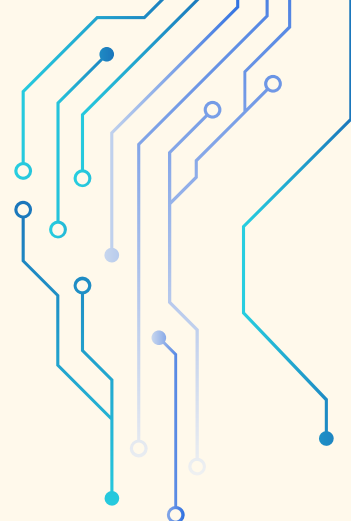




Uso responsable de IA para política pública: Manual de formulación de proyectos

Gabriela Denis
María Paz Hermosilla
Claudio Aracena
Roberto Sánchez Ávalos
Natalia González Alarcón
Cristina Pombo



Uso responsable de IA para política pública: Manual de formulación de proyectos

Gabriela Denis
María Paz Hermosilla
Claudio Aracena
Roberto Sánchez Ávalos
Natalia González Alarcón
Cristina Pombo

Septiembre de 2021

<https://www.iadb.org/>

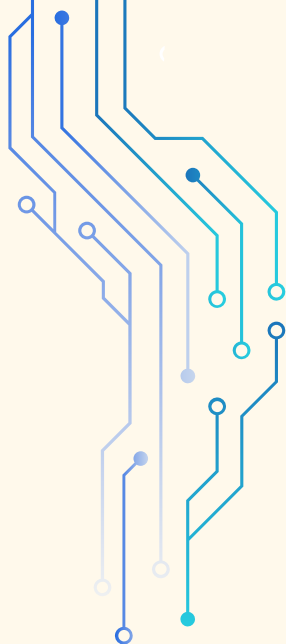
Copyright © 2021 Banco Interamericano de Desarrollo. Esta obra se encuentra sujeta a una licencia Creative Commons IGO 3.0 Reconocimiento-NoComercial-SinObrasDerivadas (CC-IGO 3.0 BY-NC-ND) (<https://creativecommons.org/licenses/by-nc-nd/3.0/igo/legalcode>) y puede ser reproducida para cualquier uso no-comercial otorgando el reconocimiento respectivo al BID. No se permiten obras derivadas.

Cualquier disputa relacionada con el uso de las obras del BID que no pueda resolverse amistosamente se someterá a arbitraje de conformidad con las reglas de la CNUDMI (UNCITRAL). El uso del nombre del BID para cualquier fin distinto al reconocimiento respectivo y el uso del logotipo del BID no están autorizados por esta licencia CC-IGO y requieren un acuerdo de licencia adicional.

Note que el enlace URL incluye términos y condiciones adicionales de esta licencia.

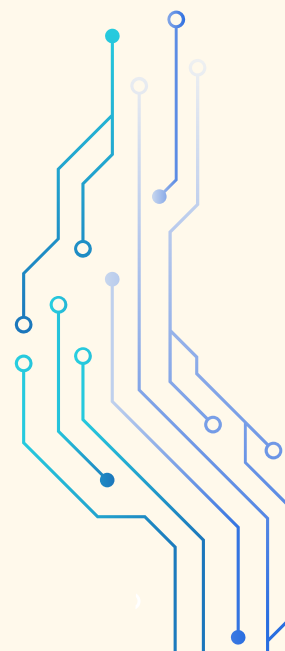
Las opiniones expresadas en esta publicación son de los autores y no necesariamente reflejan el punto de vista del Banco Interamericano de Desarrollo, de su Directorio Ejecutivo ni de los países que representa.





Agradecimientos

Expresamos nuestra gratitud a Erika Quiroz, Sebastián Elgueta, Rodrigo Moya, Claudio Collao, José Tomás Arenas y Romina Garrido, quienes proporcionaron insumos relevantes para el desarrollo del manual. Igualmente a Patricia Ardila por su labor de edición y a Alejandro Scaff por el diseño.



CONTENIDO

ACERCA DEL MANUAL

Iniciativa fAIr LAC

El Banco Interamericano de Desarrollo (BID), en colaboración con socios y aliados estratégicos, lidera la iniciativa fAIr LAC¹ mediante la cual busca promover la adopción responsable de la inteligencia artificial (IA) y de los sistemas de soporte de decisiones. Esto con el fin de mejorar la prestación de servicios sociales y crear oportunidades de desarrollo en aras de atenuar la desigualdad social. Este manual hace parte de una serie de documentos y herramientas dirigidos a guiar a los responsables² de la formulación de políticas públicas y a sus equipos técnicos en la mitigación de los retos inherentes a los sistemas de soporte de decisiones basados en IA y en la promoción de su adopción responsable (Cabrol et al., 2020).

¿Por qué este manual?

Con este manual se pretende ayudar a los responsables de formular proyectos con sistemas de soporte basados en IA a realizar su conceptualización y diseño, así como a liderar posteriormente su ejecución y monitoreo. La IA es una herramienta muy potente que puede ayudar a solucionar asuntos complejos, siempre y cuando esté contextualizada dentro del problema de política pública que se busca resolver y se aborden los temas éticos y legales que entraña la aplicación de herramientas automáticas de toma de decisiones.

¿Para quién es este manual?

Este manual está dirigido a los responsables de la toma de decisiones (gerentes, directivos o profesionales no necesariamente expertos en ciencia de datos) de instituciones públicas que lideren proyectos de IA desde la fase de diseño hasta su implementación. El documento se divide en dos partes principales: **conceptualización y diseño**, cuyas labores estarán a cargo del director de proyecto (responsable por la toma de decisiones), y **ejecución**, donde este último desempeña un rol crítico en las tareas que llevará a cabo conjuntamente con un equipo multidisciplinario compuesto por el equipo técnico de desarrolladores de modelos de IA, expertos sectoriales y especialistas jurídicos, entre otros.

Este manual complementa el manual de ciencia de datos dirigido principalmente a los equipos técnicos, disponible en <https://publications.iadb.org/es/ia-responsable-manual-tecnico-ciclo-de-vida-de-la-inteligencia-artificial>.

¹ Para más información, consulte <https://fairlac.iadb.org/>

² Por razones exclusivamente de estilo, en este documento se usa el masculino genérico inclusivo, independientemente del género gramatical. Los cargos y funciones corresponderán indistintamente a personas de cualquier sexo.

INTRODUCCIÓN

Actualmente, los sistemas de soporte de decisiones basados en inteligencia artificial (IA) permiten procesar cantidades masivas de datos para generar recomendaciones, predicciones o clasificaciones que pueden ser usadas para mejorar distintos procesos. Cada vez estamos más familiarizados con sus aplicaciones en nuestra vida cotidiana: desde sugerencias sobre posibles amistades en nuestras redes sociales o sobre series o películas que nos gustaría ver en las plataformas de transmisión continua (*streaming*), hasta avisos publicitarios destinados específicamente a responder a nuestros intereses y gustos, pasando por el empleo de técnicas de reconocimiento facial para desbloquear nuestros dispositivos electrónicos.

Sin embargo, la IA también se puede emplear para solucionar problemas de política pública, y son cada vez más las instituciones que están implementando soluciones basadas en ella con las cuales se busca generar un impacto positivo en el bienestar de la sociedad.

Las herramientas de IA, y la capacidad que tienen las computadoras modernas de procesar información, han permitido que aquellas se apliquen en los distintos quehaceres de las instituciones, solucionando tanto problemas internos como externos. A continuación se mencionan algunos de los usos de la IA en entidades públicas que se ocupan de diversos campos de actividad de la sociedad:

- Asignación de profesores o estudiantes a escuelas.³
- Prevención o tratamiento oportuno de enfermedades mediante su detección temprana.⁴
- Acoplamiento entre vacantes y candidatos a puestos de trabajo utilizando información más completa.⁵
- Respuesta automática a solicitudes de usuarios y beneficiarios de una institución.
- Focalización de entrega de subsidios y beneficios a la población destinataria.
- Supervisión de uso de permisos de aprovechamiento de fuentes de agua por medio del análisis de imágenes.
- Supervisión de la contaminación producida por las empresas con base en datos de estaciones de monitoreo.
- Focalización del patrullaje policial.
- Predicción de tiempos de desplazamiento y de espera en el transporte público.

En el documento “La inteligencia artificial al servicio del bien social en América Latina y el Caribe: panorámica regional e instantáneas de doce países” de la iniciativa fAIr LAC⁶ (Gómez, del Pozo, Martínez, y Martín del Campo, 2020) se puede encontrar una descripción detallada de algunas iniciativas de aplicación de IA para el bien social en la región.

3 Para más información acerca de la aplicación de IA en educación, véase [aquí](#).

4 Para más información acerca de la aplicación de IA en salud, véase [aquí](#).

5 Para más información sobre el uso de IA en sistemas de intermediación laboral, véase [aquí](#).

6 Para más información sobre la iniciativa, véase <https://fairlac.iadb.org/>

Sistemas de toma y/o soporte de decisiones y aprendizaje automático

La OCDE describe a los sistemas de soporte de decisiones como “sistemas computacionales que pueden, para un determinado conjunto de objetivos definidos por los seres humanos, hacer predicciones y recomendaciones o tomar decisiones que influyen en entornos reales o virtuales.” Estos sistemas están diseñados para operar con distintos niveles de autonomía (OECD, 2019).

En este manual se busca abordar los aspectos más comunes en torno al uso de los sistemas de soporte y toma de decisiones desde la perspectiva de las personas a cargo, entre ellos determinar la viabilidad del proyecto, detectar sesgos, y evaluar la posibilidad de que se produzcan resultados indeseables para la sociedad o para una institución particular.

Aunque los métodos de aprendizaje automático (AA) no son el único tipo de algoritmo que pueden utilizar los sistemas de IA, sí son los que han registrado un mayor crecimiento durante los últimos años. Se trata de un conjunto de técnicas que habilitan a un sistema para que aprenda comportamientos de manera automatizada a través de patrones e inferencias, y no de instrucciones explícitas o simbólicas introducidas por seres humanos (OECD, 2019). Se consideran dos arquetipos de uso del aprendizaje automático en los procesos decisorios ([González, Ortiz y Sánchez Ávalos, 2020](#)): los sistemas de soporte de decisiones y los sistemas de toma de decisiones.

Sistemas de soporte de decisiones: Relacionados con el concepto de inteligencia asistida o aumentada, comprenden aquellos sistemas donde la información generada por los modelos de AA se utiliza como insumo para la toma de decisiones por parte de seres humanos.

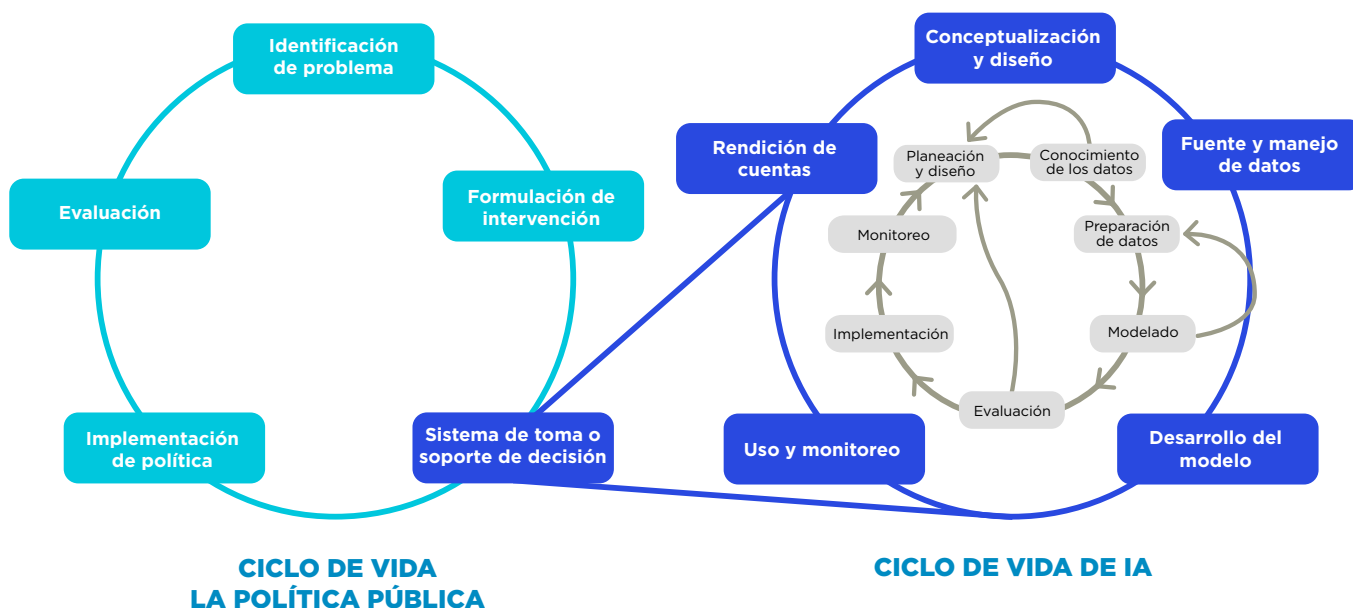
Sistemas de toma de decisiones: Relacionados con el concepto de inteligencia automatizada y autónoma, las decisiones finales y las acciones que de allí surjan se toman sin intervención humana directa. Esto significa que el sistema pasa a realizar tareas ejecutadas previamente por las personas. En muchos contextos se usa la sigla ADM para denominar estos sistemas (por las iniciales en inglés de *Automated Decision Making*).



Componentes de un sistema de IA para políticas públicas

La IA, o la creación de un sistema de toma y/o soporte de decisiones con base en ella, no reemplaza a la política pública dado que la IA por sí sola no soluciona ningún problema social. Se trata de una herramienta que se emplea durante el ciclo de vida de la política pública para proporcionar información en forma de predicción, clasificación y/o segmentación, entre otras posibilidades, en el contexto de la formulación de una intervención o acción de naturaleza social, como se observa en la figura 1:

Figura 1. Ciclo de vida de la política pública con IA



Fuente: González, Ortiz y Sánchez Ávalos, 2020.

El ciclo de vida de la política pública es una herramienta simplificada que busca representar la manera en que se deberían desarrollar las políticas, y que sirve para planificar y analizar las diferentes fases del proceso. El ciclo se inicia con la identificación y definición de un problema o asunto a resolver, para pasar luego a formular los diferentes cursos de acción. Allí el gobierno debe evaluar las alternativas propuestas, entre ellas la de mantener el statu quo, es decir, abstenerse de tomar medida alguna.

A lo largo de este proceso, la tecnología actúa como un instrumento que puede servir para desarrollar las alternativas de política pública propuestas a partir del contexto y la factibilidad de la solución. Una vez que el gobierno elige la alternativa y el instrumento para desarrollarla --por ejemplo a través de un sistema de toma o soporte de decisiones basado en IA--, se pondrá en marcha la fase de implementación. Por último, en el estadio final se busca evaluar la eficacia de la política en términos de sus objetivos, resultados e impactos previstos (Giorgi, 2017).

El ciclo de vida de la IA se activa en el momento en que se la identifica como el instrumento apropiado a incorporar dentro de la política pública que dará respuesta al problema seleccionado. Las etapas en el ciclo de vida de la IA son:

1. **Conceptualización y diseño:** Comprende los puntos claves acerca de los cuales las personas responsables por la toma de decisiones deben tener claridad antes de comenzar un proyecto.
2. **Recolección y procesamiento de los datos:** Comprende el proceso de limpieza y procesamiento de datos, así como la identificación de deficiencias y sesgos que puedan poner en riesgo el desarrollo del modelo.
3. **Desarrollo del modelo y validación:** Comprende los conceptos claves a seguir para tener sistemas robustos y validados de IA.
4. **Uso y monitoreo:** Comprende la evaluación de la herramienta una vez que ha comenzado la implementación.
5. **Rendición de cuentas:** Responde a la necesidad de proporcionar información y transparencia para fomentar la comprensión de la IA por parte de la ciudadanía.

En la interrelación de estos dos ciclos se genera un importante conjunto de retos para lograr una IA robusta y responsable, los cuales deberán ser evaluados y considerados durante el desarrollo y uso de tales sistemas. Asimismo existen retos transversales, entre ellos la transparencia, la rendición de cuentas, y la gobernanza, seguridad y protección de los datos personales. Por último están los retos relacionados con el diseño de la política pública y la definición de la intervención, y con la aplicación de criterios de necesidad y proporcionalidad en el uso de la IA, los cuales se presentan a lo largo del ciclo de vida de la IA.



El rol de la persona a cargo de la toma de decisiones de la política pública (director 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.

Figura 2. Ciclo de vida de la IA a la luz de la estructura de este manual

Fuente: Modificado a partir de González, Ortiz y Sánchez Ávalos, 2020.

En la medida en que afectan la gestión de la institución y a sus usuarios, los proyectos de IA requieren que haya una **persona a cargo de tomar las decisiones y un equipo multidisciplinario de profesionales responsables por la ejecución**. Cabe señalar que, en las distintas etapas del ciclo de vida de cada proyecto, se requerirá conocimiento de diversas áreas que pueden o no ser del dominio del equipo a cargo de implementarlo. En tales circunstancias, será necesario consultar a las personas expertas ya sea dentro de la institución o externamente.

Las dos funciones principales en la formulación y ejecución del proyecto se pueden describir de la siguiente manera:

- La persona responsable por tomar las decisiones o director del proyecto tiene a su cargo la formulación de políticas públicas y liderará el proyecto en su diseño y posterior implementación.
- El equipo técnico será el que se ocupe de analizar las fuentes de datos, desarrollar los modelos de IA y monitorear su uso.

Para acompañar el desarrollo de un sistema de toma y/o soporte de decisiones, se proponen herramientas tanto para el director del proyecto como para el equipo técnico. Estas formas (Anexos 1 a 5) deben llenarse de manera simultánea y hacen parte de un proceso de retroalimentación durante todo el ciclo de desarrollo.

Las siguientes son las herramientas para el responsable de la toma de decisiones de políticas públicas:

- **Ficha de diseño y factibilidad del proyecto:** Con esta herramienta se busca identificar los principales aspectos claves de un proyecto de IA para evaluar su viabilidad, determinar si la IA es la solución correcta para el problema y recolectar la información necesaria para el diseño, la cual deberá ser compartida con el equipo técnico y con todo el equipo multidisciplinario responsable del proyecto ([Anexo 1](#)).

- **Matriz de madurez de datos:** Permite una aproximación inicial a la calidad y relevancia de los datos que serán utilizados ([Anexo 2](#)).
- **Lista de verificación para el director del proyecto:** Herramienta que consolida las principales preocupaciones por dimensión de riesgo del ciclo de vida de IA desde de perspectiva de quien toma las decisiones ([Anexo 3](#)).

Las herramientas para el equipo técnico son:

- **Perfil de los datos:** Contiene los principales hallazgos de las bases de datos que se utilizarán con la herramienta. Se fundamenta en los análisis exploratorios de los datos por parte del equipo técnico⁷ ([Anexo 4](#)).
- **Perfil del modelo:** Descripción final de un modelo de IA que responda a los requerimientos de la política pública y que pueda llevarse a cabo de acuerdo con los datos disponibles ⁸ ([Anexo 5](#)).
- **Lista de verificación para el equipo técnico:** Herramienta que consolida las principales preocupaciones por dimensión de riesgo del ciclo de vida de IA desde la perspectiva del equipo técnico⁹.

Como se indicó al comienzo, este manual se enfoca principalmente en las herramientas arriba mencionadas correspondientes al director del proyecto, quien está a cargo de la toma de decisiones de política pública. Sin embargo, a lo largo del documento se mencionarán también las herramientas del equipo técnico, en la medida en que el director deberá supervisar su llenado por parte de aquel. Se sugiere que para ello el equipo técnico se apoye en el manual “[Uso responsable de IA para política pública: manual de ciencia de datos](#)” de González, Ortiz y Sánchez Ávalos (2020). Allí se describen las herramientas técnicas en mayor detalle, los principales riesgos que se presentan durante el desarrollo de un modelo y las medidas para mitigarlos.

A lo largo de este manual se utilizará el caso DART¹⁰—una solución de detección temprana de retinopatía diabética basada en IA— para ejemplificar tanto lo que tiene que ver con la conceptualización y diseño del proyecto como con su ejecución.

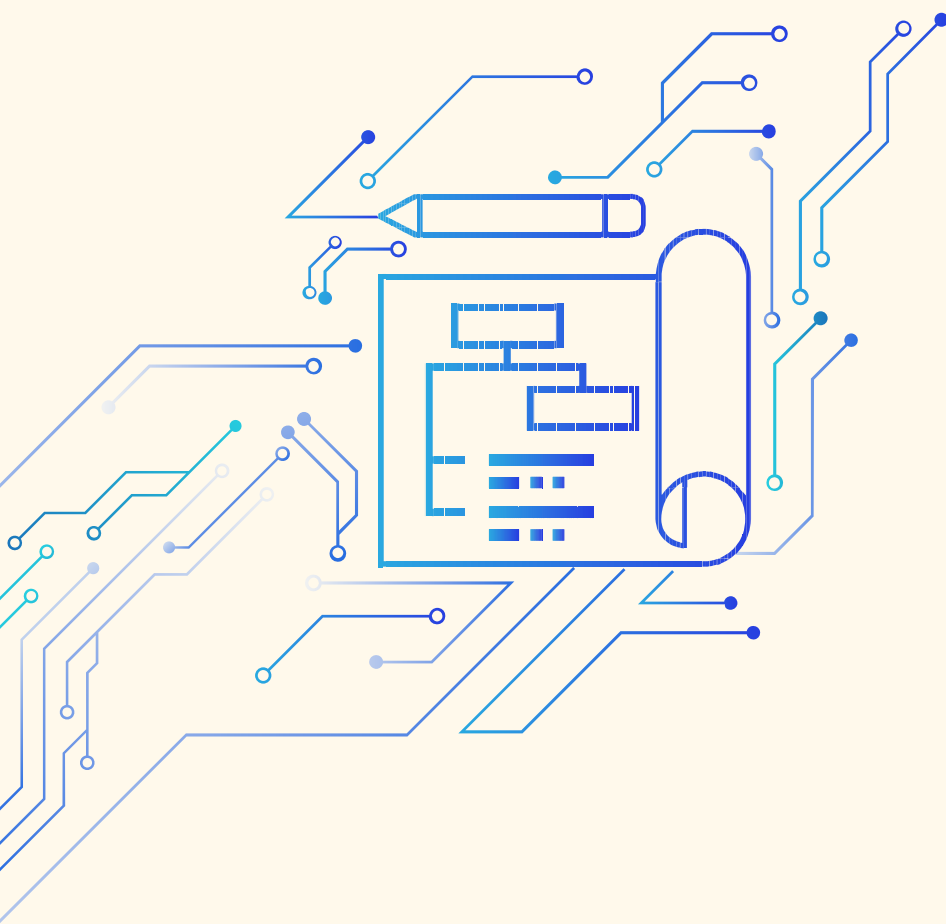
7 Parte de Manual Técnico-Ciclo de vida de la inteligencia artificial. <https://publications.iadb.org/es/ia-responsable-manual-tecnico-ciclo-de-vida-de-la-inteligencia-artificial>

8 Parte de Manual Técnico-Ciclo de vida de la inteligencia artificial. <https://publications.iadb.org/es/ia-responsable-manual-tecnico-ciclo-de-vida-de-la-inteligencia-artificial>

9 Parte de Manual Técnico-Ciclo de vida de la inteligencia artificial. <https://publications.iadb.org/es/ia-responsable-manual-tecnico-ciclo-de-vida-de-la-inteligencia-artificial>

10 DART (Chile): <https://www.teledx.org/dart/?lang=es>

PARTE 1. CONCEPTUALIZACIÓN Y DISEÑO



PARTE 1. CONCEPTUALIZACIÓN Y DISEÑO

La primera etapa en el ciclo de vida de la IA es la conceptualización y diseño del proyecto, la cual estará a cargo de su director, es decir, la persona responsable por tomar las decisiones y dirigir el esfuerzo, que será ejecutado de manera colaborativa con un equipo multidisciplinario de la institución. El liderazgo recae en esa persona en particular, dado que posee el conocimiento experto sobre el tema a tratar y la visión de la política pública que se busca implementar. Le corresponde igualmente dirigir la ejecución del proyecto a cargo del equipo técnico y aplicar la herramienta en la población destinataria.

Una buena conceptualización y diseño del proyecto asegurará su viabilidad, sostenibilidad y valor público, además de que mitigará los riesgos que conlleva la aplicación de la herramienta de IA.

En el proceso de conceptualización y diseño se pueden identificar pasos claves y preguntas esenciales a las que debe responder el responsable por la toma de decisiones antes de ejecutar un proyecto de IA.

