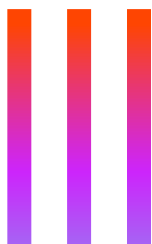




HOJA DE RUTA

Innovar con Ciencia de Datos en el sector público



Innovar con Ciencia de Datos en el sector público

Autoría

María Vanina Martínez

Directora del Programa de Ciencia de Datos e Inteligencia Artificial (PCDIA) de la Fundación Sadosky. Doctora de Ciencias de la Computación (University of Maryland) e investigadora del CONICET.

Victoria Gisel Dumas

Consultora del PCDIA de la Fundación Sadosky. Doctora en Ciencias Químicas (Universidad de Buenos Aires) y profesora de Visualización y Comunicación en la Maestría en Ciencia de Datos de la Universidad de San Andrés.

Marianela Sarabia

Consultora del PCDIA de la Fundación Sadosky. Magister en Economía Aplicada para el Desarrollo (UniTo & SciencesPo), y profesora de Políticas Públicas y Machine Learning (Universidad Nacional de Guillermo Brown).

Ivana Feldfeber Kisilevsky

Consultora del PCDIA de la Fundación Sadosky. Posgrado en Ciencia de Datos, Aprendizaje Automático y sus Aplicaciones (Universidad Nacional de Córdoba) y Directora Ejecutiva del Observatorio de Datos con Perspectiva de Género "DataGénero".

Colaboración

Manuela Cerdeiro

Fundación Sadosky

Lucas Somacal

Fundación Sadosky

Inés Roggi

Fundación Sadosky

Daniel Yankelevich

Fundar

Mariano Cervellini

Fiscalía de la Provincia de Buenos Aires

Marta Kisilevsky

Ministerio de Educación de la Nación

Gabriela Reta

Universidad Nacional de General Sarmiento

María del Mar Monti

Ministerio de Defensa de la Nación

Edición

Laura Ramírez Rivillas

Fundación Sadosky

Alejandra Stafetta

Fundación Sadosky

Federico Rey

Fundación Sadosky

Diseño

Jaqueline Schaab

Fundación Sadosky

Facundo Manini

Fundación Sadosky

Autoridades de la Fundación

Dr. Manuel Sadosky

Daniel Filmus

Presidente

Fernando Schapachnik

Director Ejecutivo

Hoja de ruta: Innovar con datos en el sector público
/ Ivana Feldfeber Kisilevsky...
[et al.]. - 1a ed. - Ciudad Autónoma de Buenos
Aires : Fundación Sadosky, 2022.
Libro digital, PDF

Archivo Digital: descarga
ISBN 978-987-48926-0-7

1. Computación. 2. Políticas Públicas. 3.
Recopilación de Datos. I. Feldfeber Kisilevsky, Ivana.
CDD 352.75



Historial de versiones

Versión

Cambios

14-11-2022

Primera versión completa publicada

Índice

Resumen Ejecutivo	5
Introducción	6
Explorando el terreno de la Innovación Pública con herramientas de Ciencia de Datos e Inteligencia Artificial	8
Primera etapa: Pensamiento de diseño para la co-creación de servicios y políticas públicas	9
Segunda etapa: ¿Qué entendemos por datos? ¿Y por Ciencia de Datos o Inteligencia Artificial?	12
Tercera etapa: El uso de datos, Ciencia de Datos e Inteligencia Artificial como un proceso de innovación pública	17
Tramo 1: Innovar con datos	23
Tramo 2: Innovar con ciencia de datos	26
Tramo 3: Mediciones de impacto e Inteligencia Artificial	28
Tramo 4: Estimaciones y predicciones con Inteligencia Artificial	29
Recomendaciones: los infaltables a la hora de usar Ciencia de Datos o Inteligencia Artificial en el sector público	34
¿Es la Inteligencia Artificial o la Ciencia de Datos la mejor herramienta para abordar el problema que quiero resolver?	35
¿Cuento con un ecosistema de actores para diseñar e implementar herramientas de Ciencia de Datos y/o Inteligencia Artificial en el sector público?	39
¿Con qué datos cuento? ¿Con qué datos me gustaría contar?	41
¿A qué nos referimos con calidad de los datos?	
¿Qué infraestructura y equipamiento necesito para realizar y mantener una herramienta de Ciencia de Datos y/o Inteligencia Artificial?	43
Una vez que tengo la herramienta ¿qué hago?	
¿Qué normativa existe sobre datos, seguridad y privacidad?	47
¿Cómo son los circuitos administrativos para usar Ciencia de Datos e Inteligencia Artificial?	
¿Cómo se trabaja hacia la transparencia en estos proyectos?	50
¿Qué son los sesgos y los riesgos? ¿Cómo puedo abordar esta problemática?	52
Conclusión	55
Checklist para implementar Ciencia de Datos e Inteligencia Artificial en el sector público	57
Referencias	60
Links de interés	61

Resumen Ejecutivo

El sector público, en sus distintos ámbitos, utiliza cada vez más sistemas de información para realizar distintas tareas y prestar servicios. Estos sistemas, la mayoría de ellos basados en datos, se utilizan para brindar información de importancia para la ciudadanía, tomar decisiones sobre asignación de recursos, agilizar los procedimientos gubernamentales, y ayudar a las y los funcionarios nacionales, provinciales y municipales a tomar importantes decisiones en el ámbito social y legal.

A pesar de la importancia de estos usos y decisiones, muchas veces las personas que tienen la responsabilidad de pensar estos sistemas provienen de otras áreas del conocimiento y no tienen la experiencia técnica para abordar todas las dimensiones que es necesario contemplar para que la construcción de estos sistemas sea tecnológica y económicamente conveniente, libre de sesgos, socialmente y éticamente responsable, y legalmente ajustada. Este documento busca brindar asistencia en esos aspectos.

Esta hoja de ruta propone navegar las alternativas e identificar las preguntas fundamentales que se presentan en el proceso de innovación con datos, específicamente, en el sector público. El documento se divide en tres secciones principales. La primera, propone enmarcar cuál es el ámbito de acción, analizando los distintos pasos en el proceso de innovación pública guiado por el pensamiento de diseño. También tiene como objetivo dar a conocer qué tecnologías de datos existen, cuáles son las particularidades que las diferencian y qué riesgos conllevan. Conocer estas especificidades resulta fundamental para poder definir las opciones más adecuadas para cada caso particular.

La siguiente sección explica dónde participan la Ciencia de Datos y la Inteligencia Artificial (CDIA) en el proceso de innovación pública, qué alternativas se presentan a lo largo de este proceso y ofrece ejemplos de proyectos locales para cada uno de los tramos. De manera transversal, todos los aspectos importantes del proceso de diseño serán re analizados a la luz del uso de los datos y las herramientas disponibles.

Finalmente, la última sección de esta hoja de ruta propone un análisis sobre los aspectos más importantes del proceso de diseño, de acuerdo a la problemática del uso de datos, y presenta una lista de recomendaciones a tener en cuenta a la hora de implementar CDIA. Este análisis, dispara a su vez, un conjunto de preguntas críticas que sirven de guía práctica al momento de enfrentar los recorridos de los proyectos de innovación en políticas públicas.

Introducción

Les damos la bienvenida a este documento que decidimos llamar hoja de ruta para implementar herramientas de Ciencia de Datos en el sector público. Seguramente se estén preguntando por qué necesitamos una hoja de ruta para estos temas.

Bueno, necesitamos este mapa, esta guía, por la misma razón por la que necesitamos una hoja de ruta cuando estamos planificando un viaje. Sin importar adonde vamos, cuando tenemos que trasladarnos por grandes distancias, o cuando cambiamos nuestro recorrido habitual y conocido, tenemos que tener un plan un poco más informado de lo que vamos a hacer: a qué hora sale el transporte, de dónde sale, adonde llega y cómo hay que hacer para ir a donde queremos ir.

Por eso, proponemos pensar en utilizar Ciencia de Datos e Inteligencia Artificial (de ahora en adelante CDIA) como un viaje a lo inexplorado que requiere preparación y lectura previa. No es un destino, sino un recorrido. No es un viaje solitario, sino colectivo, multidisciplinar. Hay quienes viajan para escaparse de sus problemas o porque piensan que cuando regresen todo se habrá solucionado mágicamente. Lejos de eso, este viaje en el cual nos vamos a embarcar no será una solución mágica a ningún problema, sino una aproximación a un mundo con mucho potencial y con muchas advertencias también.

Esta hoja de ruta tiene como objetivo ofrecer orientación a las personas a cargo de formular, monitorear o gestionar políticas o servicios públicos, para que puedan aprovechar las ventajas y conocer los riesgos que presentan las aplicaciones de CDIA en la gestión pública. Además, procura establecer un marco práctico y accesible para el uso de herramientas estadísticas, CDIA en el diseño y la gestión de políticas públicas y en la prestación de servicios públicos. En particular, nos enfocaremos en las posibilidades de iniciar, fortalecer o consolidar el uso de herramientas de CDIA en el marco de la administración pública.

El uso de CDIA permite sistematizar información, monitorear el rendimiento de los procesos de gestión, agilizar procedimientos, analizar impactos, reproducir y verificar resultados, brindar transparencia en la toma de decisiones, entre muchísimas otras aplicaciones. Para profundizar al respecto, en esta hoja de ruta contaremos cómo abordar estas temáticas pensando en las políticas y servicios públicos, tanto en las dimensiones vinculadas a la propia gestión como al diseño y la implementación o la adquisición de las mismas. Además, presentaremos algunos disparadores para problematizar el uso de CDIA, fuentes de consulta y algunos consejos e interrogantes para profundizar sobre la calidad de los datos y el impacto de estas herramientas.

A su vez, los procesos de diseño y el desarrollo de sistemas de CDIA presentan un gran potencial de escalamiento e integración transversal de la administración pública en sus distintas jurisdicciones mediante la adopción, la transferencia y la formación adecuada. Sin embargo, el uso de herramientas estandarizadas o la compra de herramientas prefabricadas por un equipo externo, carentes de un proceso consultivo e interdisciplinar que pueda pensar de manera integral el diseño o desarrollo propio, pueden generar tratamientos injustos y sesgados. Estos sesgos o problemas de implementación pueden llevar a la adquisición de un producto que no abone a solucionar el problema de fondo (social, económico, estructural) que se está intentando abordar.

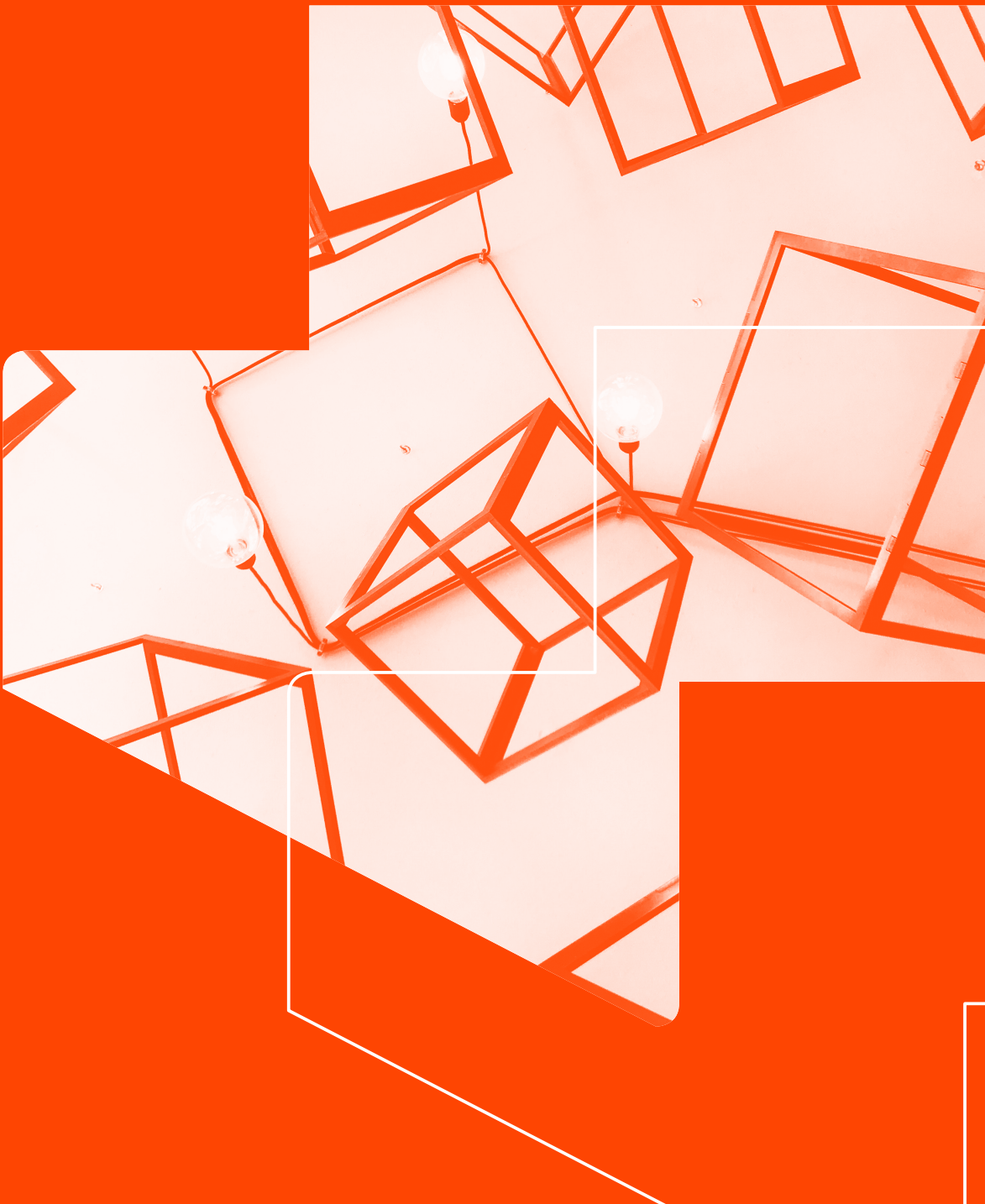
Esta hoja de ruta se propone recorrer los diversos caminos que existen en los procesos de innovación con datos e identificar las preguntas que conviene hacer si se quiere innovar con datos específicamente en la administración pública. El documento está dividido en tres secciones principales:

1. Comenzaremos enmarcando **cuál es el ámbito de acción** y analizando los distintos pasos en el proceso de innovación pública guiado por el pensamiento de diseño. Seguiremos descubriendo **qué tecnologías de datos existen** y descubriendo que no son todas lo mismo: cada una tiene sus particularidades y tal vez conlleva riesgos de distinto alcance, por lo que la primera y principal pregunta que hay que hacer es **qué opciones son las más adecuadas** para cada caso en particular.

2. Más adelante contaremos **dónde entra en juego la CDIA** en el proceso de innovación pública, qué alternativas se presentan a lo largo de este proceso y mostraremos ejemplos de proyectos locales para cada uno de los tramos del recorrido. De manera transversal, todos los aspectos importantes del proceso de diseño son re analizados desde la perspectiva del uso de datos y las herramientas disponibles.

3. Finalmente, analizaremos **cuáles son los aspectos más importantes del proceso de diseño** a la luz de la problemática del uso de datos y presentaremos una lista de preguntas “infaltables” a tener en cuenta a la hora de usar CDIA. Este análisis dispara un conjunto de preguntas críticas (checklist) que servirán de guía práctica al momento de enfrentar los recorridos que analizamos y otros nuevos que puedan ir apareciendo a lo largo de esta aventura.

Explorando el terreno de la Innovación Pública con herramientas de Ciencia de Datos e Inteligencia Artificial



Primera etapa

Pensamiento de diseño para la co-creación de servicios y políticas públicas

La creciente complejidad, la interdependencia y el dinamismo que demanda la resolución de problemas públicos o detección de necesidades en las sociedades contemporáneas ha permitido acuñar el concepto de **co-creación de servicios y políticas públicas** o innovación social para referirse a un nuevo paradigma de gobernanza colaborativa, interdisciplinaria y experimental¹. El desarrollo y la implementación de una nueva idea en forma de producto, servicio o proceso es una innovación social. Como la innovación social a la que nos referimos en esta hoja de ruta está centrada en el sector público, podemos condensar todo esto como un **proceso de innovación pública**.

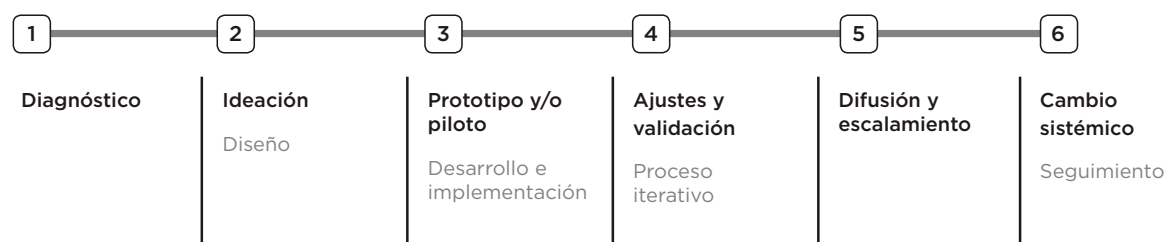
Estos nuevos y desafiantes procesos buscan integrar el mundo lógico o analítico con aquel más creativo, inductivo o experimental para analizar visiones, necesidades y conocimientos en un diagnóstico holístico. A partir de ese diagnóstico, contaremos con la posibilidad de crear una solución innovadora desde su origen o mejorar un proceso existente. ¿Cómo lo hacemos? Una primera herramienta metodológica a tener en cuenta es el **pensamiento de diseño** o *design thinking*, que nos permite desglosar el proceso de innovación pública en los siguientes pasos:

1. Observación y comprensión de la situación a modo de diagnóstico;
2. Ideación de posibles soluciones;
3. Prototipado y experimentación de la solución (o desarrollo e implementación en su fase piloto);
4. Validación, ajuste e iteración, como mecanismo para capitalizar el aprendizaje en el proceso de puesta a punto y/o mejora de la solución². En esta instancia, la documentación del proceso y su progresivo versionado son imprescindibles;
5. Difusión y escalamiento; complementado con formación y transferencia a las personas que utilizarán la herramienta;
6. Cambio sistémico.

1 Para profundizar sobre innovación social, innovación pública o innovación abierta, entre otros, ver Zurbriggen, C., Lago, M. G. (2015). Co-creando valor público. Desafíos pendientes para América Latina. Revista Iberoamericana de Ciencia, tecnología y Sociedad-CTS, 10(30), 143-171.

2 Esta misma lógica también se instrumenta en procesos de actualización y mejora como mecanismo de sustentabilidad o mantenimiento de la solución una vez implementada.

Figura 1. Una secuencia del pensamiento de diseño para la innovación pública



Fuente: Elaboración propia.

La figura 1 nos ayuda a fijar la secuencia del pensamiento de diseño. Esta secuencia nos da la posibilidad de desandar parte del camino o revisar nuestros propios pasos ante nueva información disponible o nuevas interacciones, de forma iterativa. También nos permite adoptar los mismos pasos para hacer un balance periódico de la solución generada a fin de corregir errores, mejorar procesos o incluir herramientas complementarias. Para poder hacer ese balance de forma consistente, necesitamos recordar los principales criterios que adoptamos durante el proceso: ¿por qué tomamos tal o cual decisión? ¿Por qué cambiamos aquello que no funcionaba bien en el prototipo inicial? ¿Hubo alguna restricción tecnológica o institucional que nos hizo reconsiderar otras alternativas? ¿Qué salió distinto de lo planificado? Para recordar, es recomendable documentar. Documentar el proceso también resulta útil para trascender a los equipos y garantizar transparencia.

Las soluciones innovadoras transforman la capacidad de respuesta de un grupo o de la sociedad en su conjunto: a veces, por las nuevas colaboraciones que se dan durante el proceso; otras, por las nuevas formas de hacer que propone la herramienta; o porque a la larga las herramientas terminan incidiendo sobre la forma de pensar o sobre ciertas prácticas institucionales. En otras palabras, si esa nueva idea se transforma en un modelo, una norma, un movimiento o una intervención para cubrir ciertas necesidades sociales, mejorando alternativas preexistentes o generando algo nuevo, estaríamos ante una transformación social: es aquí cuando la innovación genera valor público.

