

INAP

ISSN en trámite

Documentos del INAP

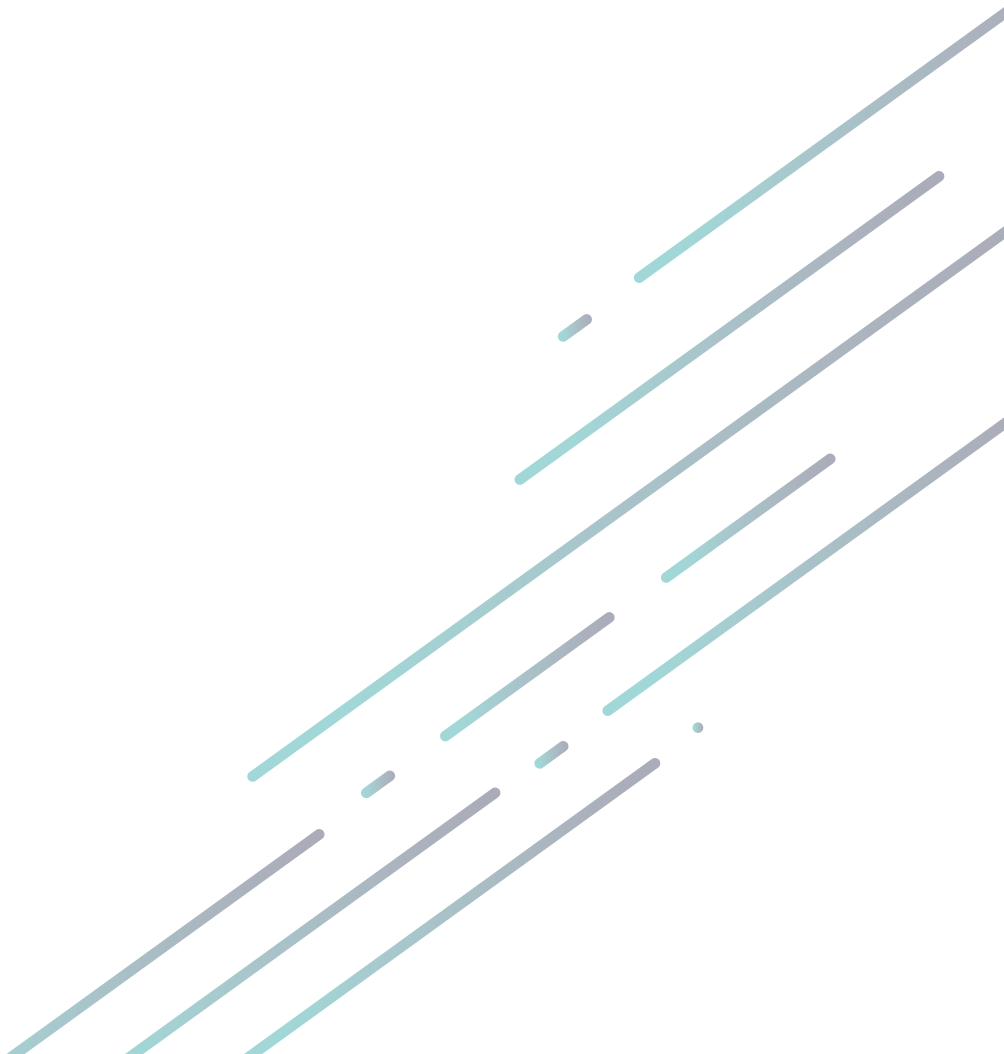
Año 1 • N.º 2 | 2024

Los desafíos de la inteligencia artificial
en las políticas de integridad

Martín Astarita



Ministerio de Desregulación
y Transformación del Estado
República Argentina



Documentos del INAP

Año 1 • N.º 2 | 2024

Los desafíos de la inteligencia artificial en las políticas de integridad

Martín Astarita*

* El autor forma parte del equipo de investigación «Inteligencia artificial en la Administración Pública».

