de 2021). Si de lo que estamos hablando es de metatecnologías que constituyen un mundoambiente, que aceleran el procesamiento y la gestión de lo social, y que pueden ser de alto riesgo en áreas de experiencia críticas para la población como el acceso a la salud, al empleo, a la educación, nos encontramos ante sistemas sociotécnicos complejos. Y por ende, la perspectiva que necesitamos va más allá de las recomendaciones voluntarias, que ofician como “códigos de buenas prácticas”.

Es muy posible que como dice Brent Mittelstadt en su artículo “Principles alone cannot guarantee ethical AI” (2019), un enfoque de ética profesional –sostenido en una dudosa similitud entre la práctica de la A y la práctica de la medicina-- sea insuficiente. No alcanza con establecer principios: “Se necesitan estructuras de rendición de cuentas vinculantes y altamente visibles, así como procesos claros de implementación y revisión a nivel sectorial y organizacional” (2019: 4).

Las iniciativas de ética de la IA hasta ahora han producido principios y declaraciones de valores que prometen guiar la acción, pero en la práctica brindan pocas recomendaciones específicas y no abordan normas y políticas fundamentales. Siempre según Mittelstadt, entre las tareas necesarias una de ellas es licenciar a los desarrolladores de IA de alto riesgo, definiendo requisitos claros de confiabilidad y reputación. “Puede ser necesario establecer el desarrollo de la IA como una profesión con una categoría equivalente a otras profesiones de alto riesgo --asegura--. Es una rareza regulatoria que otorguemos licencias a profesiones que brindan un servicio



público, pero no a la profesión responsable de desarrollar sistemas técnicos para aumentar o reemplazar la experiencia humana y la toma de decisiones dentro de ellos. Los riesgos de las profesiones autorizadas no se han disipado, sino que han sido desplazados a la IA. Para analizar los importantes desafíos que enfrentan, las iniciativas podrían dirigirse inicialmente a los desarrolladores de diseño inclusivo, revisión ética transparente, documentación de modelos y conjuntos de datos, y auditoría ética independiente” (2019, 9-10).

De manera afín, Ariel Vercelli (2023) señala la necesidad de no confundir “los problemas éticos de las IA con las políticas y las regulaciones”, ya que “¿qué poder jurídico-político tienen las recomendaciones sobre ética de la IA de la UNESCO? Siempre son útiles, pero se tratan sólo de meras recomendaciones”. Para este investigador argentino, “es necesario comprender que estas posiciones éticas podrían no ser el mejor enfoque para avanzar sobre políticas públicas y regulaciones nacionales”. En cambio, sugiere “superar los comités de ética y ofrecer a la sociedad una discusión más amplia, abierta y democrática sobre IA”. Desde su perspectiva, el peligro de las IA es que “es que profundicen las injusticias del mundo real: desigualdades sociales, económicas, jurídico-políticas y ambientales”. De allí que se requieren procesos de co-construcción entre tecnologías y regulaciones; esto es, “desarrollar tecnologías (incluso para fines regulativos: control y gestión del tiempo, el espacio y las conductas)” (Vercelli, 2023: 213-214)

En síntesis: para afrontar las IA generativas desde las ciencias sociales y humanas, para que su gobernanza sea eficaz, no es suficiente con un enfoque desde la *ética profesional* de la IA sino que es preciso promover un enfoque desde la *ética organizacional* de la IA y desde el *pensamiento sistémico*, que busca establecer procedimientos de revisión transparentes, estructuras de rendición de cuentas vinculantes, documentación de modelos y conjuntos de datos, auditoría independiente. En suma: defensas en profundidad a lo largo del sistema para que este sea más seguro y confiable.



## MARCOS NORMATIVOS EN MATERIA DE IA ANÁLISIS DE LOS MARCOS NORMATIVOS DE EE.UU. Y LA UNIÓN EUROPEA

En este apartado presentaremos las principales características de las normativas y otros documentos relativos al uso de sistemas IA propuestas en la Unión Europea y en los Estados Unidos, con especial énfasis en los marcos de referencia destinados a la gestión de los riesgos, lo cual incluye las actividades destinadas a dirigir y controlar una organización con respecto al efecto de la incertidumbre sobre el logro de sus objetivos.<sup>18</sup>

Los documentos analizados son, para la Unión Europea, el Reglamento del Parlamento Europeo y del Consejo de la UE por el que se establecen normas en materia de inteligencia artificial (Ley de Inteligencia Artificial, Bruselas, del 21 de abril de 2021) y las Enmiendas aprobadas por el Parlamento Europeo el 14 de junio de 2023 sobre aquella propuesta de Reglamento. Y para los Estados Unidos, la *Executive Order* 13859 de febrero de 2019, por la que se crea la Iniciativa Nacional de Inteligencia Artificial (*National Artificial Intelligence Initiative act*, NAI, de 2020, que entró en vigencia el 1° de enero de 2021) y el Marco de Gestión del Riesgo de IA (*Artificial Intelligence Risk Management Framework*, AI RMF 1.0) de enero de 2023, emitido por el Departamento de Comercio en el marco de la Iniciativa NAI.

### BREVE DESCRIPCIÓN ACERCA DE LA GESTIÓN DE RIESGOS

El proceso llamado gestión de riesgos consiste en un conjunto de actividades coordinadas para dirigir y controlar los sistemas sociotécnicos con relación a sus consecuencias no deseadas y los modos de mitigarlas. Sintéticamente, un proceso de gestión de riesgos consiste en:

1. Definición del alcance, el contexto y los criterios para considerar el riesgo.
2. Identificación del riesgo, también denominado identificación de peligros.
3. Evaluación del riesgo: probabilidad, severidad, exposición.
4. Valoración del riesgo: tolerable, tolerable con medidas de mitigación o no tolerable.
5. Tratamiento del riesgo, o implementación de medidas de mitigación.
6. Seguimiento y monitoreo, también denominado garantía de la seguridad operacional.

Al implementar un sistema de Inteligencia Artificial (IA), los diseñadores e implementadores tienen dos desafíos básicos: que el sistema satisfaga los requerimientos de producción (para lo cual fue diseñado) y proteger al sistema, sus operadores y usuarios de los riesgos durante su ciclo de vida.

En cuanto al segundo desafío, los diseñadores e implementadores de IA se enfrentan a diferentes tipos de riesgos: de seguridad operacional o *safety*, como accidentes e incidentes propios del sistema; de seguridad y protección o *security*, como delitos,

---

<sup>18</sup> Cf. *Risk management principles and guidelines*, ISO 31000:2018, IDT.



sabotajes, *hackeos*: incidentes o accidentes por intervención maliciosa; riesgos financieros (bancarrota), riesgos políticos, entre otros. Por lo cual cabe aclarar que la gestión de riesgos se refiere al control de los peligros potenciales que puedan presentar los sistemas de IA, en particular los referidos a la *safety* y, en cierta medida, los referidos a la *security*.

## DOCUMENTOS DE LA UNIÓN EUROPEA (UE) Y DE LOS ESTADOS UNIDOS (EEUU)

La Unión Europea emitió en abril de 2021 una Ley general que regula el desarrollo y la incorporación de sistemas de IA, cuyas últimas modificaciones fueron aprobadas en junio de 2023. Su denominación es Reglamento del Parlamento Europeo y del Consejo de la UE por el que se establecen normas generales en materia de inteligencia artificial, que luego deben ser incorporadas y adaptadas por cada país de la Unión a sus legislaciones específicas.

Este Reglamento está compuesto por doce títulos, algunos de los cuales poseen capítulos. En total suman 85 artículos. El esquema general es como sigue:

Título I. Disposiciones generales.

Título II. Prácticas de IA prohibidas

Título III. Sistemas de IA de alto riesgo

Capítulo 1. Clasificación de los sistemas de IA como sistemas de IA de alto riesgo

Capítulo 2. Requisitos para los sistemas de IA de alto riesgo

Capítulo 3. Obligaciones de los proveedores y usuarios de sistemas de IA de alto riesgo y de otras partes

Capítulo 4. Autoridades notificantes y organismos notificados

Título IV. Obligaciones de transparencia para determinados sistemas de IA

Título V. Medidas de apoyo a la innovación

Título VI. Gobernanza

Capítulo 1. Comité europeo de IA

Capítulo 2. Autoridades nacionales competentes

Título VII. Base de datos de la UE para sistemas de IA de alto riesgo

Título VIII. Seguimiento posterior a la comercialización, intercambio de información, vigilancia de mercado

Capítulo 1. Seguimiento posterior a la comercialización

Capítulo 2. Intercambio de información sobre incidentes y fallos de funcionamiento

Capítulo 3. Ejecución

Título IX. Códigos de conducta

Título X. Confidencialidad y sanciones

Título XI. Delegación de poderes y procedimientos del comité

Título XII. Disposiciones finales



El Reglamento europeo ubica, en su primera página, la gestión de riesgos como proceso fundamental para lograr una IA fiable. Bajo el subtítulo “Contexto de la propuesta razones y objetivos”, enuncia:

Los mismos elementos y técnicas que potencian los beneficios socioeconómicos de la IA también pueden dar lugar a nuevos riesgos o consecuencias negativas para personas concretas o la sociedad en su conjunto (p. 1).

De allí que adopta un pensamiento sistémico acorde al marco teórico utilizado por las industrias de alto riesgo y organizaciones de gran confiabilidad (*High Reliability Organization*, o HRO, por su sigla en inglés), que consideran a las consecuencias negativas (accidentes e incidentes) como propiedades emergentes del sistema.<sup>19</sup>

La perspectiva de los Estados Unidos es diferente. En este país, la principal norma relacionada con los sistemas de IA no adopta principalmente la perspectiva del riesgo, sino que se dirige a asegurar el liderazgo de los EE.UU. en relación con la investigación y el desarrollo de sistemas de IA. Dicha norma se denomina *National Artificial Intelligence Initiative Act* (Iniciativa Nacional de Inteligencia Artificial). Fue sancionada en 2020 y entró en vigencia el 1° de enero de 2021. A diferencia del Reglamento europeo, los propósitos de esta Iniciativa son:

- Asegurar el liderazgo continuado de los Estados Unidos en la investigación y desarrollo de inteligencia artificial;
- Liderar el mundo en el desarrollo y uso de sistemas confiables de inteligencia artificial en los sectores públicos y privados;
- Preparar a las fuerzas de trabajo tanto presentes como futuras de los Estados Unidos para la integración con los sistemas de inteligencia artificial a lo largo de todos los sectores de la sociedad; y
- Coordinar la investigación, desarrollo y demostración de las actividades en marcha de inteligencia artificial entre las agencias civiles, el Departamento de Defensa y la Comunidad de Inteligencia para asegurar que cada una informa el trabajo de las otras.(p.3)

Como puede verse, esta Iniciativa tiene como misión promover el liderazgo estadounidense en el desarrollo de IA. En ese marco, encomienda al National Institute of Standards and Technology (Instituto Nacional de Estándares y Tecnologías, NIST), dependiente del Departamento de Comercio, la elaboración de un marco general y voluntario para la gestión de riesgos. El hecho de que se trate de un marco voluntario y no una norma obligatoria ya deja clara la diferencia de enfoque.