Además, la creación de valor se potencia cuando las personas involucradas (en la gestión de la herramienta, quienes van a usarla, quienes pueden estar afectados por sus resultados, etc.) participan en el proceso de innovación pública desde su inicio. Siguiendo la figura 1, es importante que las personas involucradas participen en el **diagnóstico**, en la **ideación** y en la **validación** de la solución innovadora. La posibilidad de integrar esta diversidad y multiplicidad de miradas y necesidades en el diseño, permite mejorar la experiencia de las personas usuarias y, por lo tanto, brindar una mejor usabilidad a la herramienta. A su vez, la participación de las personas involucradas en las instancias intermedias, allana la fase de difusión e, incluso, permite encontrar una forma más amena de compartir y comunicar el propósito de la herramienta innovadora en la política o el servicio público.

El abordaje que proponemos desde el pensamiento de diseño se diferencia de aquellos cuyo equipo técnico idea y desarrolla una solución de forma aislada de su contexto de uso y aplicación, dado que en ese camino se pueden soslayar u omitir detalles relevantes sobre la cultura institucional y social, los recursos tecnológicos, los circuitos administrativos, etc. Y de esa forma, el destino de la herramienta innovadora puede tener una utilidad limitada o, directamente, nunca ser usada, desestimando los esfuerzos y los recursos que el sector público y la sociedad toda han destinado a su creación.

Esta mención no invalida que el desarrollo del **prototipo** y la implementación de las primeras versiones de la herramienta (de prueba o piloto³), requiere de conocimientos técnicos específicos. Estos conocimientos también son imprescindibles para evaluar la adquisición de una herramienta cerrada y, en caso necesario, su adaptación. La otra instancia que demanda requerimientos técnicos específicos es la de **validación** y de **monitoreo**. El punto es que al llevar adelante el proceso de innovación pública de forma **colaborativa** con las personas involucradas, logramos converger en un proceso de **co-creación**. En este caso, dado que nos abocamos al sector público y, progresivamente vamos a profundizar sobre las herramientas de CDIA, para dar lugar a la **co-creación de herramientas de CDIA para la innovación pública**.

Es interesante destacar que la co-creación nos invita a abrazar la diversidad y la interdisciplinariedad para darle mayor robustez y aplicación a las herramientas innovadoras que decidamos desarrollar. ¿Por qué? Porque no sólo incluye especificidades propias de la política pública y cierto abordaje de la democracia participativa, sino que también incorpora roles claves del análisis funcional para codificar requerimientos propios del diseño de sistemas de información, software y aplicaciones basadas en datos.

El proceso de innovación pública puede observarse en dos planos: uno más general, a partir de la **implementación** de una política o un servicio público, y uno más específico, a partir de herramientas que permiten mejorar la gestión y/o los resultados de esa política o servicio. En este último caso se encuentran las herramientas de CDIA en las que se centra este documento. Estas herramientas constituyen (o pueden constituir) un pilar fundamental de información u optimización para el diseño, la gestión y/o el monitoreo de políticas y servicios públicos. De hecho, podemos pensar los procesos de innovación como capas de una cebolla, en las capas más externas encontramos la política o el servicio público y en las internas, las herramientas y los procesos con los que se lleva adelante. Es aquí donde podemos situar las herramientas de CDIA. Sin embargo, hay un diálogo entre ambas capas o procesos. En la tercera etapa (sección 2.3), abordamos los usos de las herramientas de CDIA como soporte para la toma de decisiones en materia de política y servicios públicos. Allí planteamos un recorrido completo para la innovación pública a partir de casos de aplicación de CDIA.

Hoy en día, la información disponible puede orientar la toma de decisiones en pos de generar respuestas más comprensivas a las problemáticas sociales. También es posible generar herramientas sistémicas para monitorear y ajustar esos procesos de toma de decisiones. La pregunta que emerge, entonces, es ¿cómo generar innovación social y mejorar la capacidad de respuesta del Sector Público a partir de herramientas analíticas cada vez más asertivas? Una respuesta posible es **incorporar herramientas de CDIA en los procesos de toma de decisión**. Vayamos a la próxima etapa para abordar estos conceptos.

³ En la jerga tecnológica, esto se refiere a demostrador tecnológico o producto mínimo viable (MVP por sus siglas en inglés).

Segunda etapa

¿Qué entendemos por datos? ¿Y por Ciencia de Datos o Inteligencia Artificial?

¿De qué hablamos cuando nos referimos a datos? ¿Qué es la Ciencia de Datos (CD)? ¿Y la Inteligencia Artificial (IA)? Sí, el nombre de esta etapa es en honor a esas preguntas frecuentes. Veamos las respuestas poco frecuentes o, al menos, aquellas sobre el contexto, la implementación y las implicancias del uso de CDIA que muchas veces quedan omitidas por quienes instrumentan o desarrollan herramientas de CDIA.

Muchas veces cuando idealizamos conceptos o tecnologías, no terminamos de entender bien cuál es su alcance real, qué pueden y qué no pueden hacer, y cómo se diferencian entre sí. Aunque también necesitamos comprender desde dónde hablan quienes desarrollan soluciones de CDIA para poder formular requerimientos y acompañar el proceso. Es por eso que decidimos pararnos en esta hoja de ruta desde el punto de vista y las potencialidades de la Ciencia de Datos, que es un concepto que engloba a todas aquellas prácticas que podemos explorar cuando queremos usar datos en el sector público: desde la planificación de su recolección, hasta su visualización y comunicación oficial.

Datos

El Banco de Desarrollo de América Latina reúne varios escritos para definir a los datos como “cifras o hechos sencillos, discretos y objetivos que representan eventos que ocurren, se estructuran, capturan, cuantifican y transfieren con facilidad” (Davenport y Prusak, 2000). Estos hechos, de acuerdo con el contexto, son presentados de tal forma que puedan ser comprendidos, interpretados y comunicados por un ser humano o procesados por una máquina para servir de antecedente en la obtención de una conclusión (Guillén et al., 2015; ODLIS, 2020). Aunque los datos describen sólo una parte de lo que sucede (...), son un recurso de gran relevancia que, cuando se combinan con tecnologías digitales, generan oportunidades para promover cambios sociales, políticos y económicos (Nugroho et al., 2015)”⁴. Para comprender mejor el uso de datos en el sector público, te recomendamos estudiar el caso que presentamos más adelante en el recuadro 3: “¿Qué tipo de información podemos conseguir para tomar decisiones?” donde se explica cómo la información disponible sobre el COVID 19 se transforma en datos para comprender mejor cuándo debía aislarse a la población, la tasa de mortalidad y la edad de los individuos, entre otras cosas.

¿Qué es la Ciencia de Datos o Data Science?

La Ciencia de Datos (CD) puede pensarse como un campo interdisciplinario que hace uso de diferentes tecnologías y métodos para procesar y analizar los datos para identificar patrones y tendencias, así como predecir fenómenos futuros.

4 CAF (2021). Experiencia: Datos e Inteligencia Artificial en el sector público. Caracas: CAF. Retrieved from <http://scioteca.caf.com/handle/123456789/1793>

Este campo interdisciplinario está conformado por disciplinas como la lógica formal, la matemática y la estadística que permiten construir una secuencia de pensamiento lógico matemático y brindan cierto nivel de abstracción a través del modelado de problemas, métodos de inferencia y esquemas de validación.

Por otro lado, la ciencia de datos tiene un gran vínculo con las ciencias de la computación, la ingeniería en informática y los sistemas de información, que aportan conocimientos en torno a los sistemas y las tecnologías de la información, la resolución de problemas vinculados a la automatización y el desarrollo de algoritmos.

En la intersección de estas dos dimensiones (matemática-computación), a partir de la integración de sus prácticas y conocimientos se establecen los campos más novedosos de inteligencia artificial, aprendizaje automático supervisado (supervised machine learning) y aprendizaje profundo (deep learning), de los cuales hablaremos en profundidad más adelante. Los desarrollos provistos por este campo son algoritmos, modelos y/o procesos de decisión⁵ optimizados con un sistema de métricas o indicadores que validan el rendimiento de cada modelo frente a un problema planteado en un dominio del conocimiento específico.

Por último, es fundamental vincular lo técnico, computacional, matemático y estadístico con un área del conocimiento determinada y/o las experiencias en un dominio específico ya que propician una acertada contextualización para el uso de las herramientas aplicadas que introducimos previamente. ¿Por qué? Porque no concebimos pensar las herramientas de CDIA separadas del contexto o del dominio en el que se desarrollan. Es decir que es necesario pensarlas desde un punto de vista socio-tecnológico.

¿Qué es la Inteligencia Artificial?

El término Inteligencia Artificial (IA) es amplio e impreciso. Podríamos decir que no hay una única definición, pero para nuestro caso nos sirve pensar en la IA como sistemas que pueden, para un determinado conjunto de objetivos definidos por personas, generar predicciones, recomendaciones o decisiones que pueden influenciar ambientes reales o virtuales. Estos sistemas están diseñados para operar con diversos grados de autonomía. La IA puede incluir una amplia gama de métodos y herramientas, como el aprendizaje automático, el reconocimiento facial y el procesamiento del lenguaje natural.

La IA debe entenderse como un campo de investigación y aplicación que incluye más que enfoques técnicos. Este área debe pensarse en su diseño desde la interdisciplina y se construye desde la lógica, la estadística, la filosofía, las ciencias de la computación, las matemáticas, neurociencias, la lingüística, la psicología, la economía y otras ramas de estudios sociales. Vincula lo técnico, lo científico, las prácticas sociales, gubernamentales e industriales.⁶

5 Para más información sobre estos tres conceptos, se puede consultar la página 23 de CAF (2021). Experiencia: Datos e Inteligencia Artificial en el sector público. Caracas: CAF. <http://scioteca.caf.com/handle/123456789/1793>

6 Véase: <https://thepublicvoice.org/ai-universal-guidelines/>
<https://en.unesco.org/artificial-intelligence>

<https://oecd.ai/en/ai-principles>

<https://www.gov.uk/government/collections/a-guide-to-using-artificial-intelligence-in-the-public-sector>

Algoritmos

Los algoritmos pueden pensarse como una serie de pasos o estructuras que pueden transformar información de entrada en un producto o información de salida. Un sistema algorítmico es un sistema que usa uno o más algoritmos, en general como parte de un software para producir resultados que pueden ser utilizados para la toma de decisiones.

Utilizamos una definición funcional de un sistema algorítmico, como un sistema que utiliza el razonamiento automatizado para ayudar o reemplazar un proceso de toma de decisiones que de otro modo sería realizado por los seres humanos. Es importante señalar que todos los sistemas algorítmicos incluyen diferentes tipos de intervención humana, ya sea en la fase de diseño o en la forma en que se utilizan finalmente. En nuestro análisis, tenemos en cuenta tanto los contextos técnicos como los sociales, culturales, jurídicos e institucionales en los que se insertan los algoritmos, como determinantes importantes de la forma en que se utilizan y gobiernan estos sistemas.⁷

Modelos de Aprendizaje Automático

Cuando hablamos de modelos de aprendizaje automático (ML, por *Machine Learning*) nos referimos a un campo más comúnmente asociado con la actual explosión de la IA. El aprendizaje automático es un conjunto de técnicas y algoritmos que pueden utilizarse para “entrenar” a un programa informático de modo que reconozca automáticamente patrones en un conjunto de datos. Muchas herramientas diferentes caen bajo el paraguas del “aprendizaje automático”. Aunque hay excepciones, el ML suele utilizar “características” o “variables” (por ejemplo, la ubicación de las estaciones de bomberos en una ciudad, datos de cámaras de vigilancia, atributos de las personas acusadas de delitos) tomados de un conjunto de “datos de entrenamiento” para que el modelo de ML aprenda estos patrones sin que nadie les diga explícitamente cuáles son esos patrones. El aprendizaje automático también incluye modelos que históricamente se han llamado simplemente “estadísticas” tales como las regresiones lineales, logísticas.⁸

Big Data o Macrodatos

La big data o los macrodatos son grandes conjuntos de datos que se asocian a características como volumen, variedad, velocidad, veracidad, entre otras, estas conforman las dimensiones de los datos. En general los gobiernos cuentan con grandes volúmenes de datos, que según el manual de Experiencia⁹, requieren nuevas formas de procesamiento para permitir una mejor toma de decisiones, la generación de conocimientos y la optimización de procesos.

7 Ada Lovelace Institute, AI Now Institute and Open Government Partnership. (2021). Algorithmic Accountability for the Public Sector. Available at: <https://www.opengovpartnership.org/documents/algorithmic-accountability-public-sector/>

8 Traducción propia del reporte de Ada Lovelace Institute, AI Now Institute and Open Government Partnership. (2021). Algorithmic Accountability for the Public Sector. Available at: <https://www.opengovpartnership.org/documents/algorithmic-accountability-public-sector/>

9 CAF (2021). Experiencia: Datos e Inteligencia Artificial en el sector público. Caracas: CAF. Retrieved from <http://scioteca.caf.com/handle/123456789/1793>

En estos últimos años los avances en nuevas tecnologías han permitido que los gobiernos tengan a disposición grandes volúmenes de datos digitalizados. Con la llegada de estos datos, también surgen interrogantes sobre su disponibilización, utilización, el almacenamiento en servidores, la seguridad y la responsabilidad a la hora de utilizarlos. Con el auge de los macrodatos se avanzó en la utilización de estas grandes bases de datos para identificar patrones que antes no podían ser analizados por la falta de tecnologías disponibles. Hoy en día con el nivel de procesamiento de las computadoras actuales y la posibilidad de “alquilar” espacio en la nube¹⁰, el trabajo estadístico y de inteligencia artificial con *big data* en el sector público son herramientas de alcance gubernamental.

Creemos que estas herramientas innovadoras pueden ayudar a mejorar procesos existentes, automatizar otros y ayudarnos a comprender mejor las problemáticas de nuestra región (recuadro 1).

RECUADRO 1

¿Querés saber qué está sucediendo en materia de datos, transparencia y ética a nivel gubernamental en América Latina?

Con lo que respecta a los datos y su nivel de apertura, el Barómetro Global de Datos¹¹ es de gran relevancia, ya que mide la prevalencia y el impacto de las iniciativas de datos abiertos en todo el mundo, analizando sus tendencias mundiales. Esta iniciativa proporciona datos comparativos sobre países y regiones para clasificar a los gobiernos en cuanto a su preparación, implementación e impacto de los datos abiertos.

Para saber más sobre el uso de CDIA por parte de gobiernos en América Latina, existen diversos documentos e investigaciones que analizan el uso de estas herramientas en la región y sobre los cuales nos hemos apoyado para elaborar esta hoja de ruta.

La Web Foundation en su documento “Algoritmos e Inteligencia Artificial en América Latina”¹² describe los casos de Argentina y Uruguay y las diferentes aplicaciones de IA en el sector público de ambos países. El manual del Banco de Desarrollo de América Latina junto a otros organismos, titulado “Experiencia: datos e Inteligencia Artificial en el sector público”¹³ aborda de manera exhaustiva todas las diferentes dimensiones del uso de datos y de IA por parte de gobiernos latinoamericanos, en particular en el apartado “Adopción de la Inteligencia Artificial en América Latina”.



10 Con esto nos referimos a alquilar capacidad de almacenamiento o de procesamiento a ciertas compañías que proveen estos servicios.

11 Véase <https://globaldatabarometer.org/>

12 Ortiz Freuler, J. and Iglesias, C. (2018). Algorithms e Inteligencia Artificial en Latin America: Un Estudio de implementaciones por parte de Gobiernos en Argentina y Uruguay, World Wide Web Foundation. En: https://webfoundation.org/docs/2018/09/WF_AI-in-LA_Report_Spanish_Screen_AW.pdf

13 CAF (2021). Experiencia: Datos e Inteligencia Artificial en el sector público. Caracas: CAF. Retrieved from <http://scioteca.caf.com/handle/123456789/1793>

A su vez, desde las recomendaciones éticas del uso de estas herramientas existen documentos internacionales tales como las recomendaciones de UNESCO¹⁴, OECD¹⁵ y la regulación europea de datos (GDPR)¹⁶, entre otros. La ética y la responsabilidad son fundamentales a la hora de trabajar con tecnologías y herramientas que van a afectar de alguna manera u otra a la sociedad.

Estas herramientas generan de manera más o menos autónoma información que usamos para tomar decisiones, o incluso en el caso de las herramientas de IA ellas mismas pueden tomar decisiones que afectarán a la sociedad. Indefectiblemente tenemos que considerar los aspectos éticos y de derechos humanos vinculados al trabajo con tecnologías en el sector público.

Si nos enfocamos en los documentos éticos y responsables de la región, podemos consultar los manuales del BID de la serie “Uso responsable de IA para política pública”¹⁷, en particular el Manual de proyectos y el Manual de Ciencia de Datos. Recomendamos la “Caja de herramientas humanísticas”¹⁸ desarrollada por el Grupo GIFT de Filosofía, este documento surge en el marco de GuIA, un espacio para el desarrollo de documentos que abordan la ética de la IA, su gobernanza y aplicaciones en América Latina y el Caribe.

14 Véase <https://en.unesco.org/artificial-intelligence>

15 Véase <https://oecd.ai/en/ai-principles>

16 Véase https://ec.europa.eu/info/law/law-topic/data-protection_es

17 BID (2021). Uso responsable de IA para política pública: Manual de formulación de proyectos. En <https://publications.iadb.org/publications/spanish/document/Uso-responsable-de-IA-para-politica-publica-manual-de-formulacion-de-proyectos.pdf>

18 Pedace, K., Balmaceda, T., Lawler, D., Pérez, D., Zeller, M. (2019) Caja de herramientas humanísticas. Taller-Inteligencia Artificial en América Latina y el Caribe: ética, gobernanza y aplicaciones Lugar: Montevideo.

Tercera etapa

El uso de datos, Ciencia de Datos e Inteligencia Artificial como un proceso de innovación pública

En esta sección les proponemos integrar el recorrido que nos trajo hasta aquí. Ya está más claro el panorama, ¿no? Ahora sabemos de qué hablamos cuando hablamos de datos o de una herramienta de CDIA vinculándola con los procesos de innovación pública y de toma de decisiones.

El hecho de instrumentar una herramienta innovadora nos invita, primero, a estructurar y planificar el proyecto que le dará origen dentro de una dinámica institucional y/o programática específica. Desde la perspectiva del pensamiento de diseño, notamos que la identificación del problema/necesidad o las opciones de mejora que se encuentran en el diagnóstico pueden nutrir el diseño y el propósito de la herramienta que ideemos, prototipemos, mejoremos y, posteriormente, implementemos. De esta forma, la herramienta (el **qué y el cómo**) es una alternativa para responder al propósito (la necesidad o el problema a solucionar, el **para qué**).

También será necesario especificar resultados e impactos esperados tras la implementación de la herramienta de CDIA. Esta especificación en la fase temprana del proyecto permitirá monitorear la evolución del mismo y evaluar resultados e impactos efectivos después. Plasmar la **conexión entre el inicio y el final del proyecto** (o del proceso) es crucial para entender y comunicar la visión general en la que se vincula el uso de la herramienta innovadora y los distintos pasos que se detallan a lo largo del proyecto. En otras palabras, la **comunicación transversal** y el **registro documental** del proyecto son prioritarios para vincular las premisas y los objetivos iniciales, registrar los aprendizajes o ajustes producto del conocimiento fáctico, para dar cuenta de los avances a través de hitos verificables y evaluar asertivamente los resultados y los impactos esperados¹⁹.

Para monitorear y/o evaluar el proyecto vinculado a la herramienta de CDIA, será deseable generar datos verificables considerando el ciclo de vida del proyecto. A tal fin, desde la fase de ideación o diseño, los **indicadores** (recuadro 2) pueden constituir un termostato sobre la evolución y el impacto del proyecto. Hay indicadores orientados a captar:

Los **componentes** y las **actividades** implicadas a lo largo de la ejecución del proyecto; vinculando tanto las fases de preparación y planificación como las de implementación y ejecución.

19 Tal como las fases de diseño, planificación, monitoreo y evaluación propias de programas y políticas, los proyectos para la implementación de soluciones innovadoras también pueden contar con ellas. Para profundizar en cada una de ellas, ver: Ortigón, E., Pacheco, J. F., y Prieto, A. (2005). Metodología del marco lógico para la planificación, el seguimiento y la evaluación de proyectos y programas. CEPAL.