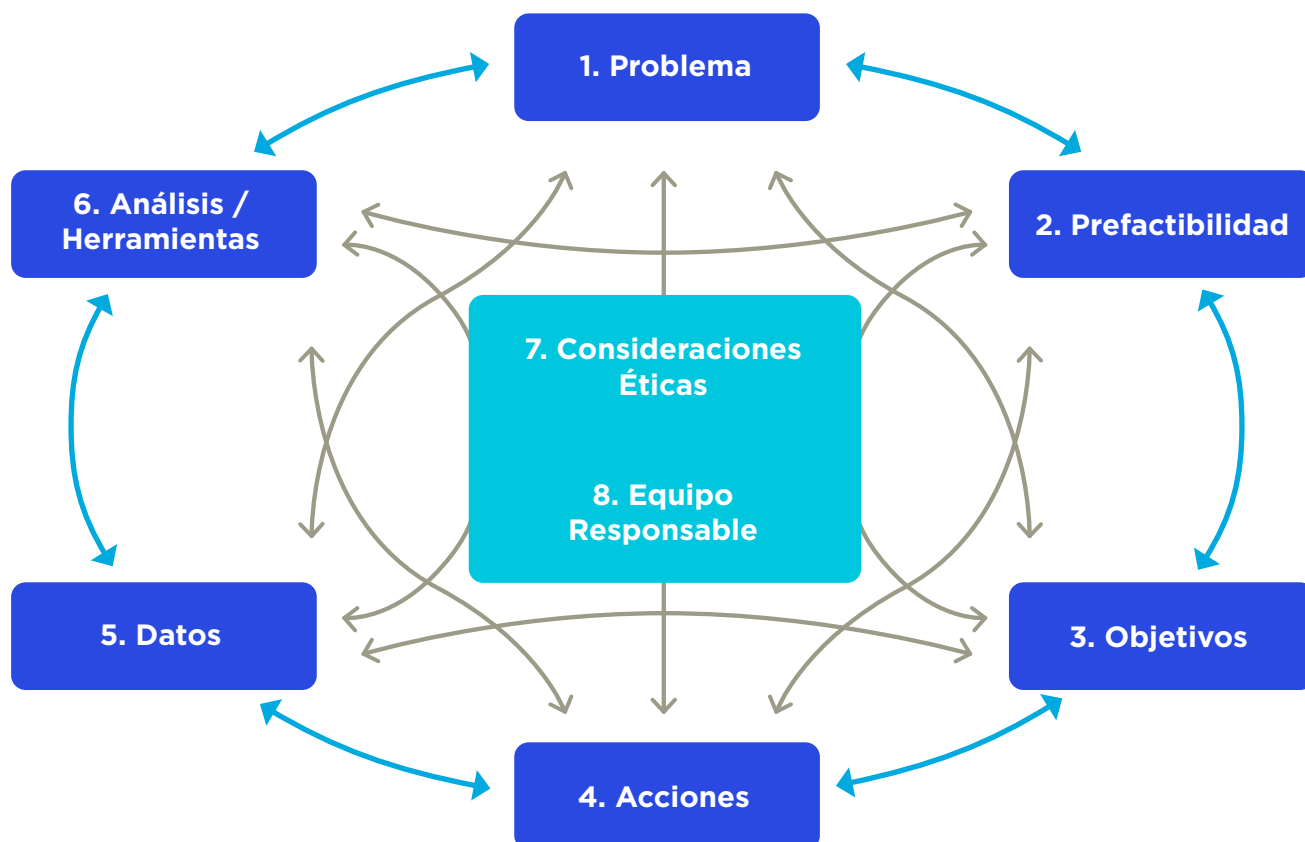
La metodología de conceptualización que se propone en este manual es una adaptación de la Guía de Alcance de Proyectos de Ciencia de Datos creada por Rayid Ghani desde el Centro de Ciencia de Datos y Políticas Públicas de la Universidad de Chicago. Los conceptos aquí descritos fueron adaptados para Latinoamérica por el GobLab de la Universidad Adolfo Ibáñez, Chile.

- 1. Definición del problema:** El primer paso en todo proyecto es definir claramente el problema de política pública al cual se quiere responder con la implementación de una herramienta de toma y/o soporte de decisiones basada en IA.
- 2. Análisis de prefactibilidad:** Luego de definir el problema y determinar si la IA es la herramienta correcta para apoyar la solución, y antes de continuar con el proyecto, hay otras preguntas claves a las que se deberá responder (véase el numeral 1.2 más abajo). El objetivo de este paso es asegurar la viabilidad del proyecto para no desperdiciar los recursos escasos de la institución.
- 3. Definición de objetivos:** Una vez que el proyecto se declara factible, es el momento de fijar los objetivos y sus métricas o indicadores, los cuales servirán para medir sus logros. Tales métricas deben reflejar el impacto esperado de la aplicación de la herramienta en la población destinataria. El logro de los objetivos deberá ayudar a solucionar el problema identificado.
- 4. Descripción de acciones:** Las acciones son las actividades que realiza la institución pública y que materializarán la respuesta de política pública para resolver el problema. Pueden ser parte de programas de política pública que apuntan a abordar temas específicos, o procesos habituales de la institución como son las contrataciones, los pagos o la atención al público. Si bien estas acciones por lo general existen independientemente del sistema de IA, esta última contribuirá a transformarlas con el fin de lograr los objetivos del proyecto.
- 5. Mapeo de datos:** Se debe investigar si existen los datos necesarios y suficientes para llevar a cabo el proyecto, si la institución tiene acceso a las bases de datos o si se necesitarán convenios para obtenerlos. Un proyecto de IA se puede basar en datos tanto internos como externos, públicos o privados. Cabe notar que, en la etapa de diseño y conceptualización, no es necesario elaborar un análisis detallado de los datos; este será realizado por el equipo técnico una vez que se decida seguir adelante con el proyecto.

- 6. Definición del análisis y sus herramientas pertinentes:** En esta etapa, el director del proyecto debe identificar preliminarmente qué tipo de análisis se requerirá para solucionar el problema. El tipo de análisis --o la herramienta a implementar-- dependerá de la naturaleza de aquel y ayudará a mejorar los procesos de atención o respuesta necesarios. En esta etapa se trata de lograr una aproximación inicial que posteriormente debe consensuarse con el equipo técnico de proyecto.
- 7. Consideraciones éticas, legales y de gobernanza:** Incluso antes de comenzar con la ejecución del proyecto, el director deberá tener claros los desafíos éticos y legales que pueden surgir durante la implementación. Esto le permitirá adelantarse a posibles situaciones que lo puedan poner en riesgo y tomar las medidas de mitigación del caso.
- 8. Conformación del equipo responsable:** En su calidad de director del proyecto, la persona responsable por la toma de decisiones debe conformar el equipo que se encargará de llevarlo a cabo. Los proyectos de IA no solo comprometen a quien decide y al equipo técnico; se requiere la participación de una variedad de áreas de la institución (por ejemplo, el equipo legal) e incluso de instituciones externas (por ejemplo, quienes poseen bases de datos útiles para el proyecto).

El proceso de conceptualización y diseño de un proyecto debe ser iterativo. Aunque la idea es comenzar con una definición sólida del alcance del problema, esto puede cambiar si, por ejemplo, la institución no tiene la capacidad necesaria para actuar sobre él, o debe replantearse si los datos requeridos no están disponibles o son insuficientes.

Figura 3. Proceso de conceptualización y diseño



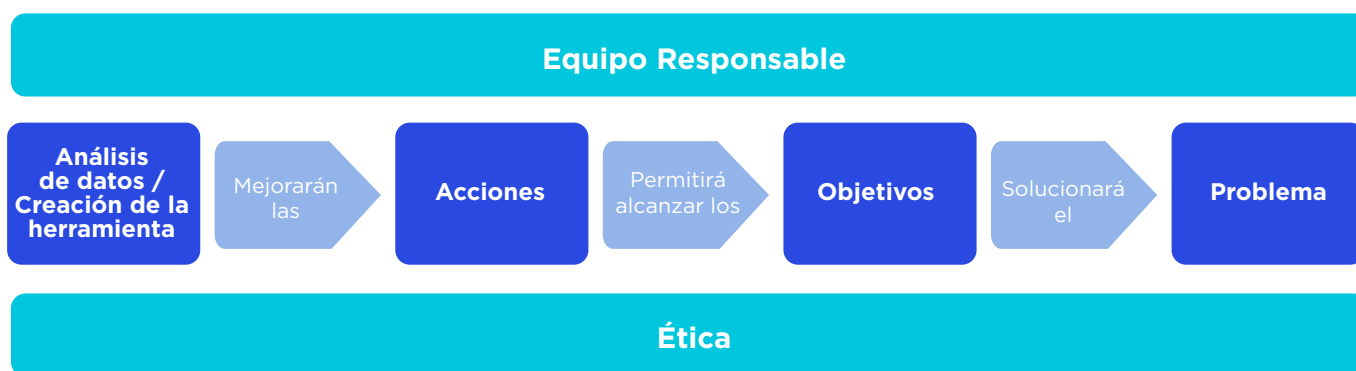
Fuente: Elaboración propia.

Una vez formulado el proyecto, se debe poder visualizar una relación lineal entre los pasos de la siguiente manera (Figura 4):

1. El análisis o herramienta dependerá de los datos disponibles.
2. El resultado del análisis o de la herramienta deberá mejorar las acciones de la institución.
3. La mejora de las acciones permitirá alcanzar el objetivo definido y el estado deseado de la situación.
4. Alcanzar el objetivo ayudará a solucionar el problema.

Las consideraciones éticas acompañan todo el proceso.

Figura 4. Relación de las etapas de conceptualización y diseño



Fuente: Elaboración propia.

A continuación se detallan las fases esenciales en la etapa de conceptualización y diseño. En el [Anexo 1](#) de este manual se encuentra la ficha de diseño y factibilidad que debe completar el director del proyecto y en la cual se basan los pasos que aquí se describen. Cada una de estas fases se ilustrarán con el ejemplo del proyecto DART de Chile sobre uso de la IA en la prevención de la ceguera, completando las subsecciones que correspondan en la ficha de diseño y factibilidad de un proyecto.

Teoría



1.1 Definición de problema

El primer paso en la conceptualización y diseño de un proyecto que involucre IA es definir claramente el problema de política pública que una institución busca resolver; en su calidad de tal, también hace parte del ciclo de vida de una política pública. La persona a cargo dirigir el proyecto y tomar las decisiones deberá tener el conocimiento experto sobre el tema a abordar y será quien priorice su resolución y comunique el problema al equipo técnico.

Las instituciones públicas deben afrontar diariamente múltiples asuntos de diferente naturaleza, como por ejemplo tiempos de respuesta, asignación de recursos, distribución de tareas, mejora en la focalización de beneficios sociales, entre otras. En esta etapa es importante acotar al máximo el problema que se quiere solucionar con la implementación de la herramienta de IA. El asunto priorizado puede afectar tanto a usuarios o beneficiarios externos a la organización como a quienes se desempeñan dentro de ella.

A continuación se listan las motivaciones que a menudo se aducen para iniciar un proyecto apoyado en IA:

“Tengo muchos datos y quiero hacer algo con ellos.”

“Quiero hacer un modelo predictivo.”

“Tenemos este software y debemos utilizarlo.”

“Me gustaría utilizar la tecnología X en la institución.”

Sin embargo, estas razones no justifican que se emprenda un proyecto basado en IA. Si bien es cierto que las herramientas de IA pueden ser altamente beneficiosas en términos de ahorro de tiempo y aumento de eficiencia dentro de la institución, también son costosas en términos de los recursos y del tiempo que exige su desarrollo. En tal sentido, es clave identificar claramente un problema prioritario que la entidad se vea abocada a resolver y para el cual una herramienta basada en IA represente un valor agregado.

Una vez definido el problema, es de suma importancia tener claro cuál es el estado actual de respuesta de la institución, identificar sus limitantes y áreas de oportunidad, y determinar de qué manera el sistema de soporte o toma de decisiones basado en IA podría mejorar el statu quo. Una buena práctica recomendada es estudiar experiencias comparadas de implementación de herramientas similares en otras instituciones o en otros países. De esta forma se tendrá información sobre los retos y desafíos que se debieron afrontar, lo cual también contribuirá a determinar la factibilidad del proyecto.

Dentro de la definición del problema es importante que el equipo cuantifique con precisión el número de afectados y el presupuesto que exigirá la solución. Esto ayudará a dimensionar la magnitud y gravedad del problema, y por lo tanto la necesidad de resolverlo.

Finalmente, debe haber voluntad política para actuar y resolver el problema. Las entidades —sobre todo los organismos públicos— tienen que hacer frente a una multiplicidad de problemas de distinta índole y por lo tanto deben ser estas quienes prioricen el asunto a tratar para que el proyecto tenga el apoyo institucional requerido.

El tema a abordar deberá definirse de manera clara y concisa, considerando todos los factores anteriormente mencionados. El director del proyecto debe tener en cuenta que no se trata solo de comunicar el problema al equipo técnico, sino también a la población destinataria. Por eso se recomienda utilizar una descripción sencilla que transmita su importancia y mantener el foco de la definición en el problema mismo y en cómo afecta a las personas y a las organizaciones, como se observa en los ejemplos a continuación.

Ejemplos de definiciones de problemas

Nombre proyecto	Definición de problema
Desarrollo de modelos predictivos de riesgo para apoyar las decisiones de la línea telefónica de ayuda para maltrato infantil en el condado de Allegheny (Pennsylvania)¹¹	Para combatir el problema del maltrato infantil, el condado de Allegheny cuenta con dos líneas de emergencia donde se pueden denunciar estos abusos: una línea de emergencia del condado, y una línea de emergencia especializada, "Childline". Las denuncias pueden clasificarse como CPS (servicios de protección infantil) y GPS (servicios de protección general). Un hogar entra en el sistema desde el momento en que se le ordena una investigación exhaustiva. De la totalidad de las denuncias que se clasifican como GPS, un 48% fueron ingresadas al sistema desde abril de 2010 hasta mayo de 2016. El problema es que el 53% de las denuncias que no generaban una intervención se repetían en un plazo de dos años. Este indicador muestra que el sistema actual de toma de decisiones, tal como está, deja por afuera a menores que necesitan de apoyo del condado. Dado el alto flujo de llamadas, y la asimetría de información entre las personas que ingresan las denuncias y los operadores de las líneas de emergencia, estos últimos deben tomar decisiones en poco tiempo y no logran considerar toda la información disponible.
Detección del riesgo de suicidio en aplicaciones de chat	En Chile, el suicidio es la primera causa de muerte entre los jóvenes de 15 a 24 años. Según datos de 2015, cada 2,8 días se suicida un menor de edad. La fundación Todo Mejora¹² es una ONG cuyo objetivo es promover el bienestar de niños, niñas y adolescentes que sufren matoneo y comportamiento suicida, producto de la discriminación con base en su orientación sexual, identidad y expresión de género. Todo Mejora ofrece un servicio de ayuda no presencial llamado "Hora segura", el cual comenzó a través del servicio de mensajería de Facebook y cuyos objetivos son evitar el suicidio, orientar y, en casos puntuales, derivar o remitir casos a instancias más avanzadas de resolución. El problema actual es que el número de usuarios que necesitan ayuda supera por momentos a los voluntarios disponibles para esta tarea y no se logra dar prioridad al contacto con aquellos niños, niñas y adolescentes que se encuentran en alto riesgo de suicidio. Específicamente, Todo Mejora cuenta con un equipo de entre 60 y 80 profesionales, los cuales deben responder a entre 8.000 y 9.000 solicitudes de asistencia telefónica al año. Anualmente se reciben alrededor de 60% de consultas provenientes de adolescentes de la comunidad LGBTQ+ y 40% de adolescentes heterosexuales cisgénero. El 68% de las personas que acuden a la "Hora segura" habían presentado comportamiento suicida en los últimos dos meses.

¹¹ Allegheny Family Screening Tool (USA): <https://www.alleghenycounty.us/Human-Services/News-Events/Accomplishments/Allegheny-Family-Screening-Tool.aspx>

¹² Fundación Todo Mejora (Chile): <https://todomejora.org/>

“Existe el riesgo adicional de plantear los proyectos de IA desde la tecnología y no desde el problema social particular. Incluso un proyecto de IA necesario y funcional puede presentar riesgos si no se plantea la acción de política pública correcta” (Cabrol, et. al., 2020).

Errores comunes en la definición del problema

Aunque la fase de identificación del problema parece ser muy simple, es quizás la más crítica del proceso de conceptualización y donde más se comenten errores, como los siguientes:

- Incluir descripciones y conceptos excesivamente amplios y sin acotar (como por ejemplo “mejorar la toma de decisiones”, “ser más eficientes”, etc.), en lugar de enfocarse en un problema concreto de la institución.
- Definir el problema como la ausencia de una herramienta/modelo. Por ejemplo: “el problema de mi institución es que no contamos con una herramienta automática de focalización de entrega de ayuda.”
- No especificar el estado actual y/o las brechas que se deben cerrar.
- No cuantificar el número de personas afectadas y/o el presupuesto requerido para solucionar el problema.

No se olvide de...

1. **Definir** claramente el problema de política pública que se quiere solucionar, identificando y cuantificando a los grupos de personas que se ven afectadas, y determinando su impacto en el presupuesto. La definición del problema debe ser fácilmente entendible para una persona externa a la institución.
2. **Contactar** a las personas de su institución que están a cargo de abordar el problema para establecer la forma en que se hace actualmente. ¿Qué información le pueden ofrecer sobre cómo mejorar el sistema de respuesta?
3. **Investigar** la manera en que otros organismos –nacionales o extranjeros— con un problema similar han implementado una solución basada en la IA. Lo ideal sería contactarlos para informarse acerca de los desafíos y dificultades que encontraron en el trayecto.
4. **Conversar** con el personal directivo de su institución sobre la prioridad de solucionar el problema y lograr un compromiso con el proyecto al más alto nivel. Si usted mismo es un directivo de alto nivel, documente cómo esta prioridad está plasmada en los planes estratégicos de la institución.

Actividad

Concluida esta sección, complete lo correspondiente a la **definición del problema** en la ficha de diseño y factibilidad de proyectos.

FICHA DISEÑO Y FACTIBILIDAD DE PROYECTO

1 Definición del problema

¿Cuál es el problema que se busca solucionar?

Describa a la(s) población(es) afectadas por el problema (personas, grupos, entidades, etc.)

¿Cuántas personas/organizaciones/localidades/etc. se ven afectadas y en qué medida?

¿Por qué resolver este problema es una prioridad para su organización?

¿Conoce algún caso parecido de uso de IA que haya sido implementado anteriormente? ¿Cuál?

FICHA DISEÑO Y FACTIBILIDAD DE PROYECTO

EJEMPLO DART: **Uso de la inteligencia artificial en la prevención de la ceguera**

1 Definición del problema

¿Cuál es el problema que se busca solucionar?

La diabetes es una enfermedad crónica que se presenta cuando el páncreas no puede producir suficiente insulina o cuando el organismo no utiliza toda la que produce. Existen dos tipos de diabetes: la diabetes tipo 1 y la diabetes tipo 2. La primera surge como resultado de un ataque del propio sistema inmunitario, no se puede prevenir y tampoco tiene cura, mientras que la segunda resulta de estilos de vida que dan lugar al sedentarismo y a la obesidad. Aquí el páncreas sigue produciendo insulina, aunque en cantidades insuficientes, y en algunos casos se puede curar. Ambos tipos de diabetes pueden producir ceguera, insuficiencia renal, accidentes cerebrovasculares (ACV) e incluso conducir a la amputación de las piernas.

La retinopatía diabética (RD) es la enfermedad ocular más común entre quienes sufren de diabetes en el mundo. El diagnóstico de esta enfermedad requiere un examen de fondo de ojo, donde un tecnólogo médico obtiene una imagen detallada del globo ocular que luego es analizada por un especialista.

Según estimaciones realizadas, en Chile se registra un déficit anual de 39.168 horas de oftalmólogos, lo cual indica que no existe el capital humano necesario para analizar todos los exámenes de fondo de ojo (Hojman, 2014).

Describe a la(s) población(es) afectadas por el problema (personas, grupos, entidades, etc.)

Este problema afecta en lo fundamental a tres grupos:

1. Las personas diagnosticadas con diabetes que no pueden realizarse el examen de fondo de ojo una vez al año, según las recomendaciones internacionales.
2. El Ministerio de Salud, que debe gastar recursos adicionales en retinopatía diabética al no poder realizar la prevención adecuada.
3. Los oftalmólogos, cuya carga de trabajo es excesiva.

¿Cuántas personas/organizaciones/localidades/etc. se ven afectadas y en qué medida?

En 2014 había 422 millones de personas con diabetes en el mundo. Según estimaciones de la Organización Mundial de la Salud, para 2040 ese número se elevará a 600 millones. En Chile, se calcula que una de cada 10 personas tiene diabetes y que entre un 15 y 20% de estos pacientes presenta algún grado de retinopatía diabética. Según la Encuesta Nacional de Salud de 2010, solo el 34,8% de los pacientes fue examinado por un oftalmólogo.

¿Por qué resolver este problema es una prioridad para su organización?

Teledx se crea con el único objetivo de apoyar tecnológicamente la tarea oftalmológica del sistema de salud y así reducir las tasas de pérdida de visión y ceguera en la población adulta. El foco principal es la retinopatía diabética, pues existe la necesidad imperiosa de aumentar la cobertura de diagnóstico.

La Estrategia Nacional de Salud, elaborada por el Ministerio de Salud para la década 2011-2020, establece distintos objetivos, siendo uno de ellos "Incrementar la proporción de personas con diabetes controlada". La estrategia específica para lograr este objetivo consiste en mejorar el control de los pacientes diabéticos a través de una cobertura más amplia de distintos exámenes clínicos, entre ellos el de fondo de ojo.

¿Conoce algún caso parecido de uso de IA que haya sido implementado anteriormente? ¿Cuál?

Sí, se realizó una reseña de literatura de otros estudios que buscan detectar retinopatía diabética (Arenas, 2012).



Luego de definir el problema, es necesario preguntarse qué tan factible es que el proyecto se lleve a cabo, más allá de sus aspectos técnicos a cargo del equipo de especialistas. Este paso es clave antes de comenzar un proyecto, puesto que evita malgastar los escasos recursos financieros y humanos de la institución. Algunas de las preguntas a las que se debe responder en este paso son:

1. ¿Está dentro de las facultades de la entidad actuar sobre el problema? ¿Será necesario asociarse con otros organismos públicos? ¿Se cuenta con los recursos humanos y financieros suficientes para llevar a cabo el proyecto?

Este punto alude a la competencia tanto legal como normativa que se tenga de responder al problema, y a la disponibilidad de recursos humanos (internos) y financieros que podrían destinarse a producir la herramienta. Es necesario que la institución pueda responder al problema dentro del marco legal y presupuestario.

2. ¿Existen los datos relevantes (suficientes para poder cambiar la forma en que hasta el momento se ha respondido al problema)? ¿Es posible acceder a ellos? Los datos deben ser suficientes para poder crear la herramienta de IA que ayude a solucionar el problema. Por ejemplo, si el asunto a resolver es que la autoridad competente no puede predecir eventos medioambientales de mala calidad del aire, lo cual se traduce en un aumento de las consultas médicas por enfermedades respiratorias, se necesita por lo menos contar con información histórica sobre la presencia de partículas contaminantes en la atmósfera en los últimos años. Si los datos están agregados por meses, entonces es poco probable que se pueda desarrollar un modelo predictivo.

3. ¿Cuáles son los riesgos del proyecto (éticos, licencia social, implementación, etc.)? Un proyecto de un organismo público apoyado en IA buscará generar valor para la población. Sin embargo, dado que las herramientas de IA y su ejecución presentan algunos riesgos, se debe hacer un análisis que compare el efecto positivo esperado versus los riesgos potenciales de las herramientas de IA (más adelante se profundizará sobre estos temas). Algunos de estos riesgos son:

Riesgos éticos: Los sistemas automatizados de respuesta pueden reproducir los sesgos presentes en los datos, o pueden resultar sesgados como resultado de un mal modelo. Dado que existe la probabilidad de que la herramienta presente alguno de esos riesgos, corresponde al director del proyecto tomar las medidas necesarias para mitigarlos.

Licencia social: Se refiere a la aprobación que da la población destinataria al uso de la herramienta. Es decir, si esta población sabe del proyecto, ¿aprobaría el uso de la IA? Para obtener la licencia social es necesario, primero, que la población tenga claro el beneficio que obtendrá de la aplicación de la herramienta, y segundo, que haya transparencia sobre la protección de los datos, la manera de utilizar la herramienta y las medidas que se tomen para mitigar los sesgos, en caso de que se detecten.