En su texto de 2020, la NAII establecía que en no más de dos años a partir de su entrada en vigencia, el NIST debía confeccionar, y luego actualizar periódicamente en

---

<sup>19</sup> Charles Perrow define estas propiedades emergentes como “accidentes normales” o “sistémicos”, que resultan de “la interacción imprevista de múltiples fallos”. Cf. PERROW, Charles ([1984] 2009). *Normal Accidents: Living with High Risk Technologies*, edición actualizada. Princeton, N.J., Princeton University Press. p. 99.



colaboración con otras organizaciones públicas y privadas del sector, “un marco de referencias voluntario de gestión de riesgos para sistemas de IA confiables”. Ese proceso dio como resultado el Artificial Intelligence Risk Management Framework (AI RMF 1.0), Marco General para la Gestión de Riesgos de IA, publicado por el NIST el 26 de enero de 2023.<sup>20</sup> En este documento se explicita que el Marco de Gestión de Riesgo de IA o AI RMF será un documento vivo, al menos hasta 2028:

El NIST revisará el contenido y la utilidad del Marco de Referencia regularmente para determinar si una actualización es apropiada, una revisión con inputs formales de la comunidad de IA se espera que ocurra no más tarde que 2028. (p.19)

Por último el IA RMF establece que la gestión de riesgos de IA es un componente clave del desarrollo y uso responsable de los sistemas de IA. “Entender y gestionar los riesgos de sistemas de IA ayudará a potenciar la confiabilidad y cómo retorno cultivar la confianza pública”. (p.6)

## DOS PERSPECTIVAS SOBRE LA GESTIÓN DE RIESGOS

Como se mencionó, en los Estados Unidos el AI RMF no es una norma vinculante, sino un marco de referencia voluntario. Se trata de una estrategia **reactiva** con respecto a los riesgos, en la medida en que no los aborda de manera directa, ni se refiere a usos o prácticas específicas de ningún sector.

El objetivo del AI RMF es ofrecer un recurso a las organizaciones que diseñan, desarrollan, implementar o utilizar sistemas de IA para ayudar a gestionar los numerosos riesgos de la IA y promover el desarrollo y el uso confiable y responsable de los sistemas de IA. El Marco pretende ser voluntario, busca preservar derechos, no es específico de ningún sector y no se refiere a casos específicos de uso, brindando flexibilidad a organizaciones de todos los tamaños, en todos los sectores y en toda la sociedad para implementar sus enfoques. El Marco está diseñado para equipar a organizaciones e individuos [a los que se referiremos como Actores de IA] con enfoques que aumenten la confiabilidad de los sistemas de IA y ayuden a fomentar el diseño, desarrollo, implementación y uso responsable de sistemas de IA a lo largo del tiempo. (p.7).

La Unión Europea, en cambio, adoptó una estrategia **proactiva** con relación a los riesgos. En efecto, el texto entero está orientado a delimitar diferentes tipos de usos y prácticas en los que puedan participar sistemas de IA, y a prescribir cuáles usos están prohibidos, cuáles se consideran de alto riesgo, y cuáles de bajo riesgo. Establece así de manera explícita tres categorías de riesgos, y deja implícita la existencia de una cuarta categoría de prácticas no reguladas, que no requieren vigilancia (como por ejemplo en los videojuegos).

---

<sup>20</sup> NIST (2023): Artificial Intelligence Risk Management Framework (AI RMF 1.0). En internet: [www.nist.gov/itl/ai-risk-management-framework](http://www.nist.gov/itl/ai-risk-management-framework) (incluido en el anexo documental).

- Riesgos inaceptables, que se asocian a usos o prácticas prohibidas, tal como aparecen definidas en el Título II: Prácticas de inteligencia artificial prohibidas
- Riesgos altos, asociados a sistemas de IA de alto riesgo, donde se deben cumplir un conjunto de requisitos obligatorios de gestión de riesgos, definidos en el Título III; y
- Riesgos Bajos o Mínimos, al que dedica el Título IV, y con respecto a los cuales establece requisitos obligatorios de transparencia.

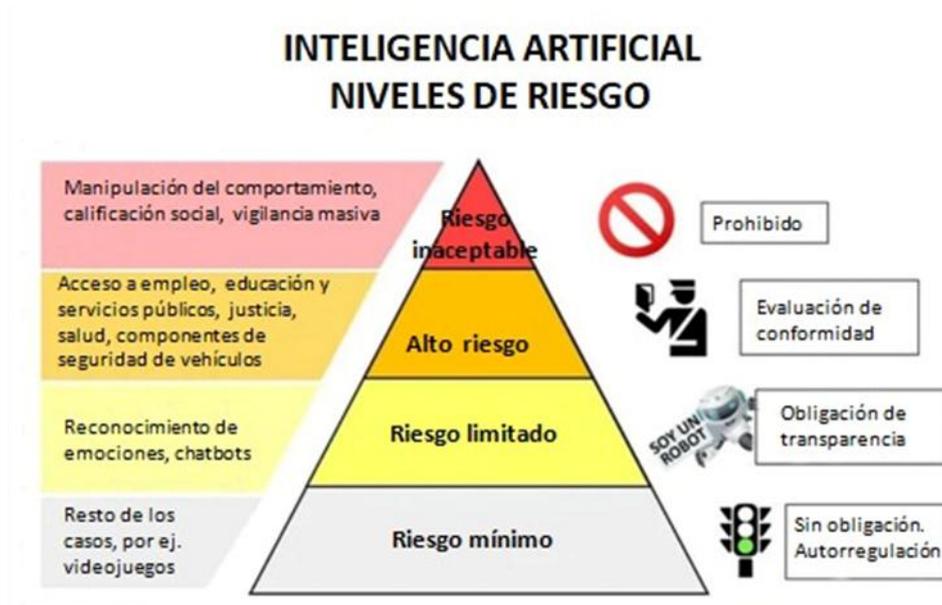


Figura 1. Niveles de riesgo de acuerdo al Reglamento Europeo

Esta diferenciación entre niveles de riesgos representa la discretización de los usos o aplicaciones y sus respectivos riesgos. Y especifica cuáles usos se deben evitar para no incurrir en delitos (Riesgo Inaceptable); cuáles deben ser obligatoriamente gestionados (Riesgos Altos), gestión que implica un complejo circuito de gestión de calidad, documentación, certificaciones, notificaciones; y cuáles deben ser transparentes para el usuario, donde los desarrolladores e implementadores deben informar al usuario que está interactuando con un sistema de IA, en el caso de Riesgos Bajos o Mínimos.

### NIVELES DE RIESGOS: ESQUEMA EUROPEO

Siguiendo el Reglamento europeo, el Título II, llamado Prácticas de inteligencia artificial prohibidas, enumera en su artículo 5 las siguientes prácticas prohibidas (se reproduce un resumen):

- Los sistemas de IA que utilicen técnicas subliminales que trasciendan la conciencia de una persona para alterar de manera sustancial su comportamiento.
- Los sistemas de IA que aproveche alguna de las vulnerabilidades de un grupo específico de personas debido a su edad o discapacidad física o mental.



- Los sistemas de IA que tengan por objeto la clasificación social, evaluando la fiabilidad de personas físicas atendiendo a su conducta social o a características personales conocidas o predichas.
- El uso de sistemas de identificación biométrica remota en tiempo real con fines de aplicación de la ley, excepto para la prevención de una amenaza importante para la vida o la seguridad física de las personas. En estos últimos casos se enumeran requisitos específicos.

En el Título III se describen los sistemas de IA de Alto Riesgo (se reproduce un resumen):

- Identificación biométrica y categorización de personas físicas
- Gestión y funcionamiento de infraestructuras esenciales
- Educación y formación profesional
- Empleo, gestión de los trabajadores y acceso al autoempleo
- Acceso y disfrute de servicios públicos y privados esenciales y sus beneficios
- Asuntos relacionados con la aplicación de la ley
- Gestión de la migración, el asilo y el control fronterizo

En referencia a los Riesgos Bajos o Mínimos, el Título IV, titulado Obligaciones de transparencia, describe en su artículo 52 la obligatoriedad de informar al usuario que está interactuando con un sistema IA. Estos sistemas son:

- Sistemas de IA destinados a interactuar con personas físicas.
- Sistema de reconocimiento de emociones o de categorización biométrica.
- Sistema de IA que genere o manipule contenido de imagen, sonido o vídeo que se asemeja notablemente a personas, objetos, lugares u otras entidades, y que pueda inducir erróneamente a pensar que son auténticos (ultra falsificación).

La Ley de la UE enuncia excepciones generales a la transparencia en sistemas autorizados por ley para fines de detección, prevención, investigación o enjuiciamiento de infracciones penales.

En cuanto a los EEUU, el AI RMF no identifica aplicaciones de alto riesgo o para las cuales sea necesario un sistema de gestión de riesgos obligatorio. En cambio, describe los desafíos que las organizaciones deben afrontar. (pág.10)

- Medición de los riesgos
- Tolerancia de los riesgos
- Priorización de los riesgos
- Integración organizacional y gestión de los riesgos.

En resumen, en la Ley europea los Estados asumen un enfoque **proactivo** ante los riesgos, identifican prácticas prohibidas, establecen mecanismos de control y obligan a contar con sistemas de gestión de riesgos y transparencia. Mientras que en los EEUU propone un marco de referencia voluntario de gestión de riesgos, con una perspectiva

reactiva, al delegar en diseñadores e implementadores la responsabilidad por eventos negativos.<sup>21</sup>

## AUTORIDADES, GOBERNANZA, RESPONSABILIDAD

### UE

En cuanto a la gobernanza de la IA, la Ley europea describe un sistema unificado para los Estados miembros que operará por medio de un mecanismo de cooperación a escala de la Unión, denominado Oficina Europea de Inteligencia Artificial (Oficina de IA). Esta oficina está pensado como “un organismo independiente [...] (con personalidad jurídica”. Sus funciones se establecen en las enmiendas aprobadas por el Parlamento Europeo el 14 de junio de 2023 (ver Anexo 1).

La Ley también se refiere a las autoridades nacionales competentes, cuya designación quedará a cargo de cada Estado miembro.

Cada Estado miembro establecerá o designará autoridades nacionales competentes con el fin de garantizar la aplicación y ejecución del presente Reglamento. Las autoridades nacionales competentes se organizarán de manera que se preserve la objetividad e imparcialidad de sus actividades y funciones. (pp. 96-97)

De la lectura del Título VI, Gobernanza, se desprende un primer mapa de actores claves (MAC):

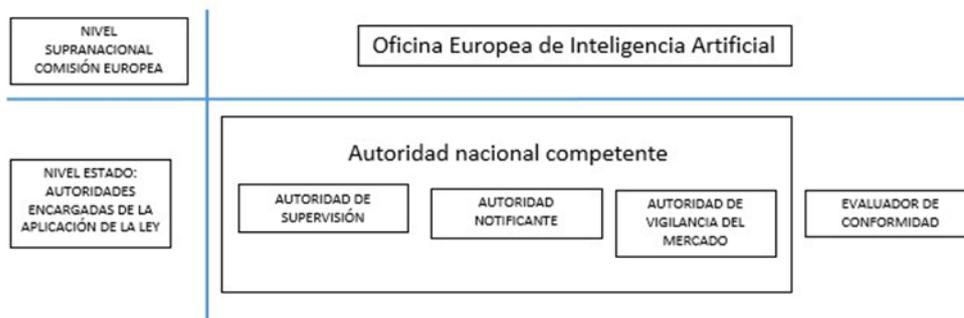


Figura 2. Mapa de Actores institucionales (UE)

En el nivel de cada Estado, la Autoridad encargada de la aplicación de la ley se define como:

- a) toda autoridad pública competente para la prevención, la investigación, la detección o el enjuiciamiento de infracciones penales o la ejecución de sanciones penales, incluidas la protección y prevención frente a amenazas para la seguridad pública; o

<sup>21</sup> Como veremos más adelante, el marco de prácticas responsables utilizado como referencia incluye tres campos: responsabilidad social, sustentabilidad y responsabilidad profesional (ver este mismo documento, página 10).



b) cualquier otro órgano o entidad a quien el Derecho del Estado miembro haya confiado el ejercicio de la autoridad pública y las competencias públicas a efectos de prevención, investigación, detección o enjuiciamiento de infracciones penales o ejecución de sanciones penales, incluidas la protección y prevención frente a amenazas para la seguridad pública. (p. 54)

La Autoridad nacional competente se define como “cualquiera de las autoridades nacionales responsables de hacer cumplir el presente Reglamento” (p. 136 de la Enmienda).