Índice

Introducción	5
1. La IA y las políticas de integridad: una relación compleja	7
A modo de conclusión	11
Referencias bibliográficas	12

Introducción

Aunque ha invadido la mayoría de los campos de acción humana y su uso forma parte cada vez más de la vida cotidiana, no resulta sencillo encontrar una definición sobre inteligencia artificial (IA) que se acepte en forma unánime (Cross, 2023).

Un punto de partida para comprender su significado es considerar la definición que aportan algunos de los sistemas creados por la propia inteligencia artificial. Por ejemplo, según Meta IA, se trata de un campo de la informática que se enfoca en la creación de sistemas capaces de realizar tareas que normalmente requieren inteligencia humana, como el aprendizaje, el razonamiento, la percepción y la comprensión. Una respuesta prácticamente idéntica brinda el *chatbot* de OpenAI.

La imitación de la inteligencia humana constituye, sin dudas, el aspecto neurálgico de las tecnologías basadas en IA (Aarvik, 2019). Por eso, tempranamente, Marvin Lee Minsky (1961) la definió como una ciencia de hacer que las máquinas hagan cosas que requerirían inteligencia si fueran hechas por hombres.

Sin embargo, en este punto es necesario distinguir entre lo que realmente es la IA en la actualidad y lo que puede ser a futuro, en caso de alcanzar todo su potencial. En este sentido, se la puede clasificar en dos categorías: la IA de propósitos específicos y la IA general.

Por un lado, la IA de propósitos específicos (*Narrow AI*), como se la conoce hoy en día, tiene un campo de aplicación estrecho, en el que se imita la inteligencia humana solo en determinadas acciones y para tareas específicas (reconocimiento de voz, de imágenes, entre otras). Es objeto de discusión si esta forma de IA puede ser considerada realmente como una forma de inteligencia humana.

Por otro lado, la IA general (*Artificial General Intelligence*) es un proyecto y como tal no existe en la actualidad. De hecho, el debate se suscita entre quienes son escépticos de la posibilidad de concretar esta aspiración y las visiones más optimistas (Antebi, 2021). La IA general hace alusión a un sistema informático que potencialmente podría llegar a replicar la capacidad cognitiva del ser humano. Ya no abocaría simplemente a la

realización de algunas tareas específicas, sino que podría cumplir con cualquier actividad intelectual de la que es capaz un ser humano, incluida, lógicamente, tener conciencia de sí misma. La IA general, en definitiva, es una máquina que se comporta igual que el cerebro humano, con funciones biológicas, psicológicas y cognitivas (Antebi, 2021).

Existen, además, distintos modelos de IA. Uno de ellos, hasta hace poco dominante, es el simbólico, vinculado con los llamados «sistemas de símbolos físicos». Se trata de un modelo *top-down* que se basa en el razonamiento lógico y la búsqueda heurística mediante el procesamiento de símbolos, conceptos y reglas definidas previamente (López de Mántaras, 2018).

El segundo modelo es el conexionista, de tipo *bottom-up*, inspirado en la neurociencia, que busca replicar el funcionamiento de las neuronas biológicas (De Spiegeleire *et al.*, 2017) mediante el aprendizaje de patrones complejos determinados por las capas de la red. A diferencia del basado en el simbolismo, este modelo no necesita de reglas o algoritmos previamente definidos, sino que descubre las reglas necesarias y ajusta sus parámetros a partir del set de datos con el que es entrenado (López de Mántaras, 2018).

Si bien la investigación sobre IA existe desde hace al menos seis décadas, su desarrollo no estuvo exento de dificultades. Cada ciclo de esperanzas que se abría con nuevos avances fue seguido de largos y desoladores inviernos en los que primó la desesperanza ante la (aparente) falta de desarrollo concreto. Sin embargo, en los últimos diez años, se asiste a un renovado interés en el tema, con el uso de modernas técnicas de aprendizaje automático, redes neuronales y el manejo computacional de enormes cantidades de conjuntos de datos (Neapolitan y Jiang, 2018; Russell, 2019).