Riesgos de implementación: Se refiere a aquellos riesgos que puedan presentarse una vez comience la ejecución por parte de la institución: el primero es que el modelo no funcione lo suficientemente bien, y que por lo tanto no haya una mejora en la labor

institucional tanto internamente como de cara al público; el segundo es que nunca se llegue a implementar por fallas en alguno de los componentes claves (compromiso de la autoridad, falta de recursos y/o resistencia al cambio por parte de las personas que deberán utilizar la herramienta); y el tercero es que se implemente pero que nunca se use.

El responder a estas preguntas permite identificar algunos aspectos claves que conducirán a la formulación y ejecución exitosas del proyecto, en la medida en que este ejercicio facilita comprender mejor sus limitaciones y riesgos.

La idea es que desde el primer momento se justifique plenamente la necesidad del proyecto, que se cuente con las capacidades para diseñarlo y ejecutarlo de manera efectiva, y que se identifiquen tanto los riesgos que aquel pueda presentar como las medidas para mitigarlos.

No se olvide de...

1. **Contar** con el compromiso de las personas que ejecutarán la herramienta e incorporarlas en el equipo de trabajo.
2. **Identificar** los posibles riesgos éticos y legales del modelo. Conversar con el equipo legal e incorporarlo al equipo del proyecto.
3. **Asegurar** los recursos monetarios y humanos que permitan sostener la herramienta en el tiempo.

Actividad



Concluida esta sección, complete lo correspondiente a **análisis de prefactibilidad** en la ficha de diseño y factibilidad de proyectos.

FICHA DISEÑO Y FACTIBILIDAD DE PROYECTO

2

Análisis de prefactibilidad

¿Está dentro de las facultades de la entidad actuar sobre el problema?
¿Será necesario asociarse con otros organismos públicos? ¿Se cuenta con los recursos humanos y financieros suficientes para llevar a cabo el proyecto?

¿Existen los datos relevantes (suficientes para poder cambiar la forma en que hasta el momento se ha respondido al problema)? ¿Se puede acceder a ellos?

¿Cuáles son los riesgos del proyecto (éticos, licencia social, implementación, etc.)?

FICHA DISEÑO Y FACTIBILIDAD DE PROYECTO

EJEMPLO DART: Uso de la inteligencia artificial en la prevención de la ceguera

2 Análisis de prefactibilidad

**¿Está dentro de las facultades de la entidad actuar sobre el problema?
¿Será necesario asociarse con otros organismos públicos? ¿Se cuenta con los recursos humanos y financieros suficientes para llevar a cabo el proyecto?**

Requiere asociarse con otros organismos, en especial con el Ministerio de Salud, para que la herramienta pueda ser implementada en los centros de salud. Existen los recursos humanos, pero no financieros, por lo cual se deberá postular para acceder a algún tipo de financiamiento.

¿Existen los datos relevantes (suficientes para poder cambiar la forma en que hasta el momento se ha respondido al problema)? ¿Se puede acceder a ellos?

Sí, pero se debe contactar a un mayor número de centros de salud para solicitar más imágenes de exámenes y así aumentar el tamaño de la muestra (N) disponible.

¿Cuáles son los riesgos del proyecto (éticos, licencia social, implementación, etc.)?

Los resultados de los exámenes de fondo de ojo son datos personales, y por eso los pacientes deben dar su consentimiento para que se les practiquen. Uno de los principales riesgos del proyecto se origina en la implementación, ya que se requiere obtener el compromiso de los centros de salud, tecnólogos médicos y oftalmólogos que van a interactuar con el modelo para que utilicen la herramienta de manera efectiva. Dado que el proyecto busca disminuir los tiempos de espera, con lo cual se reducirá también el riesgo de ceguera, es altamente probable que la población general otorgue licencia social.



Los objetivos se entenderán como el estado al que se quiere llegar en un asunto determinado. Por ejemplo, si en un momento dado un organismo público se demora en promedio cinco días en responder a solicitudes ciudadanas, un objetivo del proyecto apoyado en IA podría ser reducir el plazo de respuesta a tres días. Así pues, los objetivos expresan avances concretos de una situación de deficiencia a una de satisfacción. Por lo tanto, cualquier indicador que se declare en los objetivos debe estar también descrito en la definición de problema.

Una buena metodología para definir objetivos es la que se conoce por la sigla en inglés **SMART**, lo cual significa que los objetivos deben ser **e**Specíficos, **M**edibles, **A**lcanzables, **R**ealistas y con **T**emporalidad definida.

Como se indicó más arriba, uno de los errores más comunes es definir como objetivo “crear un sistema automático de respuesta” o “elaborar un modelo predictivo”, dado que ninguno de los dos responde realmente a la pregunta central en esta etapa: “¿Lograr el objetivo ayuda a solucionar el problema?” La respuesta en estos dos casos sería “No”, dado que el simple hecho de contar con un sistema predictivo no implica que se vaya a solucionar el problema en cuestión. En otras palabras, la eficacia de la herramienta o del modelo dependerá no solamente de cuán precisos sean, sino de la forma en que se aplique a sus beneficiarios potenciales. Las actividades auxiliares del proyecto, como por ejemplo “explorar los datos” o “entender el comportamiento histórico de los usuarios del programa,” tampoco son objetivos de aquel.

Cuando se trata de definir los objetivos, una buena práctica es preguntarse ‘para qué’: ¿Para qué se necesita un modelo predictivo? ¿Para qué se requiere implementar un chatbot en la página web? ¿Para qué se necesita saber el comportamiento de los usuarios? Al iterar constantemente este ‘para qué’ se llegará finalmente al objetivo real del proyecto. En el cuadro sinóptico a continuación se presentan tres ejemplos de proyectos con una breve descripción del problema, un objetivo inicial, y un objetivo mejorado según los parámetros discutidos en esta sección.

Caso	Problema	Objetivo inicial	Debilidad	Objetivo mejorado
Prevención de envenenamiento con plomo. ¹³	X cantidad de niños tienen algún nivel de plomo en la sangre. No existen niveles aceptables de plomo en la sangre. La fuente principal de intoxicación con plomo en los niños se encuentra en las casas antiguas cuyas paredes están recubiertas con pintura que contiene plomo. La detección del plomo se hace mediante un examen de sangre que se practica a los niños al entrar al sistema educativo. Cuando se detecta plomo, se envía un inspector a la vivienda del niño y se inician las acciones para la remoción de este metal pesado, para lo cual se dispone de subsidios estatales. Sin embargo, no se cuenta con recursos suficientes para reparar todas las viviendas recubiertas con pintura que contiene plomo en las ciudades.	Evitar el envenenamiento con plomo en los niños.	El objetivo es muy amplio. Si bien ayuda a solucionar el problema (envenenamiento con plomo en los niños), no es específico ni se relaciona con la acción concreta que puede realizar la ciudad. Tampoco se ofrece una métrica de éxito, así que no se puede saber si se logra o no.	Para el año 2019, se aspira a que el X% de las inspecciones totales detecten plomo en viviendas donde habitan madres embarazadas.
Detección de retinopatía diabética (RD). ¹⁴	Actualmente existe una demora de X meses en la entrega de los resultados de los exámenes de fondo de ojo (para establecer si un paciente sufre o no de RD). Una retinopatía diabética no diagnosticada, o con un diagnóstico tardío, puede llevar a la ceguera total. Actualmente no existe la capacidad humana (oftalmólogos y tecnólogos médicos) necesaria para cumplir con un tiempo prudente de entrega de los resultados de los exámenes de fondo de ojo.	Analizar los exámenes de fondo de ojo de todos los pacientes del sistema de salud.	El objetivo no se refiere al problema, que es la demora en el análisis de los exámenes de fondo de ojo. Esto elevaría las tasas de retinopatía diabética que, sin tratamiento, aumentarían los casos de ceguera.	Disminuir, en un plazo de dos años, el tiempo de entrega de resultados para aquellos pacientes con alta probabilidad de tener retinopatía diabética en un X%.

¹³ Lead poisoning prevention (USA): <http://www.datasciencepublicpolicy.org/projects/public-health/poison-prevention/>

¹⁴ DART (Chile): <https://www.teledx.org/dart/?lang=es>

Caso	Problema	Objetivo inicial	Debilidad	Objetivo mejorado
Fiscalización de importaciones	Se estima que actualmente el X% de las importaciones que ingresan al país es fraudulento y no se somete a control de aduanas.	Elaborar un sistema predictivo de importaciones fraudulentas.	La solución técnica no es un objetivo.	Aumentar el porcentaje de incautaciones de importaciones fraudulentas sobre el total de importaciones fiscalizadas en el siguiente año fiscal.

Si bien el logro de los objetivos conduce a solucionar los problemas identificados y permite llegar al estado deseado, no se debe olvidar que los proyectos adolecen de limitaciones que pueden afectarlos y poner en riesgo las metas establecidas. Dentro de las limitaciones típicas se encuentran la falta de recursos humanos y/o financieros, y la resistencia al cambio, entre otros. Identificar las posibles limitaciones permitirá que el director del proyecto se adelante al surgimiento de estos obstáculos en la etapa de implementación.

No se olvide de...

1. **Definir** objetivos en relación con el problema de política pública que se quiere resolver.
2. **Fijar** objetivos **SMART**: eSpecíficos, Medibles, Alcanzables, Realistas y con Temporalidad definida.

Actividad



Concluida esta sección, complete lo correspondiente a **objetivos** en la ficha de diseño y factibilidad de proyectos.

FICHA DISEÑO Y FACTIBILIDAD DE PROYECTO

3

Objetivos

	Objetivo	Limitaciones
1		
2		

FICHA DISEÑO Y FACTIBILIDAD DE PROYECTO

EJEMPLO DART: Uso de la inteligencia artificial en la prevención de la ceguera

3 Objetivos

	Objetivo	Limitaciones
1	Disminuir el tiempo de entrega de los resultados para aquellos pacientes con alta probabilidad de tener retinopatía diabética en un X% en un plazo de dos años.	Existe una cantidad limitada de oftalmólogos que pueden analizar los exámenes de fondo de ojo.
2	Disminuir el gasto público asociado con el tratamiento de la retinopatía diabética en un 50% en un plazo de 10 años.	Los centros de atención primaria de salud tienen que alcanzar las metas de cobertura del examen anual de fondo de ojo para pacientes diabéticos, con el fin de detectar a tiempo y prevenir la retinopatía.



Luego de definir los objetivos del proyecto, cabe preguntarse: ¿Qué está haciendo actualmente la institución para resolver el problema? ¿De qué forma se podría mejorar su intervención con la implementación de la herramienta de IA? Es en esta etapa cuando se deben identificar claramente los procesos de atención o de respuesta actuales de la institución frente al asunto detectado y que podrían mejorarse con la implementación de análisis o herramientas basados en IA. Algunas acciones de las instituciones son:

- **Implementación de productos o programas estratégicos:** Entrega de subsidios de vivienda, asignación de subvenciones de alimentación para estudiantes, planificación de fiscalizaciones de importaciones, etc.
- **Procesos de soporte de la institución:** Atención a usuarios, contratación de personal, planificación de rutas de transporte, etc.

En esta etapa es importante identificar con precisión quiénes son las personas que llevan a cabo estas actividades en la actualidad, ya que probablemente deberán ser incluidas en el equipo del proyecto, y mediante qué canales se las está ejecutando.

Esto ya había sido esbozado en la pregunta 2 del análisis de prefactibilidad, numeral 1.2: ¿Existe la capacidad para actuar sobre el problema? ¿Está dentro de las facultades de la organización? ¿Existen los recursos humanos y financieros requeridos para actuar? Aquí se espera que el director del proyecto analice en profundidad las acciones que está llevando a cabo la institución actualmente para resolver el asunto en cuestión.

La importancia de definir correctamente esas acciones deriva del hecho de que una institución ejecuta multiplicidad de intervenciones y/o programas, pero no todas ellas tienen relación con el problema identificado. **En esta etapa se deben definir solamente aquellas acciones que podrían tener un impacto en aquel.** Por ejemplo, si lo que se debe solucionar es el exceso de tiempo transcurrido entre el momento en que llega un derecho de petición por parte de un ciudadano y la respuesta de la entidad pública competente, una acción relevante sería determinar cuáles son los cuellos de botella durante la revisión de la solicitud o identificar si el sistema de notificaciones está teniendo demoras.

En otras instancias, un mismo problema puede concernir a distintas instituciones, pero la forma en que se le dé respuesta dependerá de la misión de cada cual. Por ejemplo, la lucha contra la pobreza es un problema muy amplio y por ende distintos ministerios se involucrarán de variadas formas en su solución. Un Ministerio de Desarrollo, por ejemplo, puede entregar bonos a las familias más vulnerables durante los meses de invierno para ayudar a costear los mayores gastos que trae esa estación del año (aumento de visitas al médico y mayores precios de los alimentos, por ejemplo), mientras que un Ministerio de Educación puede enfocarse en el acceso universal a la educación apelando a la movilidad social que conlleva una mayor educación. Por lo tanto, los sistemas de respuesta diferirán dependiendo de las capacidades y facultades de cada institución.

Otro caso sería el de un proyecto que busque aumentar el porcentaje de incautaciones de importaciones fraudulentas. Aquí la acción principal es la fiscalización, lo cual exige que se explique de manera clara cómo se toma actualmente la decisión acerca de qué importaciones se someten a ello. Por

ejemplo, es posible que cada día se decida qué cargas inspeccionar con base en un historial pasado de cumplimiento del importador, o que los inspectores a cargo tomen la decisión en el momento, sin que exista una mayor planificación. Es importante tener en cuenta cómo opera actualmente el sistema, y cómo podría afectar el cambio a las personas a cargo de gestionarlo. En este ejemplo, ¿qué pasaría si los inspectores reciben un bono por cada carga revisada? Habría entonces un incentivo para fiscalizar una gran cantidad de carga, independientemente del resultado. Por lo tanto, un modelo predictivo que disminuya el número de cargas a inspeccionar puede despertar resistencia en el equipo responsable de la tarea, ya que obligaría a sus integrantes a revisar menos importaciones en busca de mayor eficiencia, lo que a su vez reduciría los bonos. Por lo tanto, para asegurar la viabilidad del proyecto es esencial conocer a fondo la modalidad actual con la que la institución aborda el problema e incorporar en el equipo del proyecto a alguna persona responsable de esas acciones (jefe de fiscalizaciones, por ejemplo) para que llame la atención sobre posibles debilidades.

Uno de los errores más comunes que se cometen en esta etapa es concentrarse en tareas auxiliares del proyecto como recopilar información, explorar los datos o vincular distintas bases de datos. Si bien estas actividades deberán estar al servicio de aquel, no deben convertirse en su actividad principal. Para evitar este error es necesario volver al asunto a resolver y recordar la cadena de relación entre las distintas etapas. Así pues, un mejor sistema de respuesta (informado por el análisis o la herramienta) deberá ayudar a alcanzar los objetivos, los cuales a su vez contribuirán a solucionar el desafío. La definición de las acciones, junto con el contexto en el que se realizan y las personas involucradas, es clave para lograr el éxito del proyecto en la etapa de implementación. Sin embargo, surgen a veces imprevistos que dificultan el cumplimiento de los objetivos y que será necesario abortar para que no se conviertan en obstáculos.



Ejemplo: No todo está previsto

En Tailandia, con la ayuda de Google se implementó un sistema automático de detección de retinopatía diabética mediante análisis de fondo de ojo. El modelo funcionaba muy bien en el laboratorio, pero al implementarlo sucedió que los establecimientos en donde se tomaban los exámenes no tenían la iluminación suficiente. Así pues, las imágenes resultaban de peor calidad que las que se habían utilizado para entrenar el modelo y no se podía entregar un resultado confiable sobre la presencia de esta enfermedad. Se requirió entonces que los pacientes acudieran a centros asistenciales con mejores condiciones, lo cual ocurría pocas veces debido a las dificultades de desplazamiento. Por otro lado, la mala calidad de conexión a internet hacía que la espera por el resultado fuera muy larga. El problema en este ejemplo es que Google no incorporó en la etapa de conceptualización del proyecto a las enfermeras —responsables de ejecutar la acción— y por lo tanto no tenían conocimiento de los problemas del sistema de atención para incorporar acciones que pudieran mitigar los efectos de cambios en iluminación o de mala conexión a internet antes de que se implementara el proyecto (Heaven, 2020).

Si no se tiene la capacidad de actuar sobre el asunto prioritario seleccionado, y dado el proceso iterativo de formulación de un proyecto, es posible que se deba volver nuevamente al problema para su completa reformulación. Por ello es tan importante el análisis de prefactibilidad, pues evita que esta situación se presente más adelante.

No se olvide de...

1. **Conversar** con las personas que se ocupan actualmente del asunto a resolver para determinar cómo se toman las decisiones y la manera en que podrían resultar afectadas por el cambio que signifique la implementación de la herramienta de IA.
2. **Comunicar** efectivamente las implicaciones y beneficios de la implementación de una herramienta basada en IA para lograr el compromiso de quienes finalmente la utilicen.
3. **Incorporar** en el equipo del proyecto a por lo menos una de las personas que estaba a cargo de responder al problema antes de poner en marcha el sistema de IA.

Actividad



Concluida esta sección, complete lo correspondiente a **descripciones de acciones** en la ficha de diseño y factibilidad de proyectos.

FICHA DISEÑO Y FACTIBILIDAD DE PROYECTO

4 Descripción de acciones

Problema a resolver	Acción 1	Acción 2
Acciones		
¿Quién ejecuta la acción?		
¿Sobre quién o qué se está realizando la acción?		
¿Con qué frecuencia se toma la decisión de realizar esta acción?		
¿Qué canales se están usando o se pueden usar para realizar esta acción?		
Otra información útil sobre la acción		

FICHA DISEÑO Y FACTIBILIDAD DE PROYECTO

EJEMPLO Uso de la inteligencia artificial en la
DART: prevención de la ceguera

4 Descripción de acciones

Problema a resolver	Acción 1	Acción 2
Acciones	Toma de examen de fondo de ojo	Diagnóstico de retinopatía diabética por parte de un oftalmólogo
¿Quién ejecuta la acción?	Tecnólogos médicos en centros asistenciales de salud primaria	Oftalmólogo
¿Sobre quién o qué se está realizando la acción?	Personas diagnosticadas con diabetes	Personas diagnosticadas con diabetes
¿Con qué frecuencia se toma la decisión de realizar esta acción?	Anualmente por persona	Luego de la toma del examen
¿Qué canales se están usando o se pueden usar para realizar esta acción?	Centros asistenciales que cuentan con los insumos necesarios	Centros asistenciales
Otra información útil sobre la acción		



La realización del análisis o herramienta de IA está supeditada a la existencia de los datos que los alimenten. Por ello, incluso antes de comenzar el director del proyecto debe explorar la disponibilidad y calidad de los datos internos y externos que se puedan requerir. Con respecto a estos últimos, se deberá consultar si se puede acceder a ellos y cuál es el costo. Es posible que algunas herramientas de IA requieran únicamente datos abiertos. En caso de que se necesite, en esta etapa el director deberá consultar con las áreas de la institución que recopilan o administran los datos sobre algunos aspectos claves de las bases de datos. No es necesario explorar a fondo, pues esta tarea corresponderá al equipo técnico en la etapa de ejecución.

Los datos deben ser suficientes para el análisis y deben tener la granularidad requerida. Una herramienta útil para saber qué tan madura (buena) es una base de datos es la matriz de madurez de datos. En el [Anexo 2](#) de este manual se presenta una adaptación del Marco de Madurez de los Datos (Data Maturity Framework) de la Universidad de Chicago¹⁵. Con esta herramienta se busca ayudar al director del proyecto a que analice, para cada uno de los aspectos sobre los cuales se pretende actuar, cuán avanzada se encuentra cada una de las bases de datos que va a utilizar en la herramienta de IA. Es posible que en este paso se deba preguntar directamente a los responsables de los datos sobre sus características y problemas.

A continuación se ofrece una explicación de las características utilizadas en la matriz.

Acceso: Independientemente de si una base de datos es interna o externa, la facilidad de acceso será un habilitador importante del proyecto de IA.

Almacenamiento: Las instituciones tienen cantidades masivas de datos, pero estos pueden estar almacenados de distintas formas. Si bien lo ideal es que los datos se encuentren en una base de datos digital, muchas veces el director del proyecto se encontrará con registros en formato físico. Esto exigirá un paso previo a la implementación durante el cual se digitalicen los datos, de modo que sean legibles mediante el uso de algún software y utilizables en el análisis pertinente. Dado que esto puede demorar la ejecución del proyecto, es de suma importancia determinar el estado de los datos con anticipación a ese paso.

Integración: Si bien las instituciones suelen tener una gran cantidad de datos, muchas veces cada una de las bases se trata de manera separada e independiente de las demás. El escenario ideal es que exista interoperabilidad entre las bases de datos (tanto internas como externas), para que así sea más fácil acceder a la información requerida y unificar las variables o las etiquetas claves.

Relevancia: La relevancia de los datos se refiere a su pertinencia para el problema a tratar.

Calidad: Se refiere a la posibilidad de que la base de datos se encuentre incompleta o tenga errores. Una base de datos que no contenga los datos de un sujeto u observaciones sobre este (región, barrio o edad, entre otras) puede tener importantes consecuencias en el modelo resultante, dado que este podría estar sesgado hacia una parte de la población. Una base de datos con variables faltantes conlleva retos importantes puesto que el equipo técnico, liderado por el director del proyecto, tendrá que tomar la decisión de o bien eliminar la observación por completo o imputarle datos. A menudo

¹⁵ <http://dsapp.uchicago.edu/resources/datamaturity/>

las bases de datos también tienen errores de digitación, los cuales ocurren en el momento de recopilar o transcribir la información. Aquí nuevamente el equipo técnico deberá decidir, junto con el director del proyecto, qué hacer con esos errores para no comprometer la calidad de la herramienta final. El escenario ideal sería que las bases de datos no tuvieran ninguno de los problemas descritos anteriormente, lo cual es poco habitual.

Frecuencia: Se refiere a la regularidad con la que se recopilan los datos, no a la de los datos mismos (observaciones anuales, mensuales o diarias). Los datos se pueden recopilar una única vez en caso de que sea una política puntual, por ejemplo, o se pueden recopilar mensualmente, diariamente o en tiempo real, es decir, a medida que van surgiendo. Aquí lo importante es determinar cuál es la política de recopilación de información, más que saber exactamente cuántas veces se los ha recopilado.

Granularidad: Se refiere al nivel de detalle de los datos. Se pueden tener observaciones a nivel de país, región, barrio, evento (cada vez que ocurre) o individuo. Una base más madura tendrá un nivel de granularidad mayor, es decir, contará con observaciones cada vez más específicas. Tener conocimiento sobre el nivel de granularidad de las bases de datos es clave para el éxito del proyecto. Por ejemplo, un modelo predictivo de deserción escolar necesitará datos de cada uno de los alumnos; no es suficiente contar con información por zonas geográficas, ya que esta no permitiría una correcta focalización de las acciones de prevención, que podrían ser a nivel individual.

Historial: Los datos de una observación pueden ir cambiando con el tiempo. Por ejemplo, una persona puede estar soltera a los 25 años pero 10 años después estar casada, por lo cual en la nueva recolección de datos la variable 'estado civil' será distinta. Existen diversas maneras de almacenar estos historiales: la más deficiente es limpiar el historial de los datos antiguos para reemplazarlo con la nueva información, mientras que la más avanzada es guardar los nuevos datos registrando su fecha y relacionándolos con los anteriores.

Privacidad: Más allá de que se gestionen datos personales o sensibles, el uso, acceso y manipulación de este tipo de información deberán estar regulados por una política de protección. Establecer una política de privacidad implica regular quiénes (funciones) tienen acceso a qué tipo de datos y para qué tipo de proyectos. Esto se puede lograr mediante un proceso ad hoc o de un *software* (generalmente un sistema administrador de base de datos) que se encargue de manejar la política de acceso.

Documentación: Usualmente las bases de datos utilizan códigos en las variables; en estos casos es imperativo contar con un libro de códigos o diccionario de datos que explique qué significa cada una de estas variables. Asimismo, una documentación adecuada deberá contener información sobre los metadatos del conjunto de datos: su descripción, la manera en que se extrajeron y si tienen relación con otros conjuntos de datos, entre otros aspectos.

Cabe recalcar que no porque una base de datos se califique como básica en la matriz de madurez de los datos, significa que sea insuficiente para entrenar la herramienta. Sin embargo, sí es necesario tener en cuenta que el equipo técnico tendrá que invertir más tiempo y recursos en su manejo. Esto podría significar que el director del proyecto deba poner en pausa momentánea la elaboración de la herramienta de IA para implementar primero políticas más avanzadas de recolección, almacenamiento y manejo de datos, para luego continuar con la implementación.