La Autoridad nacional de supervisión se define como “una autoridad pública a la que un Estado miembro asigna la responsabilidad de ejecutar y aplicar el presente Reglamento, coordinar las actividades encomendadas a dicho Estado miembro, actuar como el punto de contacto único para la Comisión, y representar al Estado miembro en cuestión en el consejo de administración de la Oficina de IA” (p. 136 de la Enmienda).

Por Autoridad notificante se entiende “la autoridad nacional responsable de establecer y llevar a cabo los procedimientos necesarios para la evaluación, designación y notificación de los organismos de evaluación de la conformidad, así como de su seguimiento” (p. 52).

La Autoridad de vigilancia del mercado designa a “la autoridad nacional que lleva a cabo las actividades y adopta las medidas previstas en el Reglamento (UE) 2019/1020” (p. 53).

El Organismo de evaluación de la conformidad es un “organismo independiente que desempeña actividades de evaluación de la conformidad, entre las que figuran la prueba, la certificación y la inspección” (p. 52).

## **EEUU**

En el caso de EEUU, al elegir lo que se denomina un enfoque reactivo, el Estado no define obligatoriedad de sistemas de gestión de riesgos ni mecanismos de control sino que brinda herramientas para que el enfoque proactivo sea responsabilidad de las organizaciones desarrolladoras y/o de implementación, considerando la totalidad de los marcos legales vigentes.

El AI RMF no define actores claves del Estado. En cambio, define las responsabilidades profesionales en el desarrollo de sistemas de IA. “Las prácticas responsables de IA pueden ayudar a alinear las decisiones referidas al diseño, desarrollo y usos bajo el objetivo y valores intencionados de un sistema de IA. Los conceptos nucleares de una IA responsable enfatizan la centralidad humana, la responsabilidad social y la sustentabilidad.” (p. 1, Sumario Ejecutivo).

Para la definición del marco de prácticas responsable utiliza como referencia las siguientes normas ISO:



- Responsabilidad social: responsabilidad de la organización por los impactos de sus decisiones y actividades en la sociedad y el ambiente a través de comportamiento transparente y ético (ISO 26000:2010)
- Sustentabilidad: el estado del sistema global, incluidos los aspectos ambientales, sociales y económicos, en el cual las necesidades del presente son satisfechas sin comprometer la capacidad de las generaciones futuras de satisfacer sus propias necesidades (ISO/IEC TR 24368:2022)
- Responsabilidad profesional: Un abordaje que apunta a asegurar que los profesionales que diseñan, desarrollan y despliegan sistemas de IA y aplicaciones o productos o sistemas basados en IA, reconozcan su posición única de ejercer influencias sobre las personas, la sociedad, y el futuro de la IA (ISO/IEC TR 24368:2022)

De esta manera la gestión de riesgos de IA representa una herramienta para conducir a usos y prácticas responsables, promoviendo que las organizaciones y sus equipos internos –quienes diseñan, desarrollan y despliegan IA– posean pensamiento crítico sobre el contexto y el potencial de los impactos positivos y negativos.<sup>22</sup> Y como se verá enseguida, asume una perspectiva colaborativa:

“Dentro del AI RMF, todos los actores de la IA trabajan juntos para gestionar los riesgos y lograr los objetivos de una IA confiable y responsable”, afirma el texto en su apartado 2. Audiencia (p. 9).

Cabe destacar, como se dijo antes, que la AI RMF no señala una autoridad del Estado encargada de la supervisión, notificación y/o vigilancia del mercado exclusivamente para IA. En este documento se considera que el formato de revisiones para la mejora de la gestión de riesgos debe incorporar los aportes de los diferentes actores en cada una de las etapas del ciclo de vida de la IA, como aparece descrito en la página 11 del AI RMF.<sup>23</sup>

Al mismo tiempo, señala cuatro momentos clave de la gestión del riesgo:

- Gobernar
- Mapear
- Medir
- Gestionar

---

<sup>22</sup> La OCDE ha desarrollado un marco para clasificar las actividades del ciclo de vida de la IA según cinco dimensiones sociotécnicas clave, cada una con propiedades relevantes para la política y la gobernanza de la IA, incluida la gestión de riesgos [OCDE (2022) Marco de la OCDE para la clasificación de sistemas de IA — OCDE]. El AI RFM retoma esta clasificación, tal como se verá en seguida. Cf. OECD (2022), "OECD Framework for the Classification of AI systems", OECD Digital Economy Papers, No. 323, OECD Publishing, Paris, <https://doi.org/10.1787/cb6d9eca-en>.

<sup>23</sup> En el AI RMF, se establecen siete dimensiones en el ciclo de vida de la IA: desde (1) la evaluación del contexto de aplicación, en la que se planifica y diseña el sistema, pasando por (2) la recolección de datos; (3) el modelado de IA, (4) la verificación y validación de ese modelado; (5) el contexto de despliegue de tareas y resultados, incluido su testeo; (6) la operación y el monitoreo, que incluye auditoría y evaluación del impacto; hasta llegar a (7) el contexto de la población y el Planeta, que incluye a los usuarios finales y los estudios de impacto ambiental y social. Ver Fig. 4.



Estas funciones representan un ciclo iterativo de actividades propuestas a las organizaciones, que comienzan con el mapeo, luego sigue la medición, en tercer lugar la gestión, y en el centro, la función de gobernar actúa de manera transversal a las demás. Estas funciones, una vez establecida la primera iteración cíclica, son repetidas en el tiempo y deberían ajustarse entre sí, ya que se encuentran estrechamente relacionadas. (Ver más en detalle en el Anexo 2).

## ACCIDENTES, INCIDENTES, NOTIFICACIONES

### UE

La Ley europea contempla la notificación de incidentes y fallos de funcionamiento. En el Título I, artículo 3, apartado 44 se define “Incidente grave” como todo incidente o defecto de funcionamiento de un sistema de IA que, directa o indirectamente, tenga, pueda haber tenido o pueda tener alguna de las siguientes consecuencias:

- (a) el fallecimiento de una persona o daños graves para su salud,
- (b) una alteración grave de la gestión y el funcionamiento de infraestructura crítica,
- (c) una vulneración de derechos fundamentales protegidos en virtud de la legislación de la Unión,
- (d) daños graves a la propiedad o al medio ambiente.<sup>24</sup>

En cuanto a fallos de funcionamiento de sistemas de alto riesgo, se deberá notificar aquellos que constituyan “incumplimiento de las obligaciones en virtud del Derecho de la Unión destinadas a proteger los derechos fundamentales”. Dicha notificación se hará ante las autoridades de vigilancia del mercado del Estado miembro en el que se haya producido el incidente.

En el Título VIII Capítulo 2, artículo 62, llamado Notificación de incidentes graves, se obliga a los proveedores e implementadores de sistemas de IA que hayan identificado un incidente grave, “notificar cualquier incidente grave [...] a la autoridad nacional de supervisión de los Estados miembros donde se haya producido dicho incidente”.

Se establece para esto un plazo máximo de 72 horas “después de que el proveedor o, en su caso, el implementador tenga conocimiento de dicho incidente grave”. Tras la recepción de la notificación la Autoridad nacional de supervisión debe informar a las autoridades u organismos públicos nacionales. Posteriormente se elaborarán orientaciones de las obligaciones establecidas.

---

<sup>24</sup> El texto citado incluye lo corregido por la Enmienda de junio de 2023.

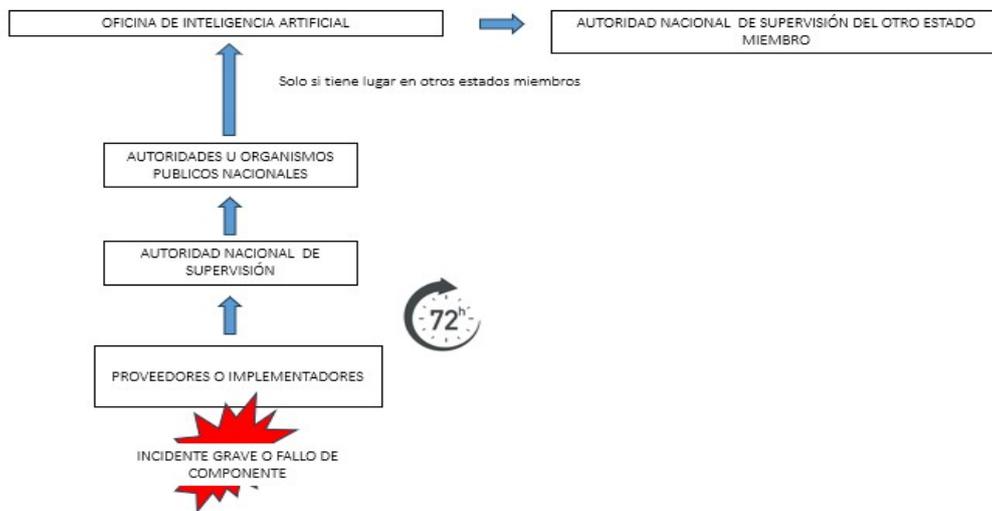


Figura 3. Notificación de incidentes graves (UE)

Las autoridades nacionales de supervisión deben adoptar medidas adecuadas en un plazo de siete días a partir de la fecha en que reciban la notificación. Si el incidente tiene lugar o es probable que tenga lugar en otros Estados miembros, la autoridad nacional de supervisión notificará a la Oficina de IA y a las autoridades nacionales de supervisión pertinentes de dichos Estados miembros.

## EEUU

En cuanto a incidentes o accidentes, en el AI RMF 1.0, este marco no define incidentes de tecnología de IA. Sin embargo, se afirma que su determinación, búsqueda, reporte y notificación, así como la respuesta ante este tipo de eventos, es una de las necesidades funcionales que las organizaciones deben contemplar a la hora de gestionar sus riesgos.

## CICLO DE VIDA Y ACTORES DE LA IA

### UE

El Reglamento europeo no presenta una descripción gráfica del ciclo de vida de un sistema IA. Sin embargo a través del artículo 3 “Definiciones” y del mapa de actores institucionales descrito en la Figura 2, se puede deducir el siguiente ciclo de vida:

1. Evaluación de la conformidad
2. Introducción en el mercado
3. Puesta en servicio
4. Comercialización
5. Seguimiento posterior a la comercialización
6. Recuperación de un sistema de IA
7. Retirada de un sistema de IA

Por otro lado, si bien la ley europea no presenta un mapa de actores en forma de gráfico, del conjunto de las organizaciones que conforman y/o influyen en lo que podríamos llamar la implementación de una IA segura (*safe*) puede deducirse esta descripción. Ese mapa puede enriquecerse tomando como referencia el sistema de gobernanza descrito en las pp. 8 y 9 del presente informe, junto con las definiciones del artículo 3 en cuanto a los siguientes actores: operador, implementador, representante autorizado, importador y distribuidor (ver Figura 4).