La **finalidad** y el **propósito**, orientados a capturar las instancias previas a la implementación del proyecto y a los resultados posteriores a su implementación.

RECUADRO 2

Criterios para la definición de indicadores²⁰

- El sentido de un indicador es claro.
- Existe información disponible o se puede recolectar fácilmente.
- Un indicador es tangible y se puede observar.
- La tarea de recolectar datos está al alcance de la dirección del proyecto y no requiere expertos para su análisis.
- Un indicador es lo bastante representativo para el conjunto de resultados esperados.
- Un criterio adicional que debe ser evaluado es que los indicadores sean independientes. A su vez, es esperable no asignar una relación de causa-efecto entre el indicador y el objetivo que se evalúa.

Hasta aquí, dimos cuenta de conceptos y abordajes para instrumentar herramientas innovadoras a partir de CDIA en el sector público. A continuación, nos enfocamos en las posibles aplicaciones de distintas herramientas basadas en datos. El final de la historia (o, de este apartado), efectivamente, será la implementación de una herramienta compleja de CDIA.

En algunos casos, este tipo de herramientas puede estar constituido por un proceso secuencial de creciente complejidad, como el caso de las Historias Clínicas Electrónicas (HCE) y Gestión epidemiológica basada en Inteligencia Artificial y Ciencia de Datos (ARPHAI) que planteamos de forma articulada en el [recuadro 6](#). En otros casos, puede estar constituido por un conjunto de modelos y predicciones que brindan resultados de forma conjunta, tal como el Sistema de Alerta Temprana del Servicio Meteorológico que presentamos en la [tabla 4](#).

Para no abrumarnos, vamos a fragmentar el recorrido completo en **4 tramos posibles**. Cada tramo, se va a enfocar en una herramienta de CDIA concreta y, por lo tanto, permitirá visibilizar cuáles son los puntos de partida y/o los avances graduales en materia de construcción de capacidades de CDIA. Es un esquema que integra los principales ejes de **un proceso de innovación pública basado en herramientas de CDIA**, a continuación recorreremos cada uno de estos 4 tramos ([figura 2](#)):

Podemos **innovar con datos**, a partir de sistematizar datos existentes o captar nuevos, ya sea desde un aplicativo o un relevamiento. Denominamos **tramo 1** a esta instancia.

20 Ortegón et al, pág. 85. Para profundizar sobre la definición de indicadores referidos a la eficiencia de políticas públicas, ver: Hinzte, J. (2012). La evaluación de la efectividad de las organizaciones públicas (y la construcción de sus indicadores). Buen Gobierno, (12), 58-70.

Podemos **innovar con herramientas de ciencia de datos**, a partir del análisis de datos, la visualización o el monitoreo. En este último caso, el monitoreo, es importante definir previamente cuáles serán los indicadores o las variables sujetas a ser observadas. Denominamos **tramo 2** a esta instancia.

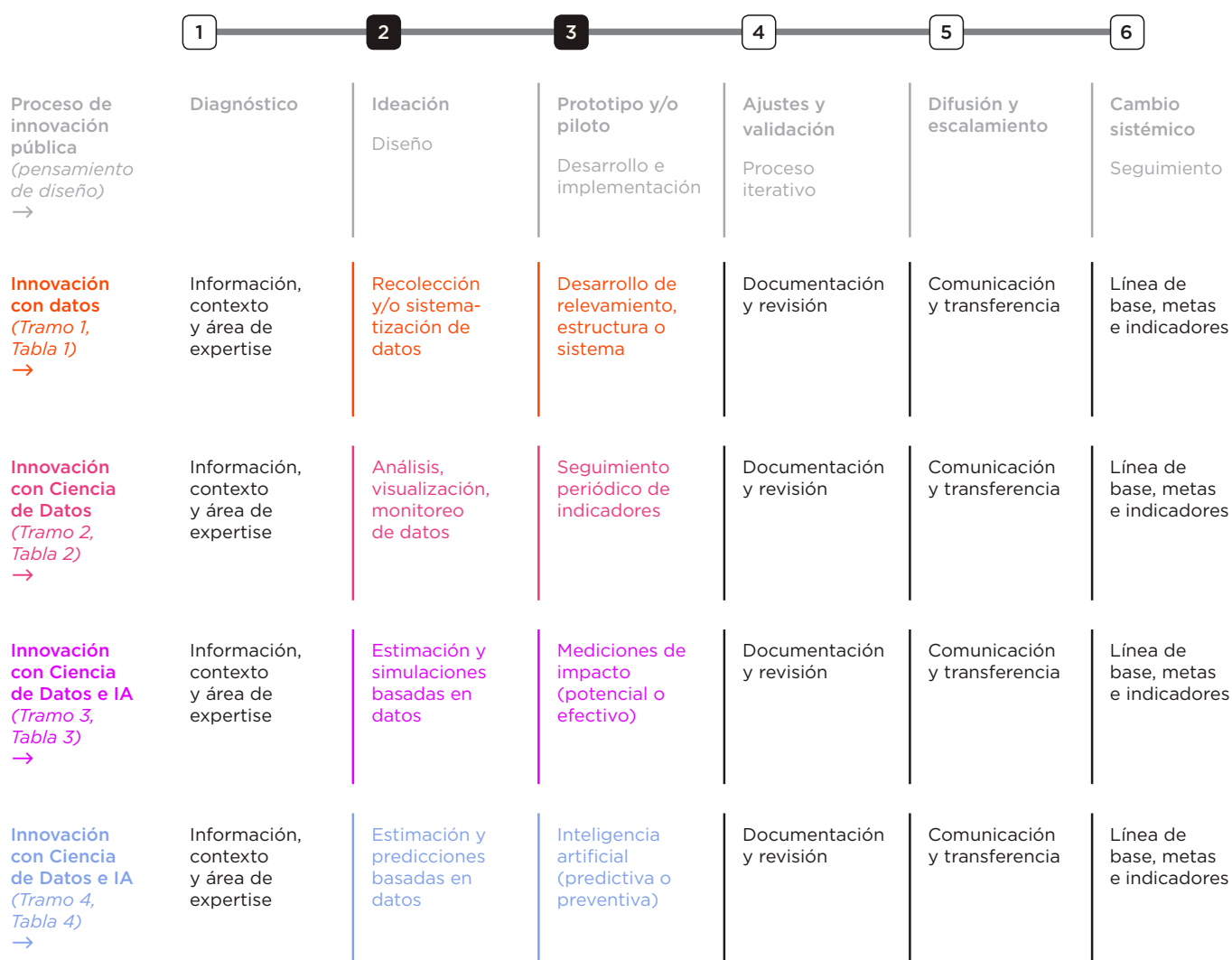
También podemos **innovar con herramientas de ciencia de datos e inteligencia artificial**. Este eje, por un lado, integra o acumula los conceptos de los dos ejes previos y, por otro, presenta **dos tramos, que pueden ser alternativos o secuencialmente complementarios**:

Estimar indicadores y/o simular impactos a partir de datos ya disponibles y la construcción de escenarios posibles. A fines expositivos, denominamos **tramo 3** a esta instancia.

Estimar y predecir patrones a partir de información ya observada. A fines expositivos, denominamos **tramo 4** a esta instancia.

Es importante notar que el tramo 1 y el tramo 2, en ese orden, representan procesos fundantes del tramo 3 y el tramo 4, se puede transitar el tramo 4 sin necesidad de haber pasado por el tramo 3. Asimismo, es posible, implementar el tramo 4 de forma previa a implementar el tramo 3, conforme los requerimientos de la solución propuesta.

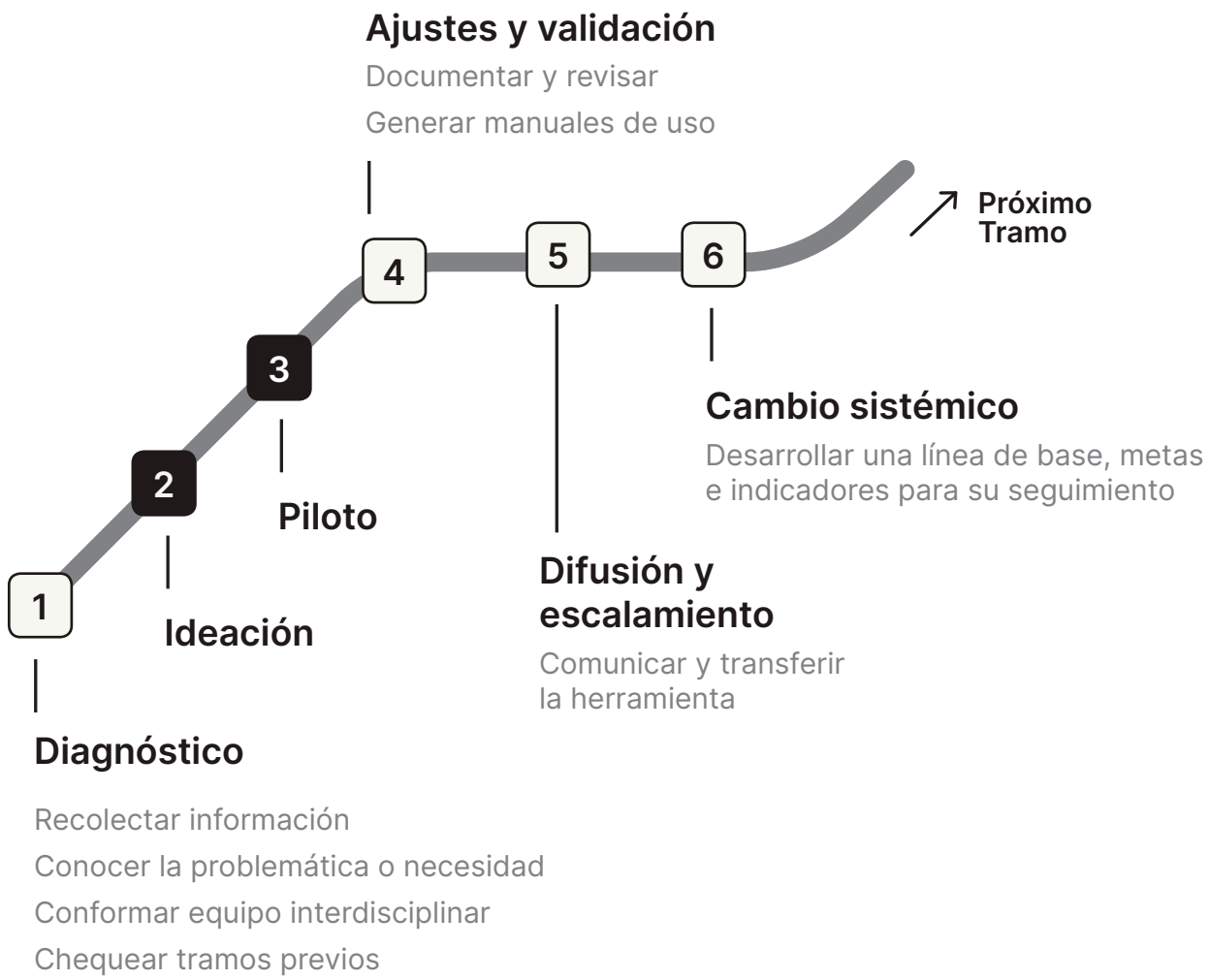
Figura 2. Uso de CDIA en servicios y políticas públicas basados en datos



Fuente: Elaboración propia. || Nota: (1) + (2) son pasos necesarios para (3) y (4).

Proceso de Innovación Pública

¿Cómo encontrar soluciones innovadoras en el sector público?
En este mismo contexto, ¿cómo integrar el mundo lógico o analítico con el creativo, inductivo o experimental? Proponemos como metodología para la innovación de servicios y políticas públicas el pensamiento de diseño, porque nos da la posibilidad de desandar parte del camino o revisar nuestros propios pasos ante nueva información.

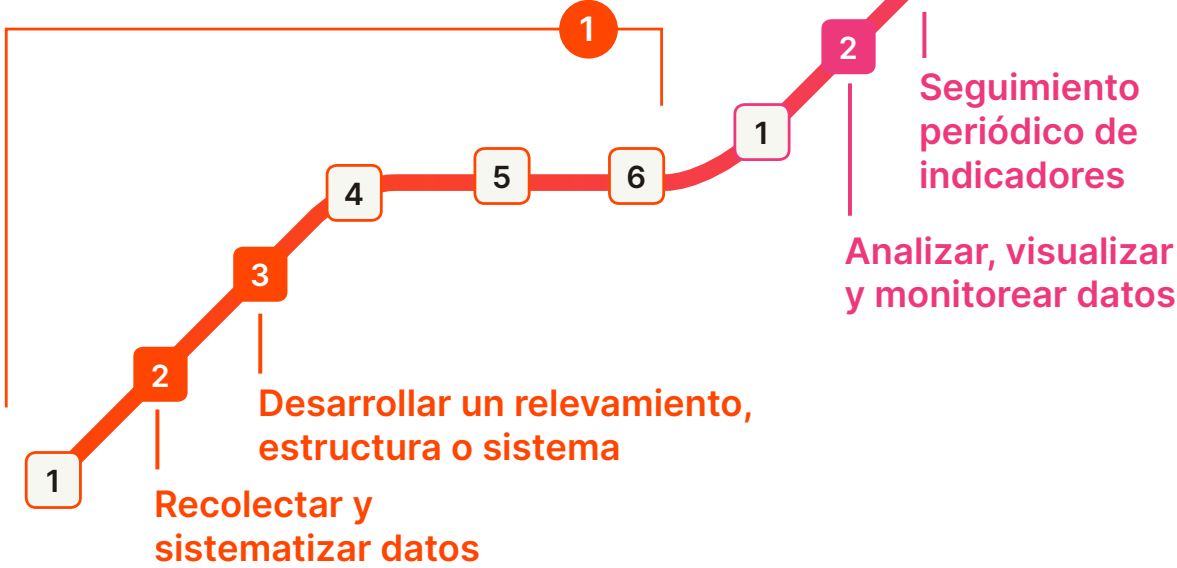


¿Cómo integramos los datos o las herramientas de Ciencia de Datos e Inteligencia Artificial a los procesos de innovación pública?

- Estructuramos y planificamos el proyecto. La herramienta (el qué y el cómo) es una alternativa para responder al propósito (la necesidad o el problema a solucionar, el para qué).
- Especificamos resultados e impactos esperados. Nos permitirá monitorear la evolución y evaluar los resultados e impactos efectivos después.
- Plasmamos la conexión entre el inicio y el final del proyecto. Comunicación transversal y registro documental.
- Generamos datos verificables.
- Planteamos metas e indicadores.

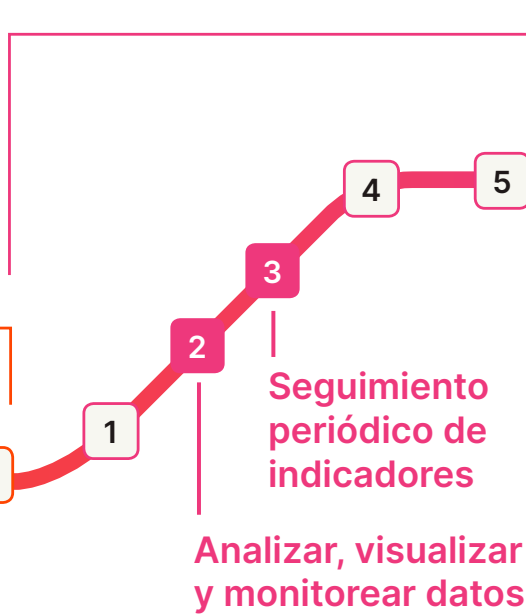
Innovar con datos

Tramo 1



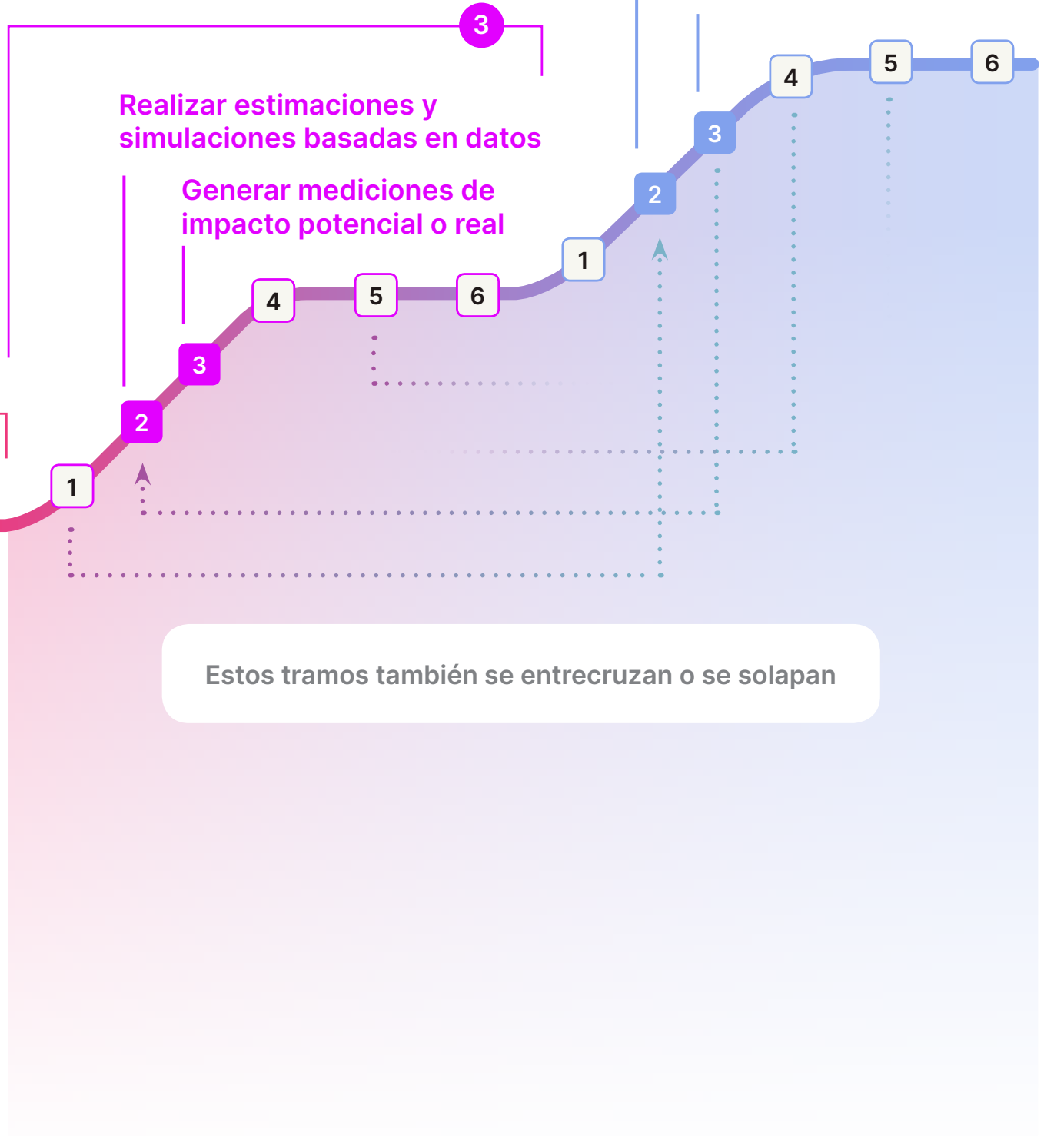
Innovar con Ciencia de Datos

Tramo 2



Innovar con Ciencia de Datos e Inteligencia Artificial

Tramos 3 y 4



Para evaluar qué tipo de herramienta es más idónea para abordar un problema o cubrir una necesidad, es recomendable adoptar el pensamiento de diseño y la gestión de la información. La [figura 2](#) nos da una pauta de la magnitud del desafío de instrumentar las herramientas más complejas de CDIA; este varía según la experticia de las personas y las instituciones que inician el recorrido. Por ejemplo, quienes cuenten con información de base, podrán comenzar el recorrido directamente en el tramo 2, en lugar de hacerlo desde el tramo 1 (ver recuadros 3 y 4).

RECUADRO 3

¿Qué tipo de información podemos conseguir para tomar decisiones?