Puntualmente, lo que dio mucho impulso es la aparición de la denominada «IA generativa», que produce contenido original a partir de datos existentes. Esta tecnología utiliza algoritmos y redes neuronales avanzadas para aprender de textos, imágenes, videos, sonido y de su combinación.

Este nuevo auge involucra numerosas subdisciplinas que incluyen el procesamiento del lenguaje natural, la visión computacional, los agentes autónomos, la inferencia mecánica, el aprendizaje automático estadístico y la robótica. Su uso sin precedentes

y la forma en que está profundamente arraigada en casi todos los aspectos de la sociedad anuncian la cuarta revolución industrial (Agrawal *et al.*, 2018).

La Administración Pública, por su parte, no ha quedado al margen de la revolución de la IA. En distintos países se constata un aumento progresivo en su uso (Consejo Latinoamericano de Administración para el Desarrollo [CLAD], 2023; Criado, 2019), desde la recopilación y el análisis de grandes volúmenes de datos hasta el diseño, la ejecución y la evaluación de políticas públicas, así como también una interacción más fluida con la ciudadanía (Ospina Díaz y Zambrano Ospina, 2023).

1. La IA y las políticas de integridad: una relación compleja

Las nuevas tecnologías basadas en IA, naturalmente, no han pasado desapercibidas entre los especialistas en materia de transparencia e integridad. Aunque su desarrollo intenso es reciente y no es posible calibrar con precisión sus efectos venideros, lo cierto es que, desde finales del siglo xx, las estrategias anticorrupción se han visto inmersas en el debate por las constantes innovaciones tecnológicas, entre otras, la aparición de internet y la revolución digital. Y en estas discusiones, incluida la que concierne estrictamente a la IA, es posible identificar básicamente tres posiciones.

Los optimistas representan la primera posición. Son quienes se inclinan por rescatar los aspectos positivos de las innovaciones tecnológicas en el diseño de políticas públicas vinculadas con la transparencia y la integridad. La premisa es que, si la corrupción es un asunto con motivaciones humanas, la automatización puede eliminar sesgos, irregularidades y espacios de discrecionalidad, además de aumentar las posibilidades de monitoreo (Nurkey *et al.*, 2022). Complementariamente, se sostiene que, mientras que las anteriores estrategias anticorrupción, al ser *top-down* o verticalistas, tuvieron poco éxito (Mungiu-Pippidi & Dadasov, 2017), el uso de IA puede propiciar mecanismos anticorrupción de tipo *bottom-up* (enfoques ascendentes, con la participación ciudadana como factor clave), donde, por ejemplo, el uso de herramientas algorítmicas (*bots*) sirva para automatizar las denuncias de corrupción en redes sociales.

En la vereda opuesta, se encuentran los denominados «pesimistas», quienes advierten que las nuevas tecnologías suponen un aumento en las posibilidades de cometer actos de corrupción con altos niveles de sofisticación que, por las características tecnológicas involucradas, resultan muy difíciles de detectar y controlar.

Finalmente, una tercera posición, adoptada en este trabajo, es que las tecnologías en general, al menos por ahora, no cambian en lo que hace a la IA: son un instrumento, es decir que el aspecto clave en todo caso es la oportunidad y el modo en que se usan. Hacia el final de este documento se ofrecen más detalles al respecto.

En las siguientes dos secciones se hace una sucinta revisión de la literatura especializada y se rescatan beneficios, oportunidades, riesgos y amenazas de la IA.

1.1 Beneficios y oportunidades

La IA constituye una herramienta cada vez más frecuente en la gestión gubernamental, sobre todo en América Latina, región en la que la corrupción suele estar al tope de las preocupaciones ciudadanas (Organización para la Cooperación y el Desarrollo Económico y Banco de Desarrollo de América Latina [OCDE/CAF], 2022), esta nueva tecnología se utiliza con especial énfasis en el diseño y la aplicación de herramientas de transparencia e integridad.