Las definiciones de cada uno de estos actores aparecen en la p. 51 y son como sigue:

El Operador reúne al proveedor, el implementador, el representante autorizado, el importador y el distribuidor.

El Proveedor es toda persona física o jurídica, autoridad pública, agencia u organismo de otra índole que desarrolle un sistema de IA o para el que se haya desarrollado un sistema de IA con vistas a introducirlo en el mercado o ponerlo en servicio con su propio nombre o marca comercial, ya sea de manera remunerada o gratuita.

El Implementador se refiere a toda persona física o jurídica, autoridad pública, agencia u organismo de otra índole que utilice un sistema de IA bajo su propia autoridad, salvo cuando su uso se enmarque en una actividad personal de carácter no profesional.

El Representante autorizado, a toda persona física o jurídica establecida en la Unión que haya recibido el mandato por escrito de un proveedor de un sistema de IA para

El Importador es toda persona física o jurídica establecida en la Unión que introduzca en el mercado o ponga en servicio un sistema de IA que lleve el nombre o la marca comercial de una persona física o jurídica establecida fuera de la Unión.

El Distribuidor es toda persona física o jurídica que forme parte de la cadena de suministro, distinta del proveedor o el importador, que comercializa un sistema de IA en el mercado de la Unión sin influir sobre sus propiedades.

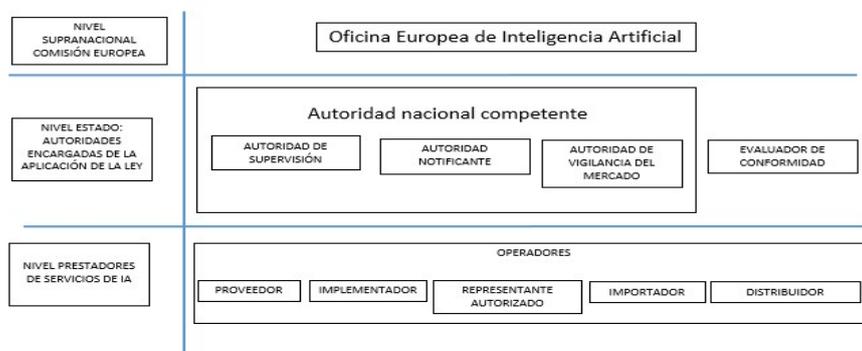
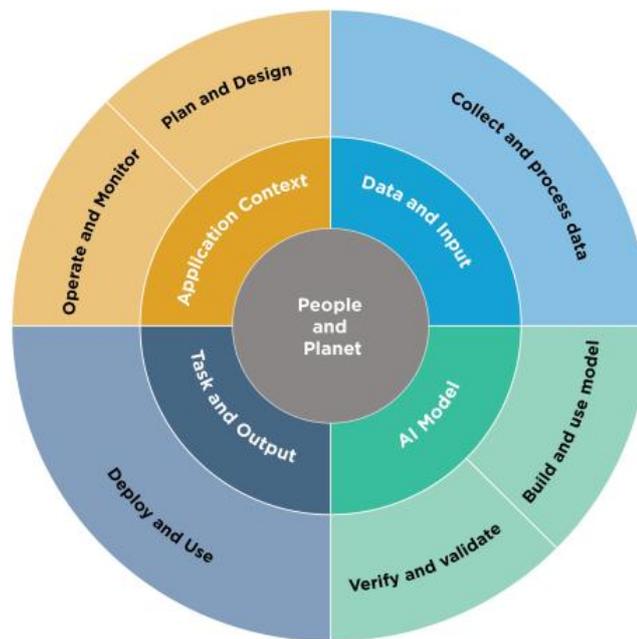


Figura 4. Actores clave (UE)

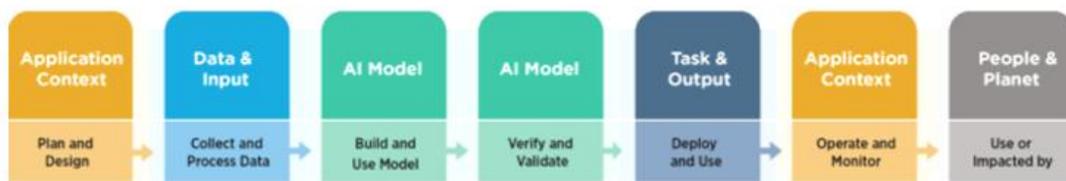
## EEUU

En el AI RMF 1.0 estadounidense se identifica un ciclo de vida pretendido, definido por etapas --donde cada una representa un conjunto de actividades-- y por dimensiones, que representan tanto objetivos intermedios como la participación de diversos actores que inciden en ellas. Si bien se identifica una progresión lineal principal del desarrollo del ciclo de vida, se reconoce la interconexión tanto entre actores y objetivos como entre actividades.

## GENERALIZADO



## PRETENDIDO



Comenzando por la dimensión del **Contexto de Aplicación**, la primera etapa es la **Planificación y Diseño**. Esta consiste en articular y documentar el concepto y objetivos del sistema, las asunciones iniciales y el contexto considerado, que incluye requisitos legales y normativos, así como las consideraciones éticas evaluadas.

La siguiente etapa se encuentra en la dimensión de **Datos y Entradas**, con la **Recolección y Procesamiento de Datos**. Esto consiste en recolectar, validar y



documentar los metadatos y características del conjunto de datos, bajo la observación de los objetivos y las consideraciones legales y éticas establecidas.

Luego, se debe pasar a la dimensión del **Modelado de IA**, a través de la etapa del **Modelado de Construcción y Uso**. Esto se basa en la creación o selección de los algoritmos, así como de los modelos de entrenamiento a emplear.

Dentro de la misma dimensión, la siguiente etapa es la de **Verificación y Validación**. Esto requiere la verificación y validación del modelo, la calibración y la interpretación de los resultados del mismo. En este punto podemos comprender que existe un acoplamiento estrecho con la primera etapa de Planeamiento y Diseño, ya que, si bien todas las etapas anteriores se ven afectadas por las consideraciones tomadas al principio del ciclo de vida, la utilidad de los resultados obtenidos en esta etapa se ve afectada de manera directa por las asunciones del comienzo del desarrollo.

**En un primer análisis, podría afirmarse que esta etapa aparece como la de mayor evaluación crítica, ya que es la que presenta la mayor necesidad de revisión de las expectativas de los actores sobre el sistema de IA, al estar en el punto previo al despliegue en condiciones reales de esta tecnología.**

Esta cuarta etapa podría resultar el principal estadio de defensa en el ciclo de vida de las IA: su participación en la generación de la percepción de la confiabilidad del sistema parece crítica, al tener incorporadas las tareas que se ven estrechamente relacionadas a la búsqueda de los efectos no deseados. Pese a que todas las etapas requieren y pueden realizar una gestión de riesgos, parecería ser en esta donde las actividades se asemejan más a necesidades usualmente asociadas a ese proceso, y por lo tanto pueden llegar a generar una dificultad de comprender que se trata de dos procesos independientes. Los resultados obtenidos en esta etapa serán fuertemente influyentes en la validación y medición de las siguientes etapas, así como en la revisión a posteriori de las etapas anteriores.

Luego de la dimensión de **Modelado de IA**, aparece la dimensión de **Tarea y Output**, con la etapa de **Despliegue y Uso**. En esta se ven involucradas las pruebas piloto, el chequeo de compatibilidad con sistemas que se encuentran en uso e interactuarán con el sistema que se está desplegando (gestión del cambio y/o de las migraciones de sistemas tecnológicos), verificación del cumplimiento regulatorio, gestión del cambio organizacional, y evaluación de la experiencia del usuario.

Pasando a la dimensión de **Contexto de Aplicación**, aparece la etapa de **Operación y Monitoreo**. Consiste en la operación del sistema de IA, evaluación continua de recomendaciones e impactos (tanto intencionados como no intencionados), bajo la perspectiva de objetivos, requisitos legales y regulatorios, y consideraciones éticas.

En la dimensión de **Personas y Planeta**, se encuentra la etapa de **Uso e Impacto**. Esta considera las perspectivas del uso del sistema y/o tecnología, monitoreo y evaluación de impactos, búsqueda de la mitigación de impactos y la advocación por los derechos.



Todas las etapas tiene asociado un requisito propio de Testeo, Evaluación, Verificación y Validación (TEVV) que puede llevarse a cabo. Cada etapa tiene enfoques distintos relacionados con sus objetivos puntuales y que se deben contrastar con los resultados de estos requisitos.

El AI RMF 1.0 señala que, *de manera ideal*, los actores que lleven a cabo los TEVV deberían ser distintos de los que realizan las tareas definidas en cada etapa, sin embargo no afirma que esto deba ser así de manera obligatoria. No señala que deban ser distintos de los propios actores participantes de la etapa, ni que deban ser ajenos a las áreas de la organización que desarrolla las tareas del ciclo de vida, ni tampoco a la propia organización.

El apéndice A describe a los diferentes actores reconocidos por el documento y las etapas del ciclo de vida donde suelen ejercer actividades, así como actores genéricos que representan disciplinas que son transversales a los dominios. Un ejemplo de esto son los Factores Humanos, donde se destaca la importancia de la participación de expertos en esta área en todos los dominios reconocidos para un sistema de IA.

## CONCLUSIONES PARCIALES

Del análisis surge que el principal objetivo de la Iniciativa Nacional de los Estados Unidos con respecto a la IA es liderar el desarrollo de IA en el mundo. Con respecto a los riesgos es una estrategia reactiva, en la medida en que no los aborda de manera directa, sino que encomienda al NIST, dependiente del Departamento de Comercio, la elaboración de un marco general y voluntario para la gestión de riesgos. Ese proceso dio como resultado el Artificial Intelligence Risk Management Framework (AI RMF 1.0). En este documento es un marco voluntario y no una Ley o norma obligatoria, lo cual deja claro el énfasis en la responsabilidad privada.

La ley de la Unión Europea, en cambio, adopta claramente una estrategia proactiva con relación a los riesgos, orientada a delimitar diferentes tipos de usos y prácticas en los que puedan participar sistemas de IA. Establece tres categorías de riesgo. Una es la de los riesgos inaceptables, lo que significa que hay usos de IA que están prohibidos: la manipulación maliciosa del comportamiento, la calificación social y la vigilancia masiva. Otros riesgos se consideran altos y deben ser obligatoriamente gestionados; esto implica un complejo circuito de gestión de calidad y riesgo, documentación, certificaciones, notificaciones. Ejemplos de esta segunda categoría son la identificación biométrica y la categorización de personas, la gestión y el funcionamiento de infraestructuras esenciales; la educación y la formación profesional; el empleo, la gestión de los trabajadores y el acceso al autoempleo; el acceso a servicios esenciales, o la aplicación de la ley. Una tercera categoría es la de los riesgos mínimos, en los que la UE exige transparencia para con el usuario: desarrolladores e implementadores deben informar al usuario que está interactuando con un sistema de IA.



La Ley europea es exhaustiva al describir la red institucional a cargo de la gestión de esos riesgos, e incluye la obligación de informar accidentes o incidentes de IA en no más de 72 horas desde su ocurrencia a las autoridades nacionales y a las autoridades de la Unión.