Si los datos necesarios para crear la herramienta no existen, o si la granularidad es insuficiente, entonces será imposible resolver el problema seleccionado a través del uso de inteligencia artificial. Esto exigirá, o bien reorientar la solución del problema o volver a la etapa de definición para revisar en detalle la disponibilidad de datos en otros casos similares que han usado IA.

No se olvide de...

1. **Identificar** claramente todas las bases de datos necesarias para llevar a cabo el proyecto.
2. **Conversar** con los responsables acerca de las características claves de las bases de datos para evaluar su nivel de madurez y si son suficientes para realizar el proyecto.
3. En caso de requerir datos de otras instituciones, **tramitar** tan pronto como sea posible los convenios necesarios para asegurar un acceso oportuno.

Actividad



Concluida esta sección, complete lo correspondiente a **mapeo de datos** en la ficha de diseño y factibilidad de proyectos.

FICHA DISEÑO Y FACTIBILIDAD DE PROYECTO

5

Mapeo de datos

¿De que datos dispone internamente?

¿Qué datos puede obtener de fuentes externas, privadas o públicas?

	Fuente de datos 1	Fuente de datos 2
¿Qué contiene?		
¿Qué nivel de granularidad?		
¿Con qué frecuencia se recopila y/o actualiza una vez que se captura?		

¿Tiene identificadores únicos y confiables que se puedan vincular con otras fuentes de datos?		
¿Quién es el responsable interno de los datos?		
¿Cómo se almacena?		
Comentarios adicionales		

FICHA DISEÑO Y FACTIBILIDAD DE PROYECTO

EJEMPLO DART: Uso de la inteligencia artificial en la prevención de la ceguera

5 Mapeo de datos

¿De que datos dispone internamente?: No se dispone de datos.

¿Qué datos puede obtener de fuentes externas, privadas o públicas?

	Fuente de datos 1	Fuente de datos 2
¿Qué contiene?	Imágenes de examen de fondo de ojo	Ficha médica del paciente, sexo, edad, fecha de diagnóstico, etc.
¿Qué nivel de granularidad?	Individual	Individual
¿Con qué frecuencia se recopila y/o actualiza una vez que se captura?	En tiempo real	En el momento del examen.
¿Tiene identificadores únicos y confiables que se puedan vincular con otras fuentes de datos?	Sí, nombre y RUT (Rol Único Tributario)	Sí, nombre y RUT
¿Quién es el responsable interno de los datos?	Centros asistenciales (Ministerio de Salud)	Centros asistenciales (Ministerio de Salud)
¿Cómo se almacena?	Imágenes	Base de datos
Comentarios adicionales		



Una vez identificados los datos disponibles, y dependiendo de la naturaleza del problema, llega el momento de pensar en los análisis o herramientas de IA que serán necesarios para solucionarlo. El rol del director del proyecto es identificar los que considere más adecuados, teniendo en cuenta que estos deberán informar y mejorar el proceso de respuesta actual, lo cual a su vez debe ayudar a alcanzar el objetivo, es decir, materializar la solución propuesta.

La IA permite realizar varios tipos de análisis —descriptivos, predictivos y de detección— y ese es precisamente su valor: crear la posibilidad de producir análisis que antes no se podían conducir con un simple procesador de datos.

A continuación se presentan algunos ejemplos sobre el uso de IA en América Latina y el Caribe según la tarea desempeñada por el sistema de IA, con base en las categorías de la OECD (OECD, próximo a publicarse): ¹⁶

Reconocimiento: Se basa principalmente en una categorización de imágenes, textos y videos, por medio de la identificación de sus características claves, entre otras cosas. En América Latina existen algunos ejemplos de proyectos en ejecución: DART (Chile), que como ya se vio, detecta la retinopatía diabética analizando los exámenes de fondo de ojo; Dymaxion Lab (Argentina), que monitorea los asentamientos humanos, las inundaciones y el uso de suelo a través de imágenes satelitales; y MIDIS-Detección temprana de anemia (Perú), que detecta casos de anemia mediante el análisis de una fotografía de la conjuntiva ocular.

Detección de eventos: El objetivo aquí es detectar patrones y anomalías. Un ejemplo es LAURA (Brasil), una plataforma de IA que se emplea para identificar a aquellos pacientes con deterioro clínico que tengan mayores probabilidades de sufrir sepsis. Con ello el equipo clínico cuenta con la información necesaria sobre los pacientes que le permita focalizar su atención en los casos de mayor riesgo.

Predicción: El objetivo es predecir un estado futuro basado en datos históricos. Algunos ejemplos en la región son NotCo (Chile), una empresa que utiliza un algoritmo para predecir los ingredientes idóneos para crear productos de origen vegetal que por lo general son de origen animal (mayonesa, leche, etc.); Traive (Brasil), un sistema crediticio alternativo que predice el desempeño futuro del solicitante; y Carabineros de Chile, que utiliza modelos de predicción del delito para pronosticar donde es más probable que este ocurra, información que a su vez se utiliza en la programación de los turnos del personal.

Personalización: El objetivo es desarrollar perfiles de usuarios que, con base en los datos generados por sus propias acciones, vayan mejorando en el tiempo. Las redes sociales y los servicios de transmisión continua (streaming) utilizan esta herramienta para hacer sugerencias personalizadas. Un ejemplo es Livox (Brasil), un programa que permite a las personas con discapacidades verbales comunicarse por medio de una selección de imágenes. Con el tiempo la aplicación comienza a sugerir imágenes probables a los usuarios con base en sus interacciones pasadas.

¹⁶ Los casos aquí listados son una selección de los que se registran en el documento “[La inteligencia artificial al servicio del bien social en América Latina y el Caribe: panorámica regional e instantáneas de doce países](#)” de la Iniciativa fAIrLAC del Banco Interamericano de Desarrollo (Gómez et al., 2020).

Apoyo a la interacción: Esta es una tarea que guarda relación con el apoyo a la interacción entre humanos y máquinas (chatbots, asistentes virtuales y otros). En la región existen algunos ejemplos como son Amanda Care (Argentina), una asistente virtual de acompañamiento de pacientes durante sus tratamientos médicos, y SpeakLiz (Ecuador), un software que transforma el lenguaje de señas a voz.

Optimización enfocada en objetivos: Se trata de optimizar los procesos dentro de una institución a través de simulaciones de escenarios. Por ejemplo Kilimo, en Argentina, tiene como objetivo que los productores agrícolas en ese país optimicen el uso del agua en sus actividades. Para ello recopila datos del campo y los combina con información histórica e imágenes satelitales en un software patentado mediante el cual proporcionan: 1) estimaciones del consumo hídrico del cultivo a siete días; 2) consejos periódicos resumidos en lenguaje sencillo para los productores sobre la cantidad de riego que se requiere para lograr sus objetivos de producción; y 3) información relacionada con el riego para fines de inteligencia empresarial.

Razonamiento con estructuras de conocimiento: Mediante este tipo de análisis se establece una relación causal entre los datos disponibles y un evento futuro no existente. A diferencia de la predicción simple, este tipo de análisis se enfoca en la causalidad entre variables. Un ejemplo de ello es el Portal Telemedicina (Brasil), a través del cual se busca predecir diagnósticos médicos con base en los historiales de los pacientes con base en un análisis causal. Asimismo es una plataforma de colaboración entre las clínicas de diagnóstico y los equipos médicos.

Los tipos de análisis que aquí se resumen son solamente una selección de casos de uso de IA en la región. La idea es que sirvan de insumo para cuando el director del proyecto le comunique al equipo técnico cuál de ellos podría ser más apropiado para resolver el problema de política pública que se busca atender. Es importante destacar que es usual que el sistema de IA usado en un proyecto combine distintas tareas. Tal es el caso de los llamados sistemas compuestos que detectan eventos anómalos y, a partir de allí, hacen predicciones.

Los análisis basados en IA utilizan igualmente instrumentos diversos, entre ellos las regresiones, los árboles de decisión y las redes neuronales. La decisión sobre la técnica de modelamiento que se emplee la tomará el equipo técnico, cuyos miembros deberán explicar al director del proyecto de qué manera se hizo el análisis y cuales fueron los errores y dificultades encontrados, antes de adoptar la decisión acerca de cuál será la técnica a implementar con la herramienta. Estos aspectos se profundizarán en el numeral 2.2 correspondiente al Desarrollo y validación del modelo.

Tareas de un sistema de IA	Posible técnica de análisis*
Reconocimiento	• redes neuronales • máquinas de soporte vectorial (SVM por su sigla en inglés)
Detección de eventos	• redes neuronales • máquinas de soporte vectorial
Predicción	• regresión lineal • árboles de regresión • redes neuronales
Personalización	• filtros colaborativos

Apoyo a la interacción	• redes neuronales
Optimización enfocada en objetivos	• programación lineal, no lineal, dinámica, etc.
Razonamiento con estructuras de conocimiento	• redes bayesianas

*Listado no exhaustivo.

No se olvide de...

1. **Identificar** el tipo de herramienta y/o análisis más adecuado para solucionar el problema de política pública seleccionado. No es necesario que decida desde un comienzo la técnica de análisis específica a utilizar en el modelamiento.
2. **Reunirse** con el equipo técnico (a cargo del modelamiento) para discutir el problema de política pública y consensuar el tipo de análisis requerido.
3. **Incluir** al equipo técnico dentro del equipo multidisciplinario del proyecto.

Actividad



Concluida esta sección, complete lo correspondiente a **análisis/herramienta** en la ficha de diseño y factibilidad de proyectos.

FICHA DISEÑO Y FACTIBILIDAD DE PROYECTO

6 Análisis/Herramienta

	Análisis 1	Análisis 2
Tipo de análisis (descripción, predicción, detección, cambio de comportamiento)		
Propósito del análisis (p. ej., entender el comportamiento histórico de las personas, estimar el riesgo de enfermedad de un paciente)		
¿Qué acción respaldará este análisis?		
¿Cómo validará este análisis usando datos existentes? (P. ej., usar datos históricos, ejecutar un ensayo aleatorio controlado)		

FICHA DISEÑO Y FACTIBILIDAD DE PROYECTO

EJEMPLO DART: Uso de la inteligencia artificial en la prevención de la ceguera

6 Análisis/Herramienta

	Análisis 1
Tipo de análisis (descripción, predicción, detección, cambio de comportamiento)	Detección
Propósito del análisis (p. ej., entender el comportamiento histórico de las personas, estimar el riesgo de enfermedad de un paciente)	Obtener de manera automática los casos anómalos que pueden traducirse en retinopatía diabética.
¿Qué acción respaldará este análisis?	Acción 2 (véase numeral 6 arriba correspondiente a “Descripción de acciones”). El oftalmólogo ahora contará con información filtrada en el momento de diagnosticar al paciente.
¿Cómo validará este análisis usando datos existentes? (P. ej., usar datos históricos, ejecutar un ensayo aleatorio controlado)	Entrenamiento y evaluación con datos históricos de exámenes de fondo de ojo y con el diagnóstico correspondiente.



En esta etapa es necesario hacer un diagnóstico de las consideraciones éticas, legales y de gobernanza del proyecto. Esto permitirá dimensionar sus riesgos y planificar adecuadamente según su complejidad. Cabe advertir nuevamente que los algoritmos y reglas automatizadas para la toma de decisiones pueden reproducir sesgos y que por ello la ciudadanía se muestre reacia a su implementación. También es posible que en los modelos haya grupos de la población subrepresentados que por ello resulten perjudicados con la implementación de la herramienta. fAIr LAC ha identificado los siguientes retos transversales a todo proyecto de IA:

Protección de datos personales: Al implementar un sistema de toma y/o soporte de decisiones basado en IA, es probable que se trabaje con datos personales¹⁷ sensibles tales como nivel de ingresos, diagnósticos médicos u otra información de salud e identificadores únicos como el documento nacional de identidad (DNI) o los datos biométricos.

El director del proyecto debe asegurarse de que el tratamiento que se le dé a dicha información cumpla con la normativa nacional de protección de datos y que se guíe por las buenas prácticas internacionales.

El tratamiento de los datos personales debe tener siempre una base de licitud, o sea, un sustento que dé legitimidad jurídica a lo que se va a hacer con ellos. Dependiendo del tipo de proyecto, y de la normativa de datos personales en cada país, la base de licitud puede variar. Lo más común es el consentimiento informado, así como las leyes orgánicas de las agencias públicas, el interés legítimo, el interés vital o la existencia de un contrato. Por eso es importante estar familiarizados con la ley de protección de datos del país en cuestión.¹⁸

Una buena práctica es contar con el consentimiento de las personas en estudio para que se utilicen sus datos, siempre que esto sea posible. Se entiende por consentimiento la “manifestación de la voluntad, libre, específica, inequívoca e informada del titular a través de la cual acepta y autoriza el tratamiento de los datos personales que le conciernen” (Red iberoamericana de protección de datos, 2017).

Para que el consentimiento sea informado, los individuos deben conocer como mínimo la identidad del responsable del tratamiento y los fines para los cuales se destinarán los datos personales. Igualmente deben estar al tanto de la política de seguridad en el manejo y almacenamiento de sus datos, de los beneficios asociados con su uso, y de los riesgos potenciales y medidas de mitigación.

Una buena práctica es anonimizar o seudonimizar las observaciones personales; sin embargo, el director del proyecto no puede olvidar que los “datos personales no se limitan a nombres y apellidos, sino que además incorporan cualquier elemento que pueda conducir a la identificación del sujeto determinado” (Buenadicha et al., 2019). Por ello no solamente se debe tener cuidado en eliminar las personalizaciones obvias en los datos, sino también codificar las posibles rutas y cruces de datos que permitirían identificar a un individuo.

¹⁷ Los datos personales son aquellos que permiten identificar a una persona de manera directa o indirecta (Red Iberoamericana de protección de datos, 2017).

¹⁸ Para más información sobre las leyes de protección de datos personales en los distintos países de América Latina, véanse Bojalil y Vela-Treviño (2018).

Tanto el conocimiento acerca de la utilización de datos personales y sensibles, como la necesidad de contar con un consentimiento informado, deben ser tomados en cuenta en el proceso de conceptualización y diseño del proyecto. Si durante su implementación cambia el uso de los datos, esto deberá ser debidamente comunicado a los individuos para que den su consentimiento para el nuevo uso.

Responsabilidad y transparencia: El director del proyecto será el responsable final de la implementación de la herramienta de IA y quien deberá comunicarle a la ciudadanía sus beneficios, riesgos y las correspondientes medidas de mitigación. Por responsabilidad se entiende la disposición a rendir cuentas por las acciones o actividades propias, creaciones o personas a cargo, aceptando las consecuencias de estos actos (Oliver, 1994). En el caso de la IA, implica responsabilidad por las decisiones que toma el sistema y por las afectaciones que este pueda ocasionar. Para ello, el director del proyecto debe crear mecanismos simples para que los afectados puedan presentar solicitudes de revisión de las decisiones o quejas con respecto al funcionamiento de la herramienta.

En cuanto a la transparencia algorítmica, se trata de un principio ético esencial en la IA porque funciona como un habilitador de otros principios: permite saber qué datos se utilizan, quiénes los utilizan, cómo se utilizan y cómo afectan las decisiones de política pública (Sangüesa, 2018). Cuando un sistema es transparente se puede saber, por ejemplo, si se protegen los datos y si los resultados son equitativos.

A menudo, los sistemas de toma y/o soporte de decisiones basados en IA suelen verse como cajas negras: la población en general no sabe cuáles de sus datos se están usando, cómo se emplean, qué significa que ahora se tomen decisiones con base en un sistema automático, y cuál es el impacto final que esto tendrá en la sociedad. Aquí el rol del director del proyecto es suministrar a la ciudadanía la información necesaria y suficiente para que comprenda los mecanismos generales de la herramienta, y para que tenga seguridad sobre la protección de sus datos y sobre la forma en que se tomarán las decisiones.

Más allá de la transparencia del modelo mismo, será el director del proyecto quien tenga la responsabilidad de informar a la ciudadanía en general y a los actores involucrados sobre los distintos pasos y decisiones tomados en la etapa de implementación e incluso durante la evaluación del modelo en el terreno. Una buena práctica es ir documentando todas las decisiones adoptadas para llevar registro e informar de manera adecuada cuando así se requiera.

Interpretabilidad y explicabilidad: Este reto surge de la necesidad de comprender los modelos de soporte y toma de decisiones basados en la IA. Dado que, como ya se indicó, los algoritmos suelen verse como cajas negras, aquí el rol del director del proyecto consiste en explicar a la población el funcionamiento general de la herramienta.

La explicabilidad y la transparencia algorítmica de un modelo, así como la cantidad de información que se pueda dar sobre el funcionamiento de un algoritmo, dependerán de su nivel de opacidad (Burrell, 2016; Buenadicha et al., 2019). Existen tres niveles:

Opacidad intencional: Habrá situaciones donde el algoritmo no podrá ser explicado a la ciudadanía porque pondría en riesgo la efectividad del modelo mismo. Por ejemplo, si se aplica un modelo para predecir la evasión de impuestos, la ciudadanía no debe saber cómo funciona específicamente, ya que algunos podrían intentar “engañar” al modelo. Esta opacidad intencional también puede presentarse cuando hay motivos de propiedad intelectual.

Opacidad analfabeta: Ocurre cuando las personas a las que se explica el algoritmo no comprenden cómo funciona por una falta de conocimiento en los temas técnicos. En este caso se recomienda capacitar a los funcionarios de primera línea de la institución para que puedan explicar, en un lenguaje claro, el funcionamiento del sistema.

Opacidad intrínseca: En ocasiones los algoritmos utilizados son sumamente complejos, por lo cual es difícil evidenciar una relación causal entre valores entrantes y salientes. Si existe este tipo de opacidad, resulta imposible explicar el modelo.

Algunas herramientas son más explicables (más fáciles de entender para la población en general) que otras. Por eso el director del proyecto deberá informar claramente al equipo técnico sobre las exigencias de explicabilidad a la luz de las regulaciones nacionales y de la definición del problema, dado que todo ello impactará el modelo que el equipo desarrolle. Si bien no siempre es necesario dar toda la información, la ciudadanía debe estar al tanto de cómo se están tomando las decisiones de política pública.

Además de los retos descritos anteriormente, desde la fase de diseño de un proyecto es importante preguntarse sobre la posibilidad de que la implementación de la herramienta dé lugar a discriminaciones negativas. Estas se pueden originar en:

Sesgos: El error del sistema es la diferencia entre el valor predicho resultado del modelo y el valor real de la variable que se está estimando. Si el error es sistemático en una dirección o en un subconjunto específico de datos, se llama **sesgo**. Por ejemplo, si una balanza siempre arroja un peso de un kilo de más, está sesgada. Asimismo, si un valor es siempre menor —como el salario de las mujeres versus el de los hombres para un trabajo equivalente—, la variable salario está sesgada. Cuando el error no es sistemático sino aleatorio, se le denomina **ruido** ([González, Ortiz y Sánchez Ávalos, 2020](#)).

El sesgo de un sistema de IA puede tener implicaciones éticas cuando se utilizan sus resultados para tomar decisiones de política pública que conduzcan a acciones que puedan considerarse injustas o prejuiciosas para algunos subgrupos de la población destinataria. Por ejemplo, un modelo o herramienta de IA podría favorecer o perjudicar constantemente a un mismo sector de la población. Existen distintos tipos de sesgos, algunos causados por problemas intrínsecos a los datos. Están los **sesgos históricos** o estados indeseados, los cuales se registran cuando en el mundo se presentan patrones que no se quieren reproducir o propagar en el modelo; los **sesgos de representación**, los cuales se producen cuando la información está incompleta, ya sea por atributos faltantes, por el diseño de la muestra o por la ausencia total o parcial de subpoblaciones; y los **sesgos de medición**, los cuales se originan en el uso u omisión de variables que van a ser utilizadas en los modelos (Harini Suresh, 2019).

Para entrenar una herramienta de IA se requieren datos históricos con los cuales esta irá “aprendiendo” y mostrando resultados. Corresponde al director del proyecto, dado su conocimiento experto acerca del problema y de la situación actual, prever qué sectores de la población podrían resultar subrepresentados en los datos o qué problemas de sesgo podrían afectarlos. Los sesgos se pueden mitigar con alguna medida de reparación o bien con una calibración de la herramienta de IA por parte del equipo técnico en la etapa de ejecución.

Si los datos se encuentran sesgados, la herramienta también podría estarlo. Una herramienta

sesgada pone en riesgo la viabilidad de un proyecto de IA. Por ejemplo, Amazon desarrolló una herramienta de contratación basada en IA, pero debió retirarla porque favorecía significativamente el enganche de hombres. Esto sucedía porque la industria tecnológica ha estado históricamente dominada por ellos, lo que hacía que los datos históricos de contrataciones se encontraran sesgados. Dada la discriminación arbitraria de la herramienta, tuvo que ser desechada (Dastin, 2018).

Desigualdades en el proceso a intervenir: Es probable que la situación actual, sin proyecto de IA, registre desigualdades entre diferentes subgrupos de la población. Por ejemplo, las tasas de suicidio son mayores entre los jóvenes LGBTQ+ que entre la población juvenil en general. Por el lado de la violencia intrafamiliar, la mayor proporción de víctimas son mujeres. El director del proyecto debe estar al tanto de estas líneas de referencia para incorporar las consideraciones del caso y analizar si una herramienta de IA puede o no ayudar a resolver el problema en cuestión.

Subgrupos de la población para los cuales se quiere garantizar igualdad: Si bien es cierto que una base de datos podría estar libre de sesgos, suele suceder que algunos grupos minoritarios de la población pierden representación en el modelo de IA. Esto por cuanto, al utilizar datos históricos, el modelo dará el mismo peso a todos los individuos u observaciones, lo que de por sí conduce a que haya menor representación de aquellos grupos que constituyen un porcentaje menor de la población. En su calidad de experto en el problema de política pública que se va a abordar, el director del proyecto es quien tiene claridad sobre la manera en que la implementación de la herramienta de IA puede afectar de forma distinta a los grupos minoritarios, y quien por lo tanto deberá asegurar equidad en la intervención e implementación de la herramienta definiendo grupos o atributos protegidos. Por ejemplo, si se quiere implementar un sistema de detección de una enfermedad, no sería aceptable que este detecte la patología con distintas tasas de precisión para hombres y mujeres. O si en una cierta institución se fuera a aplicar un modelo automático de promociones, esto podría favorecer a las personas con más experiencia frente a los jóvenes.

Corresponderá al director del proyecto informar al equipo técnico sobre las discriminaciones que puedan originarse en los datos, identificando claramente los subgrupos de población para los cuales se quiere garantizar igualdad y tomando las medidas que garanticen justicia algorítmica (es decir, que el algoritmo no perjudique a algunos subgrupos de la población y que los resultados sean equitativos para todos). Actualmente existen métricas que permiten determinar las inequidades que puedan surgir en la etapa de modelamiento. El director del proyecto deberá solicitar al equipo técnico que conduzca distintas pruebas con las cuales se pueda medir el nivel de justicia algorítmica en la herramienta (siguiendo definiciones como paridad demográfica, equidad de posibilidades o justicia contrafactual, entre otras).¹⁹ Dado que existirán casos donde no sea posible intervenir la herramienta, se podrán adoptar medidas de mitigación de sesgos y discriminación, ya no desde la IA sino desde la política pública.

Además de los puntos arriba señalados, el director del proyecto deberá preguntarse en esta fase si este último cuenta con **licencia social**. Por licencia social se entiende la aceptación, por parte de la ciudadanía, de la implementación de la herramienta de IA para un sistema de toma y/o soporte de decisiones (Data Futures Partnership, 2017). No basta solamente con cumplir con los marcos legales de los países; se requiere un paso adicional que consiste en obtener la aceptación de la población. En un principio la IA fue ampliamente acogida, dado que significaba un aumento en los rendimientos

¹⁹ Para más detalles, véanse González, Ortiz y Sánchez Ávalos (2020).

y una disminución en los tiempos de trabajo, traslado, etc. Sin embargo, en épocas más recientes las personas se han vuelto cada vez más conscientes de los posibles efectos adversos de la IA en sus vidas, por lo cual resulta clave obtener licencia social si se quiere implementar un proyecto de tal naturaleza, independientemente si lo hace una institución pública o una privada (Hewitt, 2019). Para saber si un proyecto cuenta con licencia social, su director podrá referirse a estudios de opinión sobre el valor público que se espera generar con la herramienta versus los riesgos que podría conllevar, y hacer reuniones con grupos de interés y comunidades donde se realizará su implementación para conocer sus percepciones al respecto.