En el campo de las contrataciones públicas, donde, según Rose Ackerman y Palifka (2016), se dan los principales mecanismos de corrupción, en el último tiempo, han florecido iniciativas de esta naturaleza. Pero, sobre todo, hay muchas esperanzas cifradas en el potencial que puede tener la IA generativa en las distintas fases del proceso de compras públicas, desde la asistencia en la elaboración de pliegos a los órganos de contratación, la elaboración de ofertas de los licitadores, la valoración de ofertas, el control y la eliminación de sesgos, y la detección de irregularidades (Benítez Palma y García Rodríguez, 2024).

También suelen ponderarse las ventajas que supone el uso de IA para detectar operaciones comerciales fraudulentas, que muchas veces están vinculadas de manera directa o indirecta con casos de corrupción. En la actualidad, existen distintos

modelos predictivos que emplean fundamentalmente dos técnicas. Una de ellas, el modelo supervisado, consiste en suministrarle al algoritmo grandes conjuntos de datos acerca de operaciones realizadas, identificando también aquellas que han sido objetadas como sospechosas o fraudulentas. Así, el algoritmo es entrenado para identificar patrones de irregularidades y enseguida puede comenzar a enviar alertas. La otra técnica, el modelo no supervisado, está basada en que el algoritmo es quien se encarga de detectar anomalías a partir del suministro de una gran cantidad de datos no estructurados.

Asimismo, la IA puede promover la transparencia y la rendición de cuentas mediante sistemas de identificación de riesgos y alertas tempranas, sobre la base de sistemas *bottom-up*, en los que resulta decisiva la participación ciudadana. En el marco de lo que se ha definido como «gestión pública inteligente» (Criado, 2016), la creciente digitalización de los servicios públicos permite una interacción más fluida y cercana del gobierno con sus ciudadanos, con la personalización de los procesos (enfoque *citizen-centric*) y la posibilidad de efectuar casi en tiempo real denuncias de corrupción.

El punto nodal de las variadas herramientas de este profundo proceso de transformación digital (datos abiertos, gobierno abierto, gobierno electrónico, etc.) es que implican el reemplazo de las intervenciones humanas por plataformas tecnológicas en la provisión de bienes y servicios públicos. Así, se ha señalado que los países con mayores niveles de desarrollo de gobierno electrónico suelen exhibir menores niveles de corrupción (Gallego, 2021).

En síntesis, el uso de la IA es ponderado por sus beneficios —reales o potenciales— en el diseño y la aplicación de políticas de transparencia e integridad tanto en cuanto a la prevención como al control y la sanción de hechos vinculados con la corrupción.

1.2 Riesgos y amenazas

En este apartado se efectúa una revisión sintética de un conjunto de debilidades, riesgos y amenazas que la literatura especializada ha señalado en relación con la aplicación de IA en materia de integridad y lucha contra la corrupción.

Si bien, en comparación con sus potenciales beneficios, el lado negativo de la IA ha sido menos explorado, en el último tiempo esta tendencia comenzó a revertirse. Así, surgió el concepto de inteligencia corrupta, esto es, el abuso de los sistemas de IA por parte de quienes tienen el poder para su beneficio privado (Köbis *et al.*, 2022).

En primer lugar, la IA, como toda tecnología, es un instrumento que puede ser utilizado con distintos fines, no siempre benéficos. Así como se advirtió en el apartado anterior que tiene altas potencialidades para mejorar la eficacia en la lucha contra la corrupción, del mismo modo debe advertirse que las posibilidades de recurrir a la IA para la comisión de sofisticadas prácticas irregulares se han multiplicado merced a estas nuevas tecnologías. Cabal expresión de ello son las operaciones comerciales que, en un marco de creciente digitalización, dan lugar a la proliferación de fraudes y acciones corruptas, y que, por su velocidad y magnitud, no siempre son fáciles de detectar y sobre todo de prevenir.