En cuanto a la perspectiva argentina, sobre la base de las Recomendaciones emitidas en junio por la Secretaría de Innovación Pública, entendemos que el texto acierta en identificar la necesidad de hacer un seguimiento de la IA a lo largo de todo el ciclo de vida, desde su concepción hasta su reciclado o descarte. Incorpora los valores de alineación o alineamiento sugeridos por las Recomendaciones de la UNESCO; identifica los momentos de concientización, diseño, verificación, validación, implementación, operación y mantenimiento, así como la necesidad de establecer siempre un responsable humano en última instancia. Con todo, por un lado, al ser una recomendación voluntaria, su alcance es restringido. Y por otro, no establece instancias de monitoreo ni de investigación de accidentes e incidentes, tal como sí están comenzando a sugerir organizaciones supranacionales como la OCDE. En nuestras recomendaciones sugerimos incorporar el análisis y la investigación de incidentes y accidentes de IA para robustecer el ecosistema de IA, volverlo más fiable para la sociedad y, además, disponer de equipos de expertos “bilingües” o “políglotas” actualizados de monitoreo de estas tecnologías.

En cuanto a las autoridades a cargo de estudiar y dirigir el desarrollo de IA, en el inicio de la investigación advertimos, en confrontación con las iniciativas estadounidense y europea, una dispersión institucional,<sup>25</sup> algo que fue advertido también por las autoridades. En efecto, en septiembre de 2023 se creó la Mesa Interministerial sobre IA, por Decisión Administrativa 750/2023, con el objetivo de diseñar una estrategia integral “para el avance y aplicación de la IA en diversos sectores de la economía y sociedad, considerando un marco ético y de desarrollo sostenible”.

Con todo, es pensable que la iniciativa que más podrá contribuir a ordenar de manera estable la política pública argentina con relación a IA sea el Centro Argentino Multidisciplinario para la Inteligencia Artificial (CamIA), en incubación incipiente a partir del Programa de apoyo a las exportaciones de la Economía del Conocimiento (EDC) aprobado en junio de 2023.

---

<sup>25</sup> Solo en la órbita del Poder Ejecutivo encontramos nueve áreas con iniciativas en torno a la IA, no coordinadas entre sí: la Subsecretaría de Tecnologías de la Información, dependiente de la Secretaría de Innovación Pública; la Secretaría de Asuntos Estratégicos, Presidencia de la Nación; el Consejo Económico y Social; la Subsecretaría de Políticas en Ciencia, Tecnología e Innovación, Ministerio de Ciencia, Tecnología e Innovación (Mincyt); la Agencia Nacional de Promoción de la Investigación, el Desarrollo Tecnológico y la Innovación (Agencia I+D+i); la Fundación Sadosky, con sede en el Mincyt; el Centro Interdisciplinario de Estudios en Ciencia, Tecnología e Innovación (CIECTI); la Secretaría de Economía del Conocimiento, en el Ministerio de Economía; y el Instituto Nacional de Tecnología Industrial (INTI), Ministerio de Economía.



## IMPACTO DE LA IA GENERATIVA Y LOS LLM EN EL MERCADO DE TRABAJO RELEVAMIENTO DE BIBLIOGRAFÍA INTERNACIONAL<sup>26</sup>

Desde hace algunos años, pero en forma acelerada desde la salida al mercado de ChatGPT en noviembre de 2022, emergieron un conjunto de investigaciones que están dando forma a una naciente literatura que se enfoca en la estimación y análisis de los potenciales impactos que los Modelos de Lenguajes Grandes (LLM por su sigla en inglés) y más en general las IA generativas podrían tener sobre el mercado de trabajo, principalmente en los EE.UU. (Agrawal et al., 2022; Brynjolfsson et al. 2023; Eloundou et al, 2023; Felten et al., 2023; Hatzius, J., et al. 2023; Bommasani et al., 2022; Noy & Zhang, 2023; Peng et al., 2023; Boyang, Zongxiao y Zaho, 2023; Zarifhonarvar, 2023; OIT, 2023).

A diferencia de la literatura precedente, estos trabajos sugieren que es posible que los LLM realicen una variedad de tareas consideradas “no rutinarias”, como la codificación de software, la escritura persuasiva y creativa, la investigación, el diseño gráfico, etc. Dado que muchas de estas tareas actualmente las realizan trabajadores que se han beneficiado de oleadas anteriores de adopción de tecnologías digitales, la expansión de la IA generativa puede provocar importantes cambios en la relación entre tecnología, empleo, desigualdad y productividad (Brynjolfsson et al, 2023). A su vez, estas transformaciones pueden ocurrir a un ritmo dramático dada la veloz adopción y las bajas barreras de acceso y uso que poseen estas tecnologías.

En tal sentido, en lo que respecta a la productividad laboral, Peng et al. (2023) realizan un estudio en el que contrataron ingenieros de software mediante la plataforma upwork en EEUU para una tarea de codificación específica, y encuentran que aquellos a los que se les da acceso a IAG completan la tarea dos veces más rápido que quienes no lo usan, y que los desarrolladores juniors muestran los mayores incrementos de productividad.

Del mismo modo, Noy y Zhang (2023) estudian el impacto en tareas de escritura profesional, y encuentran que aquellos con acceso a ChatGPT completan las tareas de escritura con un incremento de la productividad en torno a 50%. Además, muestran que ChatGPT comprime la distribución de la productividad, siendo los trabajadores menos calificados los que más se benefician del uso. Brynjolfsson y otros (2023) estudian la introducción de un asistente conversacional basado en IAG en atención al cliente, y encuentran que el acceso a la IAG aumenta la productividad, medida por los problemas resueltos por hora, en un 14% en promedio con mayor impacto en los trabajadores *juniors* y poco calificados. Ya en el nivel de firma, Eisfeldt y otros (2023) estudian el impacto de las IAG en el valor de las firmas que cotizan en bolsa en EE.UU en función del grado de exposición de su fuerza laboral a las IAG, y encuentran que el efecto del lanzamiento de ChatGPT en los valores de las empresas con mayor exposición fue significativo, generando una diferencia en los rendimientos de

---

<sup>26</sup> El trabajo de relevamiento bibliográfico se realizó en el marco de la postulación a la convocatoria PISAC 2023, bajo la dirección del Dr. Mariano Zukerfeld.



aproximadamente 0,4 % diario, lo que se traduce en más del 100% sobre una base anualizada.

Por tales motivos, estas investigaciones son cruciales. Sin embargo, aún no se dispone de estudios similares en Argentina.

## **APROXIMACIÓN A LA PROBLEMÁTICA EN EL MERCADO DE TRABAJO NACIONAL**

Los estudios mencionados se basan en un enfoque de tareas. En tal sentido, uno de los principales obstáculos para avanzar en esta dirección radica en que el Clasificador Nacional de Ocupaciones (CNO) utilizado por la encuesta Permanente de Hogares (EPH) no posee una desagregación de tareas laborales ni es compatible con clasificadores internacionales que posean esa desagregación, como la Clasificación Internacional Uniforme de Ocupaciones, también conocida por sus siglas en inglés ISCO, a 4 dígitos u O\*net, por Occupational Information Network.<sup>27</sup> Por tal motivo, para evaluar los potenciales impactos de la IA en el mercado de trabajo nacional es necesario construir un sistema de información que describa y clasifique tareas laborales que sea compatible con el CNO.

Se encuentra en proceso la construcción de una base de datos de tareas laborales adaptada al CNO, con similar estructura a las bases de datos de tareas laborales de O\*net e ISCO rev. 4, con asistencia de IA siguiendo la metodología propuesta por OIT (2023).<sup>28</sup>

Por otro lado, es posible realizar una estimación de la distribución del empleo por grupo ocupacional del CNO en el mercado de trabajo argentino en base a la EPH (1er trimestre 2023). Y a partir de este insumo, realizar una estimación preliminar de la exposición del mercado de trabajo argentino a los LLM en base a EPH (1er trimestre 2023) y la bibliografía internacional.<sup>29</sup>

En efecto, la estructura ocupacional del mercado de trabajo argentino se encuentra concentrada en dos tipos de actividades con distinto grado de exposición a los LLM. Por un lado, ocupaciones de servicios físicos o de producción industrial/artesanal con bajos niveles de exposición a las IA generativas: Ocupaciones de la construcción; de venta física; de empleo doméstico; de limpieza fuera de los hogares; de producción industrial; de transporte de pasajeros, de salud y sanidad. En total, estas ocupaciones representan un 71% del total de empleo. Por otro lado, se observa también una concentración en ocupaciones administrativas o profesionales con elevados niveles de exposición a las IA generativas: Ocupaciones de venta directa online, ocupaciones de venta indirecta, Ocupaciones de la gestión administrativa, planificación y control de gestión, ocupaciones de la educación y la investigación, Ocupaciones de gestión

---

<sup>27</sup> Actualmente, el CNO solo posee cruces a 2 dígitos con el ISCO.

<sup>28</sup> Se trata de un proyecto conjunto de investigadores del Centro Interdisciplinario de Estudios en Ciencia, Tecnología e Innovación (CIECTI), el Observatorio de Empleo y Dinámica Empresarial (OEDE) y el Centro Interinstitucional de Datos de la UBA (CID).

<sup>29</sup> La estimación fue realizada en el marco de un trabajo en curso para el CIECTI.

presupuestaria, contable y financiera. En conjunto estas ocupaciones representan un 29% del empleo.



### OCUPACIONES (Clasificador Nacional de Ocupaciones)

**72: construcción**

**30: comercialización directa**

**80: producción artesanal**

**10: gestión administrativa, planificación y control de gestión**

**41: educación e investigación**

**34: transporte**

**30b: compraventa por medios telefónicos o informáticos**

**56: servicios de limpieza no domésticos**

**40: salud y sanidad**

**55: servicios domésticos**

**20: gestión presupuestaria, contable y financiera**

**57: cuidado y atención de personas**

**53: establecimientos de servicios de gastronomía**

**05: directivos de pequeñas empresas y microempresas**

**82: reparación de bienes de consumo**

**47: servicios de vigilancia y seguridad civil**

**11: gestión jurídico-legal**

**36: almacenaje de insumos, materias primas, mercaderías e instrumentos**

**58: otras ocupaciones de servicios varios**

**06: directivos de medianas empresas privadas productoras de bienes y/o servicios**

**81: producción de software**

En el caso de las ocupaciones con mayor exposición, lo que vemos es que muchas de ellas --gestión administrativa, planificación y control de gestión, educación e investigación, gestión presupuestaria, contable y financiera— son ocupaciones en las que los estados nacional y provinciales han invertido muchos recursos de formación en las últimas tres décadas. En este sentido, el desafío que se presenta es el de seleccionar y priorizar áreas fundamentales --tanto en el sector público como en el



privado-- para emprender procesos controlados de transformación digital, capaces de aprovechar las potencialidades de las IA sin generar efectos traumáticos en el nivel de ocupación.

## **ESTIMACIÓN DE IMPACTO EN LA PRODUCTIVIDAD Y EN LOS SALARIOS EN EL SECTOR DE SOFTWARE ARGENTINO**

A fines de brindar un ejemplo en un sector concreto, se relevó la estimación de impacto en la productividad y salarios por el uso de LLM en el sector de Software Argentino en base a encuesta sectorial.<sup>30</sup>

El ámbito de la programación de software es señalado por la literatura como aquel donde se concentran los principales impactos de las IAG, y herramientas como ChatGPT o GitHub Copilot son ampliamente utilizadas. A su vez, se presume que su uso implica notables incrementos de productividad horaria. De hecho, en lo que respecta a la productividad laboral, Peng y otros (2023) realizan un estudio en el que contrataron ingenieros de software mediante la plataforma *upwork* en EE.UU. para una tarea de codificación específica, y encuentran que aquellos a los que se les da acceso a GitHub Copilot (una IAG asistente del desarrollo de software basada en GPT3) completan la tarea dos veces más rápido que quienes no lo usan, y que los desarrolladores *juniors* muestran los mayores incrementos de productividad, de forma tal que se comprime la distribución de la productividad.