Gobernanza y seguridad: Los retos de gobernanza y seguridad son transversales a los proyectos de IA, y guardan relación con las regulaciones y con la seguridad de la infraestructura en la que se desarrolla la herramienta. Aquí se identifican tres retos:

Normas y regulaciones: El alcance de un proyecto de IA se enmarcará dentro de las normas y regulaciones propias de los países donde se esté implementando, y del sector de la política pública en el cual se use la herramienta. Es de esperar que el proyecto se rija por las leyes de protección de datos personales y de acceso a la información. Cabe notar que hay sectores específicos, como por ejemplo el de la salud, que se encuentran regulados por normas especiales y que varios países tienen leyes antidiscriminación. Por eso, cuando quiera que se vaya a implementar un sistema de IA, se requiere hacer una recopilación de todas las normas pertinentes. Contar con esta información en la etapa de conceptualización contribuirá a la viabilidad de la implementación. Se recomienda acudir al departamento legal de la institución e incluirlo en el equipo del proyecto.

Ciberseguridad: Los sistemas de IA se entrenan con datos, de modo que tanto los datos personales como los resultados de los algoritmos deben estar debidamente protegidos para evitar filtraciones de información sensible perteneciente a los ciudadanos. Por eso el director del proyecto debe establecer protocolos de seguridad para el almacenamiento y manejo de datos, y también para cuando ocurran robos de información.

Penetración ilegal de la IA y ataques adversarios: Más allá del robo de información sensible, el sistema puede ser objeto de ataques que busquen confundir al algoritmo, donde el atacante se hará pasar por usuario de la herramienta. Aquí también deberán instalarse sistemas de seguridad para evitarlos, y protocolos de respuesta en caso de que se produzcan.

No se olvide de...

1. **Realizar** un inventario de las leyes del país aplicables a su proyecto para cumplir con los requerimientos legales.
2. **Contar** con los permisos legales necesarios para utilizar información que puede ser personal y/o sensible.
3. **Determinar** todos los riesgos éticos que puedan surgir en la implementación de la herramienta.
4. **Definir** claramente aquellos grupos de población a los que se busca garantizar igualdad en la intervención (justicia algorítmica).
5. **Evaluar** si el proyecto cuenta con licencia social, es decir, si la población acepta el uso de IA que se propone.

Actividad



Concluida esta sección, complete lo correspondiente a **consideraciones éticas y legales** en la ficha de diseño y factibilidad de proyectos.

FICHA DISEÑO Y FACTIBILIDAD DE PROYECTO

7

Consideraciones éticas y legales

Proporcionalidad	¿Cree que un sistema de ciencia de datos/IA es el medio adecuado para resolver el problema? ¿Por qué? ¿Ha evaluado otras alternativas?	
	¿Qué impactos negativos podría tener su proyecto? Estudie casos de uso similares identificados en la sección "Definición del problema".	
Licencia social	¿Cree que los usuarios o afectados encontrarán aceptable el uso de datos planteado para resolver el problema? ¿Por qué?	
	Si la población del proyecto se entera del uso de la IA para los propósitos indicados, ¿lo aprobará? ¿Por qué?	
Protección de datos	¿Está trabajando con datos personales y/o sensibles identificables a nivel individual? ¿Cuáles?	
	¿Ha identificado la justificación o base legal para trabajar con esos datos?	
	¿Ha identificado las regulaciones que podrían incidir en el proyecto?	
	¿Se requieren mecanismos (de acceso, eliminación o rectificación, por ejemplo) para garantizar la calidad de los datos personales?	

Transparencia	¿Qué actores interesados deberían estar al tanto del proyecto? (responsables por la formulación de políticas públicas, trabajadores de primera línea, organizaciones de la sociedad civil, organismos públicos, personas que se verán afectadas por las acciones, etc.). Mencione organizaciones/ personas específicas.	
	¿Ha considerado algún mecanismo para que los actores interesados se comuniquen con la institución para obtener información sobre el proyecto?	
	¿Será necesario explicar los mecanismos de toma de decisiones o los análisis a implementar? ¿Por qué?	
Discriminación/ equidad	¿Qué desigualdades estructurales hay en el proceso o entorno donde se inserta el proyecto?	
	¿Existen grupos específicos (vulnerables) para los que se busca garantizar la equidad en los resultados o la protección de sus derechos? (P. ej., por género, edad, localización, clase social, nivel educativo, origen urbano o rural, etnia).	
	¿Qué sesgos cree que podrían tener los datos?	
Rendición de cuentas	En caso de ocurrir un requerimiento de información respecto del proyecto ¿quién se encarga de responder?	
	¿Quién es responsable si el sistema se equivoca?	
	¿Tiene previsto mecanismos de monitoreo, control y evaluación? ¿Cómo se documentarán y qué periodicidad tendrán?	

FICHA DISEÑO Y FACTIBILIDAD DE PROYECTO

EJEMPLO DART: Uso de la inteligencia artificial en la prevención de la ceguera

7 Consideraciones éticas y legales

Proporcio- nalidad	¿Cree que un sistema de ciencia de datos/IA es el medio adecuado para resolver el problema? ¿Por qué? ¿Ha evaluado otras alternativas?	Sí, dado que no es viable elevar sustancialmente el número de oftalmólogos en el corto plazo y que el proyecto posibilita una ampliación de la cobertura. Esto contribuirá a su vez a aumentar la detección temprana de la retinopatía diabética y a prevenirla, reduciendo eventualmente los gastos del Ministerio de Salud. Asimismo, se requiere de una herramienta que permita aumentar de manera significativa el número de exámenes de fondo de ojo que se reportan. La opción de formar más oftalmólogos es una solución de muy largo plazo.
	¿Qué impactos negativos podría tener su proyecto? Estudie casos de uso similares identificados en la sección "Definición del problema".	Un posible riesgo es que la calidad de las imágenes sea insuficiente para realizar el análisis. Un impacto adverso podría ser que los resultados de los exámenes fueran falsos negativos y que por ello no fueran analizados por un oftalmólogo. Lo anterior causaría un daño a los pacientes cuyos resultados son incorrectos y sí sufren de esa enfermedad.
Licencia social	¿Cree que la población encontrará aceptable el uso de datos planteado para resolver el problema? ¿Por qué? Si la población del proyecto se entera del uso de la IA para los propósitos indicados, ¿lo aprobará? ¿Por qué?	Sí, porque los pacientes obtendrán un claro beneficio en cuanto a la oportunidad de que les revisen sus exámenes, lo cual aumenta la cobertura del chequeo anual y libera el tiempo de los oftalmólogos, quienes ahora podrán enfocarse en los casos más extremos. En cuanto a una posible resistencia de oftalmólogos y tecnólogos médicos, esta podría originarse en el hecho de que deberán cambiar el método de realizar los exámenes y diagnosticar los casos. Es importante incluirlos en el proyecto para lograr su compromiso con la implementación. Sí, puesto que reporta un claro beneficio para todos dado que libera recursos humanos y financieros de los centros de salud y focaliza los existentes en los casos que los necesitan. De todas maneras es importante informar sobre los beneficios de largo plazo del uso de la herramienta en los centros de salud, dada la inversión inicial que debe realizar el Ministerio del ramo. Sin embargo, puede existir una cierta resistencia a que el resultado de un examen sea producido por un algoritmo. Es necesario entonces incorporar el factor humano y que un especialista confirme los casos de retinopatía diabética encontrados automáticamente, para que luego el equipo de salud comunique el diagnóstico al paciente. En el momento de la implementación será necesario ajustar las expectativas de los pacientes, dado que la herramienta solo ha sido diseñada para detectar la retinopatía diabética y no otras enfermedades oculares.

Protección de datos	¿Está trabajando con datos personales y/o sensibles identificables a nivel individual? ¿Cuáles?	Sí, son datos sensibles porque se refieren al estado de salud de las personas.
	¿Ha identificado la justificación o base legal para trabajar con esos datos?	Sí, la protección de datos personales según la ley 19.628. Ley de derechos y deberes del paciente (N°20.584) y su reglamento, CSAN, DFL 1 Salud de Chile.
	¿Ha identificado las regulaciones que podrían incidir en el proyecto?	Sí, la protección de datos personales según la ley 19.628. Ley de derechos y deberes del paciente (N°20.584) y su reglamento, CSAN, DFL 1 Salud de Chile.
	¿Se requieren mecanismos (de acceso, eliminación o rectificación, por ejemplo) para garantizar la calidad de los datos personales?	Los pacientes deberán poder acceder a las imágenes de sus exámenes de fondo de ojo y a los resultados de la revisión del oftalmólogo en caso de que los requieran, por ejemplo para asistir a una consulta médica privada. De otro modo podrían negarse a que les practiquen el examen.
Transparencia	¿Qué actores interesados deberían estar al tanto del proyecto? (responsables por la formulación de políticas públicas, trabajadores de primera línea, organizaciones de la sociedad civil, organismos públicos, personas que se verán afectadas por las acciones, etc.). Mencione organizaciones/ personas específicas.	Las jefaturas de servicios de salud y los consultorios que implementen el software, el Ministerio de Salud, la Sociedad Chilena de Oftalmología (SOCHIOF), el Colegio de Tecnólogos Médicos de Chile, la Confederación Nacional de Funcionarias y Funcionarios de Salud Municipal (Confusam), la Asociación de Diabéticos de Chile (Adich) y la Fundación Diabetes Juvenil de Chile.
	¿Ha considerado algún mecanismo para que los actores interesados se comuniquen con la institución para obtener información sobre el proyecto?	Sí, por medio de reuniones periódicas con los responsables en el Ministerio de Salud. El Ministerio de Salud, por su parte, también recolecta información de percepción en los centros de salud.
	¿Será necesario explicar los mecanismos de toma de decisiones o los análisis a implementar? ¿Por qué?	Sí, aunque no el algoritmo propiamente tal por ser altamente técnico. Se deberá poder explicar su validación clínica.

Discriminación/ equidad	¿Qué desigualdades estructurales hay en el proceso o entorno donde se inserta el proyecto?	Existen desigualdades en cuanto a los factores de riesgo de diabetes (Sapunar, 2016). El riesgo de diabetes es más elevado en personas con más edad, de sexo femenino, y que viven en entornos urbanos y tienen menor nivel educativo. También se ha encontrado que hay menos riesgo en los pueblos originarios.
	¿Existen grupos específicos (vulnerables) para los que se busca garantizar la equidad en los resultados o la protección de sus derechos? (P. ej., por género, edad, localización, clase social, nivel educativo, origen urbano o rural, etnia).	No; se analizarán los exámenes de todos los pacientes diabéticos. Es posible que el algoritmo tenga distinto desempeño para algunos subgrupos de la población diabética y por ello sería recomendable realizar un análisis de disparidades.
	¿Qué sesgos cree que podrían tener los datos?	Ninguno, ya que los factores de riesgo anteriormente descritos no deberían influir en la calidad de la imagen de fondo de ojo.
Rendición de cuentas	En caso de ocurrir un requerimiento de información respecto del proyecto ¿quién se encarga de responder?	El Ministerio de Salud y Teledx
	¿Quién es responsable si el sistema se equivoca?	El Ministerio de Salud.
	¿Tiene previsto mecanismos de monitoreo, control y evaluación? ¿Cómo se documentarán y qué periodicidad tendrán?	Sí; está previsto un estudio observacional una vez que la herramienta se implemente en suficientes centros, con el fin de obtener una muestra lo suficientemente amplia. Asimismo se realizará un monitoreo periódico de la herramienta en cuanto a la calidad de las predicciones, los sesgos algorítmicos, la experiencia de los usuarios y los análisis de brechas de acceso a los exámenes.



1.8 Conformación del equipo

Una parte crítica del proceso de conceptualización y diseño del proyecto consiste en identificar a las personas de la institución —o bien externas— que deben hacer parte del equipo multidisciplinario a cargo e involucrarlas desde el comienzo. Su participación activa en la formulación contribuirá a asegurar sostenibilidad de la herramienta durante la implementación.

A lo largo de todo este proceso, el director del proyecto ha ido conversando con personas internas y externas a la institución sobre la iniciativa propuesta e incorporando sus aprendizajes en el diseño. Por ejemplo, si se trata de un proyecto de optimización de fiscalizaciones aduaneras, seguramente se habrá reunido con la jefatura a cargo de ello para entender a fondo cómo se llevan a cabo actualmente, cómo se decide qué inspeccionar, cuánto tiempo demora y qué problemas se encuentran en el terreno. También es probable que haya contactado al departamento de estadísticas para solicitar información sobre los datos existentes y analizar su madurez. Todas estas personas deben formar parte del equipo del proyecto, ya que el impacto los atañe a ellos también. En el caso del equipo a cargo de las fiscalizaciones, serán sus integrantes quienes utilicen la herramienta de IA, por lo cual es importante que tengan pleno conocimiento de las dificultades actuales y futuras en la implementación. El departamento de estadísticas seguramente estará a cargo de administrar internamente los datos, y por ello serán sus expertos los encargados de suministrárselos al equipo técnico.

A continuación se describen algunos cargos de los posibles integrantes de un proyecto de IA y sus funciones. Esta lista es solamente una sugerencia y por lo tanto no se aborda en detalle la conformación del equipo técnico, pues esto dependerá de la naturaleza del proyecto.

Director del proyecto (responsable por la toma de decisiones)	Realiza el rol principal en la etapa de conceptualización y diseño y está a cargo de supervisar todos los pasos que dé el equipo técnico en la etapa de ejecución.
Referente del proceso	Participa en el proceso donde se implementará la herramienta y de IA; por su experiencia y conocimiento es respetado por sus pares. Su conocimiento experto y reconocimiento contribuirán a la adopción de la tecnología.
Responsables de los datos (internos y/o externos)	Suministran información sobre calidad de los datos necesarios para la implementación de la herramienta. Determinan las necesidades en cuanto a su recolección, almacenamiento y manejo. Aplican su conocimiento experto para detectar posibles anomalías en los datos.
Analistas de datos	Realizan el análisis exploratorio de los datos y desarrollan la herramienta. En caso de que no existan estas capacidades dentro de la institución, el director del proyecto deberá contratar este servicio con una institución externa.
Equipo legal	Vela por el cumplimiento de las normas legales presentes en cada país en cuanto a protección de datos personales, transparencia y regulaciones sectoriales.

Línea directiva de la institución	Apoya y promueve el desarrollo de la herramienta. Es clave para que el proyecto sea viable y pueda ser implementado.
Equipo que ejecuta la acción	Tiene el conocimiento experto sobre la forma en la cual se responde actualmente al problema y sobre cómo podría mejorarse. Presenta inquietudes sobre la forma en que los funcionarios van a interactuar con la herramienta y las posibles consecuencias de ello.
Comunicaciones	Comunica interna y externamente los aspectos claves de la implementación de la herramienta.
Departamento de estudios o evaluación	Contribuye a diseñar los aspectos metodológicos de la evaluación del proyecto.

No se olvide de...

1. **Considerar** a todos los actores relevantes al conformar el equipo.
2. **Lograr** el compromiso de los integrantes del equipo con el proyecto.
3. **Asegurar** que todos los miembros del equipo participen en la conceptualización y diseño del proyecto.

Actividad



Concluida esta sección, complete lo correspondiente a **conformación del equipo** en la ficha de diseño y factibilidad de proyectos.

FICHA DISEÑO Y FACTIBILIDAD DE PROYECTO

8 Conformación del equipo

Organización/ departamento	Nombre/rol de la contraparte	Descripción de la participación deseada

FICHA DISEÑO Y FACTIBILIDAD DE PROYECTO

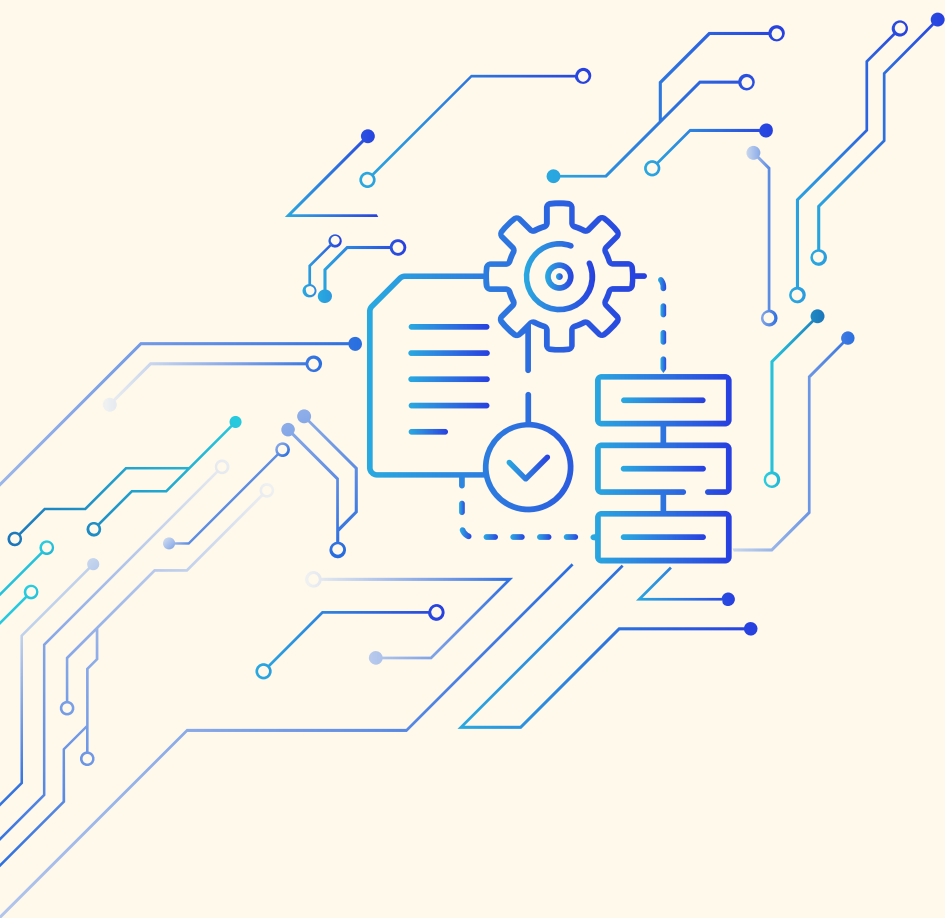
EJEMPLO DART: Uso de la inteligencia artificial en la prevención de la ceguera

8 Conformación del equipo

Organización/ departamento	Nombre/rol de la contraparte	Descripción de la participación deseada
Teledx	Analistas de datos	A cargo de la creación de la herramienta de análisis de imágenes
Ministerio de Salud	Subsecretaría de redes asistenciales	Implementación de la herramienta en los centros de salud
Ministerio de Salud	División jurídica	Asesoría de cumplimiento normativo del proyecto
Teledx/Ministerio de Salud	Equipo de capacitación	Capacitación a oftalmólogos y tecnólogos médicos sobre el uso de la herramienta



PARTE 2. EJECUCIÓN

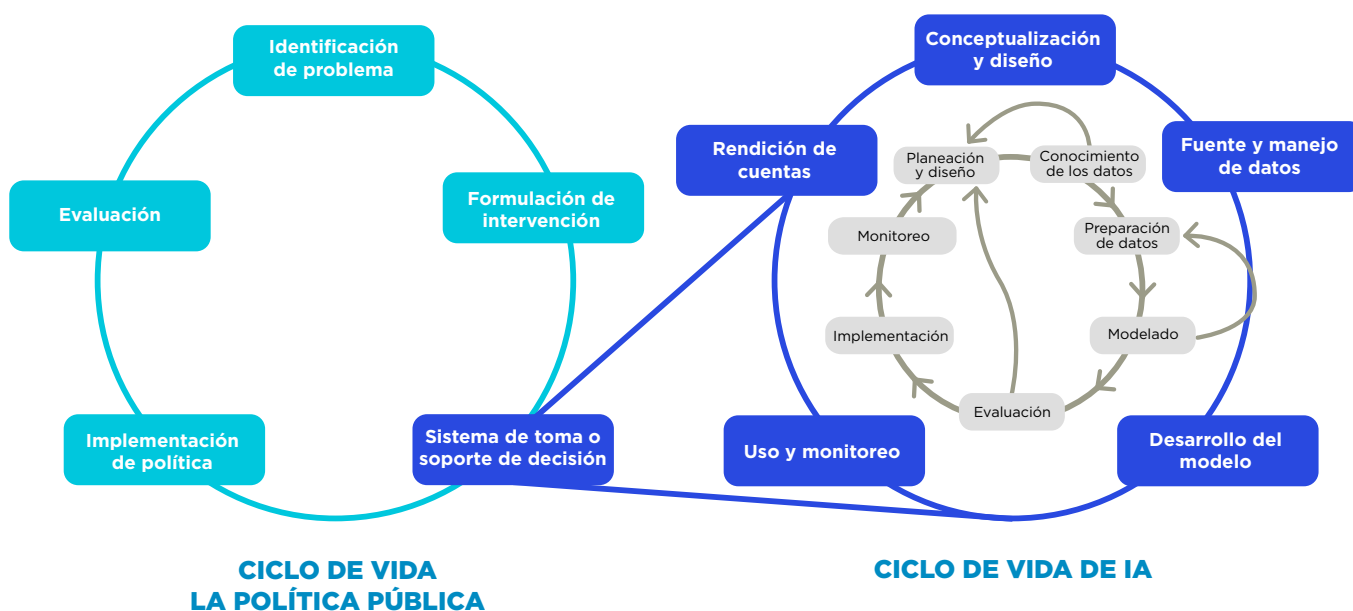


PARTE 2. EJECUCIÓN

Una vez que el proyecto ha sido formulado, y que la institución ha decidido llevarlo adelante, es el momento de ejecutarlo. ¿Cómo se lleva a cabo? ¿En qué aspectos es necesario enfocarse durante esta fase? ¿Hasta qué punto deben llegar la transparencia y la rendición de cuentas sobre las especificidades del modelo?

En este manual se describen los retos que implica un proyecto de IA tanto para su director como para el equipo técnico, según el ciclo de vida de aquella. Estos retos también pueden entenderse a través de las diferentes etapas que el equipo responsable debe desarrollar, donde el director desempeña una serie de funciones clave (figura 5):

Figura 5. Conexión entre la política pública y el desarrollo de la IA



Fuente: Fuente: González, Ortiz y Sánchez Ávalos, 2020.

1. Recolección y procesamiento de datos: En la etapa de conceptualización y diseño, el director del proyecto ya hizo el mapeo de las distintas bases de datos y determinó su grado de madurez. En esta etapa el equipo técnico iniciará un análisis exploratorio de los datos para obtener más información e identificar los posibles retos que puedan surgir durante su uso.

2. Desarrollo y validación del modelo: Una vez que se han definido las variables a utilizar y que se han entendido los desafíos que presentan los datos, es el momento de que el equipo técnico comience a desarrollar el modelo utilizando técnicas de modelamiento específicas. En la etapa de conceptualización y diseño, el director del proyecto debe haber comunicado al equipo técnico el objetivo del análisis: “¿Qué tipo de análisis (o herramienta) es más apropiado para solucionar el problema”? En este punto, el equipo técnico deberá presentar distintas opciones de modelamiento al director del proyecto, y de manera conjunta decidirán cuál es la tecnología más adecuada acorde

con criterios técnicos y de equidad. Es posible que durante el desarrollo y validación del modelo, el equipo técnico identifique nuevas necesidades de datos, por lo que se deberán realizar iteraciones del modelo integrando las nuevas fuentes de información. Este proceso se deberá repetir hasta que el o los modelos generados sean satisfactorios para el equipo técnico y el director del proyecto.

3. Uso y monitoreo: En el momento de implementar la herramienta de IA en la población destinataria, o en un subconjunto de ella, se deberá medir su impacto teniendo en cuenta los objetivos del proyecto. También será necesario implementar un sistema periódico de monitoreo para subsanar errores que pueden surgir durante su implementación y destinar los recursos necesarios para asegurar el uso de la herramienta a lo largo del tiempo.

4. Rendición de cuentas: Siendo el director del proyecto la persona que en últimas está a cargo de la implementación de la herramienta de IA, es su responsabilidad documentar, junto con el equipo técnico, todo el proceso de desarrollo para poder justificar frente a la ciudadanía y otras partes interesadas las decisiones tomadas a lo largo de todo el ciclo de vida de la IA.

En los pasos 1 y 2 arriba descritos, será el equipo técnico el principal responsable del modelo, dado que estará a cargo del modelamiento de los datos. Sin embargo, corresponderá al director del proyecto supervisar cada uno de los pasos y decisiones del equipo técnico. Al inicio de cada paso se presenta el rol del director del proyecto.