Asimismo, una cuestión especialmente sensible se vincula con la manipulación de los algoritmos con fines corruptos, es decir, cuando el código o los datos de entrenamiento no son representativos o están incompletos, lo que deriva en resultados con sesgos a favor de un determinado grupo en desmedro de otro. Los expertos utilizan el concepto de *captura algorítmica* para referirse a este proceso (Rahwan, 2022), que puede producirse por una decisión consciente al preparar el código o hacer la carga de datos, o bien de manera involuntaria al crear una base de datos para el entrenamiento que esté sesgada por prejuicios o factores de discriminación subyacentes (Köbis *et al.*, 2022).

Más en general, se habla de la caja negra de la IA, es decir, de la dificultad para hacer explicable de qué manera la IA aprende determinadas habilidades o por qué toma las decisiones que toma. Ello ocurre especialmente con el modelo conexionista (ver introducción, que resulta opaco y poco transparente (Caversan, 2020).

Otra de las cuestiones señaladas se relaciona con que los trabajadores del sector público suelen tener pocas competencias en el manejo de IA, por lo que es usual que los gobiernos recurran a proveedores externos. El riesgo es que la decisión de

política pública, en un tema tan sensible como el de la integridad y la lucha contra la corrupción, quede en manos de empresas privadas, lo que ha sido denominado críticamente como «externalización de la responsabilidad» (Ceva y Jiménez, 2022).

Finalmente, otra crítica es que, si bien la IA puede analizar con rapidez y eficiencia grandes volúmenes de datos, aún hoy carece de habilidades blandas, como la empatía, la persuasión y, en general, habilidades interpersonales que poseen los seres humanos y que son fundamentales para comprender las motivaciones de un acto de corrupción (Muntanga, 2023).

A modo de conclusión

Han sido expuestos de manera sucinta beneficios y oportunidades, así como riesgos y amenazas de la IA en materia de transparencia, integridad y corrupción a partir de una revisión de literatura especializada. Se trata de un primer eslabón con el fin de sentar las bases para una nueva agenda de investigación que pueda reunir progresivamente evidencia empírica por medio de estudios de caso.

Del análisis efectuado surge que la disponibilidad de tecnologías basadas en IA, a pesar de sus evidentes ventajas y potencialidades, no se traduce automáticamente en una mejora de la transparencia y la integridad en la gestión pública.

Son determinantes, en este sentido, el modo en que se usa y se regula la IA, así como el contexto en el que se utiliza. Distintos organismos internacionales destinan cada vez más esfuerzos por construir una ética de la gestión pública que sitúe en el centro el debate sobre usos y riesgos de la IA en la Administración Pública. Esto no significa importar acríticamente modelos concebidos para otros contextos. Es necesario diseñar una estrategia de desarrollo de las habilidades y las potencialidades de la IA para la gestión pública que considere las particularidades locales.

Referencias bibliográficas

- Aarvik, P. (2019). *Artificial Intelligence - A Promising Anti-Corruption Tool in Development Settings?* (U4 Report 2019:1). CMI Chr. Michelsen Institute.
- Agrawal, A., Gans, J. & Goldfarb, A. (2018). *Exploring the Impact of Artificial Intelligence: Prediction versus Judgment* (Working Paper 24626). National Bureau of Economic Research.
- Antebi, L. (2021). *Artificial Intelligence and National Security in Israel*. Institute for National Security Studies.
- Benítez Palma, M. y García Rodríguez, E. (2024). El uso de la IA generativa en la contratación pública para detectar prácticas irregulares. *Revista Auditoría Pública*, (83), 83-92.
- Caversan, F. (2020). *Symbolism Versus Connectionism In AI: Is There A Third Way?* Forbes Technology Council.
- Centro Latinoamericano de Administración para el Desarrollo (2023). *Carta Iberoamericana de Inteligencia Artificial en la Administración Pública*.
- Ceva, E. & Jiménez, M. C. (2022). Automating Anticorruption? *Ethics and Information Technology*, 24, 24-48.
- Criado, I. (2016). Las administraciones públicas en la era del gobierno abierto. Gobernanza inteligente para un cambio de paradigma en la gestión pública. *Revista de Estudios Políticos*, (173).
- Criado, I. (2019). *Inteligencia artificial y administración pública. Incorporando nuevas capacidades para la innovación pública gobernando con algoritmos*. Ministerio de Ciencia, Innovación y Universidades de Madrid.
- Cross, P. (2023). *Artificial Intelligence: A Threat to Middle-Class, White-Collar Jobs?* Macdonald-Laurier Institute.
- Gallego, J. (2021). ¿Puede la digitalización del Estado ayudarnos en la lucha contra la corrupción? CAF.
- De Spiegeleire, S., Maas, M. & Sweijts, T. (2017). *Artificial Intelligence and the Future of Defense: Strategic Implications for Small and Medium-Sized Force Providers*. Hague Centre for Strategic Studies.