Aquí vemos resultados parciales de un estudio todavía en curso, en el que se analiza el uso de ChatGPT por parte de los programadores argentinos en base a una encuesta sectorial de 5500 casos, con el objetivo de evaluar su impacto en los niveles salariales y entender cómo su uso varía en función de la experiencia laboral de los programadores. Entre los principales resultados, se encuentra que un 73% de la muestra declara haber utilizado al menos una vez herramientas de IA para la codificación (como ChatGPT o GitHub Copilot) y casi un 20% lo utiliza de manera muy frecuente. Por otro lado, la frecuencia de uso y el nivel de experiencia están inversamente relacionados, lo que indica que los programadores menos experimentados son más propensos a integrar ChatGPT en sus prácticas laborales. Este fenómeno plantea interrogantes sobre cómo las herramientas de IAG están redefiniendo los paradigmas de aprendizaje y desarrollo de habilidades en los trabajos intensivos en conocimiento no rutinarios.

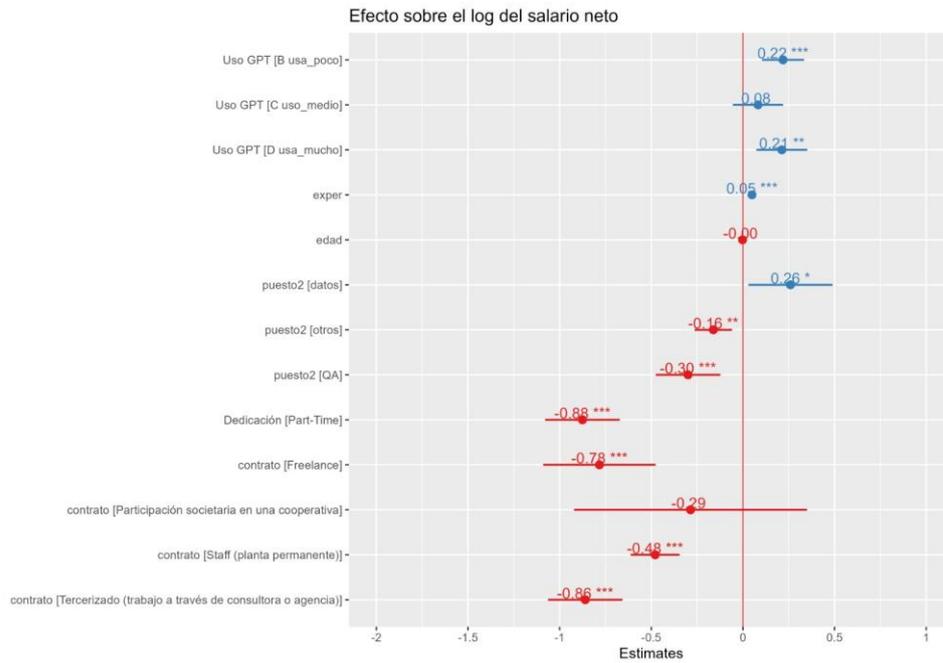
Además, se identificaron brechas salariales positivas asociadas al uso de ChatGPT para todos los niveles de experiencia, sugiriendo que la adopción de esta herramienta podría estar vinculada a una remuneración superior. Más aún, este efecto se mantiene incluso tras controlar por variables relevantes como edad, nivel educativo, puesto de trabajo, tipo de contrato, entre otras, como se observa en el gráfico a continuación, que muestra los coeficientes de regresión log-lineales sobre el salario neto de un set de variables entre las que se encuentran el uso de ChatGPT.

---

<sup>30</sup> La estimación fue realizada en el marco de un trabajo en curso para el CIECTI.



Como puede observarse, existe una prima salarial de aproximadamente el 20% para aquellos programadores que hacen un uso poco o muy intensivo de ChatGPT en comparación con aquellos que no lo utilizan.



Fuente: elaboración propia (A. Rabosto) en base a sysarmy 2023



## **IMPACTO DE LA IA GENERATIVA Y LOS LLM EN LAS INDUSTRIAS CULTURALES Y CREATIVAS: LA PERSPECTIVA DE LOS ACTORES**

Una vez realizada la delimitación del sector y de los agentes que conformarían el universo de análisis, se seleccionó el sector de las **industrias culturales y creativas** y se entrevistaron diferentes agentes clave: representantes sindicales de sectores de la producción, posproducción y distribución de contenidos audiovisuales, locutores y comunicadores; empresarios pertenecientes a productoras locales en alianza con plataformas OTT (por sus siglas en inglés, *over-the-top*, también llamadas plataformas de transmisión libre); asesores legales: un abogado especializado en negociación sindical que presta servicios a plataformas audiovisuales globales que operan en el país y una abogada especializada en propiedad intelectual; representantes de entidades de gestión de derechos de autor -cuyas obras se distribuyen en plataformas OTT y comprenden actividades como guión, composición, arreglos e intérpretes musicales, coreógrafos e intérpretes dancísticos, actores y extras-; asociaciones profesionales de sectores audiovisuales, artes visuales y trabajadores creativos, como artistas pertenecientes a organizaciones menos formalizadas.

Se elaboró una guía de preguntas en relación a los objetivos establecidos en el proyecto, y una vez realizadas, se transcribieron las entrevistas en un documento que forma parte de la documentación del proyecto (ver en anexo: [docs.google.com/document/d/1ZKNIWOD6sw3QBEt911cnrXJRwr73lJxGV6V2QaNx\\_x0/edit?usp=sharing](https://docs.google.com/document/d/1ZKNIWOD6sw3QBEt911cnrXJRwr73lJxGV6V2QaNx_x0/edit?usp=sharing)).

Para este informe se sistematizaron los materiales reunidos, se realizó una caracterización conceptual y socioeconómica de perfiles laborales y actividades y agentes que conforman el caso de estudio, y se sistematizaron las principales percepciones relevadas en el discurso de los propios actores.

### **CARACTERIZACIÓN DE PERFILES LABORALES Y ACTIVIDADES PRODUCTIVAS EXPLORADAS**

Se realizaron diez entrevistas a representantes de sectores de trabajo de actividades culturales y creativas. Esto incluye a representantes sindicales de todo personal obrero, técnico y administrativo que realiza actividades de producción y posproducción de contenidos audiovisuales, representantes sindicales de guionistas de contenidos audiovisuales y representantes sindicales de intérpretes musicales y vocales; referentes de asociaciones profesionales de artistas visuales y de diseño gráfico; entidades de gestión de derechos colectivos que abarcan intérpretes, autores, productores y compositores musicales, intérpretes actorales y bailarines, y autores de cine y televisión.

En materia normativa, estas actividades se encuentran abarcadas por convenios colectivos de trabajo –tal es el caso de técnicos (CCT N° 235/75; 223/75 y 131/75), actores (CCT N° 357/75 y la ley del actor N° 27.203) e intérpretes musicales estables (CCT N° 53/75)– y por la ley de propiedad intelectual N° 11.723 –que abarca a



intérpretes y bailarines, autores de cine y televisión, las actividades musicales mencionadas, artistas visuales y diseñadores gráficos—.

El sector productivo se enmarca en la actividad de la producción de contenidos audiovisuales para plataformas OTT, que en la actualidad son las que generan la mayor cantidad de puestos de trabajo e inversión privada a la producción local (Del Bono & Bulloni, 2018). De manera específica, el estudio se abocó a las actividades que nuclean a empresas productoras de televisión y cinematografía, que se organizan en entidades empresariales por tipo de actividad.

Dentro de la cadena de valor de la producción audiovisual local, se contemplaron dos grupos: aquellas empresas que se ocupan de las etapas de la creación, la realización, la producción propiamente dicha y el desarrollo de contenidos, y las empresas que proveen servicios de tecnología en equipamiento para la realización de contenidos audiovisuales. Esta actividad cuenta con un instituto de fomento y promoción, el Instituto Nacional del Cine y las Artes Audiovisuales. Está regulada por la ley de propiedad intelectual, la ley de cine N° 17.741 y la ley de servicios de comunicación audiovisual N° 26.522.

La producción de contenidos audiovisuales es reconocida como actividad industrial desde el año 2013 y también es alcanzada como parte de las actividades de la ley N° 27.506 de régimen de promoción de la economía del conocimiento. Con la consolidación de las empresas de plataformas audiovisuales, la industria local y su tradicional modelo de negocios ha sufrido grandes transformaciones en los últimos años. La actividad que hasta hace poco tiempo se encontraba mayormente subvencionada y regulada por el Estado, ha comenzado a ser permeada cada vez más por la inversión privada (González, 2022).

Se incluyeron además, perfiles de asesores corporativos en materia jurídica que brindan servicios a empresas de plataformas OTT que operan en Argentina y abogados/as especialistas en derechos de propiedad intelectual que asesoran a entidades de gestión de derechos colectivos de las actividades laborales contempladas.

Todas estas actividades y ocupaciones se caracterizan por su intermitencia y eventualidad debido a la dinámica de la producción audiovisual que se realiza por proyectos, de manera discontinua. Lo mismo sucede con la producción de obras visuales y de diseño gráfico. A su vez, la producción y el trabajo se encuentran encadenados a procesos de organización y estándares de procedimientos internacionales. Frente a la complejidad jurídica, impositiva y capacidad de auditar en el exterior, la realización de gestiones por el cobro de regalías por derechos de autor en el caso de creadores, como así también la contratación de servicios a terceros, en el caso de productoras, estas organizaciones mantienen filiación con empresas, sindicatos y entidades de gestión de derechos de autor de diferentes países.

## **PERCEPCIONES RELEVADAS**



### 1. AMBIVALENCIA, NO UNIFORMIDAD, CONTEXTO COMPLEJO

La información relevada hasta el momento arroja un cuadro complejo que no admite lecturas lineales. Al mismo tiempo, los efectos que comienzan a percibirse entre las actividades en las que impacta la utilización de IA generativas (en adelante, IAG), son ambivalentes, heterogéneos y tanto suscitan preocupación como revelan aspectos positivos señalados por las personas entrevistadas.

La implementación de las IAG no se da de manera uniforme entre las actividades relevadas. Una primera cuestión de contexto que aparece de la mano de empresas productoras de contenidos y servicios tecnológicos es la situación económica. Por un lado, los flujos de capitales que invierten en el país en un contexto inflacionario debilita la rentabilidad de las empresas locales. A su vez, también tiene su correlato para importar bienes tecnológicos que permiten actualizar equipamientos para la oferta de servicios. Esto tiene como consecuencia que la incorporación de estas tecnologías es relativamente lenta. Sin embargo, los testimonios de asesores legales de empresas de plataformas OTT advierten que la llegada es inminente y puede acontecer “en cualquier momento”.