En algunos casos esta información estará disponible para ser recopilada, pero en otros casos tendremos que recolectar nuevos datos. También podemos encontrarnos con una situación en donde haya datos pero no podamos o no corresponda usarlos.

Viajemos en el tiempo: recordemos los primeros meses de la Pandemia del COVID-19, cuando sólo había información sobre la rápida tasa de contagios. Una de las necesidades era conocer de manera urgente de qué se trataba el virus y cómo afectaba a la población. Tener más información permitiría mejorar las políticas ante la emergencia de salud pública.

En línea con ello, se instrumentaron mecanismos estandarizados de reporte que podrían haberse iniciado como simples registros de atención médica (más adelante volveremos sobre este ejemplo). Esta información unificada permitió dar cuenta de la mayor tasa de mortalidad en función de la edad y otras condiciones de riesgo, a medida que globalmente también se avanzaba en generar conocimiento sobre el virus. En función del **seguimiento y la evolución de estos indicadores**²¹, se endurecía o se flexibilizaba el aislamiento obligatorio. A su vez, estos datos permitieron monitorear y proyectar la necesidad creciente de testeo durante los principales brotes y flexibilizar o endurecer la política de salud según el caso. Asimismo, se triangulaba esta información con la capacidad de respuesta en las unidades de terapias intensivas (UTI), para dar cuenta de una mirada más comprehensiva del sector salud²².

21 En este caso, la referencia a indicadores difiere de la previa sobre el ciclo de vida del proyecto que da origen a la herramienta innovadora de CDIA.

22 La visualización de datos está disponible en <https://www.argentina.gob.ar/salud/coronavirus-COVID-19/sala-situacion>.

RECUADRO 4

¿Qué política o servicio público se puede implementar si necesitamos generar cierto impacto en una situación determinada?

Continuemos pensando en el escenario de la pandemia. La necesidad de mitigar el impacto del coronavirus en la población es una prioridad. Para ello, se identifican dos dimensiones a trabajar: 1) reducir los contagios (ya sea por aislamiento o mejora en su identificación a través de testeos) y 2) brindar atención de salud adecuada. Dado que la misma es crítica en los cuadros más complejos, es necesario identificar riesgos; i) comorbilidades para atender complejidades mayores y ii) disponibilidad en las unidades de terapia intensiva (UTI) frente a crecimiento exponencial de casos.

Dada la inexistencia de vacunas en las primeras fases de la pandemia, el aislamiento surgió como una alternativa para mitigar el contagio y, por lo tanto, reducir el riesgo de contraer el virus. Asimismo, esta medida también permitía mitigar posibles colapsos del sistema de salud (brindar atención tan asertiva como posible) a medida que se monitoreaba la información disponible y se generaban proyecciones sobre escenarios factibles. En ese marco, se generaban nuevas opciones (o ajustes) de política, contrastando con escenarios alternativos frente a la inacción o ausencia de una medida, entre otras. A su vez, la identificación de comorbilidades luego sirvió para priorizar necesidades y gestionar el esquema de vacunación.

Los 4 tramos, de forma progresiva, dan una idea de la complejidad que ganamos a medida que recorremos un nuevo tramo. Sin embargo, no siempre se verifica esa linealidad en la realidad. Estos tramos también se entrecruzan o se solapan. A continuación detallamos los pasos del proceso de innovación pública comunes a los 4 tramos; posteriormente, profundizaremos sobre los pasos distintivos de cada tramo posible bajo la denominación de innovadora que lo identifica:

1. Diagnóstico: **información, contexto y área de expertise,**
2. Ideación: este paso será detallado a continuación según la complejidad de cada tramo,
3. Prototipo y prueba (piloto): este paso será detallado a continuación según la complejidad de cada tramo,
4. Ajuste y validación: **documentación y revisión,**
5. Difusión y escalamiento: **comunicación y transferencia,**
6. Cambio sistémico: **línea de base, metas e indicadores**²³.

23 En este caso, nos referimos a los indicadores o resultados que habilitará la herramienta de CDIA una vez implementada. En el tramo 2 de nuestro recorrido, cuando se refiere a analizar y/o visualizar datos, variables o indicadores, se refiere a aquellos datos que serán monitoreados para dar cuenta de la política o del servicio público.



Tramo 1: Innovar con datos

Ya intuimos que los **datos** emergen como un pilar troncal de las herramientas de CDIA. En ellos se sustenta la toma de decisiones basadas en información. Además, los datos posibilitan la construcción de diagnósticos empíricos, líneas de base y escenarios alternativos. Cuanto mayor es la **completitud**, la **cobertura** y la **diversidad** de los datos con los que trabajamos, tendremos la posibilidad de delimitar mejor nuestro universo de análisis. Como se amplía la especificidad del conocimiento, ese mix de datos también nos permitirá incrementar la **calidad** y la **robustez** de los diagnósticos y de las aplicaciones que diseñemos, brindando una mayor **precisión** para dar respuestas o tomar decisiones basadas en información. Por lo tanto, como un primer paso será necesario construir los cimientos: tendremos que delimitar cuál será el espectro de datos con el que vamos a trabajar, garantizar el **acceso** (regular o esporádico, según nuestra necesidad), delinear los objetivos del trabajo y diseñar el mecanismo de abordaje.

Pero, ¿cómo nos familiarizamos con el uso de datos? Podemos empezar por perderle el miedo a las hojas de cálculo y las estructuras de datos. Podemos esquematizar nuestro proceso de trabajo como una secuencia de pasos y, en cada uno de ellos, preguntarnos qué tipo de información podemos obtener, si ese aporte contribuye a nuestro propósito y cómo extraemos esa información para construir una base de datos o, directamente, una fuente de información. Podemos relevar qué información vinculada a nuestro problema está disponible, si es relevante y si complementa otras fuentes.

Si aún predominan los procesos analógicos, puede ser útil esquematizarlos y derivar toda la información posible a una planilla de cálculo. Si los procesos en los que trabajamos ya están digitalizados o son nativos digitales contamos con una gran ventaja porque los datos ya están **digitalizados, estandarizados y/o integrados**. A mayor familiaridad con procesos digitales y datos sistematizados, menor será el tiempo y el esfuerzo invertido hasta el próximo paso.

Los sistemas informáticos cuyo diseño integra tanto la digitalización de procesos de gestión como la consecuente sistematización, recopilación y almacenamiento de la información, por su parte, constituyen claros ejemplos del tramo 1 que, en la [figura 2](#), remite a **Innovar con datos**. Estos sistemas promueven, en general, la unicidad y la trazabilidad de la información y de las transacciones que allí se realizan, regulando también su accesibilidad y velando por su seguridad. En la [tabla 1](#) presentamos varios ejemplos de sistemas de gestión con datos nombrados y estructurados que hoy se encuentran ampliamente difundidos en el sector público argentino, tales como el Sistema Integrado de Gestión Financiera (eSIDIF), el de Gestión de Expedientes Electrónicos (GDE) y la Historia Clínica Electrónica (HCE). Es interesante destacar que, si bien estos sistemas de gestión no fueron pensados con fines estrictamente analíticos, una vez que se consolida su uso, se constituyen como una fuente de información crítica para la formulación de desarrollos más complejos o ulteriores innovaciones. Al respecto, más adelante ([tabla 4](#), [recuadro 6](#)) veremos cómo la HCE constituye una base fundamental para el desarrollo de ARPHAI.

Tabla 1. Innovación con datos: Recolección y sistematización de datos

Sistemas de información como herramienta de
carga de datos y sistematización

VER DOCUMENTO ↘

De los ejemplos previos, se destaca lo siguiente: 1) la modularización y la instrumentación de metodologías ágiles en lo que respecta al desarrollo y la revisión o ampliación de la herramienta y 2) tanto en la fase de diseño como en la de uso o implementación, la lógica de pensamiento de diseño resulta crucial para encarar el propósito de toda comunicación, proceso de transformación, priorización de tareas y funciones y objeto de la formación específica.

Como resultante del proceso de recolección y/o recopilación y sistematización de información empírica, las bases y los conjuntos de datos sistematizados y disponibles publicados en los sitios web del Instituto Nacional de Estadísticas y Censos (INDEC)²⁴ o de Datos Argentina²⁵ también constituyen ejemplos virtuosos de este tramo. A fin de generar un primer acercamiento a los datos, es recomendable aproximarse a ellos a través de la **metadata**, el **diseño de registro** o el **diccionario de datos** (ver recuadro 5). En caso que no exista metadata unificada sobre los datos recopilados con los que trabajamos, construirla puede ser un paso propio de este primer tramo.

RECUADRO 5

Metadata, diseño de registro y diccionario de datos.

Como venimos diciendo, es importante dejar en claro cuál es el objetivo del relevamiento o la sistematización de los datos. Si es un relevamiento institucional y/o ha habido algún antecedente o fase piloto, también es interesante detallarlo. Es crítico especificar el período de recolección, su última actualización, el alcance de su cobertura y detalles del relevamiento, ya que también brinda una primera impresión espacio-temporal de los datos. Este esquema remite a la validez general del conjunto de datos en función de su propósito de origen y constituye la base de lo que se denomina **metadata**²⁶. También permite simplificar la decisión de uso por parte de personas externas al equipo que los crea.

En resumen, podemos concebir la metadata como el relato documentado de la innovación basada en datos desarrollada en el tramo 1. Podemos resumir el proceso documentando lo siguiente:



24 <https://www.indec.gob.ar>

25 <https://www.datos.gob.ar/>

26 Para detalles sobre datos abiertos, ver: <https://datosgobar.github.io/paquete-apertura-datos/perfil-metadatos/>.

- Fundamentos: ¿por qué? ¿antecedentes?
- Objetivo: ¿para qué?
- Alcance y cobertura: ¿período? ¿territorio? ¿población? ¿es la fase piloto?
- Tipo de datos: ¿qué es (un dataset, una nueva sistematización, un nuevo relevamiento)? Si se trata de un nuevo dataset o de un relevamiento, una buena práctica resultante es incluir, respectivamente, un diccionario de datos o el diseño de registro.

Por otro lado, se habla de **diccionario de datos**²⁷ por una analogía con los diccionarios impresos, en donde se indica la entrada, la referencia o el nombre de cada variable o característica y en su definición se detalla su descripción y/o sus rangos o categorías. En algunos casos, también se suma el tipo de dato que contiene el campo, por ejemplo, si es un número entero o una cadena de texto.

En los grandes relevamientos como censos o encuestas con marco muestral, se puede encontrar un **diseño de registro**²⁸ y estructura de las bases de datos asociado a dicho relevamiento. Aquí se da cuenta de las preguntas y las secuencias lógicas de respuestas al formulario así como de la limpieza, la estandarización y el preprocesamiento de los datos previas a su publicación. En este caso se pueden incluir algunas codificaciones y/o imputaciones.

27 Para profundizar al respecto, ver: Guía para la publicación de datos en formatos abiertos - Paquete de Apertura de Datos de la República Argentina (datosgobar.github.io). Aquí podés encontrar un ejemplo:

28 Ver el diseño de registro de la Encuesta Permanente de Hogares: (Microsoft Word - EPH_dise\361o_registro.docx) (indec.gob.ar)



Tramo 2: Innovar con Ciencia de Datos

Cuando ya contamos con información estructurada y sistematizada, será posible iniciar un análisis descriptivo o una visualización con distintos niveles de complejidad de la misma. Esta instancia es útil para identificar qué es medible en función del propósito que atienda la política o el servicio público. En función de eso, se construyen indicadores vinculados a dicha política o servicio, a fin de llevar a cabo el monitoreo, verificar resultados, evaluar impactos y/o generar reportes, etc.

Si, por ejemplo, además contamos con un diccionario de datos desde el inicio del proceso (tramo), tendremos un panorama general antes de trabajar con nuestro conjunto de datos; en caso negativo, tendremos que ir construyéndolo como parte del proceso inicial de trabajo. De hecho, este punto resulta indispensable para la construcción de bases de datos que permitan acceder a información sistematizada.

Tabla 2. Innovación con ciencia de datos: Análisis, visualización y monitoreo de datos

Sistematización de datos, análisis,
visualización y monitoreo

[VER DOCUMENTO](#) ↘

En la tabla 2 nos referimos a **Innovar con ciencia de datos** (tramo 2, [figura 2](#)) a partir de dos casos de sistematización, análisis, visualización y monitoreo de datos: Instituciones Abiertas e Inteligencia Artificial (IA2) y el monitoreo del COVID-19, casos registrados en la República Argentina.

Como podemos observar en los primeros dos tramos sobre los usos de CDIA para la innovación pública ([figura 2](#)), resulta crucial **sistematizar los datos** con los que vamos a trabajar. En línea con los [recuadros 3](#) y [4](#), cuando se trata de datos estructurados, se destaca la importancia de establecer criterios relevantes (variables o características) a incluir en el dataset a medida que se va construyendo o nutriendo de información.

Por otra parte, cuando partimos de un conjunto de datos no estructurados²⁹ (como se explica en la tabla 2 sobre la experiencia de IA2 en el Juzgado Penal, Contravencional y de Faltas N°10 de la Ciudad Autónoma de Buenos Aires), terminaremos estructurando los datos en función de nuestro objetivo o propósito de análisis. En ese proceso inicial de identificar qué es relevante o aporta información, iremos seleccionando datos y estructurando categorías, niveles de agregación, etc.

²⁹ Cuando hablamos de datos no estructurados nos referimos a aquellos que no están disponibles en tablas o sets de datos, y se encuentran en texto libre, por ejemplo en un pdf de una resolución judicial. En este caso, nos referimos a texto libre. En otros casos, puede referirse a datos mixtos o no clasificados (videos, imágenes, sonido, texto, números, etc.).

Además, el **monitoreo de los datos** en sí y de resultados de los modelos, nos permite mejorar, ajustar o reforzar acciones durante la implementación de la política o la prestación del servicio según las metas y los objetivos que se plasmaron en el diseño y la planificación³⁰.

30 En caso de interés en metodologías como la matriz de marco lógico, ver esta guía de la CEPAL.



Completando el recorrido: Innovar con Ciencia de Datos e Inteligencia Artificial

Tramo 3: Mediciones de impacto e Inteligencia Artificial

Sumando un nuevo tramo a nuestro recorrido, podemos **Innovar con CDIA** a partir de **estimaciones y simulaciones** (figura 2, tramo 3). El **análisis o la estimación de impacto** puede orientarse a mensurar impactos potenciales (ex-ante) o efectivos (ex-post) de una política, un programa o un servicio público. Este caso de aplicación de CDIA, permite estimar y/o simular el impacto de ciertas intervenciones. A partir del análisis de datos, la caracterización del entorno y los objetivos de política se delinean escenarios viables o reales, respectivamente. Luego se instrumentan, por ejemplo, microsimulaciones o modelos basados en agentes para simular o mensurar la incidencia de una intervención de política y analizar sus impactos.

Cuando se establece una línea de base o punto de partida previo a la intervención y se recolectan datos tras la intervención de la política, se habla de evaluación de impacto ya que se construye evidencia para analizar su desempeño: su foco reside en contrastar situaciones en dos puntos del tiempo, al inicio, con una línea de base, y con los valores efectivos a posteriori, tras la intervención. De esa forma, se contrastan los resultados esperados con la situación real, incluyendo los imprevistos, que de haberse presentado, hubieran afectado significativamente los resultados una vez transcurrido el tiempo previsto. En la tabla 3 nos referimos a **Innovar con ciencia de datos e inteligencia artificial para la medición de impactos** (el tercer tramo de la figura 2), se presentan los casos de análisis de impacto para la Asignación Universal por Hijo (AUH) y una Evaluación de Impacto Ambiental (EIA).

Tabla 3. Innovación con ciencia de datos e inteligencia artificial: Estimaciones y simulaciones basadas en datos

[Análisis y evaluación de impacto](#)

[VER DOCUMENTO](#) ↘



Completando el recorrido: Innovar con Ciencia de Datos e Inteligencia Artificial

Tramo 4: Estimaciones y predicciones con Inteligencia Artificial

Llegamos al último trayecto de nuestro recorrido, **Innovar con CDIA** a partir de **estimaciones y predicciones** (figura 2, tramo 4). Cuando se incorporan reglas de decisión basadas en datos, podemos hablar de algoritmos como reglas lógicas programadas. Hay algunas más sencillas como, por ejemplo, determinar si una persona es mayor de edad en función de superar un umbral mínimo o, a modo de filtro, chequear si se cumple determinada característica con un verdadero o falso (el denominado operador booleano).

Cuando tenemos la información necesaria y lo suficientemente depurada, podemos pensar en aplicar **inteligencia artificial**. La IA es una versión más compleja de algoritmos que, a partir de un sistema de cálculo (procesamiento) veloz e iterativo pueden computar características de los datos (aprender) e identificar y/o reconocer patrones. Si el propósito del algoritmo es identificar patrones, hablamos de **aprendizaje no supervisado** o exploración y explotación de datos³¹. Si, en cambio, el propósito del algoritmo es reconocer patrones sobre datos preexistentes y aplicar este aprendizaje sobre datos nuevos mediante un proceso de inferencia, hablamos de **aprendizaje supervisado** (para mayor detalle, [ir al Glosario](#)). Es importante recordar que el **aprendizaje automático** o *Machine Learning* está integrado tanto por el aprendizaje supervisado como por el no supervisado, aunque en muchos casos, se lo refiere como análogo al supervisado.

³¹ También se conoce como data mining (sí, como en el mismísimo caso de la actividad minera, aquí hablamos de iminería de datos!).

En la tabla 4 presentamos herramientas más complejas de CDIA como son los sistemas de detección temprana y pronósticos para la gestión y la prevención del riesgo, en particular, el Sistema de Alerta Temprana por Temperaturas Extremas (SAT), la Gestión epidemiológica basada en CDIA (ARPHAI) y el Reconocimiento Facial de personas (SI-FACE). Es destacable que en estos casos, se presentan uno o varios sets de información que se utilizará como insumo crítico, modelos prototípicos preexistentes y/o el desarrollo de una herramienta previa en la cual se apalanca la innovación.

Tabla 4. Innovación con ciencia de datos e inteligencia artificial: Estimaciones y predicciones basadas en datos

Sistemas de detección temprana y pronósticos para la gestión y la prevención del riesgo como soluciones complejas

[VER DOCUMENTO](#) ↘

A modo de reflexión, es posible trazar un **recorrido completo y secuencial de acceso a información sistematizada e implementación de herramientas de CDIA** para la toma de decisiones basadas en información y conocimiento. Aquí miramos una serie de posibilidades de innovación pública a la luz de los sistemas de información, el análisis de datos y el conocimiento específico del dominio, incluyendo, la operacionalización del flujo de trabajo con el uso de CDIA con creciente complejidad e integración. En resumen, el uso de CDIA permite sistematizar información, monitorear el rendimiento y los procesos de gestión, agilizar procedimientos, analizar impactos, reproducir y verificar resultados, brindar transparencia en la toma de decisiones, entre muchísimas otras aplicaciones. Así se inicia un círculo virtuoso de aprendizaje reproducible y optimización con un gran potencial de escalamiento e integración transversal de la administración pública en sus distintas jurisdicciones mediante la adopción, la transferencia y la formación adecuada.

El inicio de un proyecto puede tener múltiples aristas. Ahora ¿queda un poco más claro de qué va cada paso vinculado a la implementación de herramientas de CDIA en el sector público? El conjunto de procesos que acabamos de plantear, cómo se van enlazando con la fase siguiente y la secuencia de tramos posibles que proponemos en la [figura 2](#) ayudan a identificar en qué tramo posicionamos al proyecto que nos gustaría llevar adelante. Si volvemos a la [figura 2](#), también contaremos con elementos para contextualizar el desafío e idear un flujo de trabajo con pasos intermedios que permita dimensionar los avances progresivos para no colapsar en el intento. Es decir para que la ansiedad o el agobio no pasen por alto las instancias de desarrollo y aprendizaje intermedio (adquisición y adaptación) que son esenciales para un uso adecuado de las herramientas de CDIA. En esta línea, el [recuadro 6](#) vincula una serie de etapas y ejemplos que presentamos en el primer tramo y en el último: la sistematización de historias clínicas electrónicas y el caso de ARPHAI.