- Köbis, N., Starke, C. & Gill, J. (2022). *The Corruption Risks of Artificial Intelligence*. Transparency International.
- López de Mántaras, R. (2018). El futuro de la IA: hacia inteligencias artificiales realmente inteligentes. En *¿Hacia una nueva Ilustración? Una década trascendente*. BBVA.
- Minsky, M. (1961). Steps Toward Artificial Intelligence. *Proceedings of the IRE*, 49(1), 8-30.
- Mungiu-Pippidi, A. & Dadasov, R. (2017). When do Laws Matter? *Crime, Law and Social Change*, 68(1), 1-16.
- Muntanga, K. (2023). *Artificial Intelligence – Is It the Ultimate in the Fight against Corruption?* IACA Alumni Magazine.
- Neapolitan & Jang (2018). *Artificial Intelligence: With an Introduction to Machine Learning* (Second Edition). Chapman & Hall.
- Nurkey, A., Mukasheva, A. & Yedilkhan, D. (2022). Models and Methods of Digital Mechanisms in Anticorruption, their Advantages and Disadvantages, and Applications. IOP Conf. Ser.: Mater. Sci. Eng. 1216 012015.
- Organización para la Cooperación y el Desarrollo Económico y Banco de Desarrollo de América Latina (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.
- Ospina Díaz, M. y Zambrano Ospina, K. (2023). Gobierno digital e inteligencia artificial: una mirada al caso colombiano. *Administración y Desarrollo*, 53(1), 1-34.
- Rahwan, I. (2022). Prevent Algorithm Capture. Evil AI Cartoons. <https://www.evilaicartoons.com/archive/prevent-algorithm-capture>
- Rose Ackerman, S. y Palifka, S. (2016). *Corrupción y gobierno: Causas, consecuencias y reformas*. Marcial Pons.
- Russell, S. (2019). *Human Compatible: AI and the Problem of Control*. Penguin.

Documentos del INAP

Año 1 - N.º 2 2024

Instituto Nacional de la Administración Pública (INAP)

Av. Roque Sáenz Peña 511, Ciudad Autónoma de Buenos Aires, República Argentina

C. P.: C1035AAA - Tel.: (+54 11) 6065-2310

Correo electrónico: dinvesti@jefatura.gob.ar

ISSN en trámite

Editor responsable

Cecilia Mendoza

Por equipo Dirección de Gestión del Conocimiento, Investigación y Publicaciones

Redacción

Martín Astarita

Edición y corrección

Laura Scisciani

Eugenia Caragunis

Maquetado web

Natalia Baez Becker

Ciro Paroli

Por equipo Comunicación INAP

Revisión editorial

Patricia Iacovone

Diseño y diagramación editorial

Roxana Pierri

Lucía Maillo Puente

Federico Cannone

Las ideas y planteamientos contenidos en la presente edición son de exclusiva responsabilidad de su/s autor/es y no comprometen la posición oficial del INAP.