### 2. TEMPORALIDADES DIVERGENTES

En consonancia con el eje de la aceleración temporal, agentes vinculados a la gestión colectiva de derechos coinciden con que el ritmo que imprime la implementación de las IAG no es acorde a los tiempos que lleva registrar obras, contar con mecanismos de trazabilidad y control en la explotación de creaciones en plataformas OTT. Estos sectores además advierten desafíos para poder hacer cumplir lo establecido en la ley de propiedad intelectual que protege a toda creación humana y pone límites al reconocimiento de este derecho a las empresas, más allá de la sucesión de derechos, el reconocimiento de una creación intelectual es hacia la persona física. Dicha norma implica el reconocimiento como autor/a de la obra, la autorización de su reproducción o utilización para ser modificada, como así también la remuneración por la explotación de obra. Entre las principales cuestiones, las entidades reclaman por un lado, poder amplificar derechos en el ámbito de la esfera de la producción audiovisual digital y los servicios de *streaming* que ya se encuentran consolidados, es y sigue siendo una disputa que no se origina con el advenimiento de la IAG tanto en el plano nacional como cuando estos grupos de creadores ofrecen servicios en producciones internacionales. Por otro, aparece la cuestión de intérpretes actorales, vocales, coreógrafos, locutores que son creados de manera artificial -en la jerga del sector, sintética-. Estos personajes que bailan, actúan, simulan voces humanas son creados con IAG y a su vez, desarrollan creaciones hechas por la IAG. Lo cual, esto permite que las empresas puedan ofrecer servicios en este tipo de actividades para la creación de contenidos audiovisuales.

### 3. FORMACIÓN TRANSVERSAL: NUEVOS DISPOSITIVOS, ANTIGUAS CURRÍCULAS

Parte de los desafíos que emergen de las entrevistas se vinculan con la formación de capacidades profesionales, donde la programación dejó de ser una disciplina específica para ser parte de las herramientas que los perfiles entrevistados incorporan para el desarrollo de muy diferentes procesos productivos, creativos y culturales. Esto supone



para las instituciones educativas la necesidad de actualizar sus programas para que incluyan el manejo de *softwares* de manera transversal a las disciplinas artísticas, creativas y técnicas dentro de la oferta académica. Abre un nuevo paradigma en el diseño y la planificación de contenidos de materias, en la búsqueda de perfiles docentes y, al mismo tiempo, requiere inversión en equipamiento tecnológico para el dictado de clases.

#### 4. MENOS PUESTOS, RECONVERSIÓN DE TAREAS: EL CASO DE LOS GUIONISTAS

En lo que respecta a actividades como la redacción de guiones se observan casos de pérdida de puestos de trabajo y reconversión de actividad, allí donde se implementa IAG para la producción audiovisual. Se reemplaza la actividad de un grupo de guionistas y en su lugar se contrata un solo guionista experto quien realiza tareas que se asemejan más a las de un editor, que ajusta y mejora lo creado con IAG, que a un trabajo creativo tradicional.

#### 5. TRATAMIENTO DE IMÁGENES: OPTIMIZACIÓN DEL TIEMPO Y AMPLIFICACIÓN DE OPCIONES CREATIVAS

Por su parte, quienes realizan actividades técnicas vinculadas a la manipulación y tratamiento de imágenes y sonido afirman utilizar las IAG como herramientas creativas que les permite contar con nuevas opciones y optimizar tiempo de trabajo, ya que simplifica etapas del proceso.

#### 6. EL DESAFÍO A LOS DERECHOS DE AUTOR

En tanto las entidades que gestionan derechos de autor advierten desafíos para poder hacer cumplir lo establecido en la ley de propiedad intelectual que protege a toda creación humana. En efecto, el reconocimiento de una creación intelectual es hacia la persona física del creador o la creadora. La ley sanciona el derecho moral de dicha persona a ser reconocida como autor o autora de la obra en cuestión; quien tiene potestad para autorizar su reproducción o su utilización para ser modificada, y también puede ceder el derecho a reproducir y explotar su obra, pero el derecho moral es inalienable.

#### 7. VOCES Y ROSTROS SINTÉTICOS

Entre las principales cuestiones las entidades reclaman, por un lado, poder amplificar derechos en el ámbito de la esfera de la producción audiovisual digital y los servicios de *streaming* que ya se encuentran consolidados. Esta es una disputa que no se origina tanto con la llegada de las IAG al país como cuando estos grupos de creadores ofrecen servicios en producciones internacionales. Por otro, aparece la cuestión de los intérpretes actorales, vocales, coreógrafos, locutores que son creados de manera artificial -en la jerga del sector, la voz sintética-. Estos personajes que bailan, actúan, simulan voces humanas son creados con IAG y, a su vez, desarrollan creaciones hechas por IAG. Lo cual permite que las empresas puedan ofrecer servicios en este tipo de actividades para la creación de contenidos audiovisuales.

### A MODO DE SÍNTESIS: LA PERSPECTIVA DEL SECTOR



En lo que respecta a las actividades alcanzadas bajo convenios colectivos de trabajo, el fortalecimiento del diálogo social tripartito y la regulación convencionada es el modo en que el impacto de la IAG podrá adoptarse en condiciones justas y seguras. Entre las principales cuestiones, destacan la necesidad de ser consultados, de contar con información y mecanismos de mediación y arbitraje.

En tanto, las actividades protegidas por el derecho de autor bregan por la adopción de respuestas globales para garantizar el cobro de regalías por la explotación de sus obras en diferentes partes del mundo. A su vez, también estos grupos demandan ser parte de los debates para que sus intereses se encuentren contemplados en una regulación.

A su vez, señalan la importancia de crear áreas especializadas de evaluación ética, capaces de monitorear e investigar los usos de las IAG. Estas áreas deberán estar integradas por perfiles “híbridos” o “bilingües”, capaces de comprender las distintas disciplinas involucradas y que puedan moderar aspectos como la utilización de *datasets* basados en obras de creación humana, la protección de datos personales, la transparencia en cuanto al origen de los contenidos, los sesgos, entre otros riesgos y desafíos vinculados al uso de la IAG. Entre las principales demandas, las entidades expresan la necesidad de que artistas y creadores puedan ser parte de los debates para que sus intereses se encuentren contemplados en una regulación y por otro lado, también surgió la recomendación de la creación de una ley de IAG.



## **ANTECEDENTES DEL PROYECTO CENTRO ARGENTINO MULTIDISCIPLINARIO DE INTELIGENCIA ARTIFICIAL (CAMIA)**

En el marco del relevamiento sobre iniciativas vigentes para promover las potencialidades y hacer frente a los desafíos y los riesgos de la IA generativa, se identificó la existencia de un proyecto en curso para desarrollar un Centro Argentino Multidisciplinario de Inteligencia Artificial (CamIA). Se describen aquí los antecedentes de ese proyecto y su actual situación.

### **ANTECEDENTES DE CAMIA**

Por resolución del 26 de noviembre de 2021,<sup>31</sup> y en virtud de que el decreto N° 802/2020 había establecido como objetivo de la Dirección Nacional de Gestión del Conocimiento el “desarrollar y monitorear en el Sector Público el uso de tecnologías que constituyen iniciativas prioritarias para el desarrollo de la economía del conocimiento tales como la inteligencia artificial, cadenas de bloques y otros proyectos que contribuyan a consolidar la soberanía tecnológica argentina en la revolución 4.0”, el entonces Secretario de Asuntos Estratégicos Gustavo Beliz creó el Programa de Inteligencia Artificial.

Lo hizo dentro de la órbita de la Dirección Nacional de Gestión del Conocimiento, de la Secretaría de Asuntos Estratégicos de la Presidencia de la Nación, con el objetivo de brindar apoyo al Consejo Económico y Social (CEyS) “para el desarrollo de actividades vinculadas a la promoción de habilidades tecnológicas relativas a la inteligencia artificial” (Resolución 90, 2021).

A partir de estas funciones del CES, el 5 de abril de 2022 se presentó formalmente el proyecto de un Centro Argentino Multidisciplinario de Inteligencia Artificial (CamIA),<sup>32</sup> en un acto celebrado en la Casa Rosada en el que además se anunció la adhesión de la Argentina al Pacto Global de Inteligencia Artificial (Global Partnership on Artificial Intelligence o GPAI, por sus siglas en inglés) impulsado por los gobiernos de Francia, Canadá y Japón en el marco de la Organización para la Cooperación y el Desarrollo Económico (OCDE), al que habían adherido hasta el momento 19 estados. El GPAI es una iniciativa lanzada en el año 2000 con participación de diversos actores, entre ellos científicos, industriales, sociedad civil, gobiernos, organismos internacionales, y representantes del mundo académico, con el objetivo de cooperar, acercar la teoría y la práctica de la IA, y financiar investigaciones y actividades relativas a las prioridades

<sup>31</sup> Ver Resolución 90/2021: <https://www.boletinoficial.gob.ar/detalleAviso/primera/253666/20211130>

<sup>32</sup> <https://www.argentina.gob.ar/noticias/un-foro-para-fomentar-el-desarrollo-de-la-inteligencia-artificial-en-la-argentina> (CEyS, 2022a) y <https://www.argentina.gob.ar/noticias/el-consejo-economico-y-social-impulsa-la-creacion-de-un-centro-para-la-promocion-de>



en IA. A inicios de 2023 ya cuenta con 29 países asociados, entre ellos, Argentina, Brasil y México, por América Latina.

El 7 de junio de 2022 se realizó un segundo evento, el *Foro Internacional de Inteligencia Artificial: hacia un Centro Argentino Multidisciplinario de Inteligencia Artificial*, con la intención de discutir la arquitectura institucional de CamIA.<sup>33</sup> Se realizó junto al Team Europe de la Unión Europea, y tuvo entre sus objetivos desarrollar un diseño que pudiera servir para (a) promover una participación activa de los actores públicos y privados del ecosistema de IA; (b) generar mecanismos para la identificación y priorización de necesidades de formación (desarrollo de talentos), de investigación aplicada y de servicios a las empresas; (c) promover la transferencia tecnológica al sector productivo; y (d) garantizar la sustentabilidad financiera de mediano y largo plazo del centro (CEyS, 2022a; CEyS, 2022b).

En julio de ese mismo año, la renuncia del entonces Secretario de Asuntos Estratégicos Gustavo Beliz, quien articulaba la agenda de IA, motivó cambios institucionales, que derivaron en que la Subsecretaría de Políticas en Ciencia, Tecnología e Innovación (CTI), en conjunto con la Agencia I+D+i, pasaron a llevar adelante el proyecto y delinear el perfil del Centro.

#### **PROGRAMA DE APOYO A EXPORTACIONES DE LA ECONOMÍA DEL CONOCIMIENTO**

En el marco de esta iniciativa, se aprobó en junio de 2023 un Programa destinado a “contribuir al aumento de las exportaciones de los sectores de la Economía del Conocimiento (EDC) a través de la provisión de capital humano especializado, del desarrollo y adopción de tecnologías basadas en Inteligencia Artificial (IA) y de la promoción de su inserción internacional”.<sup>34</sup> En la presentación se menciona el objetivo de crear “un centro de inteligencia artificial” que “tendrá entre sus objetivos generar capacidades de dirección y gestión de proyectos multidisciplinarios de desarrollo tecnológico basados en IA, articular las capacidades del sistema científico y tecnológico en IA y las necesidades del sector productivo, elaborar una agenda de política regulatoria en IA, desarrollar talentos en IA y contribuir a la internacionalización del ecosistema”.<sup>35</sup>

En los documentos específicos del Programa se precisa que uno de los componentes estará destinado a la creación e incubación del Centro Argentino Multidisciplinario de Inteligencia Artificial (CamIA), encargado de (a) generar conocimiento y realizar

<sup>33</sup> <https://www.argentina.gob.ar/noticias/nuevo-foro-taller-sobre-inteligencia-artificial-junto-al-team-europe-de-la-ue> (CEyS, 2022b).