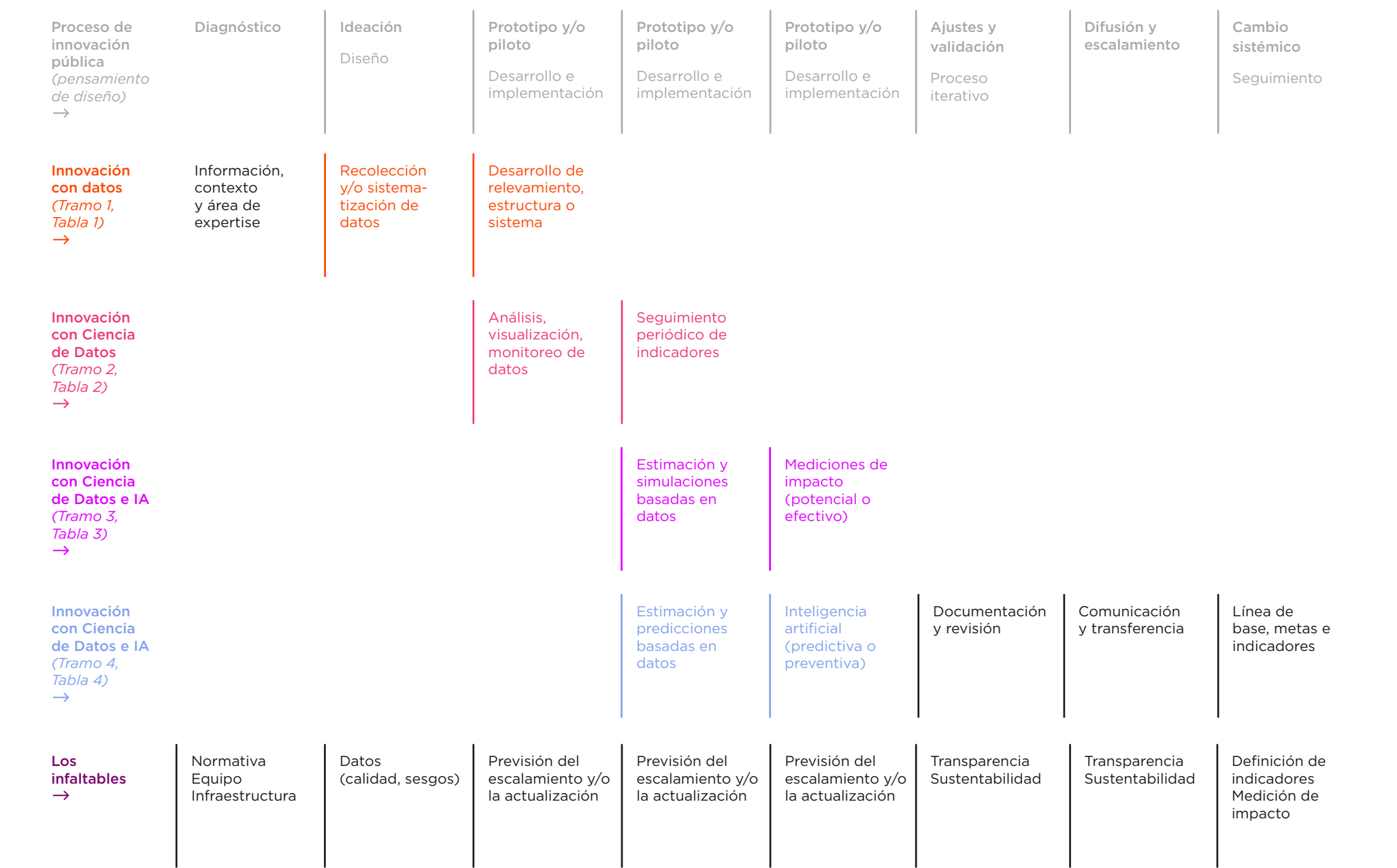
RECUADRO 6

Un ejemplo de recorrido completo en la implementación de herramientas de CDIA en el sector público

Volvamos, entonces, al caso de ARPHAI (tramo 4, tabla 4). Este proyecto como propuesta piloto se inició en un área de estudios interdisciplinarios mediante la concreción de una serie de asociaciones institucionales frente a la posibilidad de presentarse a una convocatoria internacional de fondos. Dada la información preexistente: el registro de Historias Clínicas Electrónicas por entonces ya operativo (tramo 1, tabla 1); la propuesta de ARPHAI emerge como un norte por abordar la salud de forma preventiva mientras intenta resolver un problema local cuya aplicación se orienta a identificar y prevenir los efectos del Dengue y el COVID19. Es destacable que, aún esperando contar con una cobertura mayor del territorio nacional, se ha implementado una fase piloto a partir de la información disponible, lo cual va en línea con el prototipo que abordamos desde el pensamiento de diseño para testear la solución. A la vez, ese testeo ha vehiculado correcciones a menor escala a fin de reducir tiempos y costos de instrumentación mientras se capitalizan los errores.

Complementariamente, la [figura 3](#) presenta el recorrido completo para innovar con herramientas de CDIA en el sector público, integrando los pasos troncales de cada tramo y la secuencia completa siguiendo la perspectiva de pensamiento de diseño. Cada desarrollo de un tramo funciona de base la ideación del siguiente. A su vez, a modo de índice gráfico, se vislumbran algunos puntos críticos que denominamos “los infaltables” a la hora de trabajar con datos y CDIA, los cuales abordaremos en detalle en la próxima sección.

Figura 3. Un recorrido completo para innovar con herramientas de CDIA en el sector público



Fuente: Elaboración propia. || Nota: (1) + (2) son pasos necesarios para (3) y (4).

El recorrido completo de este proceso de innovación pública pone de manifiesto que la disponibilidad y la accesibilidad a la información es un proceso de construcción colectiva que evoluciona en el día a día. Es posible que no contemos con información completa a la hora de tomar decisiones y, por lo tanto, debamos usar algunas aproximaciones o ajustar ciertas premisas mediante prueba y error. El tipo de datos necesario para la toma de decisiones puede exceder u omitir ciertas líneas de bases o indicadores propios de la gestión y el monitoreo. Como hemos visto en este recorrido, la progresión de información disponible y la posibilidad de instrumentar soluciones complejas a partir de la integración de sistemas y predicciones, permite avanzar hacia una transformación sistémica más preventiva que reactiva en términos de diseño y gestión de políticas y servicios públicos.

Hasta aquí exploramos sobre las principales conceptualizaciones, abordajes y aplicaciones de herramientas de CDIA en el sector público. Ahora, con el propósito de simplificar el abordaje de la problemática inmersa en la política pública, introducimos una serie de preguntas por responder, con una impronta de lista a verificar (o *checklist*)³² que opera como un llamado a la acción en materia de innovación pública.

CHECKLIST

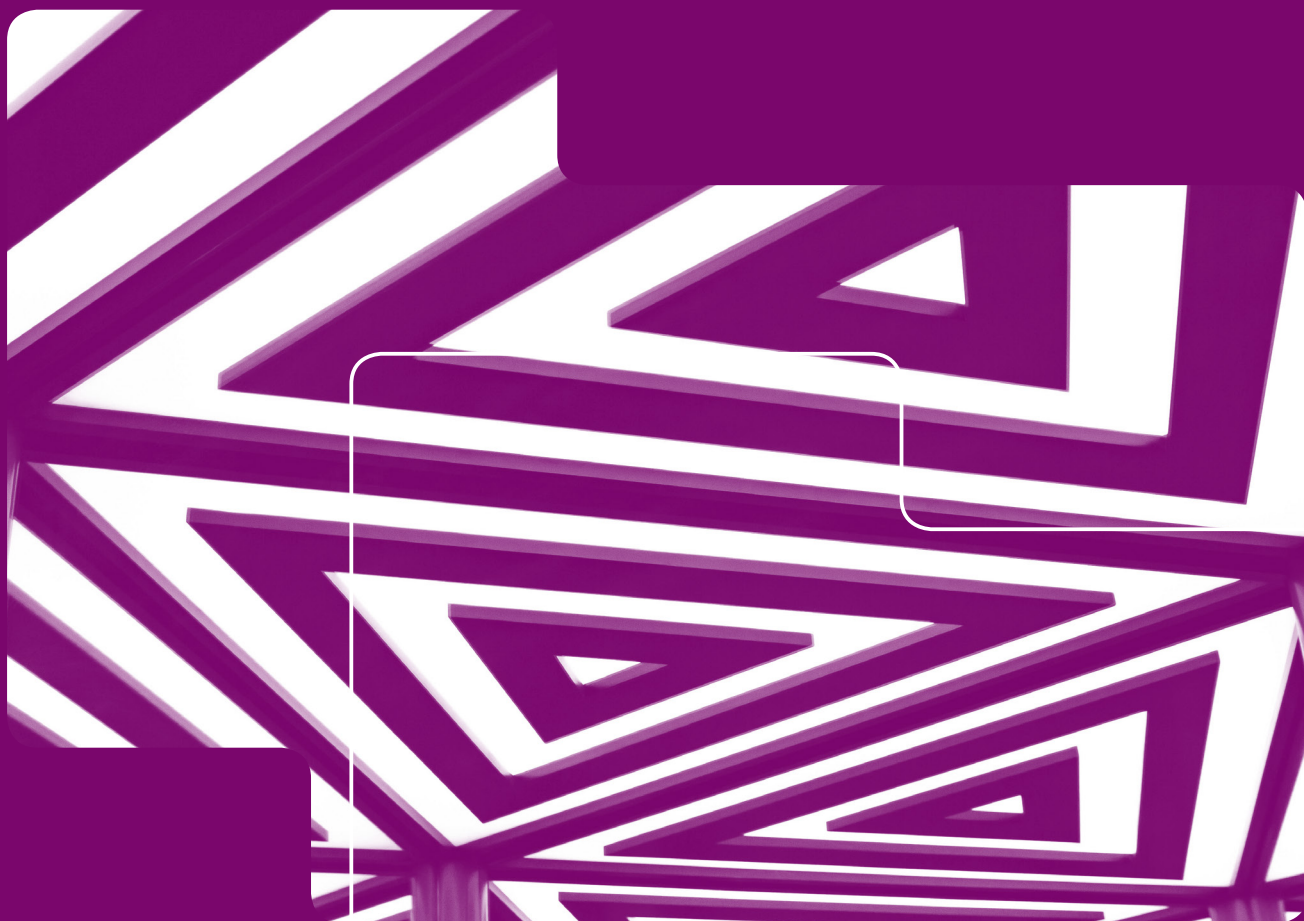
Definición del problema o la necesidad

- ☐ ¿Qué quiero resolver o mejorar?
 - ☐ ¿Cuál es el alcance de esta problemática?
 - ☐ ¿Cuál es la población destinataria?
 - ☐ ¿Tengo en claro qué impactos genera? ¿Qué impactos quiero generar?
¿Hay algún aspecto negativo entre los posibles impactos?
 - ☐ ¿Qué es lo indispensable que debo comunicar o difundir en las distintas instancias?

³² Al final del documento, encontrarás el checklist completo, que recorre preguntas abordables a partir de los conceptos que se vuelcan a continuación.



Recomendaciones: Los infaltables a la hora de usar Ciencia de Datos o Inteligencia Artificial en el sector público





En este apartado brindaremos una serie de recomendaciones para crear un proyecto que involucre CDIA. Pero antes de implementar cualquier solución o herramienta debemos preguntarnos si es efectivamente la mejor opción disponible.

PALABRAS CLAVE:

HERRAMIENTAS
TECNOLÓGICAS

PROBLEMÁTICAS
SOCIALES

SOLUCIONES
COMPLEJAS

¿Es la Inteligencia Artificial o la Ciencia de Datos la mejor herramienta para abordar el problema que quiero resolver?

Muchas veces, la necesidad de innovar y usar nuevas herramientas en la gestión pública surge desde un lugar genuino de querer mejorar los procesos actuales dentro de una dirección o secretaría, bajo la certeza de que estas herramientas pueden ayudar a optimizar procesos o darnos información nueva y valiosa que no teníamos. Muchas otras veces, esta ‘necesidad’ surge porque la CDIA está de moda aunque nadie sepa bien qué hace ni cómo funciona, o porque se piensa que es la mejor alternativa tecnológica: *“Quiero que mi dirección/secretaría utilice IA para resolver tal problema, o para predecir tal tema”*, *“Tenemos que hacer algo con estos datos, seguro que podemos usar Inteligencia Artificial con ellos”*, *“La empresa X nos propone trabajar en una solución de IA para tal problema”*, etc. En este punto es necesario **hacer foco sobre el problema que se busca resolver o la necesidad que queremos abordar y buscar las mejores herramientas posibles para lograrlo.**

Para responder a esta pregunta tan importante, debemos **identificar el problema** que estamos intentando resolver. En este sentido **es necesario aclarar que la CDIA por sí sola no resuelve problemas, sino que es una herramienta que forma parte de un conjunto de estrategias para abordar un problema, en general complejo.** Si no, caemos en lo que se suele llamar “solucionismo tecnológico” donde creemos que los problemas de la sociedad pueden resolverse con más y mejor tecnología. Lejos de ser así, para abordar problemas sociales es necesario contar con un extenso conocimiento sobre los mismos, y con varios soportes no tecnológicos también.

¿A qué nos referimos? Vamos a analizar un caso: en nuestro país en 2018 se implementó un modelo de Inteligencia Artificial que generó mucha polémica. Un ministerio provincial, en alianza con una empresa, desarrollaron una serie de algoritmos para poder predecir, con seis años de anticipación, qué niña estaría “predestinada” a tener un

embarazo adolescente y en base a estos resultados, prevenir estos embarazos³³. Si bien parece un argumento de una película de ciencia ficción, este programa fue implementado y funcionó durante un buen tiempo hasta que varias instituciones tales como el Laboratorio de Inteligencia Artificial Aplicada (LIAA) de la Universidad de Buenos Aires alertaron sobre su uso.

Según los datos compartidos por el equipo de desarrollo de esta herramienta (en particular del repositorio de código de un empleado de la empresa involucrada), la misma poseía un 86% de eficacia en sus predicciones. Los datos que utilizaron para realizar estas predicciones fueron edad, barrio, etnia, país de origen, presencia de discapacidad, agua caliente en el hogar y si el jefe de hogar abandonó estudios³⁴. La herramienta buscaba “resolver” la problemática del embarazo adolescente, sin considerar el acceso a recursos como la Educación Sexual Integral (ESI) como otra variable en el análisis, por ejemplo.

En ese sentido nos preguntamos ¿es esta la herramienta adecuada para la problemática compleja de los embarazos adolescentes? Como venimos diciendo, los problemas complejos requieren soluciones complejas, y muchas veces pensamos que la tecnología nos va a solucionar todos nuestros problemas, cuando lo que termina haciendo es generar problemas nuevos. Este sistema de predicción de embarazo adolescente, al estar entrenado con datos de la sociedad, reproduce sesgos que están entrelazados en esos datos y por lo tanto, sus predicciones estarán sesgadas. De esto hablaremos en profundidad en el apartado dedicado a los sesgos, prejuicios y riesgos a la hora de utilizar herramientas de CDIA, pero por ahora podemos pensar en los sesgos de esta herramienta, desde las afirmaciones de la investigadora Luciana Ferrer³⁵: *“En un tema tan delicado como el embarazo adolescente, sería cauteloso asumir que muchas adolescentes mintieron, sobre todo si abortaron. Esto implica que, al usar estos datos, estaremos aprendiendo información sesgada, influenciada por factores como cuáles sectores de la población tienen más facilidades para abortar o en qué sectores es más tabú quedar embarazada en la adolescencia y por lo tanto es algo que la adolescente escondería”*³⁶.

Para utilizar herramientas tecnológicas en la gestión pública, es necesario diferenciar entre los problemas técnicos, por ejemplo “predecir a partir de datos históricos la posibilidad de que una adolescente en particular quede embarazada de acá a 5 años”, y otro es el problema social o *no técnico*, que sería “reducir la cantidad de embarazos adolescentes”. A veces ambos se alinean y la solución al problema técnico aporta a resolver el problema social (en conjunto con otras herramientas), pero muchas otras veces estos problemas no se vinculan de esta manera.

Por lo tanto es fundamental realizar un análisis previo antes de implementar una herramienta de CDIA, verificar que realmente es la mejor posibilidad, que su implementación será cuidada y no generará problemas nuevos. La organización Web Foundation menciona que no solamente hay que validar la efectividad de una herramienta, sino su legitimidad, esto implica por un lado que la herramienta esté alineada con las leyes y normas actuales y con la protección de derechos humanos³⁷.

Para determinar la legitimidad de una herramienta de CDIA es necesario pensar en los resultados de la misma evaluando sus impactos, si son justos/equitativos y si no

33 Véase <https://www.pagina12.com.ar/109080-la-inteligencia-que-no-piensa>

34 Idem anterior

35 Véase <http://habla.dc.uba.ar/lferrer/>

36 Véase <https://www.pagina12.com.ar/109080-la-inteligencia-que-no-piensa>

37 Ortiz Freuler, J. and Iglesias, C. (2018). Algorithms and Artificial Intelligence in Latin America: A Study of Implementation by Governments in Argentina and Uruguay, World Wide Web Foundation. En: http://webfoundation.org/docs/2018/09/WF_AI-in-LA_Report_Screen_AW.pdf

discriminan a cierta población. También es necesario evaluar los procesos para ver si los mismos pueden ser explicados fácilmente, si están co-diseñados con la población destinataria y si son transparentes.

Podríamos concluir entonces que es necesario analizar todas las aristas de una herramienta de CDIA, ya que no son solamente los algoritmos en sí lo que la conforman, sino que se requiere de un sólido entramado alrededor de la CDIA para generar herramientas que duren en el tiempo.

Los datos que poseemos, su calidad, la infraestructura requerida, el equipo técnico, la normativa vigente, su medición y diálogo son variables a tener en cuenta, más allá de los modelos algorítmicos en sí mismos. Estas variables deben estar planificadas como parte estructural del desarrollo para generar herramientas exitosas.

A través de las siguientes preguntas podremos obtener el panorama completo de lo que implica herramientas de CDIA para abordar problemas en el sector público, desterrando la idea de que trabajar con Ciencia de Datos o Inteligencia Artificial es solamente usar algoritmos.

CHECKLIST

Herramientas tecnológicas y el uso de CDIA

- ☐ ¿Se puede vincular esta problemática con las herramientas tecnológicas que quiero utilizar? ¿De qué manera?
 - ☐ ¿Sería algo novedoso? ¿O una mejora de algo existente?
 - ☐ ¿Cómo me ayudaría la herramienta ideal a mejorar la gestión, la política pública o el servicio?
 - ☐ ¿La herramienta ideal implica modificar algún proceso de trabajo o circuito de gestión? ¿Sobre qué/quienes impactaría en este último caso? ¿Cómo podríamos trabajar esta transición?
- ☐ ¿Tengo datos sobre la problemática? ¿Cuáles necesitaría sistematizar y/o recolectar?
- ☐ ¿La herramienta ideal utilizará CD, aprendizaje automático o la IA?
 - ☐ ¿Serían las mejores herramientas para ayudar a resolver este problema?
 - ☐ ¿Qué otras herramientas o abordajes podría utilizar?
 - ☐ ¿Es posible que esta herramienta sea usada de manera incorrecta y genere efectos no deseados? ¿Cuáles? ¿Son evitables?
- ☐ ¿Pienso en desarrollar o en adquirir y adaptar la herramienta de CDIA?
 - ☐ En caso de adquirirla, ¿existe documentación y acompañamiento para su implementación? ¿Quién del equipo la probaría o la adaptaría?

RECUADRO 7

Consideraciones para compras y contrataciones de herramientas de CDIA en el sector público

Si bien en este documento nos estamos enfocando en la construcción de herramientas de CDIA queremos hacer hincapié en que en muchas ocasiones estas herramientas pueden estar construidas de antemano y se nos ofrecerán como productos cerrados, cajas negras que fueron desarrolladas por un equipo externo (en muchos casos consultoras o empresas que ofrecen productos con CDIA incluida para la administración pública). Sabemos que mantener equipos técnicos es difícil a nivel estatal por la competitividad de los sueldos y la alta demanda, por lo tanto en muchas ocasiones será necesario contratar equipos externos. Si bien pareciera que no hay mucho por hacer en estos casos, definitivamente podemos verificar un sinnúmero de temas antes de implementar esas herramientas.