En el manual de [González, Ortiz y Sánchez Ávalos \(2020\)](#), al cual se ha hecho referencia a lo largo de este texto, se encuentran dos materiales dirigidos al equipo técnico que se incluirán a continuación para promover una comunicación fluida entre este y el director del proyecto: el perfil de datos y el perfil del modelo de IA. Uno y otro apoyarán la documentación de las decisiones del proyecto en aras de la transparencia y la rendición de cuentas.

1. Perfil de los datos:

Contiene los principales hallazgos de las bases de datos que se utilizarán con la herramienta. Se fundamenta en los análisis exploratorios de los datos realizados por el equipo técnico bajo la supervisión del director del proyecto, con quien además discutirá tales hallazgos.

2. Perfil del modelo:

Descripción final de un modelo de IA. Contiene aspectos críticos para todo el ciclo de vida de la IA.

Para ambas herramientas se explican las funciones del director del proyecto en cada uno de los aspectos claves de las etapas de ejecución.



Rol del director del proyecto

Durante la fase correspondiente a la recolección y procesamiento de los datos, el director del proyecto estará a cargo de:

- contextualizar los hallazgos del equipo técnico en la etapa exploratoria de los datos;
- proporcionar los antecedentes del proceso de recolección de los datos y el significado de las variables, conectando al equipo técnico con las personas de la organización que se requieran para una mejor comprensión de los datos;
- validar la formulación y utilidad de la variable objetivo;
- tomar decisiones sobre imputaciones de datos o eliminación de variables, ponderando y registrando lo que se obtiene versus lo que se pierde (*trade-offs*);
- proponer mejoras a los procesos de gobernanza de los datos del proyecto que puedan contribuir al logro de los objetivos y faciliten la implementación del modelo; y
- hacer ajustes a la definición del problema, a los objetivos y al plan de ejecución del proyecto según los resultados de esta etapa.

En esta instancia, el equipo técnico deberá realizar, en primer lugar, un análisis exploratorio de los datos. Si bien es de esperar que el director del proyecto tenga un conocimiento básico de las bases de datos y su madurez —según lo registrado en el punto correspondiente al Mapeo de Datos (1.5 de la primera parte)—, es en esta fase cuando el equipo técnico procede a hacer un análisis detallado y completo de aquellos utilizando herramientas estadísticas convencionales.

Aquí el rol del director del proyecto es actuar como puente entre el equipo técnico y las áreas que generan o administran los datos, suministrando información sobre la metodología de recolección, la frecuencia de actualización e incluso el diccionario de variables. Esto con el fin de hacer más fluida la comunicación entre ambas partes, en caso de que surjan dudas metodológicas o conceptuales.

En segundo lugar, cuando se realiza el análisis exploratorio, el equipo técnico puede encontrarse con situaciones consideradas anómalas. Aquí, dado que el director del proyecto es la persona que tiene el mayor conocimiento especializado sobre el problema de política pública que se busca resolver y que puede conectarse con especialistas sectoriales que tengan la experticia del caso, deberá ayudar a identificar los hallazgos potenciales en los datos que pueden ser producto de algún contexto sectorial desconocido para el equipo técnico.

En esta etapa del proyecto de IA pueden surgir diversos retos relacionados con la calidad de los datos disponibles, algunos de las cuales se enumeran a continuación. Estos retos también se consignan en el manual de González, Ortiz y Sánchez Ávalos (2020) dirigido al equipo técnico²⁰, aunque aquí se abordan desde la perspectiva del director del proyecto, quien como responsable final de la herramienta deberá solucionarlos.

20 <https://publications.iadb.org/es/ia-responsable-manual-tecnico-ciclo-de-vida-de-la-inteligencia-artificial>

Calidad y relevancia de la base de datos

A través del mapeo de datos realizado en la etapa de conceptualización y diseño, el director del proyecto ya tiene una aproximación a los datos que le permite determinar si están completos, si son relevantes y si cuentan con la granularidad necesaria para el análisis. Sin embargo, en su exploración el equipo técnico puede encontrar retos adicionales relacionados con la calidad de la base de datos, como por ejemplo:

Estados indeseables o subóptimos en datos recolectados. Si los datos están sesgados, la herramienta entrenada con ellos también lo estará, con lo cual se pone en riesgo su implementación. Para identificar estos sesgos, el director del proyecto solicitará al equipo técnico –con el cual ya habrá discutido oportunamente las implicaciones adversas de contar con una herramienta sesgada-- no solo que realice un análisis exploratorio de los datos en su conjunto, sino que también tome en cuenta, con base en el concepto de justicia algorítmica definido previamente, a aquellos subgrupos de la población que puedan resultar eventualmente afectados.

Las variables para las cuales se dispone de datos no son las ideales. Puede suceder que no se disponga de los datos necesarios sobre la variable objetivo que requiere el proyecto, es decir, aquella que se busca explicar o predecir (Greene, 2003). En tal caso se deberá tomar la decisión de reemplazarla por una variable sustituta (proxy) para la cual exista la información necesaria.

Por ejemplo, si se requiere medir la experiencia laboral y no se tiene ese dato en la base, pero sí se cuenta con la edad y los años de escolaridad del individuo, entonces una variable sustituta podría ser una combinación de las dos anteriores. Aunque con ello se busca aproximarse al máximo a la variable objetivo, no es lo mismo y en consecuencia su uso puede afectar la calidad de las predicciones. En este ejemplo la variable sustituta no identifica períodos de cesantía o desempleo, viajes al exterior prolongados e incluso permisos parentales, lo cual es especialmente complejo para las mujeres (dependiendo de la normativa nacional sobre permisos maternos). Por lo tanto, en caso de no contar con la variable ideal, el director del proyecto deberá comunicar al equipo técnico la necesidad de elaborar una variable proxy, si es posible.

Por su parte, el equipo técnico deberá mostrar los resultados de esa variable explicando claramente sus ventajas y desventajas. Corresponderá al director del proyecto decidir cuál será la variable sustituta a utilizar o si se debe replantear el problema de acuerdo con los datos disponibles.

Información incompleta acerca de la población destinataria

Tal como se vio en el numeral 1.5 sobre Mapeo de Datos en este documento, puede haber atributos faltantes o incompletos en la base de datos. Aquí el reto principal para el director del proyecto será aprobar el tratamiento se va a dar a dichas observaciones, a partir de las recomendaciones del equipo técnico. A continuación se resumen otras situaciones problemáticas de las cuales el director debe estar consciente:

Muestras probabilísticas y naturales pertinentes. Si la base de datos se alimenta de encuestas o entrevistas, el director del proyecto deberá saber que dichos datos corresponden a un subgrupo de la población. Existen dos extremos posibles de muestreo que guardan relación con la aleatoriedad de la selección de la muestra²¹. Una muestra no aleatoria puede conllevar

21 Para más información, véanse González, Ortiz y Sánchez Ávalos (2020, p. 18).

un sesgo de participación, por lo cual el director del proyecto debe tener claro que puede no ser representativa de la población general. Al analizar la forma de muestreo, el director del proyecto podría decidir modificar la forma en la cual se capturan los datos, lo que pondría en pausa el proyecto hasta que aquellos cumplan con los requisitos definidos.

Atributos faltantes o incompletos. A través del análisis exploratorio realizado se pueden detectar situaciones adicionales de atributos faltantes o incompletos. Las razones para ello pueden ser varias: problemas ocasionados por las personas responsables de ingresar los datos; fallas de sistema; problemas de transcripción y/o falta de información sobre ese atributo para el sujeto en cuestión. Cuando faltan atributos o están incompletos, el director del proyecto deberá preguntar al equipo técnico sobre la naturaleza de estos errores y si son aleatorios o no. Dependiendo de su naturaleza, y de la cantidad de errores, se deberá tomar la decisión de imputar datos o bien de eliminar observaciones de los datos, disminuyendo el tamaño de la muestra (ambos tienen ventajas y desventajas).

Comparación causal

El director del proyecto debe tener muy claro que los métodos tradicionales de aprendizaje automático no establecen relaciones causales entre las variables utilizadas (Varian, 2014). Sin embargo, sabrá igualmente que existen técnicas para intentar establecer causalidad en un análisis, lo cual deberá ser comunicado oportunamente al equipo técnico para que este decida cuál es la herramienta más apropiada para ello. En caso de que se requiera de un análisis de causalidad, entonces deberán recopilarse los datos necesarios según la metodología de causalidad elegida.



Al analizar en detalle los datos disponibles e identificar sus limitaciones, el director del proyecto y el equipo técnico pueden tomar medidas a futuro para mejorar la forma en la que se recopila y almacena la información. Por ejemplo, si no hay datos sobre la variable objetivo deseada, se puede gestionar un cambio en la captura de los datos.

Actividad

Concluida esta fase, complete junto con el equipo técnico lo correspondiente a la **recolección y procesamiento de datos en el Perfil del Modelo, así como en el Perfil de Datos.**

A continuación se presentan el Perfil de Datos, y la sección de recolección y procesamiento de datos del Perfil del Modelo, destacando el rol del director del proyecto en la elaboración de estas herramientas, que deberán ser completadas por el equipo técnico²².

PERFIL DEL MODELO	
Recolección y procesamiento de datos	
Datos de entrenamiento ²³	Conjunto de datos usados y etiquetado
	Pasos de preprocesamiento o preparación de datos
	Sesgos y deficiencias potenciales según el caso de uso (2)

Rol del director del proyecto

Revisar, junto con el equipo técnico, los resultados del análisis exploratorio de los datos. Explicar, con base en conocimiento experto, las anomalías encontradas (si las hay) y los resultados del análisis, y contextualizar los principales hallazgos.

Solicitar al equipo técnico un análisis exploratorio de los datos por subgrupos de población. Adoptar medidas de mitigación de sesgos y de equidad algorítmica.

PERFIL DE DATOS
Descripción general de datos y motivación
Nombre de conjunto de datos usado
¿Qué institución creó la base de datos?
¿Con qué objetivo creó la institución la base de datos utilizada?

Rol del director del proyecto

Conectar al equipo técnico con los responsables de los datos (tanto internos como externos) para obtener información detallada sobre las bases de datos, la forma de recolección y las posibles debilidades de la información.

²² Para más información sobre ambos perfiles (de datos y del modelo), véanse González, Ortiz y Sánchez Ávalos (2020).

²³ Esta numeración corresponde a la herramienta del perfil de modelo del manual de González, Ortiz y Sánchez Ávalos (2020).

¿Qué mecanismos o procedimientos se utilizaron para recopilar los datos (por ejemplo encuesta en vivienda, sensor, software, API)?	Discutir con el equipo legal las restricciones de uso de los datos, dada la forma en que se capturaron. Por ejemplo: ¿Es necesario obtener un consentimiento informado? ¿Se trata de información pública o privada?
Número de individuos para los cuales se han recolectado datos	Dependiendo del número, discutir junto con el equipo técnico si esta muestra puede ser representativa. Si no lo es, ¿cómo afectará esto el entrenamiento de la herramienta?
Frecuencia de captura (semanal, mensual, diaria) o número de observaciones promedio por individuo. ¿Se actualizará el conjunto de datos (por ejemplo, añadiendo nuevas instancias y/o eliminando otras)?	
Controles esenciales	
Obtener documentación para cada variable dentro del conjunto de datos. Breve descripción que incluya su nombre y tipo, lo que representa, cómo se mide su valor, etc.	
Hacer un análisis exploratorio de los datos. Calcular estadísticas descriptivas identificando el porcentaje de valores faltantes y determinar la distribución de cada variable dentro de la base de datos.	Analizar los resultados del análisis exploratorio de los datos. Asegurar que la madurez de la base de datos sea la adecuada para solucionar el problema de política pública.
Analizar la cobertura espacial y temporal de los datos.	
Analizar la cobertura de los grupos protegidos (sexo, raza, edad, etc.)	Comunicar al equipo técnico qué grupos serán considerados protegidos.
Describir las dimensiones importantes en las cuales la muestra de datos puede ser diferente a la población, en particular los sesgos de selección no medidos. Utilizar literatura relacionada con el tema e información de expertos.	Ayudar a determinar las razones que expliquen las diferencias encontradas a partir del conocimiento experto. Compartir estos insumos con el equipo técnico.
Identificar 'estados indeseables' en los datos como pueden ser los sesgos e inequidades perjudiciales para subgrupos o cualquier otro patrón que se considere subóptimo o no deseable desde el punto de vista de la política social.	Explicar, con base en conocimiento experto, aquellos sesgos de los datos que puedan surgir de un estado indeseado de recolección. Definir lo que se considera un estado subóptimo. Discutir con el equipo técnico las posibles medidas de mitigación de sesgos en los grupos protegidos.
¿Faltan valores? En caso afirmativo, explique por qué no se tiene esa información (Esto incluye información eliminada intencionalmente). Determine si la falta de datos está asociada con la variable a predecir.	Ayudar a determinar las razones que expliquen los valores faltantes (por ejemplo, una circunstancia adversa que no permitió la correcta recopilación de datos) a través del conocimiento experto.



Teoría 2.2 Desarrollo y validación del modelo

Rol del director del proyecto

Durante la fase correspondiente al desarrollo y validación del modelo, el director del proyecto estará a cargo de:

- evaluar el modelo con métricas que no solo consideren la eficiencia y la eficacia, sino también la equidad;
- decidir el modelo a implementar teniendo en cuenta los recursos humanos y financieros disponibles en la institución; y
- comprender las limitaciones del modelo y ajustar el plan de ejecución del proyecto según los resultados de esta etapa.

En esta sección se explicarán algunos conceptos básicos propios de la fase de entrenamiento de un modelo, una tarea que corresponde al equipo técnico, si bien es el director del proyecto quien en últimas decidirá cuál será el modelo a implementar.²⁴

¿Cómo se desarrolla un modelo?

En el numeral 1.6 de este manual (Herramienta/Análisis), el director del proyecto ya debe haber informado al equipo técnico sobre el tipo de herramienta o análisis basados en IA que se requerirán para solucionar el problema. Si bien en esa fase ya se introdujeron algunas herramientas específicas de la IA, la técnica precisa de modelamiento dependerá de la naturaleza del asunto a resolver y de los datos. Aquí el equipo técnico deberá ensayar con distintas opciones que luego serán presentadas al director del proyecto y al equipo multidisciplinario.

En el caso de un modelo de aprendizaje supervisado, cuando se lo alimenta con información sobre las variables, los datos se deben dividir por lo menos en dos grandes secciones (idealmente tres):

- **Datos de entrenamiento:** Son los que se usan para entrenar un modelo
- **Datos de validación:** Son los que se usan para validar el modelo según el entrenamiento previo.
- **Datos de prueba:** Son los que se mantienen ocultos hasta después de seleccionar el modelo y se usan para confirmar los resultados.

El equipo técnico estará a cargo de dividir los datos según la información disponible, pero corresponderá al director del proyecto decidir qué datos (no todos) se utilizarán en la fase de entrenamiento. Más aún, el director deberá informar al equipo técnico sobre posibles situaciones anómalas en los datos como producto de ciertas contingencias (situaciones extremas en el ámbito nacional, cambio en la definición de una variable, etc.), para que este tome las medidas del caso cuando vaya a utilizar esos datos.

Todos los modelos serán entrenados con los datos designados para ello, que luego se perfeccionarán utilizando los de validación. Por último, su funcionamiento se evaluará con los datos de prueba. Es importante que el director del proyecto tenga presente que los modelos son simplificaciones de la realidad. Por lo tanto no serán 100% precisos y tampoco se espera que lo sean. Se trata de que

²⁴ Para más información sobre modelamiento y validación, véanse González, Ortiz y Sánchez Ávalos (2020).

el modelo de IA tenga un nivel de ajuste suficiente y que pueda, por ejemplo, predecir resultados futuros, por lo cual un cierto porcentaje de error en el proceso de validación y prueba será inevitable. A continuación se describe la matriz de confusión --herramienta que ayuda a validar el desempeño de un modelo de IA durante su desarrollo--, así como el rol del director del proyecto y del equipo técnico en su implementación.

- **Matriz de confusión:** Cuando las variables objetivo son categóricas (sí/no), se pueden obtener métricas de ajuste en cuanto a los falsos positivos y falsos negativos, los cuales se obtienen según el valor predicho versus el dato real. Por ejemplo, para una clasificación binaria se tiene la siguiente matriz de confusión:

		Real	
		Positivo	Negativo
Predicción	Positivo	Verdadero positivo	Falso positivo
	Negativo	Falso negativo	Verdadero negativo

A partir de estos errores y aciertos en la predicción es posible elaborar distintas métricas de ajuste de un modelo,²⁵ tarea que corresponderá al equipo técnico. Por su parte, el director del proyecto deberá estar al tanto de los niveles de errores de clasificación en el modelo a implementar, de acuerdo con la definición del problema. Los falsos positivos ocurren cuando el modelo establece, por ejemplo, que una persona es beneficiaria de un programa pero en realidad no debería serlo. Por su parte, los falsos negativos se presentan cuando, según el modelo, una persona no debe recibir beneficios del programa cuando en realidad sí le corresponden. Más allá de las métricas que elabore el equipo técnico dependiendo de la naturaleza del problema, el director del proyecto deberá decidir qué tipo de error soporta más el proyecto. Por ejemplo, en el caso anterior de entrega de beneficios, y teniendo en cuenta que los recursos son escasos, el hecho de que una persona no reciba un beneficio que sí le corresponde (falso negativo) puede ser un error más grave que entregar beneficios a quien no los necesita (sobre todo si se puede hacer un análisis más profundo de los umbrales para ser beneficiario y se descubre que la persona en realidad se encuentra muy cerca del umbral).

Sin embargo, el que los errores producidos por el modelo sean aceptables también dependerá de los recursos de la institución. Así pues, el director del proyecto no solo deberá estar al tanto de su ocurrencia, sino que también deberá realizar un análisis costo-beneficio de ello. Siguiendo con el ejemplo anterior, si el presupuesto institucional está muy ajustado se podrían fijar límites a la entrega de beneficios cuando no corresponde (falso positivo), lo que tendrá que ser balanceado con no entregarlos cuando sí corresponde (falso negativo). Realizar este análisis costo-beneficio es clave antes de implementar el proyecto.

Nótese que estos errores son todos elaborados “en el laboratorio”; por ello es necesario tener presente que durante la implementación pueden surgir otros retos. Es posible que la herramienta sea de muy difícil ejecución o de poca aceptación entre los usuarios. Si así fuera, se podría volver al numeral 1.6 para analizar nuevamente la necesidad del proyecto, y también al inicio del numeral 2.2 para evaluar otra herramienta basada en IA cuya implementación tenga mayores probabilidades de éxito.

25 Véanse González, Ortiz y Sánchez Ávalos (2020, p. 16.).

Equidad y desempeño diferencial de predictores

Teniendo en cuenta lo anterior, el director del proyecto deberá estudiar en detalle los sesgos y desigualdades que puedan surgir en el modelo para tomar la decisión, o bien de corregirlos o de implementar políticas y medidas de mitigación.

Como se discutió en la sección anterior, los sesgos pueden aparecer en los datos y desde allí se los puede mitigar. Sin embargo, con las métricas de errores puede suceder que un modelo favorezca a un subgrupo de la población. Por ejemplo, supóngase que el modelo de detección de retinopatía diabética tiene mayor precisión para los hombres que para las mujeres, sesgo este que no necesariamente proviene de los datos sino del modelo mismo. El director del proyecto deberá tener claro este sesgo, ponerlo en conocimiento del público y mitigarlo con medidas de política pública. Una posible solución sería que, mientras se recalibra el modelo, todos los exámenes de fondo de ojo correspondientes a mujeres sean enviados a un oftalmólogo para un análisis exhaustivo. Esto con el objeto de garantizar una mayor precisión en los resultados de tales pruebas en las mujeres e ir alimentando el modelo con estos análisis.

También puede suceder que el objetivo de un proyecto sea efectivamente favorecer a un subgrupo determinado de la población. Tómese, por ejemplo, el caso de un subsidio que históricamente se ha asignado a los hombres, pero que ahora se busca que beneficie por lo menos en un 50% a las mujeres. Aquí el objetivo de la herramienta de predicción (sensibilidad, precisión y exactitud, entre otros) no se corresponde con el objetivo del proyecto, a saber, alcanzar a un 50% de mujeres beneficiarias. Debido a ello, se requiere que el director del proyecto informe efectivamente al equipo técnico sobre el objetivo de este —que puede ser distinto al del modelamiento propiamente tal— e identifique con claridad a aquellos subgrupos de la población que se busca beneficiar.

En el numeral 1.7 de este manual se señaló la importancia de que el director del proyecto esté consciente de las implicaciones éticas y legales de este último, pero también de las del modelo de AI propiamente tal. Ahora, con información más detallada acerca del modelo a implementar, el director debe volver a esa sección y constatar que no hayan surgido otros riesgos éticos, sesgos y discriminaciones como resultado del entrenamiento. En el numeral 2.1 sobre recolección y procesamiento de datos ya se discutieron los sesgos que se pueden presentar pero, como se ha indicado aquí, el modelo podría dar lugar a otros sesgos sin que los datos necesariamente se encuentren sesgados.



Será el equipo técnico el que informe al equipo multidisciplinario qué modelo se implementará, aunque corresponde al director del proyecto tomar la decisión final teniendo en cuenta la naturaleza del problema, los recursos financieros y humanos disponibles en la institución, y los riesgos y sesgos que pueda ocasionar la implementación de la herramienta.

Actividad

Concluida esta fase se recomienda revisar la sección correspondiente al **Desarrollo de Modelo de la herramienta Perfil del Modelo**, junto con el equipo técnico y de acuerdo con las funciones identificadas.

A continuación se presenta la sección sobre el Desarrollo del Modelo, destacando el rol del director del proyecto en la elaboración de esta herramienta, que deberá ser completada por el equipo técnico. Este perfil se detalla en González, Ortiz y Sánchez Ávalos (2020).

PERFIL DEL MODELO	
Desarrollo del modelo	
Modelamiento	Algoritmos usados en el entrenamiento, parámetros supuestos o restricciones
Métricas de desempeño	Métricas técnicas usadas para seleccionar y evaluar modelos
	Análisis costo-beneficio del modelo para su caso de uso
	Definición de grupos protegidos y medidas de equidad seleccionadas
Datos de validación	Conjuntos de datos usados y su etiquetado
	Pasos de preprocesamiento
	Evaluación de adaptación de datos de validación según el caso de uso
	Sesgos y deficiencias potenciales según el caso de uso
Resumen de análisis cuantitativo	Error de validación reportado
	Resumen de análisis costo-beneficio
	Reporte de medidas de equidad para grupos protegidos

Rol del director del proyecto

Realizar el análisis de costo-beneficio de la implementación de la herramienta, preguntándose si el problema de política pública se puede solucionar sin utilizar IA.

Definir medidas de equidad en los resultados (subgrupos de la población a los cuales se quiere garantizar equidad).

Establecer, junto con el equipo técnico, umbrales aceptables de sesgos con base en conocimiento experto. Elaborar medidas de mitigación basadas en la política pública, en caso de que se requieran.

Decidir, junto con el equipo técnico, qué modelo será implementado.



Teoría 2.3 Uso y monitoreo

Rol del director del proyecto

Durante la fase correspondiente al uso y monitoreo del modelo, el director del proyecto estará a cargo de:

- garantizar la capacitación de los servidores públicos que interactúan con el modelo de IA para que la herramienta sea sostenible en el tiempo;
- elaborar un manual de usuario enfocado en los funcionarios que van a interactuar con el modelo;
- establecer mecanismos de retroalimentación para las personas que interactúan con el modelo;
- diseñar e implementar procesos administrativos simples para corregir aquellos errores presentes en el modelo que afecten a los usuarios;
- establecer sistemas automatizados, o al menos periódicos, de monitoreo del modelo;
- mantener un registro de los resultados del modelo teniendo en cuenta las restricciones de acceso y seguridad;
- implementar mejoras necesarias al modelo y al proceso a partir de los hallazgos del monitoreo y la evaluación, y
- destinar los recursos necesarios para mantener la herramienta en el tiempo.

Una vez elegido el modelo más adecuado en la etapa previa, llega el momento de ponerlo en práctica. En esta fase es importante que el director del proyecto diseñe un piloto de la herramienta que permita monitorear su uso bajo condiciones acotadas, antes de desplegarla entre toda la población o en las entidades involucradas.