<sup>34</sup> <https://www.argentina.gob.ar/noticias/nuevo-programa-de-35-millones-de-dolares-para-el-desarrollo-de-la-inteligencia-artificial>

<sup>35</sup> Ídem



acciones de difusión y promoción para generar interés y demanda; (b) poner a disposición de investigadores bases de datos existentes, generar bases de datos sintéticas y *sandboxes* como bienes de libre acceso; y (c) apoyar el diseño de marcos regulatorios y políticas públicas que busquen un balance entre el fomento de la innovación y el uso seguro y responsable de la IA.

Al momento de cierre de este informe, la información disponible es que el CamIA se emplazará en las instalaciones del edificio Cero + Infinito perteneciente a la Facultad de Ciencias Exactas de la Universidad Nacional de Buenos Aires (UBA), y que será incubado por UBATEC, la Unidad de Vinculación Tecnológica de la Universidad de Buenos Aires durante un período máximo de 4 años.

Se contempla que en todas las actividades y desarrollos llevados adelante por el Centro y los proyectos que se gestionen en su marco, se considerará la inclusión de información y estrategias de equidad (*fairness*) que eviten sesgos relacionados al género, etnias y condiciones socioeconómicas, así como criterios de accesibilidad.



## BIBLIOGRAFÍA

- Albrieu, Ramiro, Rapetti Martín, Brest López Caterina, Larroulet Patricio y Sorrentino Alejo (2018): "Inteligencia artificial y crecimiento económico. Oportunidades y desafíos para Argentina", CIPPEC.
- Agrawal Ajay, Gans Joshua y Goldfarb Avi (2022). Power and prediction: The disruptive economics of artificial intelligence, Harvard Business Press.
- Ansermet François y Magistretti Pierre (2007). Plasticidad neuronal e inconsciente. Buenos Aires, Katz.
- Bommasani Rishi y Percy Liang (2022). Trustworthy Social Bias Measurement, 2212.11672, archivePrefix={arXiv}.
- Böstrom Nick (2014). Superinteligencia. Caminos, peligros, estrategias. Madrid, Teell.
- Boyang Chen, Zongxiao Wu y Ruoran Zhao (2023) From fiction to fact: the growing role of generative AI in business and finance, Journal of Chinese Economic and Business Studies, 21:4, 471-496, DOI: 10.1080/14765284.2023.2245279
- Brynjolfsson E., Li, D., Raymond, L. R. (2023). Generative AI at Work, Working Paper 31161, <http://www.nber.org/papers/w31161>, National Bureau of Economic Research, Cambridge, MA.
- Coeckelbergh Mark (2022). The Political Philosophy of IA. An introduction. Cambridge, Polity Press.
- Consejo Económico y Social (CEyS). (2022a). El Consejo Económico y Social impulsa la creación de un centro para la promoción de Inteligencia Artificial en la Argentina. Presidencia de la Nación, Argentina. Disponible en <https://www.argentina.gob.ar/noticias/el-consejo-economico-y-social-impulsa-la-creacion-de-un-centro-para-la-promocion-de>
- Consejo Económico y Social (CEyS). (2022b). Nuevo foro-taller sobre Inteligencia Artificial junto al Team Europe de la U.E. Presidencia de la Nación, Argentina. Disponible en <https://www.argentina.gob.ar/noticias/nuevo-foro-taller-sobre-inteligencia-artificial-junto-al-team-europe-de-la-ue>
- Costa Flavia (2021). Tecnoceno. Algoritmos, biohackers y nuevas formas de vida. Buenos Aires, Taurus.
- Crawford Kate (2022). Atlas de inteligencia artificial. Poder, política y costos planetarios. Buenos Aires, Fondo de Cultura Económica.
- Cheney-Lippold, John (2017). We are Data: Algorithms and The Making of Our Digital Selves. Nueva York, New York University Press.
- Couldry, Nick y Mejías, Ulises (2019). The Costs of Connection. How Data Is Colonizing Human Life and Appropriating it for Capitalism. Stanford University Press.
- Declaración de Montevideo sobre Inteligencia Artificial y su impacto en América Latina, Montevideo, 10 de marzo de 2023. En Internet: [fundacionsadosky.org.ar/declaracion-de-montevideo-fun/](http://fundacionsadosky.org.ar/declaracion-de-montevideo-fun/)
- Dupuy Jean-Pierre (1999). Aux origines des sciences cognitives. Paris, La Découverte.
- Eisfeldt et al. (2023), Generative AI and Firm Values, Working Paper 31222, <http://www.nber.org/papers/w31222>, National Bureau of Economic Research, Cambridge, MA.
- Eloundou, T., S. Manning, P. Mishkin, y D. Rock (2023): "Gpts are gpts: An early look at the labor market impact potential of large language models". En arXiv preprint arXiv:2303.10130.



Felten Edward W., Raj, Manav y Seamans, Robert (2023), Occupational Heterogeneity in Exposure to Generative AI. En SSRN: <https://ssrn.com/abstract=4414065> or <http://dx.doi.org/10.2139/ssrn.4414065>

Floridi Luciano (2011). Energy, Risks, and Metatechnology, May 03, 2011. En Internet: <https://ssrn.com/abstract=3854445>

Gardner Howard (1987). La nueva ciencia de la mente. Historia de la revolución cognitiva. Buenos Aires, Paidós.

Gómez Mont C., Del Pozo C.M., Martínez Pinto C., Martín del Campo Alcocer A.V. (2020): "La Inteligencia Artificial al servicio del bien social en América Latina y el Caribe: panorámica regional e instantáneas de doce países", Banco Interamericano de Desarrollo.

Hatzius, Jan et al (2023). The Potentially Large Effects of Artificial Intelligence on Economic Growth (Briggs/Kodnani), Goldman Sachs.

Heims Steve Joshua (1991). The Cybernetics Group. Cambridge, MIT Press.

Helbing Dirk. (2015). Societal, Economic, Ethical and Legal Challenges of the Digital Revolution: From Big Data to Deep Learning, Artificial Intelligence, and Manipulative Technologies. SSRN Electronic Journal. 10.2139/ssrn.2594352.

Hinton Geoffrey (2023). "Las máquinas serán más inteligentes que las personas en casi todo". Fundación BBVA. <https://www.youtube.com/watch?v=ag9YIHIncbM>

Kurzweil Ray ([2005] 2012). La Singularidad está cerca. Cuando los humanos transcendamos la biología. Berlín, Lola Books.

Lafontaine Céline (2004). L'empire cybernétique. Des machines à penser à la pensée machine. Paris, Éditions du Seuil.

Malik Momim (2020). "A hierarchy of limitations in machine learning". arXiv:2002.05193 (cs, econ, math, stat). En Internet: <http://arxiv.org/abs/2002.05193>.

Maturana Humberto, y Varela, Francisco (2003). El árbol del conocimiento. Las bases biológicas del entendimiento humano. Buenos Aires, Lumen / Editorial Universitaria.

Mayz Valenilla Ernesto (1993). Fundamentos de la meta-técnica. Barcelona, Gedisa.

Minsky Marvin (1974). "Inteligencia artificial". En Carnap, Rudolf y otros. Matemáticas en las ciencias del comportamiento. Madrid, Alianza.

Mitcham, Carl (1995). "Notes Toward a Philosophy of Meta-Technology", PHIL & TECH 1:1&2 Fall.

Mittelstadt Brent (2019). Principles alone cannot guarantee ethical AI. Nat Mach Intell 1, 501–507 (2019). <https://doi.org/10.1038/s42256-019-0114-4>.

NIST (2023): Artificial Intelligence Risk Management Framework (AI RMF 1.0). En internet: [www.nist.gov/itl/ai-risk-management-framework](http://www.nist.gov/itl/ai-risk-management-framework).



Noy, Shakked y Zhang, Whitney (2023), Experimental Evidence on the Productivity Effects of Generative Artificial Intelligence. En SSRN: <https://ssrn.com/abstract=4375283> or <http://dx.doi.org/10.2139/ssrn.4375283>

OECD/CAF (2022), Uso estratégico y responsable de la inteligencia artificial en el sector público de América Latina y el Caribe, Estudios de la OCDE sobre Gobernanza Pública, OECD Publishing, Paris, En Internet: <https://doi.org/10.1787/5b189cb4-es>.

OIT (2023), Generative AI and jobs: A global analysis of potential effects on job quantity and quality. Autores: Paweł Gmyrek, Janine Berg, David Bescond.

Pasquale, Frank (2015). The Black Box Society. The Secret Algorithms that Control Money and Information. Cambridge (EE.UU.), Harvard University Press.

Pasquinelli Matteo y Joler Vladan (2021). “El nooscopio de manifiesto. La inteligencia artificial como instrumento del extractivismo cognitivo”, en revista La Fuga. <https://lafuga.cl/el-nooscopio-de-manifiesto/1053>.

Penrose Roger (1996). La mente nueva del emperador. En torno a la cibernética, la mente y las leyes de la física. México, Fondo de Cultura Económica.

Peng Sida et al (2023). The Impact of AI on Developer Productivity: Evidence from GitHub Copilot. En arXiv:2302.06590, Cornwell University. <https://doi.org/10.48550/arXiv.2302.06590>

Simon Herbert y Newell Allen (1975). “Proceso de la información en el computador y en el hombre”. En Pylsychyn Zenon (comp.). Perspectivas de la revolución de los computadores. Madrid, Alianza.

Srnicek Nick (2018). Capitalismo de plataformas. Buenos Aires, Caja Negra Editora.

Tacsir Ezequiel y Tacsir Andrés (2022). “Experiencias internacionales de Centros de Inteligencia Artificial y recomendaciones”. Propuesta de Diseño, Gobernanza y Evaluación del Centro Argentino Multidisciplinario de Inteligencia Artificial (CamIA). En Internet: <https://www.iadb.org/document.cfm?id=EZIDB0000029-756715771-18>

Urban Tim (2015). “The AI Revolution: The Road to Superintelligence”, dos partes, en Internet: <https://waitbutwhy.com/2015/01/artificial-intelligence-revolution-1.html>

Van Dijck José; Poell, Thomas y De Waal, Martijn (2018). The Platform Society. Public Values in a Connective World. Oxford, Oxford University Press.

Weizenbaum, Joseph (1978). Las fronteras entre el ordenador y la mente. Madrid, Pirámide.

Vercelli Ariel (2023). “Las inteligencias artificiales y sus regulaciones: Pasos iniciales en Argentina, aspectos analíticos y defensa de los intereses nacionales”, en Revista de la Escuela del Cuerpo de Abogados y Abogadas del Estado, mayo 2023, Año 7 N° 9, Buenos Aires, Argentina (ISSN 2796-8642), pp. 195-21. En internet: [revistaecae.ptn.gob.ar/index.php/revistaecae/article/download/232/213/548](http://revistaecae.ptn.gob.ar/index.php/revistaecae/article/download/232/213/548)

Zarifhonarvar, Ali (2023). Economics of ChatGPT: A Labor Market View on the Occupational Impact of Artificial Intelligence. En SSRN: <https://ssrn.com/abstract=4350925> or <http://dx.doi.org/10.2139/ssrn.4350925>

Zuboff Shoshana (2020). La era del capitalismo de la vigilancia. Barcelona, Paidós.



**.UBA40**<sup>∞</sup>  
AÑOS DE  
DEMOCRACIA

**.UBA SOCIALES 35**  
Facultad de Ciencias Sociales