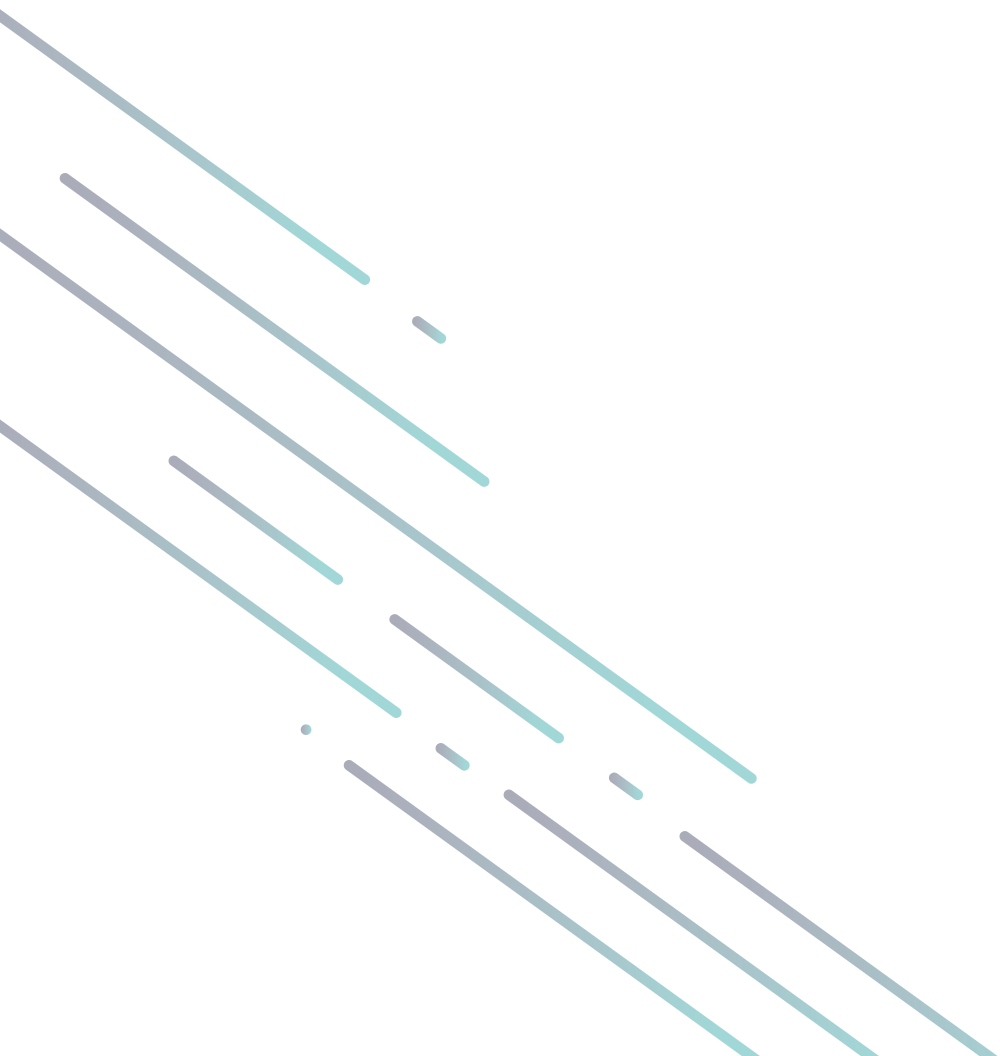
INAP no asume responsabilidad por la continuidad o exactitud de los URL de páginas web externas o de terceros referidas en esta publicación y no garantiza que el contenido de esas páginas web sea, o continúe siendo, exacta o apropiada.



El contenido de esta publicación se brinda bajo una Licencia Creative Commons Atribución-No Comercial 2.5 Argentina. Es posible copiar, comunicar y distribuir públicamente su contenido siempre que se cite al/a los autor/es individual/es y el nombre de esta publicación, así como la institución editorial. No puede utilizarse con fines comerciales.

Esta publicación se encuentra disponible en forma libre y gratuita en: publicaciones.inap.gob.ar

Septiembre 2024





**Secretaría de Transformación
del Estado y Función Pública**
Ministerio de Desregulación
y Transformación del Estado

**Subsecretaría de Desarrollo y
Modernización del Empleo Público**