Las preguntas que nos iremos haciendo y que luego conforman el checklist final, en su gran mayoría son válidas cuando hablamos de contrataciones externas y compras de productos de CDIA. Muchas preguntas del checklist apuntan a garantizar la calidad de los datos, comprender las leyes vigentes y normativas vinculadas al tema, construir transparencia en la aplicación de herramientas, identificar sesgos en los sistemas de CDIA y mantenerlos en funcionamiento. Todos estos temas aplican también cuando estamos comprando herramientas externas.

Si se busca profundizar en la contratación pública de CDIA hay tres documentos centrales que servirán de guía, ya que son lineamientos para compras y licitaciones de herramientas de Inteligencia Artificial para gobiernos. Es importante aclarar que todos estos documentos fueron elaborados en el Norte Global, por lo tanto están en inglés, con excepción del documento del World Economic Forum que está en español.

El primer documento que recomendamos fue desarrollado por el Gobierno de Reino Unido³⁸ y busca informar a las diversas áreas de la administración pública que estén considerando la utilizar tecnologías de IA para mejorar los servicios existentes o como parte de la futura transformación de los servicios. El segundo documento fue elaborado por el World Economic Forum³⁹ y contiene lineamientos para la contratación pública de Inteligencia Artificial.

Por último recomendamos el White Paper en la Ética de datos en Contratación Pública de Inteligencia Artificial⁴⁰ del Instituto Nacional de Estándares y Tecnología (NITS). Este documento es una contribución para orientar y hacer un uso estratégico de la contratación pública de la IA en la Unión Europea con el objetivo final de alcanzar objetivos compartidos en relación con la adopción y el desarrollo de una IA confiable en Europa.

38 Véase <https://www.gov.uk/government/publications/guidelines-for-ai-procurement>

39 Véase https://www3.weforum.org/docs/WEF_Adquisicion_de_IA_Guidelines.pdf

40 Véase <https://www.nist.gov/system/files/documents/2021/08/23/ai-rmf-rfi-0043.pdf>

PALABRAS CLAVE:

EQUIPO TÉCNICO

PERSONAS
EXPERTAS EN EL
PROBLEMA

FUNCIONARIOS Y
FUNCIONARIAS

POBLACIÓN
DESTINATARIA

INTERDISCIPLINA

¿Cuento con un ecosistema de actores para diseñar e implementar herramientas de Ciencia de Datos y/o Inteligencia Artificial en el sector público?

Antes de encarar cualquier proyecto de CDIA en el sector público deberemos identificar de antemano cuál es el ecosistema de actores que estará en diálogo constante con esta implementación. **Hablamos de ecosistema para comprender que este desarrollo trasciende al equipo técnico;** éste deberá ser acompañado por otras personas y equipos para lograr su objetivo. El reporte de Inteligencia Artificial en Latinoamérica de la Web Foundation⁴¹ explica que hay tres actores clave en esta etapa:

1) Funcionarios/as y personas expertas en el problema:

Estas personas serán las encargadas de decidir si la CDIA es o no la herramienta adecuada para ayudar en el problema concreto que queremos abordar en la gestión pública. Si deciden que es útil, viable y de interés para la población, fijan los objetivos y los valores que deben optimizarse. Según el manual “Uso responsable de la IA en el sector público: manual de proyectos”⁴² el rol de la persona a cargo de la toma de decisiones de la política pública (dirección del proyecto) no es desarrollar la herramienta basada en IA; le corresponde, sí, la responsabilidad de formular el proyecto, comunicarse con el equipo técnico, monitorear la mitigación de riesgos, determinar la viabilidad de la herramienta y rendir cuentas por la ejecución. Este proceso es el que estructuramos en cada tramo de la innovación pública con herramientas de CDIA basada en la secuencia propuesta por el pensamiento de diseño (figura 2).

2) Equipo técnico. Las personas expertas en CDIA:

La interdisciplina dentro de estos equipos es fundamental para generar herramientas robustas. El diálogo constante con las personas expertas en el problema, la comunicación y su participación son un engranaje clave y deberían formar parte del trabajo diario codo a codo con el equipo técnico. A su vez, las personas encargadas del proyecto a nivel política pública también tienen que ser parte de la toma de decisiones y del proceso ya que serán quienes gestionen los recursos disponibles, el alcance y aplicación de la política.

41 Ortiz Freuler, J. and Iglesias, C. (2018). Algorithms and Artificial Intelligence in Latin America: A Study of Implementation by Governments in Argentina and Uruguay, World Wide Web Foundation. En http://webfoundation.org/docs/2018/09/WF_AI-in-LA_Report_Screen_AW.pdf

42 Véase <https://publications.iadb.org/publications/spanish/document/Uso-responsable-de-IA-para-politica-publica-manual-de-formulacion-de-proyectos.pdf>

Es fundamental aclarar que el equipo técnico no puede ni debe trabajar aislado del resto de los equipos emplazados en la institución que instrumentará la herramienta de CDIA, y no puede perder de vista el objetivo del proyecto. Debe ser instruido y animado a criticar las decisiones de diseño, aun cuando no tenga las capacidades para resolver algunos cuestionamientos, por eso se habla de trabajo en equipo interdisciplinario.

Parte de este equipo técnico debe poner en práctica los valores definidos por los funcionarios y funcionarias que hacen políticas públicas y expertos del problema y debe decidir qué dimensiones incluir, qué pesos asignarles y cómo deben interactuar, en diálogo con los diversos actores del equipo. Estos puntos son críticos en la fase de diseño e implementación de la herramienta de CDIA y es indispensable que estén documentados para visibilizarlos y validarlos. Otra parte del equipo técnico estará enfocada en curar los datos necesarios para los modelos, pensar en cómo va a interactuar la herramienta con las base de datos actuales y cómo puede optimizarse esta vinculación.

3) La población destinataria:

Con respecto a este último punto tenemos que diferenciar por un lado a quiénes afectará la herramienta que estamos construyendo y también quiénes serán las personas usuarias de la misma. Pongamos un ejemplo concreto: en el caso de las Historias Clínicas Electrónicas (desarrollado con anterioridad en la tercera etapa: tramo 1, [tabla 1](#)), podemos ver que la población destinataria es la sociedad en su conjunto, ya que esta herramienta es utilizada para la digitalización de las historias clínicas de pacientes, la gestión de turnos y para hacer un seguimiento de la evolución de las historias clínicas. Por otro lado, las personas usuarias de esta herramienta son el personal de salud que trabaja con la misma. En ese sentido podemos ver que quienes se benefician o se ven afectados con los resultados de este proyecto no son las mismas personas que usarán la herramienta.

Es fundamental mencionar que en muchos casos los proyectos vinculados con innovación tecnológica en el sector público cuentan con equipos de investigación vinculados a universidades, organizaciones de la sociedad civil, empresas y fondos internacionales, que están a cargo del desarrollo, financiamiento y ejecución de estas herramientas y también son actores relevantes que pueden formar parte de los equipos de trabajo. En línea con esto, es tan importante identificar propósitos e intereses de cada grupo interviniente como los posibles riesgos que se pueden manifestar en la práctica para poder gestionarlos, por ejemplo, frente al uso de la herramienta de CDIA y la disponibilización de información crítica.

CHECKLIST

Equipo de trabajo

- ☐ ¿Mi equipo de trabajo es interdisciplinario? ¿Qué capacidades necesito reforzar?
- ☐ ¿Hay planificadas instancias de intercambio y trabajo en conjunto entre el equipo experto en la problemática y el equipo técnico?

PALABRAS CLAVE:

CALIDAD DE LOS DATOS

COMPRENSIÓN ANÁLISIS

DATOS FALTANTES

CONSISTENCIA

LIMITACIONES

¿Con qué datos cuento? ¿Con qué datos me gustaría contar? ¿A qué nos referimos con calidad de los datos?

Vamos a empezar por lo fundamental: **sin datos de calidad no puede haber herramientas integrales de CDIA** ya que éstos son una parte constitutiva tanto de la IA como de cualquier proceso de Ciencia de Datos. Por lo tanto es necesario antes que nada realizar un análisis de los datos existentes, comprenderlos en su totalidad, saber quién los recolectó, con qué propósito, qué significa la información de cada campo y ver sobre todo si hay datos que faltan y que serán necesarios para el proyecto que queremos desarrollar.

Cuando hablamos de análisis nos referimos a comprender los datos en sí. Resulta fundamental saber en qué formato se encuentran estos datos, podríamos tener dos escenarios posibles: datos no estructurados y datos estructurados. En el primer caso es necesario saber si estos datos se encuentran en información sacada de la web, en audios, videos o pdfs con textos y por lo tanto se trabajará con herramientas que ayuden a estructurar esta información. Si los datos ya están estructurados será fundamental saber qué columnas hay, qué contenido tiene cada una, qué datos faltan, cuáles tienen valores que son inconsistentes o carecen de homogeneidad. Por ejemplo, en la edad aparece que una persona tiene 200 años, o las fechas que aparecen no se condicen con las fechas de recolección o de alta del sistema, o un dato que tiene que ser numérico aparece con diferentes formatos, con puntos, comas, guiones y hay que estandarizar esa información.

Ante los datos faltantes tenemos dos opciones: la primera es verificar si éstos existen en otra base de datos e intentar cruzar las bases para agregarlos de manera correcta a nuestras bases actuales, y la segunda es pensar en alguna forma de recolectarlos para agregarlos a la base. Cuando estamos realizando este proceso es necesario verificar que los datos que estemos agregando sean consistentes y no tengan errores de continuidad con los datos existentes.

Es fundamental asegurarnos que los datos que vamos a usar en el proyecto sean de buena calidad y es necesario conocer también sus limitaciones. Cuando trabajamos con datos en el sector público, siempre estamos en una constante tensión entre la obtención de la mayor cantidad posible de datos y la protección de la privacidad de las personas y la seguridad de estos datos. Es decir que tenemos que pensar en sus procesos de apertura⁴³, anonimización⁴⁴ y seguridad⁴⁵.

43 Véase <https://datosgobar.github.io/paquete-apertura-datos/guia-abiertos/>.

44 Para comprender mejor los procesos de anonimización, recomendamos la lectura de Luvini, Paula (2022). Guía práctica para la protección de datos. Buenos Aires: Fundar. Disponible en <https://www.fund.ar>

45 Para comprender mejor el término seguridad en los datos recomendamos la lectura de este sitio <https://www.ibm.com/ar-es/topics/data-security>

Para facilitar la generación de valor a partir del procesamiento de los datos y determinar la confiabilidad en la toma de decisiones, es necesario que dichos datos cuenten con características que satisfagan las expectativas y necesidades de información de las personas usuarias, las cuales se centran básicamente en la calidad de los datos.

La calidad de los datos

Según La Guía de Fundamentos para la Gestión de Datos (DAMA, 2017), existen muchas dimensiones que permiten determinar el grado de calidad de los datos, a través de su gestión. Los datos deben **tener fácil acceso y ser fáciles de encontrar** cuando se los necesita, a su vez que un control sobre quiénes pueden verlos si son datos sensibles. Los datos deben ser **completos y suficientes** para la herramienta que los va a utilizar. También tienen que tener **consistencia** entre sí y ser **precisos y válidos**.

CHECKLIST

Conocimiento en profundidad de los datos

Si quiero implementar un modelo de aprendizaje automático o IA:

- ☐ ¿Tengo los datos que requiere el algoritmo?
- ☐ ¿Conozco en profundidad mi set de datos?
- ☐ ¿Con qué información adicional me gustaría contar?
- ☐ ¿Mis datos son de calidad?
- ☐ ¿Qué preguntas tengo que hacerme para comprender mis datos en profundidad?

PALABRAS CLAVE:

CONECTIVIDAD

SERVIDORES

DOCUMENTACIÓN

EQUIPAMIENTO

MANTENIMIENTO

¿Qué infraestructura y equipamiento necesito para realizar y mantener una herramienta de Ciencia de Datos y/o Inteligencia Artificial? Una vez que tengo la herramienta ¿qué hago?

Si estamos buscando una implementación que pueda sostenerse en el tiempo y que siga funcionando correctamente y que no dependa al 100% del equipo actual que está ejecutando el proyecto, es necesario pensar en infraestructura, equipamiento, documentación y evaluación permanente. Al pensar en una herramienta de CDIA será necesario pensar en el mantenimiento de la misma a futuro, la necesidad de una infraestructura robusta para que el proyecto tenga continuidad más allá de la permanencia del equipo.

Para implementar CDIA en el sector público es fundamental tener buena conectividad, acceso a servidores, una buena documentación del proyecto y constantes revisiones para ver si esta tecnología sigue cumpliendo el objetivo que cumplía originalmente. Pero vamos por partes.

Cuando hablamos de **equipamiento**, tenemos que tener en cuenta que los proyectos que involucren CDIA pueden requerir de un gran poder de cómputo, ya que deben realizarse operaciones con grandes volúmenes de datos y éstas pueden requerir computadoras que tengan un buen procesador, o si no tenemos esas computadoras, se va a requerir comprar espacio en la nube para implementar estos modelos. Por eso antes de salir a comprar mejores equipos (sabiendo lo difícil y/o lento que puede ser gestionar este punto en ámbitos públicos) es importante hablar con el equipo técnico o con asesores/as técnicos/as sobre qué equipamiento será necesario conseguir para desarrollar y mantener las herramientas de CDIA. Además, dado que hoy día los principales proveedores de servicios en la nube son empresas que mantienen sus servidores en el exterior, es primordial cuestionarse si es viable, deseable, o incluso legal que los datos sean alojados fuera del territorio nacional, aunque sea temporalmente.

Cuando hablamos de **almacenamiento** de nuestros datos, tal vez escuchamos hablar sobre el término servidores. En muchos casos vamos a trabajar con bases de datos pequeñas que pueden alojarse en una computadora o en la nube personal de algún usuario, pero cuando se trata de grandes volúmenes de datos/macrodatos en muchos casos será necesario utilizar servidores externos (pueden ser del Estado o pueden ser contrataciones a empresas que ofrecen estos servicios). Será fundamental planificar con anticipación dónde estarán alojados estos datos, cómo se los podrá consultar y utilizar y qué medidas de seguridad estamos implementando para resguardar datos sensibles o personales.

Con respecto a la **evaluación permanente**, es fundamental monitorear el sistema que estamos utilizando para mantener su calidad y su objetivo. Debemos ver si se requiere actualizarlo en función de ciertos cambios en la política pública, en el objetivo del proyecto o en los datos a utilizar. Para evaluar el desempeño de una herramienta desarrollada de manera eficiente es necesario diseñar métricas que nos sirvan para comparar el antes, el durante y el después, y que nos ayuden a medir en qué grado está funcionando la herramienta y sobre todo si se sigue manteniendo la calidad de los resultados. El sistema de evaluación permanente tiene que planificarse antes de implementar cualquier herramienta, debe ser una parte más a diseñar de manera interdisciplinar.

Otro ítem fundamental es la **documentación** del proyecto. Es una buena práctica dejar registrado todo lo realizado hasta el día de la fecha: los consensos del equipo, las decisiones sobre qué modelos implementar, qué datos usar, el código utilizado, comentarios en el código que expliquen qué hace y cómo. Los equipos técnicos suelen ensamblarse para resolver ciertos problemas particulares y luego se disuelven rápidamente, si éstos no generan la información necesaria para que otro equipo pueda retomar el trabajo o para que nadie sea indispensable, tendremos un problema de mantenimiento en un futuro. También dentro de la documentación es fundamental que el equipo pueda generar manuales de uso para las personas sin bagaje técnico, para que puedan usar la herramienta y navegarla sin dificultades. El apartado sobre documentación de nuestra checklist puede ser utilizado por quien tenga que mantener o comunicar estas herramientas en un futuro.

RECUADRO 8

Impacto ambiental y sustentabilidad en tecnología

Un aspecto que rara vez es mencionado cuando hablamos de tecnología es su impacto o huella ambiental. Hablar de algoritmos parece abstracto, en especial porque pensamos que “la nube” es un espacio intangible. La realidad dista mucho de esta concepción ya que tenemos que tener en cuenta varias consecuencias del uso tecnológico y en particular del almacenamiento de datos y procesamiento de algoritmos. Por un lado, los *data centers* son espacios físicos, generalmente situados en lugares alejados de los centros urbanos que utilizan grandes cantidades de agua para ser enfriados y de electricidad para funcionar.

A su vez Kate Crawford dice que “el término “inteligencia artificial” puede invocar ideas de algoritmos, datos, arquitecturas en la nube pero nada de eso puede funcionar sin minerales y recursos que constituyen el núcleo duro de la computación”⁴⁶. Por lo tanto, los minerales y metales utilizados para producir la tecnología (computadoras, dispositivos, servidores, cables de fibra óptica, etc.) tienen detrás su impacto ambiental también.



46 Crawford, K. (2021). Atlas of AI Power, Politics, and the Planetary Costs of Artificial Intelligence. Yale University Press.

Paz Peña, activista ambiental y feminista afirma a través del Instituto Latinoamericano de Terraformación que: “Las estimaciones iniciales de la intensidad computacional de los algoritmos de Inteligencia Artificial sugieren que los requisitos energéticos pueden ser enormes sin un esfuerzo concertado para mejorar la eficiencia, especialmente en el aspecto computacional”⁴⁷. Esto no quiere decir que podamos modificar cómo las grandes empresas monopólicas de tecnología administran sus recursos, pero sí informarnos sobre el impacto socioambiental de estas tecnologías. Para saber más sobre estos temas recomendamos seguir el trabajo del Instituto Latinoamericano de Terraformación⁴⁸ y de la Anatomía de un sistema de IA de Kate Crawford⁴⁹.

CHECKLIST

Infraestructura

- ☐ ¿Qué infraestructura necesito para realizar y mantener una herramienta de CDIA?
 - ☐ ¿Dónde están almacenados los datos?
 - ☐ ¿Van a ser necesarios servidores externos para alojarlos o realizar operaciones sobre éstos?
- ☐ ¿Qué medidas de seguridad estamos planificando para que estos datos estén resguardados y no haya filtraciones?
- ☐ ¿Con qué equipamiento cuento y qué necesito? ¿Las computadoras actuales permiten leer los archivos y procesar los datos que necesito?
 - ☐ ¿Estoy planificando documentar mi herramienta para que sea replicable por otros equipos? ¿Y documentar las decisiones sobre los datos y el proceso?
 - ☐ ¿Tengo pensado desarrollar un manual de uso de la misma para personas no técnicas? ¿Qué capacidades adicionales necesitamos para producir este tipo de manual (por ejemplo: expertise en didáctica, en redacción simple, etc.)?
 - ☐ ¿Cómo y dónde voy a documentar el proyecto? ¿Hay una persona del equipo designada? ¿Es un proceso colaborativo?
 - ☐ ¿Qué herramientas tengo a disposición para hacerlo?
 - ☐ ¿Tengo en cuenta que la documentación tiene que ser interdisciplinaria y no solamente del equipo técnico?
 - ☐ Si otro equipo toma la herramienta que diseñé y la documentación: ¿puede replicarla o editarla fácilmente? ¿Qué acciones debemos tomar para que si el equipo cambia, el proyecto siga funcionando?

47 Véase https://user.fm/files/v2-5f59336b831517c05a164405cee8d931/IPCC%202022_Ese%20futuro%20tecnol%C3%B3gico%20no%20existe_ILDT.pdf

48 Véase <https://terraforminglatam.net/>

49 Véase <https://anatomyof.ai/>

Impacto

- ☐ ¿Podemos prever el impacto desde el diseño de la herramienta?
- ☐ ¿Cómo podemos planificar una evaluación de impacto de la herramienta?
- ☐ ¿Está prevista la difusión o capacitación sobre la herramienta? ¿Cuenta con posibilidad de escalamiento?

Revisión y monitoreo continuos de la herramienta/proyecto

- ☐ ¿Qué instancias de revisión o monitoreo continuos estamos considerando para este proyecto?
- ☐ ¿Con cuánta periodicidad vamos a monitorear el desempeño de estas herramientas?
 - ☐ Si identifico que hay nuevos datos no contemplados en la fase de aprendizaje del modelo, ¿Puedo pautar un análisis de los nuevos datos? ¿Tengo prevista la actualización del modelo en función de eso?
- ☐ En el inicio del proyecto: ¿estamos haciéndolo de manera correcta?
- ☐ Durante el proyecto: ¿lo hemos diseñado bien?
- ☐ Después del proyecto: ¿sigue haciendo lo que queríamos que hiciera cuando lo diseñamos?