Uno de los aspectos críticos a considerar es la capacitación de los servidores públicos que van a interactuar con el modelo. Ello debe contemplar información sobre sus aspectos claves, su objetivo, la manera como funciona, y la forma en que operarán sus sistemas de retroalimentación. Si se incluyó a las personas adecuadas en la etapa de diseño y conceptualización, los equipos relacionados ya tendrán información acerca del proyecto, de modo que su participación y posterior capacitación ayudarán a mitigar el riesgo de que finalmente la herramienta nunca se utilice. Es posible que se requiera crear un equipo especial de capacitación dentro de la institución. También se sugiere elaborar un manual de usuario interno dirigido a los funcionarios que van a interactuar con el modelo. Dicho manual debe ser una versión resumida de las herramientas aquí presentadas. Se sugiere que, como mínimo, contenga:

- la definición del problema de política pública;
- la definición de los objetivos;
- la definición y justificación de la población protegida;
- el reglamento sobre el uso y almacenamiento de los datos (junto con las consideraciones legales pertinentes a cada país);
- una explicación simple sobre cómo funciona la herramienta y sobre las variables utilizadas en su entrenamiento;

- una sección de las preguntas frecuentes, formuladas en un lenguaje claro, que puedan formular las personas afectadas por la implementación de la herramienta;
- información de contacto para expresar dudas y solicitar información adicional.

Evaluación de efectividad

Esta etapa contempla el diseño de la evaluación del modelo, ya no en el laboratorio —explicado en el numeral 2.2 sobre su desarrollo y validación—, sino en el contexto donde se va a desplegar, con el fin medir su efectividad real durante la implementación del proyecto. El objetivo aquí es poder atribuir causalidad a la implementación de la herramienta de IA, para lo cual existen diversas técnicas de evaluación de impacto, entre ellas los experimentos naturales, los experimentos aleatorios controlados y los cuasi experimentos (Shadish, Cook y Campbell, 2002). De ser posible, se recomienda privilegiar métodos de experimentos aleatorios controlados, dado su nivel elevado de validez.

Para poder realizar un experimento aleatorio controlado es necesario contar con el consentimiento informado de los participantes y tomar las precauciones del caso frente a los riesgos éticos que puedan entrañar cuestiones como experimentar con seres humanos, privar a beneficiarios potenciales de una medida por ser parte de un grupo de control, y hacer asignaciones aleatorias versus por necesidad, además de otros problemas de índole legal (Shadish, Cook y Campbell, 2002). Si bien los experimentos aleatorios controlados constituyen uno de los métodos preferidos para medir el impacto, la herramienta que para ello se emplee en el proyecto dependerá de los datos disponibles, y de las restricciones éticas y legales de cada país. Será el director del proyecto quien en últimas decida y comunique al equipo técnico la forma en que se medirá la efectividad de la herramienta en el mundo real.

El éxito de la herramienta, y por lo tanto del proyecto, estará determinado por el logro de los objetivos. En tal sentido, el propósito de una evaluación de impacto **no será medir el nivel de ajuste del modelo a la realidad** sino el cumplimiento de aquellos. Una medición adecuada del cumplimiento de las metas establecidas exigirá contar con una línea de referencia clara sobre el desempeño del proceso actual (sin herramienta de IA), la cual se definirá en la etapa de análisis exploratorio de los datos. Si la herramienta no logra el impacto esperado en comparación con el estado previo, se debe revisar la implementación para constatar si tuvo fallas, o volver a la fase de desarrollo del modelo para producir uno nuevo, o bien reformular el proyecto desde la etapa de conceptualización y diseño.

Degradación del desempeño

El director del proyecto deberá estar consciente de que los resultados de un modelo se pueden degradar en el tiempo por diversas razones:

- puede cambiar la relación de las variables de entrada y salida;
- puede cambiar la forma en que se recolectan y almacenan los datos;
- es posible que los sistemas de retroalimentación usuarios-herramienta sean cerrados, lo cual quiere decir que las interacciones no dan lugar a la incorporación de datos nuevos originados en situaciones emergentes que enriquezcan el modelo.

Para mitigar estos riesgos será necesario que el director del proyecto monitoree el comportamiento de las variables de entrada y salida, y que, junto con el equipo técnico, vaya actualizando los supuestos del modelo. Se pueden establecer sistemas automáticos, o con una cierta periodicidad, para mo-

monitorear los resultados de la herramienta y el comportamiento de las métricas de error y de equidad previamente establecidas. La periodicidad dependerá de la naturaleza del problema.

Asimismo, **deberá existir un sistema de retroalimentación de experiencia de uso de la herramienta entre los usuarios, los beneficiarios finales, el equipo técnico y el director del proyecto**, quien será su responsable exclusivo. El objetivo de la implementación de un sistema de retroalimentación es contar con información oportuna acerca de las dificultades y errores que hayan detectado quienes interactúan con el sistema, con el fin de subsanarlos lo más rápido posible, bien volviendo a la etapa anterior o por medio de políticas públicas correctivas.



En esta etapa es clave que se destinen todos los recursos necesarios —humanos y financieros— a reentrenar el modelo y a capacitar a las personas de la institución que van a interactuar con él. Es importante que, al aplicar una solución basada en IA, esta sea sostenible en el tiempo y que se ajuste de acuerdo con los datos nuevos, las tecnologías emergentes y/o las nuevas interfaces de usuario.



Actividad

Al terminar esta fase proceda a revisar, junto con el equipo técnico, la sección **Uso y monitoreo de la herramienta en el Perfil del Modelo**, de acuerdo con las funciones identificadas.

A continuación se presenta la sección Uso y monitoreo del Perfil del Modelo, destacando el rol del director del proyecto en la elaboración de esta herramienta, la cual debe ser completada por el equipo técnico. Este perfil se detalla en González, Ortiz y Sánchez Ávalos (2020).

PERFIL DEL MODELO	
Uso y monitoreo	
Recomendaciones de monitoreo	Estrategia de monitoreo y mejora en producción
	Estrategias de monitoreo humano de predicciones (si aplica)

Rol del director del proyecto

Definir, junto con el equipo técnico, un programa piloto o ensayo aleatorio antes de desplegar el modelo para toda la población.

Hacer monitoreos periódicos de la herramienta junto con el equipo técnico y con los funcionarios que interactúan con ella.

Teoría

2.4 Rendición de cuentas

Rol del director del proyecto

En la fase de rendición de cuentas, el director del proyecto estará a cargo de:

- crear mecanismos permanentes y actualizados de comunicación sobre el funcionamiento del sistema de IA, para lo cual deberá usar un lenguaje claro;
- rendir cuentas periódicas acerca del monitoreo y los impactos del sistema; y
- establecer un sistema de respuesta apropiado para gestionar solicitudes individuales de resultados de la aplicación del sistema de IA.

Cabe insistir en que, entre los principios básicos de la administración pública figuran la transparencia y el derecho de acceso a la información por parte de la ciudadanía, los cuales también cubren los proyectos de inteligencia artificial a cargo del Estado.

Interpretabilidad y explicación de las predicciones

Por interpretabilidad o explicabilidad se entiende la posibilidad de que la ciudadanía en general pueda entender el funcionamiento de una herramienta de IA (Miller, 2019). Entre los beneficios que ello reporta figuran los siguientes:

- conocimiento acerca de la manera en que se va a solucionar el problema;
- obtención de licencia social;
- detección de sesgos en los algoritmos, y
- depuración y mejora del modelo.

Corresponde al director del proyecto crear mecanismos permanentes y periódicos de comunicación sobre el funcionamiento de la herramienta de IA (problemas, impactos y ajustes en el tiempo) para todos los interesados.

El uso de herramientas automáticas de toma de decisiones puede resultar extraño y confuso para la ciudadanía, lo cual puede generar resistencia a su implementación, sobre todo porque se suele pensar en los algoritmos como “cajas negras”. Cabe reiterar aquí que la explicabilidad y transparencia algorítmica de un modelo, así como la información que se pueda proporcionar sobre su funcionamiento, dependerán de su nivel de opacidad (Burrell, 2016 y Buenadicha et al., 2020), tal como se discutió en la sección 1.7 sobre consideraciones éticas, legales y de gobernanza.

Sesgos potenciales y depuración

Corresponderá al director del proyecto proporcionar a las partes interesadas la información sobre las limitaciones y sesgos de la herramienta.

Ya se ha indicado cómo el modelo y/o los datos pueden estar sesgados, lo cual requiere que el director del proyecto implemente medidas que garanticen la equidad, en caso de que no se haya logrado subsanar los sesgos en etapas previas a la ejecución.

Asimismo, para obtener y mantener la licencia social, es importante que el director del proyecto preste especial atención a los riesgos y consideraciones éticas que surjan de la implementación del modelo, que los dé a conocer y que indique claramente la manera en que se propone mitigarlos.

Explicación de predicciones individuales

Aquí el director del proyecto deberá elaborar protocolos de respuesta para aquellos casos particulares en que las personas soliciten información adicional sobre el funcionamiento y aplicación de un sistema de IA. Por ejemplo, si una persona no resulta beneficiada con un bono del gobierno debe poder solicitar información acerca del porqué de esa decisión automática. Para ello se requerirá que existan personas a cargo de revisar casos particulares y de dar respuesta satisfactoria a las inquietudes de los afectados.

Por su parte, el equipo técnico deberá evaluar la mejor forma de explicar estas predicciones individuales.

Trazabilidad

Se considera de suma importancia que el director del proyecto supervise la documentación detallada de todas las decisiones tomadas a lo largo de las distintas etapas de ejecución de la herramienta, y que tal información esté a disposición de todos los interesados, independientemente del nivel de opacidad de aquella.

El objetivo de la trazabilidad de las decisiones es mitigar los riesgos aquí identificados.

Actividad

Al terminar esta fase proceda a revisar, junto con el equipo técnico, la sección **Uso y monitoreo de la herramienta en el Perfil del Modelo**, de acuerdo con las funciones identificadas.

Al terminar esta fase, proceda a llenar, junto con el equipo técnico, la sección de rendición de cuentas del Perfil de Modelo de acuerdo con las funciones identificadas.

A continuación se presenta la sección de rendición de cuentas del Perfil del Modelo, destacando el rol del director del proyecto en la elaboración de esta herramienta, la cual debe ser completada por el equipo técnico. Este perfil se detalla en González, Ortiz y Sánchez Ávalos (2020).

PERFIL DEL MODELO	
Rendición de cuentas	
(Opcional) Explicabilidad de la predicciones	Estrategia para explicar predicciones particulares (si es necesario)
	Estrategia para entender la importancia de distintos atributos
Otras consideraciones éticas, recomendaciones y advertencias	

Rol del director del proyecto

Diseñar una política de respuesta o de revisión para atender posibles solicitudes de la ciudadanía sobre el resultado de la herramienta.

Discutir con el equipo legal los riesgos que pueden surgir en la implementación del modelo (protección de datos personales, sesgos u otros) y elaborar medidas de mitigación y de explicabilidad, en caso de que se requieran.

¿QUÉ PASÓ CON DART EN LA ETAPA DE EJECUCIÓN?

¿Cuál es el problema que se busca solucionar?

La diabetes es una enfermedad crónica que se presenta cuando el páncreas no puede producir suficiente insulina o cuando el organismo no utiliza toda la que produce. Existen dos tipos de diabetes: la diabetes tipo 1 y la diabetes tipo 2. La primera surge como resultado de un ataque del propio sistema inmunitario, no se puede prevenir y tampoco tiene cura, mientras que la segunda resulta de estilos de vida que dan lugar al sedentarismo y a la obesidad. Aquí el páncreas sigue produciendo insulina, aunque en cantidades insuficientes, y en algunos casos se puede curar. Ambos tipos de diabetes pueden producir ceguera, insuficiencia renal, accidentes cerebrovasculares (ACV) e incluso conducir a la amputación de las piernas.

Recolección y procesamiento de datos

Inicialmente, los datos utilizados para el modelo incluían 550 imágenes de fondo de ojo del Centro de Referencia de Salud (CRS) Cordillera Oriente de la comuna de Peñalolén, en Santiago de Chile.

El equipo técnico debió preprocesar las imágenes obtenidas, comprimiéndolas para estandarizar su tamaño a la luz de trabajos internacionales previos en el mismo campo.

En el momento de desarrollar el modelo se encontró que, en aquellas imágenes donde el contraste era insuficiente, la predicción de la herramienta resultaba afectada, dado que era más difícil realizar la segmentación de vasos sanguíneos en la retina, primer paso clave en la detección de la retinopatía diabética (Arenas, 2012).

El modelo se siguió reentrenando a medida que se tuvo acceso a un mayor número de exámenes de fondo de ojo y que se incorporaron las últimas técnicas en materia de IA. En un estudio observacional publicado en 2021 se utilizaron 1123 imágenes de examen de fondo de ojo (Arenas et al., 2021).

Desarrollo y validación del modelo

En la etapa inicial de desarrollo, los datos se dividieron en dos partes iguales (275 imágenes cada una) para crear los conjuntos de entrenamiento y prueba.

Se utilizó conocimiento experto de oftalmólogos de la Universidad de Chile para etiquetar las lesiones oculares presentes en las imágenes. La detección de la retinopatía diabética tiene cuatro etapas: segmentación de vasos sanguíneos en la retina, localización del disco óptico, detección de lesiones brillantes, y detección de lesiones rojas.

Al analizar los costos del modelo, se estableció que la inversión inicial implicaba un costo extra para el Ministerio de Salud debido a la adquisición de la licencia para usar el modelo de IA desarrollado por DART. Sin embargo, el beneficio comenzó a materializarse cuando se pasó de una acción correctiva a una preventiva, con el efecto que esto tiene en la salud de las personas que sufren de diabetes. En un horizonte de 10 años, se estima que los recursos monetarios destinados al tratamiento de retinopatía diabética se reducirían en aproximadamente 50% (Shokiche, 2013).

Uso y monitoreo

A través del Laboratorio de Gobierno de Chile²⁶ se logró una colaboración con la municipalidad de Recoleta, el Instituto de Salud Pública y el Ministerio de Salud para llevar a cabo un piloto en esa localidad durante 2016.

Para utilizar la herramienta de detección se creó una interfaz web que permite a los oftalmólogos interactuar con sus resultados. En esta plataforma se ingresan los datos personales del paciente, junto con el examen de fondo de ojo. Luego se envían los exámenes para ser analizados. El sistema automático estudia las imágenes y las devuelve, junto con los resultados, a la plataforma. Finalmente, los oftalmólogos ingresan a la plataforma y obtienen información de casos probables y descartados, de tal forma que puedan enfocar sus esfuerzos y tiempo en los casos probables para dar un diagnóstico definitivo (Arenas et al., 2015).

El uso de la herramienta exigió capacitar al personal de los centros de salud, a los tecnólogos médicos que toman examen de fondo de ojo y que deben subir las imágenes a la herramienta en la nube, y a los oftalmólogos de los centros de salud que reciben los resultados del tamizaje. Desde 2018 se han realizado videoconferencias sincrónicas para grupos de usuarios en todo Chile. Actualmente se está trabajando en tutoriales interactivos.

El Ministerio de Salud de Chile adoptó la tecnología como un estándar en sus centros de atención; actualmente DART se utiliza en más de 130 de ellos.

Existe una relación constante entre los centros de salud y las autoridades del Ministerio de Salud, que a 2021 estaba llevando a cabo encuestas para establecer el nivel de la satisfacción de los beneficiarios.

Con respecto al funcionamiento de la herramienta propiamente tal, se realizó un estudio observacional para evaluar las métricas de ajuste. Esto se hizo con base en una muestra de 1123 imágenes de cinco centros de atención primaria en Chile. Se detectaron 13 casos de falsos negativos, mientras que el área bajo la curva ROC (AUC) fue de 0,915. Las medidas de rendimiento superaron lo requerido por el Ministerio de Salud de Chile (Arenas et al., 2021).

Rendición de cuentas

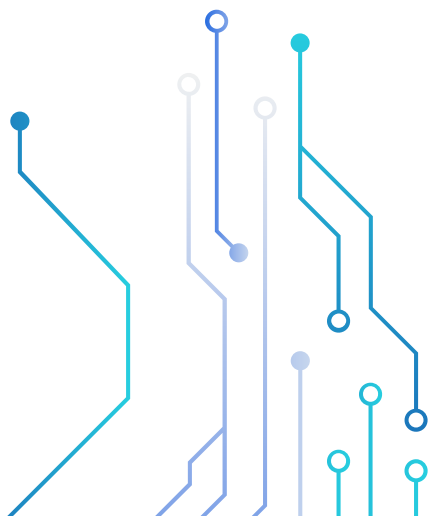
DART se inició como el proyecto de tesis de José Tomás Arenas, actual CEO de Teledx, por lo cual toda la información acerca de los datos utilizados en la primera etapa (antes de su puesta en marcha) y modelamiento inicial se encuentran en el repositorio de la Universidad de Chile.

Después de su implementación en el sistema de salud chileno, se han publicado múltiples documentos en revistas científicas internacionales y ha participado en charlas internacionales, donde se han compartido porciones significativas de información en materia de datos, procesos e incluso del funcionamiento del algoritmo. Si bien el modelo que subyace a la herramienta es sumamente complejo, se ha evidenciado el valor que tiene su aplicación en los distintos centros de salud al reducir el tiempo de espera para obtener el resultado de exámenes y los costos

²⁶ Laboratorio de Gobierno es la agencia del Estado chileno que desde 2015 se dedica a cocrear soluciones para problemas públicos prioritarios y a formar capacidades para innovar en las instituciones públicas, con el fin de mejorar los servicios que se prestan a la ciudadanía.

monetarios asociados con el tratamiento de la enfermedad. Sin embargo, se deben estudiar en detalle las posibles discriminaciones del modelo, las cuales se podría originar, por ejemplo, en las condiciones en que se practica el examen de fondo de ojo, dado que pueden afectar su lectura. Lo anterior puede ser más prevalente en algunas localidades que en otras.

Dado que el sistema actualmente se implementa bajo los auspicios del Ministerio de Salud de Chile, su funcionamiento se encuentra sujeto a las leyes de transparencia y de protección de datos personales vigentes en ese país.



CONCLUSIONES

En este manual se describen las fases necesarias para la correcta implementación de un proyecto de IA desde la perspectiva de la persona a cargo de la toma de decisiones de política pública y de dirigir el proyecto, considerando los retos que surgen en el ciclo de vida de la IA.

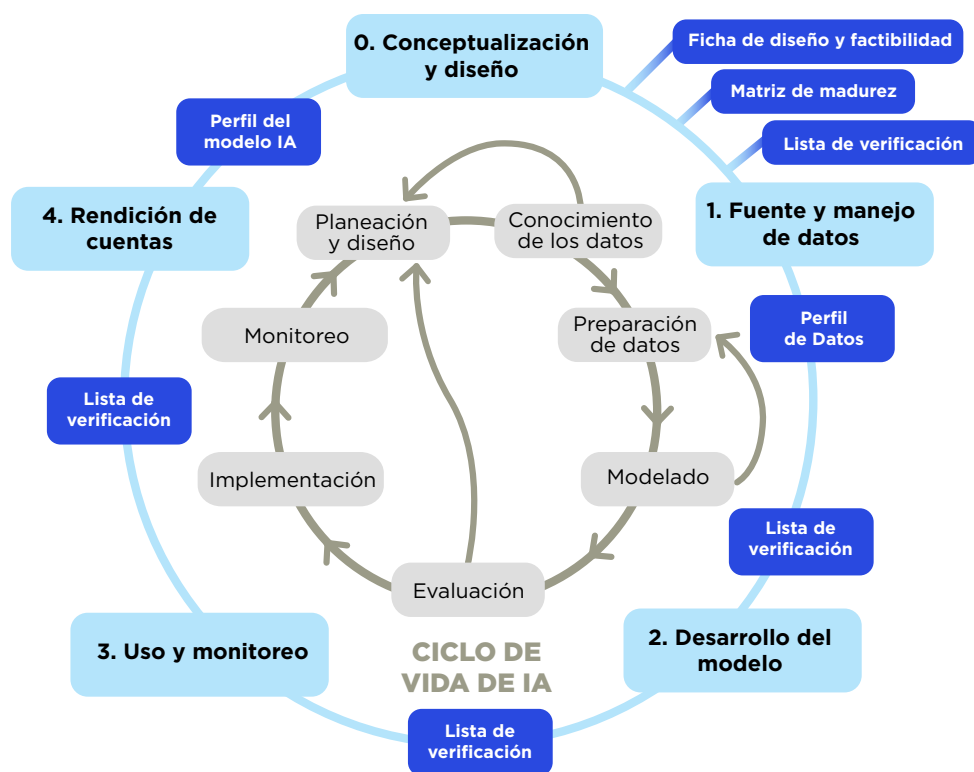
Un proyecto de toma y/o soporte de decisiones basado en IA se encuentra estrechamente vinculado con el ciclo de la política pública, en la medida en que debe solucionar problemas reales y significativos que afectan a la sociedad.

A menudo los proyectos suelen comenzar por un interés en querer utilizar los datos disponibles en la institución o una nueva tecnología. Sin embargo, no sobra insistir en que, para que un proyecto tenga valor social, debe surgir de la necesidad de resolver un problema real de política pública.

Si bien la herramienta de IA será desarrollada por el equipo técnico, corresponderá al director del proyecto tomar las decisiones y liderar su conceptualización y diseño, la comunicación con el personal técnico, el análisis de viabilidad de la herramienta y su implementación.

Cabe reiterar aquí la importancia de que la creación de un sistema de soporte y/o toma de decisiones sea producto de un proceso iterativo. Es común que los equipos deban volver sobre las etapas aquí mencionadas, incluso redefiniendo la conceptualización y diseño del proyecto a medida que van surgiendo nuevos desafíos. Esa etapa es crítica porque obliga al director del proyecto a hacerse preguntas claves que maximicen las probabilidades de éxito de este último. Una vez definida su viabilidad, comienza la etapa de ejecución con el procesamiento de datos. La figura 6 contiene las herramientas que ofrece este manual para el director del proyecto, según el ciclo de vida de la IA.

Figura 6. Ciclo de vida de la IA y herramientas del manual



Desde el primer momento, el director del proyecto deberá tener claridad sobre los riesgos potenciales de la herramienta de IA en materia de licencia social, implementación y temas éticos. También deberá asegurarse de que posee los recursos humanos y/o financieros para mantenerla y actualizarla de manera periódica. La herramienta estará en constante ajuste, a medida que cambie el contexto o que se disponga de nuevos datos. Asimismo es necesario que los funcionarios comprendan sus beneficios, la adopten y la incorporen en sus quehaceres para los fines propuestos.

Este manual presenta el ciclo de vida de la IA dentro del ciclo de vida de la política pública, que no concluye al final del ciclo de aquella. Por lo tanto, es necesario evaluar la implementación de la política, no solo desde el punto de vista de la herramienta sino también de su efectividad en términos de solucionar el problema identificado en la sociedad.

Esperamos que este manual sea de utilidad para los responsables por la toma de decisiones que estén pensando en implementar un sistema basado en IA. Este manual debe ser usado en conjunto con el manual técnico de IA responsable (González, Ortiz y Sánchez Ávalos, 2020), donde se especifican las tareas y retos del equipo técnico.