Sustentabilidad y ambiente

- ☐ ¿Conocemos el impacto ambiental de las herramientas tecnológicas que vamos a utilizar? ¿Y de los equipamientos e infraestructura?
- ☐ ¿Estamos teniendo en cuenta este impacto a la hora de elegir las herramientas a utilizar?

PALABRAS CLAVE:

DATOS PERSONALES

DATOS SENSIBLES

LEY 25.326

TRATADOS
INTERNACIONALES

ÉTICA

PRINCIPIOS OECD

¿Qué normativa existe sobre datos, seguridad y privacidad? ¿Cómo son los circuitos administrativos para usar Ciencia de Datos e Inteligencia Artificial?

Cuando hablamos de regulación y normativa sobre los datos en Argentina, es fundamental mencionar la **Ley 25.326 de Datos Personales** sancionada en el año 2000. Esta ley refiere a aquellos datos personales guardados en archivos, registros, bancos de datos públicos o privados y que están guardados para dar informes. Las imágenes y datos biométricos de las personas también son considerados datos personales. Esta ley reconoce el derecho de las personas a que sus datos personales no sean utilizados ni registrados sin su consentimiento; pedir y que se dé información sobre qué datos personales están registrados en bancos de datos públicos o privados; pedir que sus datos sean corregidos o actualizados; pedir que sean suprimidos, en los casos en que corresponda; pedir que sean guardados confidencialmente; iniciar acción judicial para conocer sus datos o exigir su rectificación, supresión, confidencialidad o actualización.⁵⁰

A su vez esta ley marca una diferenciación entre datos personales y datos sensibles:

“— Datos personales: Información de cualquier tipo referida a personas físicas o de existencia ideal determinadas o determinables.

— Datos sensibles: Datos personales que revelan origen racial y étnico, opiniones políticas, convicciones religiosas, filosóficas o morales, afiliación sindical e información referente a la salud o a la vida sexual.”

A su vez, la **Ley 17.622 de Estadística sancionada en 1968** en el Art. 10 expone que *“Las informaciones que se suministren a los organismos que integran el Sistema Estadístico Nacional, en cumplimiento de la presente ley, serán estrictamente secretos y sólo se utilizarán con fines estadísticos. Los datos deberán ser suministrados y publicados, exclusivamente en compilaciones de conjunto, de modo que no pueda ser violado el secreto comercial o patrimonial, ni individualizarse las personas o entidades a quienes se refieran.”*⁵¹ El **secreto estadístico** por lo tanto es otro aspecto a tener en cuenta cuando buscamos usar y disponibilizar datos de la población y esto puede generar tensiones entre los datos que necesitamos para llevar a cabo el proyecto y las restricciones u obstáculos que pueden poner algunos organismos para disponibilizarlos. Si bien en muchos casos algunas agencias o secretarías pueden negarse a otorgar ciertos datos, es imposible generalizar y estas decisiones dependen de cada caso en particular.

⁵⁰ Véase <https://www.argentina.gob.ar/justicia/derechofacil/leysimple/datos-personales> y <https://www.argentina.gob.ar/normativa/nacional/ley-25326-64790>

⁵¹ Véase https://www.indec.gob.ar/ftp/documentos/Ley_17622.htm

La disponibilidad de algunos datos también puede estar limitada conforme el propósito institucional que la da origen: en nuestro país, los datos de la Administración Federal de Ingresos Públicos (AFIP) y la Administración Nacional de la Seguridad Social (ANSES) consolidan **registros administrativos**, que incluyen los hoy denominados datos personales y/o datos sensibles (al momento de su creación, formalmente no se les daba este tratamiento; sin embargo, los datos que publican son a nivel agregado para preservar la privacidad).

Según el informe elaborado por el Centro de Inteligencia Artificial y Política Digital (CAIDP) en 2021⁵², la actual Ley de Protección de Datos Personales de Argentina no contiene prescripciones legales formales que reconozcan el derecho de la ciudadanía a recibir información sobre algoritmos u oponerse a una decisión basada únicamente en métodos de procesamiento automático de datos. No obstante, la Agencia de Acceso a la Información Pública (AAIP) de Argentina ha emitido una resolución en 2019 en la que reconoce que, en virtud del derecho de acceso consagrado en la actual ley de protección de datos, los titulares de los datos tienen derecho a solicitar a los responsables del tratamiento una explicación sobre la lógica utilizada por cualquier sistema que llegue a decisiones basadas en algoritmos y que pueda afectar a los ciudadanos o tener efectos jurídicos perniciosos sobre ellos.⁵³

Teniendo en cuenta esta normativa, toda herramienta de CDIA deberá estar alineada con la Ley de Protección de Datos Personales. A su vez el equipo encargado de llevar a cabo la política pública de innovación deberá estar informado sobre cambios en la legislación o sobre nuevas normativas respecto a estas temáticas. A nivel mundial cada vez son más los países que buscan regular el uso de algoritmos mediante leyes y también son muchos los países que adhieren a tratados internacionales sobre el uso responsable y ético de los datos. En ese sentido Argentina como parte del G20 firmó su adhesión a los principios de Inteligencia Artificial de la OECD/G20 (ver detalle en el próximo recuadro).

RECUADRO 9

Los principios de Inteligencia Artificial de la OECD/G20⁵⁴

La IA debe estar **al servicio de las personas y del planeta**, impulsando un crecimiento inclusivo, el desarrollo sostenible y el bienestar.

Los sistemas de IA deben diseñarse de manera que **respeten el Estado de derecho, los derechos humanos, los valores democráticos y la diversidad**, e incorporar los mecanismos y protecciones adecuadas —por ejemplo, permitiendo la intervención humana cuando sea necesario— con miras a garantizar una sociedad justa y equitativa.

Los sistemas de IA deben estar **guiados por la transparencia y una divulgación responsable de los mismos** a fin de garantizar que las personas sepan cuándo están interactuando con ellos y puedan oponerse a los resultados de esa interacción.



52 Véase <https://www.caidp.org/reports/aidv-2021/>

53 Resolución 4/2019, RESOL-2019-4-APN-AAIP (Sept. 13, 2019) y Anexo I de la Resolución 4/2019 (IF-2019-01967621-APN-AAIP).

54 Véase <https://oecd.ai/en/ai-principles>

Los sistemas de IA han de funcionar con **robustez, de manera fiable y segura** durante toda su vida útil, y los potenciales riesgos deberán evaluarse y gestionarse en todo momento.

Las organizaciones y las personas que desarrollen, desplieguen o gestionen sistemas de IA deberán responder de su **correcto funcionamiento** en consonancia con los principios precedentes.

CHECKLIST

Normativa / Leyes

- ☐ ¿Mi herramienta cumple con la normativa vigente sobre acceso y protección de datos?
- ☐ ¿Es la herramienta segura y se adecúa a las legislaciones vigentes en materia de seguridad y privacidad de la información?
- ☐ Si se trabaja con datos sensibles, ¿qué medidas de seguridad se están utilizando para protegerlos?
- ☐ ¿El uso de estos datos interfiere con los derechos de las personas cuyos datos están siendo utilizados? Si es así, ¿Hay alguna forma menos intrusiva de llegar al objetivo deseado?
- ☐ A partir de la herramienta ¿necesito generar algún registro formal? ¿o proponer alguna actualización normativa de los circuitos administrativos?

PALABRAS CLAVE:

RENDICIÓN
DE CUENTAS

REPOSITORIOS DE
ALGORITMOS

RESPONSABILIDAD

LEY DE ACCESO A
LA INFORMACIÓN
PÚBLICA

¿Cómo se trabaja hacia la transparencia en estos proyectos?

Uno de los pilares fundamentales de las políticas de innovación en el sector público es la **transparencia**. Siguiendo con el apartado anterior sobre normativas, Argentina cuenta con la **Ley 27.275 de Acceso a la Información Pública (2016)**. La misma tiene como objeto garantizar el efectivo ejercicio del derecho de acceso a la información pública, promover la participación ciudadana y la transparencia de la gestión pública. Esta ley tiene como principio la transparencia y máxima divulgación, lo que implica que toda la información en poder, custodia o bajo control del Estado debe ser accesible para todas las personas.⁵⁵

La transparencia genera confianza en el gobierno y abona a elaborar herramientas y políticas públicas más sensibles a las diversas problemáticas sociales y más colaborativas, es por esto que el Instituto Ada Lovelace, el Instituto AI Now y el Open Government Partnership⁵⁶ desarrollan el concepto de **rendición de cuentas algorítmica** (en inglés *algorithmic accountability*) y enumeran seis puntos que ayudarán a dar forma al despliegue y la aplicación efectiva de las políticas de rendición de cuentas algorítmica:

1. Debe haber incentivos institucionales claros y se deben desarrollar marcos jurídicos vinculantes que puedan apoyar la aplicación coherente y eficaz de los mecanismos de rendición de cuentas y responsabilidad en los algoritmos, con observación y participación de actores de sociedad civil y una cobertura mediática responsable.

2. Las políticas de responsabilidad algorítmica deben definir claramente los objetos de la gobernanza, así como establecer terminologías compartidas entre los departamentos gubernamentales.

3. Establecer el ámbito adecuado de aplicación de las políticas favorece su adopción. Los enfoques existentes para determinar el ámbito de aplicación, como la jerarquización basada en el riesgo, tendrán que evolucionar para evitar una aplicación insuficiente o excesiva.

4. Los mecanismos de las políticas que se centran en la transparencia deben ser detallados y tener una audiencia adecuada para respaldar la responsabilidad.

5. La participación pública apoya las políticas que satisfacen las necesidades de las comunidades afectadas. Las políticas deben dar prioridad a la participación pública como un objetivo político fundamental, con el apoyo de recursos adecuados y estrategias formales de participación pública.

⁵⁵ El acceso a la información pública sólo puede ser limitado cuando concurra alguna de las excepciones previstas en esta ley, las excepciones pueden leerse en el Artículo N°10 de la misma <http://servicios.infoleg.gob.ar/infolegInternet/anexos/265000-269999/265949/norma.htm>

⁵⁶ Ada Lovelace Institute, AI Now Institute and Open Government Partnership. (2021). Algorithmic Accountability for the Public Sector. Available at: <https://www.opengovpartnership.org/documents/algorithmic-accountability-public-sector>

6. Las políticas se benefician de la coordinación institucional entre sectores y niveles de gobernanza para crear coherencia en la aplicación y aprovechar los diversos conocimientos.

Registros o Repositorios Algorítmicos

Varios países en el mundo están comenzando a adoptar otra forma de transparencia novedosa y eficiente: los repositorios o registros algorítmicos. Estos son repositorios disponibles para que la ciudadanía se informe sobre qué algoritmos se están implementando en las diversas áreas de gobierno. Un ejemplo de este tipo de registros es el sitio “Algoritmos públicos” creado por el Gob_Lab de la Universidad Adolfo Ibáñez en Chile⁵⁷, donde se categoriza a los algoritmos según su área: Servicios Públicos Generales, Defensa, Seguridad, Economía, Medio Ambiente, Vivienda y Servicios Comunitarios, Salud, Actividades Recreativas, Cultura y Religión, Educación, Protección Social. A su vez, cada modelo algorítmico posee una ficha detallada que indica:

- A qué ministerio pertenece
- Quién lo está ejecutando
- El objetivo del modelo
- La descripción
- La fecha de inicio
- El estado actual conocido
- El financiamiento
- Las fuentes de información

Esta ficha es realmente útil para promover políticas de innovación transparentes, en especial si se trata de temas sensibles para la ciudadanía y cuando hay licitación de empresas de por medio. En línea con lo expuesto en este apartado, el Marco de Ética de los Datos del Reino Unido⁵⁸ fomenta la transparencia no sólo del modelo algorítmico, sino también de los procesos administrativos que hay detrás del sistema, incluidos los beneficios previstos, la estructura del equipo del proyecto, la publicación de los datos no personales y no sensibles utilizados en el sistema y una explicación del funcionamiento del mismo.

CHECKLIST

Transparencia

- ☐ ¿Qué acciones tengo en cuenta para garantizar la transparencia de la herramienta?
- ☐ ¿Hay instancias planificadas para explicar el uso de los datos y los resultados de los algoritmos a la población?
- ☐ ¿Cómo se podrían co-diseñar las herramientas de CDIA con la sociedad civil?
- ☐ ¿Qué partes de los desarrollos de estas herramientas y de los datos se pueden hacer públicas para que cualquier persona tenga acceso?

⁵⁷ Véase <https://www.algoritmospublicos.cl/>

⁵⁸ Véase <https://www.gov.uk/government/publications/data-ethics-framework>

PALABRAS CLAVE:

PREJUICIOS

DESBALANCE

ASIMETRÍA

DE PODER

INVISIBILIZACIÓN

¿Qué son los sesgos y los riesgos? ¿Cómo puedo abordar esta problemática?

A lo largo de esta hoja de ruta estuvimos reflexionando sobre los límites de la Ciencia de Datos e Inteligencia Artificial y por lo tanto es necesario hablar de **sesgos**. Para este fin vamos a centrarnos en dos definiciones o modos de ver la problemática de los sesgos en el mundo de los datos y la IA:

DEFINICIÓN A

En el dominio de las tecnologías digitales el uso de la palabra sesgo es difuso. Por un lado, suele ser la traducción en español de la palabra en inglés *bias*, que en ocasiones también es traducida como prejuicio, vocablo que tiene en el hablar cotidiano una connotación negativa. Por otro lado, como los algoritmos en sentido estrecho están íntimamente ligados a modelos estadísticos, algunas personas piensan los sesgos en los algoritmos tal como en ocasiones se piensan los sesgos estadísticos, esto es, como errores que se pueden detectar ex post facto en un estudio y que se producen en alguna de sus etapas, ya sea en la recolección, el análisis, la interpretación o la revisión de los datos.

Balmaceda, T. y Pedace, K. (2022)⁵⁹

DEFINICIÓN B

El sesgo de los datos se ha definido como “una distorsión sistemática en los datos” que puede medirse “contrastando una muestra de los datos con muestras de referencia extraídas de diferentes fuentes o contextos”. Esta definición deja entrever una premisa importante: *que hay un valor verdadero y absoluto en los datos y que el sesgo es una “distorsión” de ese valor*. Esta premisa motiva ampliamente los enfoques para “erradicar el sesgo” de los datos y los sistemas de aprendizaje automático. Sin embargo, argumentamos que el problema de esta suposición es que los datos nunca representan una verdad absoluta. Los datos, al igual que la verdad, son el producto de relaciones sociales subjetivas y asimétricas.

Traducción propia de Miceli, M. et al (2022)⁶⁰

59 Pedace, K., Balmaceda, T., Lawler, D., Pérez, D., Zeller, M. (2022) Pensar la tecnología digital con perspectiva de género. En: https://drive.google.com/file/d/1TI4_dn1-fxCTiL61D3BvoEcaZB47DxLQ/view?usp=sharing

60 Miceli, M., Posada, J., Yang, T. (2022) Studying Up Machine Learning Data: Why Talk About Bias When We Mean Power?. At Proceedings of the ACM on Human-Computer Interaction.

¿Pero por qué deberían importarnos estos sesgos, desbalances y prejuicios que se encuentran detrás de los sistemas de CDIA? En primer lugar porque nuestras políticas de innovación van a afectar directa o indirectamente a un sector de la población y tendrán un impacto en sus vidas. Queremos que la herramienta que vayamos a utilizar no discrimine ni deje afuera (aunque sea sin malas intenciones) a ningún sector de la población. Por ejemplo, si los datos que tenemos sobre el género de las personas son binarios (varón-mujer) y queremos trabajar con políticas hacia la población LGBTQ+, es probable que esta población no esté representada en los datos con los que vamos a entrenar nuestros algoritmos.

En segundo lugar, porque en la gran mayoría de los casos los datos que vamos a utilizar provienen de la sociedad, y como mencionamos antes, las relaciones sociales son subjetivas, están atravesadas por prejuicios y por asimetrías de poder. Hay muchas poblaciones que estuvieron históricamente subrepresentadas en los datos y esto llevó a que haya productos, políticas y programas que tuvieron grandes problemas en su desarrollo, y estas desigualdades pueden analizarse con diferentes tipos de sesgos.

A su vez, en algunos casos, los datos con los que contamos no conforman una muestra representativa de la realidad que queremos observar. Un ejemplo extremo podría ser: realizamos una encuesta para saber si a la gente le gusta contestar encuestas y obtuvimos un 80% de respuestas positivas ¿Podemos deducir que a un 80% de la población le gusta completar encuestas? La respuesta es no, y se debe a que la información que releva la encuesta va a estar indefectiblemente sesgada, representando principalmente a gente a la que le gusta completar encuestas.

Sin ir tan al extremo, podemos mencionar otros ejemplos como el estudio sobre la predicción de embarazos adolescentes, donde debemos tener en cuenta cómo se puede obtener información sobre estos embarazos ¿De qué manera registramos estos datos? ¿Todas las personas tienen la misma predisposición a brindar esta información? Si por ejemplo las mujeres de un determinado sector geográfico tienen más recursos que las de otro para ocultar la realización de un aborto, nuestros datos van a reflejar esta diferencia. Sería un error deducir que en un barrio hay más probabilidad de embarazos adolescentes que en el otro.

Sesgos de género, raza/etnia, capacidad, económicos

En el artículo “¿Por qué necesitamos datos con perspectiva de género?”⁶¹ el Observatorio DataGénero explica que los sesgos de género se encuentran cuando las mujeres y las personas LGBTQ+ son borradas de los datos o están subrepresentadas, cuando los equipos que trabajan con datos no realizan su trabajo desde una perspectiva sensible al género, o cuando se asume que el universal detrás de los datos son varones cis heterosexuales. En su libro “Mujeres invisibles” Caroline Criado-Pérez hace referencia a la representación del cuerpo masculino, como el cuerpo humano en la medicina y afirma que este supuesto persiste en la actualidad y esto se puede ver en los textos de medicina, en el imaginario social y en muchas pruebas médicas para testear nuevas drogas⁶².

61 Véase <https://datagenero.medium.com/por-qu%C3%A9-necesitamos-datos-con-perspectiva-de-g%C3%A9nero-78f75a83a5f1>

62 Criado-Perez, C. (2019). Invisible women: data bias in a world designed for men. New York: Abrams Press.

Siguiendo en la misma línea que los sesgos de género, el suponer que el cuerpo masculino es el cuerpo humano es cierto y también hay que agregarle que ese cuerpo masculino es blanco. Históricamente los cuerpos racializados fueron segregados y excluidos. D'Ignazio y Klein en su libro "Data Feminism" explican el caso de Serena Williams que tuvo "complicaciones" durante su embarazo y al hacer público este problema, muchísimas mujeres negras comenzaron a hablar sobre sus complicaciones también. Williams escribió que "las mujeres negras tienen 3 veces más probabilidad que las mujeres blancas de morir por causas relacionadas con el embarazo o el parto".⁶³

Desde una mirada interseccional podemos ver que poblaciones históricamente discriminadas o invisibilizadas pueden sufrir graves problemas si no se las tiene en cuenta a la hora de recolectar datos, hacer pruebas y generar políticas públicas. Así como asumimos que el cuerpo humano es varón blanco también asumimos que es un cuerpo sin ninguna discapacidad, un cuerpo de una persona heterosexual, de clase media, con educación y alfabetizada digitalmente. Estas suposiciones suelen dejar muchos grupos fuera: personas con discapacidad, personas que no saben usar tecnología o que no están alfabetizadas, personas pobres.

Por estos sesgos tan arraigados en la sociedad es que es necesario tener una mirada atenta y sensible ante diferentes colectivos invisibilizados y asegurarnos de su correcta representación en los datos de entrenamiento y también en la construcción del problema que estamos intentando solucionar.

CHECKLIST

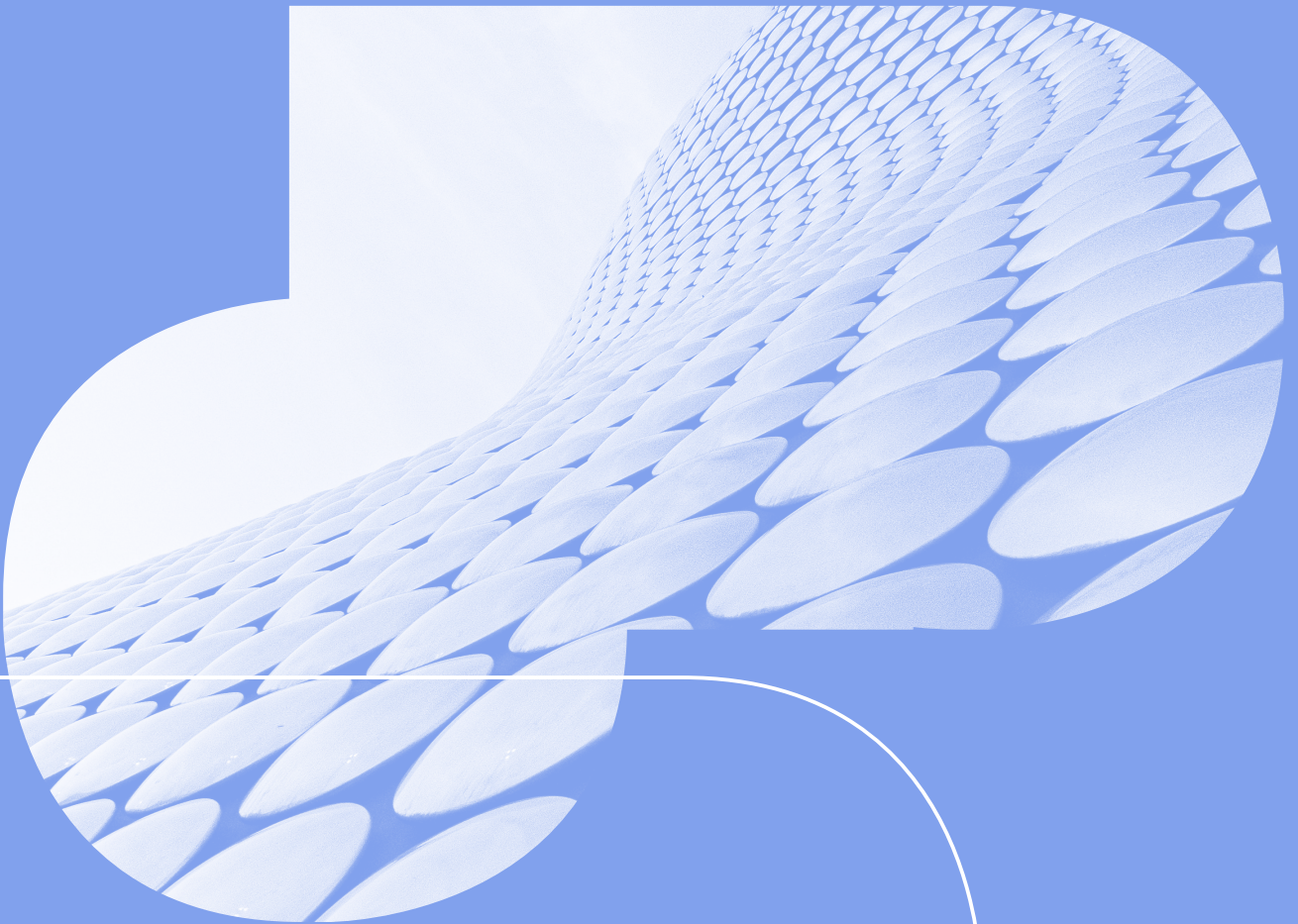
Sesgos

- ☐ ¿Puedo identificar algún sesgo en mis datos?
- ☐ ¿Existen prejuicios en mi equipo con respecto a la necesidad o problema que queremos abordar? ¿Cuáles son?
- ☐ ¿Qué riesgos encontramos en la implementación de la herramienta o ante los resultados que brinda?
- ☐ ¿Qué acciones concretas puedo realizar para mitigar y balancear los sesgos de los datos que uso y del modelo? ¿Tengo previsto exponerlos y/o validarlos por fuera de mi propio equipo?

63 D'Ignazio, C., & Klein, L. F. (2020). Data feminism. The MIT Press.



Conclusión
Checklist
Bibliografía



Conclusión

Vivimos en una era de innovación tecnológica sin precedentes. Los gobiernos poco a poco se van modernizando, van adquiriendo nuevas herramientas y van optimizando sus procesos para lograr mejores políticas públicas. La Ciencia de Datos e Inteligencia Artificial aplicadas al sector público son algunas de estas herramientas que si son bien utilizadas pueden ayudar a comprender mejor ciertos problemas, mejorar muchos procesos y darnos claridad donde antes no la había.

El objetivo de esta hoja de ruta es brindar herramientas y una guía concreta para implementar herramientas CDIA en el sector público. Propusimos una serie de pasos para facilitar la aproximación al tratamiento de los datos, la Ciencia de Datos y la Inteligencia Artificial a través de metodologías como el pensamiento de diseño y la observación participante que abren paso a los procesos de innovación pública. Con conceptos, pasos, prácticas y problemáticas comunes a un recorrido completo de CDIA, es que podemos generar innovaciones robustas, que promuevan acciones para integrar esa mayor complejidad, interdependencia y dinamismo de las problemáticas sociales. Para esto es necesario contar con la mayor disponibilidad de información y la posibilidad de generar mayor conocimiento, rapidez y asertividad en la toma de decisiones públicas. Lejos de idealizar estas herramientas, buscamos problematizarlas, abordarlas desde una mirada integral y crítica para lograr proyectos que sean de calidad, sensibles a las problemáticas sociales y que puedan perdurar en el tiempo o servir de base para los próximos pasos.

Hasta aquí llegamos con esta hoja de ruta, con estas recomendaciones y exploración hacia terrenos innovadores en el sector público. Si se van con más preguntas que antes de comenzar a leer este documento, hemos cumplido nuestro objetivo ¡hay mucho trabajo previo por hacer a la hora de emprender este viaje! Esperamos que esta guía sirva como espacio de consulta, de debate y reflexión para impulsar la agenda de innovación en el sector público, de la mano de aplicaciones de CDIA.

A continuación dejamos un listado de “infaltables” para cuando estén armando las valijas para emprender el viaje: todas aquellas preguntas que nos fuimos haciendo a lo largo de este documento y que son indispensables para pensar en herramientas de CDIA sensibles, integrales y duraderas.

Checklist para implementar Ciencia de Datos e Inteligencia Artificial en el sector público

Definición del problema o la necesidad

- ☐ ¿Qué quiero resolver o mejorar?
 - ☐ ¿Cuál es el alcance de esta problemática?
 - ☐ ¿Cuál es la población destinataria?
 - ☐ ¿Tengo en claro qué impactos genera? ¿Qué impactos quiero generar? ¿Hay algún aspecto negativo entre los posibles impactos?
 - ☐ ¿Qué es lo indispensable que debo comunicar o difundir en las distintas instancias?

Herramientas tecnológicas y el uso de CDIA

- ☐ ¿Se puede vincular esta problemática con las herramientas tecnológicas que quiero utilizar? ¿De qué manera?
 - ☐ ¿Sería algo novedoso? ¿O una mejora de algo existente?
 - ☐ ¿Cómo me ayudaría la herramienta ideal a mejorar la gestión, la política pública o el servicio?
 - ☐ ¿La herramienta ideal implica modificar algún proceso de trabajo o circuito de gestión? ¿Sobre qué/ quiénes impactaría en este último caso? ¿Cómo podríamos trabajar esta transición?
- ☐ ¿Tengo datos sobre la problemática? ¿Cuáles necesitaría sistematizar y/o recolectar?
- ☐ ¿La herramienta ideal utilizará CD, aprendizaje automático o la IA?
 - ☐ ¿Serían las mejores herramientas para ayudar a resolver este problema?
 - ☐ ¿Qué otras herramientas o abordajes podría utilizar?
 - ☐ ¿Es posible que esta herramienta sea usada de manera incorrecta y genere efectos no deseados? ¿Cuáles? ¿Son evitables?

- ☐ ¿Pienso en desarrollar o en adquirir y adaptar la herramienta de CDIA?
 - ☐ En caso de adquirirla, ¿existe documentación y acompañamiento para su implementación? ¿Quién del equipo la probaría o la adaptaría?

Equipo de trabajo

- ☐ ¿Mi equipo de trabajo es interdisciplinario? ¿Qué capacidades necesito reforzar?
- ☐ ¿Hay planificadas instancias de intercambio y trabajo en conjunto entre el equipo experto en la problemática y el equipo técnico?

Conocimiento en profundidad de los datos

- ☐ Si quiero implementar un modelo de aprendizaje automático o IA:
 - ☐ ¿Tengo los datos que requiere el algoritmo?
 - ☐ ¿Conozco en profundidad mi set de datos?
 - ☐ ¿Con qué información adicional me gustaría contar?
 - ☐ ¿Mis datos tienen calidad?
 - ☐ ¿Qué preguntas tengo que hacerme para comprender mis datos en profundidad?

Infraestructura

- ☐ ¿Qué infraestructura necesito para realizar y mantener una herramienta de CDIA?
 - ☐ ¿Dónde están almacenados los datos?
 - ☐ ¿Van a ser necesarios servidores externos para alojarlos o realizar operaciones sobre éstos?
- ☐ ¿Qué medidas de seguridad estamos planificando para que estos datos estén resguardados y no haya filtraciones?

- ☐ ¿Con qué equipamiento cuento y qué necesito? ¿Las computadoras actuales permiten leer los archivos y procesar los datos que necesito?

Documentación

- ☐ ¿Estoy planificando documentar mi herramienta para que sea replicable por otros equipos? ¿Y documentar las decisiones sobre los datos y el proceso?
- ☐ ¿Tengo pensado desarrollar un manual de uso de la misma para personas no técnicas? ¿Qué capacidades adicionales necesitamos para producir este tipo de manual (por ejemplo: expertise en didáctica, en redacción simple, etc.)?
- ☐ ¿Cómo y dónde voy a documentar el proyecto? ¿Hay una persona del equipo designada? ¿Es un proceso colaborativo?
- ☐ ¿Qué herramientas tengo a disposición para hacerlo?
- ☐ ¿Tengo en cuenta que la documentación tiene que ser interdisciplinaria y no solamente del equipo técnico?
 - ☐ Si otro equipo toma la herramienta que diseñé y la documentación: ¿puede replicarla o editarla fácilmente? ¿Qué acciones debemos tomar para que si el equipo cambia, el proyecto siga funcionando?

Impacto

- ☐ ¿Podemos prever el impacto desde el diseño de la herramienta?
- ☐ ¿Cómo podemos planificar una evaluación de impacto de la herramienta?
- ☐ ¿Está prevista la difusión o capacitación sobre la herramienta? ¿Cuenta con posibilidad de escalamiento?

Revisión y monitoreo continuo de la herramienta/proyecto

- ☐ ¿Qué instancias de revisión o monitoreo continuo estamos considerando para este proyecto?
- ☐ ¿Con cuánta periodicidad vamos a monitorear el desempeño de estas herramientas?
 - ☐ Si identifico que hay nuevos datos no contemplados en la fase de aprendizaje del modelo, ¿podemos pautar un análisis de los nuevos datos? ¿Tengo prevista la actualización del modelo en función de eso?

Algunas instancias de revisión:

- ☐ En el inicio del proyecto: ¿estamos haciéndolo de manera correcta?
- ☐ Durante el proyecto: ¿lo hemos diseñado bien?
- ☐ Después del proyecto: ¿sigue funcionando como queríamos que lo hiciera cuando lo diseñamos?⁶⁴

Sustentabilidad y ambiente

- ☐ ¿Conocemos el impacto ambiental de las herramientas tecnológicas que vamos a utilizar? ¿Y de los equipamientos e infraestructura?
- ☐ ¿Estamos teniendo en cuenta este impacto a la hora de elegir las herramientas a utilizar?

Normativa / leyes

- ☐ ¿Es la herramienta segura y se adecúa a las legislaciones vigentes en materia de seguridad y privacidad de la información?
- ☐ Si se trabaja con datos sensibles, ¿qué medidas de seguridad se están utilizando para protegerlos?
- ☐ ¿El uso de estos datos interfiere con los derechos de las personas cuyos datos están siendo utilizados? Si es así, ¿Hay alguna forma menos intrusiva de llegar al objetivo deseado?
- ☐ A partir de la herramienta ¿necesito generar algún registro formal?

⁶⁴ Véase www.gov.uk/government/publications/data-ethics-framework

Transparencia

- ☐ ¿Qué acciones tengo en cuenta para garantizar la transparencia de la herramienta?
- ☐ ¿Hay instancias planificadas para explicar el uso de los datos y los resultados de los algoritmos a la población?
- ☐ ¿Cómo se podrían co-diseñar las herramientas de CDIA con la sociedad civil?
- ☐ ¿Qué partes de los desarrollos de estas herramientas y de los datos se pueden hacer públicas para que cualquier persona tenga acceso?

Sesgos

- ☐ ¿Puedo identificar algún sesgo en mis datos?
- ☐ ¿Existen prejuicios en mi equipo con respecto a la necesidad o problema que queremos abordar? ¿Cuáles son?
- ☐ ¿Qué riesgos encontramos en la implementación de la herramienta o ante los resultados que brinda?
- ☐ ¿Qué acciones concretas puedo realizar para mitigar y balancear los sesgos de los datos que uso y del modelo? ¿Tengo previsto exponerlos y/o validarlos por fuera de mi propio equipo?

Referencias

Ada Lovelace Institute, AI Now Institute and Open Government Partnership. (2021). Algorithmic Accountability for the Public Sector. Available at: <https://www.opengovpartnership.org/documents/algorithmic-accountability-public-sector/>

BID (2021). Uso responsable de IA para política pública: Manual de formulación de proyectos. En: <https://publications.iadb.org/publications/spanish/document/Uso-responsable-de-IA-para-politica-publica-manual-de-formulacion-de-proyectos.pdf>

CAF (2021). Experiencia: Datos e Inteligencia Artificial en el sector público. Caracas: CAF. Retrieved from <http://scioteca.caf.com/handle/123456789/1793>.

Crawford, K. (2021). Atlas of AI Power, Politics, and the Planetary Costs of Artificial Intelligence. Yale University Press.

Criado-Perez, C. (2019). Invisible women: data bias in a world designed for men. New York: Abrams Press.

D'Ignazio, C. y Klein, L. F. (2020). Data feminism. The MIT Press.

González, F., Ortiz, T., Sánchez Ávalos, R. (2020). Uso responsable de la IA para las políticas públicas: Manual de ciencia de datos. BID.

Hey, J. (2004). The data, information, knowledge, wisdom chain: the metaphorical link. Intergovernmental Oceanographic Commission, 26, 1-18. ISO 690

Luvini, Paula (2022). Guía práctica para la protección de datos. Buenos Aires: Fundar. Disponible en <https://www.fundar.ar>

Miceli, M., Posada, J., Yang, T. (2022) Studying Up Machine Learning Data: Why Talk About Bias When We Mean Power?. At Proceedings of the ACM on Human-Computer Interaction.

OECD (2022). The OECD Framework for the Classification of AI systems. OECD.IA Policy Observatory.

Ortegón, E., Pacheco, J. F., Prieto, A. (2005). Metodología del marco lógico para la planificación, el seguimiento y la evaluación de proyectos y programas. CEPAL.

Ortiz Freuler, J., Iglesias, C. (2018). Algoritmos e Inteligencia Artificial en Latin America: Un Estudio de implementaciones por parte de Gobiernos en Argentina y Uruguay, World Wide Web Foundation. En: https://webfoundation.org/docs/2018/09/WF_AI-in-LA_Report_Spanish_Screen_AW.pdf

Pedace, K., Balmaceda, T., Lawler, D., Pérez, D., Zeller, M. (2022) Pensar la tecnología digital con perspectiva de género. En: https://drive.google.com/file/d/1TI4_dn1-fxCTiL61D3BvoEcaZB47DxLQ/view?usp=sharing

Pedace, K., Balmaceda, T., Lawler, D., Pérez, D., Zeller, M. (2019) Caja de herramientas humanísticas. Taller-Inteligencia Artificial en América Latina y el Caribe: ética, gobernanza y aplicaciones Lugar: Montevideo.

Peña P. (2022) Ese Futuro tecnológico no existe. Instituto Latinoamericano de Terraformación. En: https://user.fm/files/v2-5f59336b831517c05a164405cee8d931/IPCC%202022_Ese%20futuro%20tecnol%C3%B3gico%20no%20existe_ILDT.pdf

Peña, P., Varon, J. (2019). Decolonising AI: a transfeminist approach to data and social justice, en Global Information Society Watch, disponible en <https://www.giswatch.org/node/6203>

UK Government. (2021). Department for Digital, Culture, Media and Sport. Data Ethics Framework. (2021) Available at: <https://www.gov.uk/government/publications/data-ethics-framework>

Zurbriggen, C., Lago, M. G. (2015). Co-creando valor público. Desafíos pendientes para América Latina. *Revista Iberoamericana de Ciencia, tecnología y Sociedad-CTS*, 10(30), 143-171.

Links de interés

Anatomía de sistemas de IA <https://anatomyof.ai/>

Barómetro Global de Datos <https://globaldatabarometer.org/>

Framework de ética de datos UK

<https://www.gov.uk/government/publications/data-ethics-framework>

Guías universales de la IA <https://thepublicvoice.org/ai-universal-guidelines/>

Guía de IA de la UNESCO <https://en.unesco.org/artificial-intelligence>

Guía para la publicación de datos abiertos del GCBA https://datosgobar.github.io/paquete-apertura-datos/guia_abiertos/#estructura-y-caracteristicas-de-los-datos-tabulares

Guía para el uso de IA en el sector público UK <https://www.gov.uk/government/collections/a-guide-to-using-artificial-intelligence-in-the-public-sector>

Instituto Latinoamericano de Terraformación <https://terraforminglatam.net/>

Perfil de Aplicación Nacional de Metadatos para Datos Abiertos de GCBA <https://datosgobar.github.io/paquete-apertura-datos/perfil-metadatos/>

¿Por qué necesitamos datos con perspectiva de género? <https://datagenero.medium.com/por-qu%C3%A9-necesitamos-datos-con-perspectiva-de-g%C3%A9nero-78f75a83a5f1>

Principios éticos de IA de la OECD <https://oecd.ai/en/ai-principles>

Reporte rendición de cuentas algorítmicas del Instituto Ada Lovelace <https://www.adalovelaceinstitute.org/report/algorithmic-accountability-public-sector/>

Recomendaciones éticas de IA de UNESCO <https://unesdoc.unesco.org/ark:/48223/pf0000381137/PDF/381137eng.pdf.multi>

Footnotes

1 CAF (2021). Experiencia: Datos e Inteligencia Artificial en el sector público. Caracas: CAF. Retrieved from <http://scioteca.caf.com/handle/123456789/1793>

2 Para más información sobre estos tres conceptos, se puede consultar la página 23 de CAF (2021). Experiencia: Datos e Inteligencia Artificial en el sector público. Caracas: CAF. <http://scioteca.caf.com/handle/123456789/1793>

3 Véase: <https://thepublicvoice.org/ai-universal-guidelines/>

<https://en.unesco.org/artificial-intelligence>

<https://oecd.ai/en/ai-principles>

<https://www.gov.uk/government/collections/a-guide-to-using-artificial-intelligence-in-the-public-sector>

4 Ada Lovelace Institute, AI Now Institute and Open Government Partnership. (2021). Algorithmic Accountability for the Public Sector. Available at: <https://www.opengovpartnership.org/documents/algorithmic-accountability-public-sector/>

5 Traducción propia del reporte de Ada Lovelace Institute, AI Now Institute and Open Government Partnership. (2021). Algorithmic Accountability for the Public Sector. Available at: <https://www.opengovpartnership.org/documents/algorithmic-accountability-public-sector/>

6 CAF (2021). Experiencia: Datos e Inteligencia Artificial en el sector público. Caracas: CAF. Retrieved from <http://scioteca.caf.com>

Autoría de las fotos

Fotos de Ash @ModernAfflatus, Olga Lioncat y Oleg Magni. <https://www.pexels.com/>



Innovar con
Ciencia de Datos
en el sector
público

Fundación
SADOSKY

 Ministerio de Ciencia,
Tecnología e Innovación
Argentina

INNOVACIONPUBLICACONDATOS.FUNDATIONSADOSKY.ORG.AR