REFERENCIAS

- Arenas Cavalli, J. T. A. (2012). Diseño y desarrollo de un sistema para la detección automática de retinopatía diabética en imágenes digitales. Disponible en <http://repositorio.uchile.cl/handle/2250/104406>
- Arenas-Cavalli, J. T., Abarca, I., Rojas-Contreras, M., Bernuy, F. y Donoso, R. (2021). Clinical validation of an artificial intelligence-based diabetic retinopathy screening tool for a national health system. *Eye*, 1-8.
- Arenas-Cavalli, J. T., Ríos, S. A., Pola, M. y Donoso, R. (2015). A web-based platform for automated diabetic retinopathy screening. *Procedia Computer Science*, 60, 557-563.
- Bojalil, P. Vela-Treviño, C. (2019). Despuntan las reformas en materia de protección de datos en América Latina. Conocimiento abierto. Disponible en <https://blogs.iadb.org/conocimiento-abierto/es/proteccion-de-datos-gdpr-america-latina/>. Consultado el 29 de marzo de 2021.
- Buenadicha, C., Galdon, G., Hermosilla, M. P., Loewe, D. y Pombo, C. (2019). La gestión ética de los datos. Por qué importa y cómo hacer un uso justo de los datos en un mundo digital. Disponible en https://publications.iadb.org/publications/spanish/document/La_Gesti%C3%B3n_%C3%89tica_de_los_Datos.pdf.
- Burrell, J. (2016). How the machine ‘thinks’: Understanding opacity in machine learning algorithms. *Big Data & Society*, 3(1), 2053951715622512.
- Cabrol, M., González, N., Pombo, C. y Sánchez, R. (2020). Adopción ética y responsable de la Inteligencia Artificial en América Latina y el Caribe. Disponible en https://publications.iadb.org/publications/spanish/document/fAlr_LAC_Adopci%C3%B3n_%C3%A9tica_y_responsable_de_la_inteligencia_artificial_en_Am%C3%A9rica_Latina_y_el_Caribe_es.pdf
- Dastin, J. (2018). Amazon Scraps Secret AI Recruiting Tool that Showed Bias Against Women. Disponible en <https://www.reuters.com/article/us-amazon-com-jobs-automationinsight/amazon-scraps-secret-ai-recruiting-tool-that-showed-bias-against-womenidUSKCN1MK08G> Consultado el 19 de enero de 2021.
- Data Futures Partnership (2017). A Path to Social Licence: Guidelines for Trusted Data Use.
- Gebbru, T., Morgenstern, J., Vecchione, B., Wortman, J., Wallach, H., Daumé, H., & Crawford, K. (2018). Datasheets for Datasets. Obtenido de <https://arxiv.org/pdf/1803.09010.pdf>
- Giorgi, S. (2017). How to improve the evaluation of complex systems to better inform policymaking Learning from evaluating Defra’s Reward & Recognition Fund. Fellowship Report, Julio. Disponible en <https://www.cecan.ac.uk/sites/default/files/2018-01/Guidance%20Report%20%20RRF%20Fellows-hip%20Final.pdf>
- Gómez, C., del Pozo, C., Martínez, C. y Martín del Campo, A. (2020). La inteligencia artificial al servicio del bien social en América Latina y el Caribe: Panorámica regional e instantáneas de doce países. Disponible en <https://publications.iadb.org/publications/spanish/document/La-inteligencia-artificial-al-servicio-del-bien-social-en-America-Latina-y-el-Caribe-Panor%C3%A1mica-regional-e-instant%C3%A9neas-de-doce-paises.pdf>

González, F., Ortiz, T. y Sánchez Ávalos, R. (2020). IA Responsable: Manual técnico: Ciclo de vida de la inteligencia artificial. Disponible en <https://publications.iadb.org/publications/spanish/document/IA-Responsable-Manual-tecnico-Ciclo-de-vida-de-la-inteligencia-artificial.pdf>

Greene, W. H. (2003). Econometric Analysis. India: Pearson Education.

Harini Suresh , J. V. (2019). A Framework for Understanding Unintended Consequences of Machine Learning. MIT. Obtenido de <https://arxiv.org/pdf/1901.10002.pdf>

Heaven, W.D. (2020). Google's medical AI was super accurate in a lab. Real life was a different story. Disponible en <https://www.technologyreview.com/2020/04/27/1000658/google-medical-ai-accurate-lab-real-life-clinic-covid-diabetes-retina-disease/> Consultado el 29 de enero de 2021.

Hewitt, T. (2019). Canada is a leader in AI research – but that's not enough. Disponible en <https://www.theglobeandmail.com/opinion/article-canada-is-a-leader-in-ai-and-we-must-lead-in-research-on-how-we-want/> Consultado el 19 de enero de 2021.

Hojman Cano, I. (2014). El mercado de las especialidades médicas de anestesiología y oftalmología en Chile. Disponible en <http://repositorio.uchile.cl/handle/2250/130678>

Miller, T. (2019). Explanation in Artificial Intelligence: Insights from the Social Sciences. Artif. Intell., 267, págs. 1-38.

Mitchell, M., Wu, S., Zaldivar, A., Barnes, P., Vasserman, L., Hutchinson, B., . . . Gebru, T. (2019). Model Cards for Model Reporting. Obtenido de <https://arxiv.org/abs/1810.03993>

Morrison, J. (2014). The Social License. How to keep your organization legitimate. Palgrave Macmillan, ISBN: 9781137370716.

OECD. (2019). Artificial Intelligence in Society. París: OECD Publishing.

OECD (Próximo a publicarse). OECD Framework for the Classification of AI Systems. París: OECD Publishing.

Oliver, D. (1994). Law, politics and public accountability. The search for a new equilibrium. Public Law, 238-238.

Red Iberoamericana de Protección de datos personales (2017). Estándares de protección de datos personales. Disponible en <https://www.redipd.org/es/documentos/estandares-iberoamericanos>

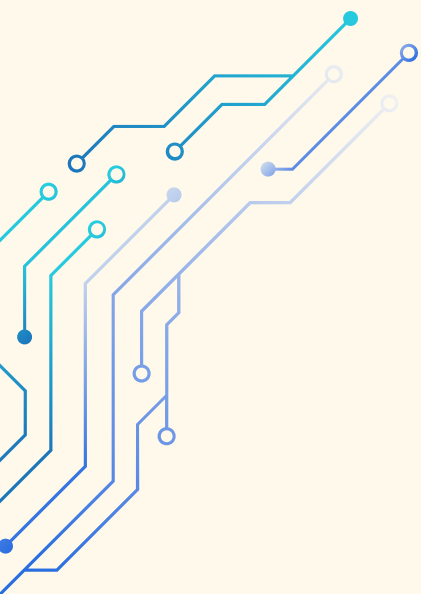
Sangüesa, R. (2018). Inteligencia artificial y transparencia algorítmica: "It's complicated". BiD: textos universitaris de biblioteconomia i documentació, No. 41, diciembre. <<http://bid.ub.edu/es/41/sanguesa.htm>>. Disponible en <http://bid.ub.edu/es/41/sanguesa.htm#:~:text=La%20transparencia%20de%20datos%20y,vital%20de%20quien%20reclama%20esta> Consultado el 29 de enero de 2021.

Sapunar, J. (2016). Epidemiología de la diabetes mellitus en Chile. Revista Médica Clínica Las Condes, 27(2), 146-151.

Shadish, W. R., Cook, T. D., y Campbell, D. T. (2002). Experimental and quasi-experimental designs for generalized causal inference. Boston: Houghton Mifflin. ISBN: 0395615569. Disponible en <https://www.alnap.org/system/files/content/resource/files/main/147.pdf>

Shokiche Vega, D. A. (2013). Estudio del problema de salud pública asociado a la patología oftalmológica retinopatía diabética en Chile y dimensionamiento del potencial impacto de detección basado en tecnología para abordar este problema.

Varian, Hal R. (2014). Big Data: New Tricks for Econometrics. *Journal of Economic Perspectives*, 28 (2): 3-28.



ANEXOS

- Anexo 1.** Ficha de diseño y factibilidad del proyecto
- Anexo 2.** Matriz de madurez de los datos
- Anexo 3.** Lista de verificación para el director del proyecto
- Anexo 4.** Perfil de los datos
- Anexo 5.** Perfil del modelo

Anexo 1. Ficha de diseño y factibilidad de proyecto

1 Nombre del proyecto:

2 Nombre de la organización:

3 Definición del problema:

¿Cuál es el problema que se busca solucionar?

Describa a la(s) población(es) afectadas por el problema (personas, grupos, entidades, etc.)

¿Cuántas personas/organizaciones/localidades/etc. se ven afectadas y en qué medida?

¿Por qué resolver este problema es una prioridad para su organización?

¿Conoce algún caso parecido de uso de IA que haya sido implementado anteriormente? ¿Cuál?

4 Análisis de prefactibilidad

¿Está dentro de las facultades de la entidad actuar sobre el problema?
¿Será necesario asociarse con otros organismos públicos? ¿Se cuenta con los recursos humanos y financieros suficientes para llevar a cabo el proyecto?

¿Existen los datos relevantes (suficientes para poder cambiar la forma en que hasta el momento se ha respondido al problema)? ¿Se puede acceder a ellos?

¿Cuáles son los riesgos del proyecto (éticos, licencia social, implementación, etc.)?

5 Definición de objetivos:

Los objetivos por lo general se expresan en términos de mejorar, maximizar, aumentar o disminuir, mitigar y/o reducir un resultado. El objetivo debe ser medible, lo cual exige establecer una métrica o indicador que refleje los avances. Alcanzar el objetivo debería ayudar a solucionar el problema. La solución técnica (p. ej., un modelo predictivo) no es el objetivo.

Las limitaciones típicas se relacionan con el presupuesto, la falta de capital humano, las restricciones legales, la voluntad política y la licencia social.

Tenga en cuenta que, cuando hay objetivos contrapuestos, puede tener que sacrificar algo para obtener un beneficio.

	Objetivo	Limitaciones
1		
2		
3		

6 Descripción de acciones:

Las acciones son aquellas actividades que las instituciones realizan o pueden realizar en relación con un determinado problema, como por ejemplo los programas específicos que llevan a cabo según su misión frente a la sociedad o sus procesos habituales de funcionamiento (contrataciones, atención a usuarios, pago de salarios, etc.).

Estas acciones se podrán mejorar cuando la institución cuente con la información generada por el proyecto de ciencia de datos. Asimismo deberán tener una conexión con los resultados generados por el sistema de IA y ayudar a alcanzar los objetivos planteados (sección anterior).

	Completar		
	Acción 1	Acción 2	Acción 3
Acción P. ej., taller de sexualidad responsable para estudiantes de 13 años/Entrega de métodos anticonceptivos en la enfermería escolar Nota: Se debe explicar cada acción en una casilla separada			
¿Quién ejecuta la acción? P. ej., equipo de psicología y psicopedagogía de cada establecimiento y/o equipo de enfermería escolar.			
¿Sobre quién o qué se está realizando la acción? P. ej., el alumnado de los establecimientos que actualmente cursan 7º. grado y/o el alumnado general.			
¿Con qué frecuencia se toma la decisión de realizar esta acción? P. ej., anualmente/mensualmente			
¿Qué canales se están usando o se pueden usar para realizar esta acción? P. ej., de manera presencial			
Otra información útil sobre la acción			

7 Mapeo de datos

Si se quiere que la institución alcance su objetivo, los datos tienen que conectarse con las acciones que van a respaldar. Los proyectos típicos de IA usan datos administrativos como fuente primaria y la mejoran con otras fuentes de datos del dominio público (censos, otros datos abiertos, etc.). La asociación con el sector privado o con organizaciones sin fines de lucro puede contribuir a obtener los datos que faltan a nivel interno.

¿Con qué datos se cuenta internamente?

	Completar		
	Fuente de datos 1	Fuente de datos 2	Fuente de datos 3
Nombre P.ej., sistema de alta hospitalaria			
¿Qué contiene? Describe los atributos con el mayor detalle posible (P. ej., registros de admisión y alta de hospitales a nivel nacional, con datos sociodemográficos del paciente, diagnóstico, días de internación, tipo de seguro de salud, información del doctor.			
¿Qué nivel de granularidad? P. ej., transacción, persona, organización, ubicación			
¿Con qué frecuencia se recopila y/o actualiza la información una vez que se captura? P. ej., en tiempo real, a diario, semanalmente, mensualmente, anualmente, de manera ocasional			
¿Tiene identificadores únicos y confiables que se pueden vincular con otras fuentes de datos? P. ej., RUN, SSN, DNI, según el país.			
¿Quién esta a cargo de los datos? P. ej. el departamento de registros del hospital			
¿Cómo se almacena? P. ej. en una base de datos, PDF, Excel, SPSS			
Comentarios adicionales			

¿Qué datos puede obtener de fuentes externas privadas o públicas?

	Completar		
	Fuente de datos 1	Fuente de datos 2	Fuente de datos 3
Nombre P.ej., Registro de la calidad del aire			
¿Qué contiene? Describa los atributos con el mayor detalle posible (P. ej. concentración de contaminantes --tales como material particulado de distintos tamaños-- en el aire.			
¿Qué nivel de granularidad? P. ej., estación de monitoreo geolocalizada por hora			
¿Con qué frecuencia se recopila/actualiza, una vez que se captura? P. ej., diaria			
¿Tiene identificadores únicos y confiables que se pueden vincular con otras fuentes de datos? P. ej. Código de la estación de monitoreo			
¿Quién es responsable de los datos? Ministerio del Medio Ambiente			
¿Se requieren acuerdos legales para el intercambio y/o acceso a la información?			
¿Cómo se almacena? P. ej. base de datos descargable mediante un API en un portal de datos abiertos.			
Comentarios adicionales			

En un mundo ideal, ¿existen datos adicionales que tengan relevancia para este problema y que le gustaría obtener? (Encuestas, circuito cerrado de televisión, registros telefónicos, ADN, variedad de frecuencia o granularidad para datos disponibles actualmente, etc.)

8 Análisis/Herramienta

Los proyectos típicos de IA incluyen una combinación de varios análisis, dependiendo de las necesidades y particularidades de cada proyecto. Los análisis son herramientas, no el objetivo del proyecto.

Elija el o los análisis adecuados para el problema correcto.

- Los análisis o herramientas elegidos deben mejorar las acciones actuales en respuesta al problema.
- Debe someter a prueba el o los análisis, y el proceso de validación debe coincidir con su objetivo.

	Completar		
	Análisis/ Herramienta 1	Análisis/ Herramienta 2	Análisis/ Herramienta 3
Tipo de análisis/ herramienta P.ej., descripción, predicción, detección, cambio de comportamiento			
Propósito del análisis P. ej., entender el comportamiento histórico de las personas; estimar el riesgo de enfermedad de un paciente; identificar las acciones que lograrían disminuir la sobrepesca.			
¿Para que tipo de acciones se utilizará la información generada a partir de este análisis? P. ej., inspección de barcos pesqueros industriales y artesanales			
¿Cómo validará este análisis usando datos existentes? P. ej., usando datos históricos, realizando un ensayo aleatorio controlado, etc.			

9 Consideraciones éticas y legales

Proporcionalidad	¿Cree que un sistema de ciencia de datos/IA es el medio adecuado para resolver el problema? ¿Por qué? ¿Ha evaluado otras alternativas?	
	¿Qué impactos negativos podría tener su proyecto? Repase casos de uso similares identificados en la sección "Definición del problema."	
Licencia social	¿Cree que los usuarios o afectados encontrarán aceptable el uso de datos planteado para resolver el problema? ¿Por qué?	
	Si la población objetivo del proyecto se entera, ¿lo aprobará? ¿Por qué?	
	¿Ha identificado la justificación o base legal para trabajar con esos datos?	
	¿Ha identificado las regulaciones que podrían incidir en el proyecto?	
	¿Será necesario contar con mecanismos para garantizar la calidad de los datos personales, como por ejemplo mecanismos de acceso, eliminación o rectificación?	
Transparencia	¿Qué actores interesados deberían estar al tanto del proyecto? Entre los actores interesados suelen figurar los responsables por la formulación de políticas, los trabajadores de primera línea, las organizaciones de la sociedad civil, los organismos públicos, las personas que se verán afectadas por las acciones, etc. Mencione organizaciones y/ o tipo de personas específicas.	
	¿Ha considerado algún mecanismo para que las partes interesadas se comuniquen con la institución en torno al proyecto?	
	¿Será necesario explicar los mecanismos de toma de decisiones o el análisis que se va a realizar? ¿Por qué?	

Discriminación/eqüidad	¿Qué desigualdades estructurales hay en el proceso y/ o en el entorno donde se inserta el proyecto?	
	¿Existen grupos específicos (vulnerables) para los que se quiere garantizar la equidad de los resultados o la protección de derechos? P. ej., grupos por género, edad, localización, clase social, nivel educativo, origen (urbano o rural), etnia	
	¿Qué sesgos cree que podrían tener los datos?	
Rendición de cuentas	En caso de que se solicite información sobre el proyecto ¿quién se encarga de proporcionarla?	
	¿Quién es responsable si el sistema se equivoca?	
	¿Se tienen previstos mecanismos de monitoreo, control y evaluación? ¿Cómo se documentarán y qué periodicidad tendrán?	
	¿Se tienen previstos mecanismos de formación para que el equipo a cargo comprenda las responsabilidades, así como las obligaciones legales y éticas que conlleva el proyecto?	

10 Conformación del equipo

Generalmente, los proyectos de inteligencia artificial exigen la vinculación de diversos profesionales de la misma entidad pública, aunque también de otras organizaciones relacionadas. Participan los responsables de los datos, de la infraestructura de TI y del problema o proceso en cuestión, así como expertos en analítica de datos, del área legal y de comunicaciones. Agregue los renglones que requiera en el siguiente cuadro.

Organización/ departamento	Descripción de la participación deseada	Nombre/rol de la contraparte
Departamento de TI	Proporcionar infraestructura de datos	Jefe del Departamento de TI
Organismo de estadísticas	Proporcionar datos de población	Jefe del Departamento de Estadísticas

Esta ficha de trabajo fue desarrollada originalmente por el Center for Data Science and Public Policy de la Universidad de Chicago. Para más información sobre nuestros programas y trabajo, por favor visite <http://datasciencepublicpolicy.org> o escribenos a info@datascienceforsocialgood.org

Esta versión de la hoja de trabajo ha sido actualizada a través de una colaboración entre el GobLab UAI, Carnegie Mellon University y el Instituto Tecnológico de Monterrey.

El GobLab UAI es el laboratorio de innovación de la Escuela de Gobierno de la Universidad Adolfo Ibáñez. Su misión es contribuir a la innovación en políticas públicas para beneficiar a la sociedad. Trabaja con organismos públicos, organizaciones de la sociedad civil e investigadores para lograr políticas públicas más eficaces, eficientes y equitativas mediante la ciencia de datos. Para obtener más información, visite <https://goblab.uai.cl> o envía un correo electrónico a goblab@uai.cl.



Attribution ShareAlike (CC BY-SA)

Anexo 2. Matriz de madurez de los datos

Matriz de madurez de datos²⁷

Categoría	Áreas	Deficiente	Básico	Intermedio	Avanzado
Cómo se almacena la información	Acceso	Solo se accede desde la aplicación que recolecta los datos.	Se puede acceder a los datos, pero con un software especial y específico.	Los datos están en formatos accesibles como CSV, JSON, XML o una base de datos. accesible de forma remota.	Los datos están en formatos accesibles y se puede acceder a ellos mediante una API.
	Almacenamiento	Papel	PDF o imágenes	Archivos de texto	Bases de datos
	Integración	Los datos solo están en el sistema donde se recolecta la información.	Los datos se exportan ocasionalmente y se integran de manera ad-hoc	Centro único de datos con actualización automática	Tanto los datos internos como los externos están integrados en la base de datos.
Qué información se recolecta	Relevancia	Los datos son irrelevantes para el problema. Por ejemplo, quieres ver la probabilidad de terminar la universidad pero no tienes los datos acerca de quién se gradúa.	Algunos de los datos son relevantes, pero son insuficientes porque faltan componentes claves. Por ejemplo, volviendo al caso anterior, solo tienes asistencia a clases.	Se tiene información útil y relevante, pero no completa. Por ejemplo, en el caso anterior hay datos demográficos y notas anuales, pero no hay información de apoyo extracurricular.	Existe toda la información relevante y es suficiente para resolver el problema sin hacer transformaciones relevantes.
	Calidad	Faltan filas (observaciones)	Faltan columnas (variables claves de ciertas observaciones)	Están todos los datos; solo hay errores menores, como por ejemplo de digitación.	No hay problemas de falta de datos o errores de digitación; las bases están limpias.
	Frecuencia	Solo una vez	Anual	Frecuente	Tiempo real
	Granularidad	Agregado a nivel de ciudad	A nivel de cuadra o código postal	A nivel individual (persona o dirección)	Detalle a nivel de evento/hitos
	Historia	No hay historial; los datos antiguos se borran.	Se guardan los datos históricos pero las actualizaciones se sobrescriben.	Se guarda la información histórica ocupando una marca temporal.	Todos los datos se mantienen y se relacionan con los datos anteriores con un modelo integrado.
Otros	Privacidad	No hay política de privacidad	La política no permite usar ningún dato	Se necesita aprobación ad-hoc para el uso de los datos.	Está definido el acceso a los datos y se controla la privacidad de los mismos para preservar la privacidad de los individuos.
	Documentación	No existe documentación digital o metadata. Los códigos de las variables no están documentados.	Se define un diccionario de datos que explica las variables y categorías.	Existe un diccionario de datos y disponibilidad de metadata.	Existe diccionario de datos, metadata, y adicionalmente los supuestos, sesgos y los datos que no se están obteniendo.

²⁷ Attribution ShareAlike (CC BY-SA). Adaptada del Data Maturity Framework de la U. de Chicago <http://dsapp.uchicago.edu/resources/datamaturity/>

Categoría	Áreas	Deficiente	Básico	Intermedio	Avanzado
Cómo se almacena la información	Acceso				
	Almacenamiento				
	Integración				
Qué información se recolecta	Relevancia				
	Calidad				
	Frecuencia				
	Granularidad				
	Historia				
Otros	Privacidad				
	Documentación				

Anexo 3. Lista de verificación para el director del proyecto

La siguiente lista de verificación contiene las principales tareas y funciones del director del proyecto (responsable por la toma de decisiones) a lo largo del ciclo de vida de la IA.

Conceptualización y diseño de las políticas públicas

- ☐ Definir claramente el problema de política pública que se quiere solucionar, identificando y cuantificando a los grupos de personas que se ven afectadas, y determinando su impacto en el presupuesto. La definición del problema debe ser fácilmente entendible para una persona externa a la institución.
- ☐ Contactar a las personas de su institución que están a cargo de abordar el problema para establecer la forma en que se hace actualmente. ¿Qué información le pueden ofrecer sobre cómo mejorar el sistema de respuesta?
- ☐ Investigar la manera en que otros organismos –nacionales o extranjeros— con un problema similar han implementado una solución basada en la IA. Lo ideal sería contactarlos para informarse acerca de los desafíos y dificultades que encontraron en el trayecto.
- ☐ Conversar con el personal directivo de su institución sobre la prioridad de solucionar el problema y lograr un compromiso con el proyecto al más alto nivel. Si usted mismo es un directivo de alto nivel, documente cómo esta prioridad está plasmada en los planes estratégicos de la institución.
- ☐ Confirmar que la lista de verificación del equipo técnico de esta fase se haya completado.

Ciclo de vida

Recolección y procesamiento de datos

- ☐ Contextualizar los hallazgos del equipo técnico en la etapa exploratoria de los datos.
- ☐ Proporcionar los antecedentes del proceso de recolección de los datos y el significado de las variables, conectando al equipo técnico con las personas de la organización que se requieran para una mejor comprensión de los datos.
- ☐ Validar la formulación y utilidad de la variable objetivo.
- ☐ Tomar decisiones sobre imputaciones de datos o eliminación de variables, ponderando y registrando lo que se obtiene versus lo que se pierde (trade-offs).
- ☐ Proponer mejoras a los procesos de gobernanza de los datos del proyecto que puedan contribuir al logro de los objetivos y faciliten la implementación del modelo.
- ☐ Hacer ajustes a la definición del problema, a los objetivos y al plan de ejecución del proyecto según los resultados de esta etapa.
- ☐ Confirmar que la lista de verificación del equipo técnico de esta fase se haya completado.

Desarrollo de modelos

- ☐ Evaluar el modelo con métricas que no solo consideren la eficiencia y la eficacia, sino también la equidad.
- ☐ Decidir el modelo a implementar teniendo en cuenta los recursos humanos y financieros disponibles en la institución.
- ☐ Comprender las limitaciones del modelo y ajustar el plan de ejecución del proyecto según los resultados de esta etapa.
- ☐ Confirmar que la lista de verificación del equipo técnico de esta fase se haya completado.

Uso y monitoreo

- ☐ Garantizar la capacitación de los servidores públicos que interactúan con el modelo de IA para que la herramienta sea sostenible en el tiempo.
- ☐ Elaborar un manual de usuario enfocado en los funcionarios que van a interactuar con el modelo.
- ☐ Establecer mecanismos de retroalimentación para las personas que interactúan con el modelo.
- ☐ Diseñar e implementar procesos administrativos simples para corregir aquellos errores presentes en el modelo que afecten a los usuarios.
- ☐ Establecer sistemas automatizados, o al menos periódicos, de monitoreo del modelo.
- ☐ Mantener un registro de los resultados del modelo teniendo en cuenta las restricciones de acceso y seguridad.
- ☐ Implementar mejoras necesarias al modelo y al proceso a partir de los hallazgos del monitoreo y la evaluación.
- ☐ Destinar los recursos necesarios para mantener la herramienta en el tiempo.
- ☐ Confirmar que la lista de verificación del equipo técnico de esta fase se haya completado.

Rendición de cuentas

- ☐ Crear mecanismos permanentes y actualizados de comunicación sobre el funcionamiento del sistema de IA, para lo cual deberá usar un lenguaje claro.
- ☐ Rendir cuentas periódicas acerca del monitoreo y los impactos del sistema.
- ☐ Establecer un sistema de respuesta apropiado para gestionar solicitudes individuales de resultados de la aplicación del sistema de IA.
- ☐ Confirmar que la lista de verificación del equipo técnico de esta fase se haya completado.

Anexo 4. Perfil de los datos

El perfil de los datos es un análisis exploratorio que proporciona información para evaluar la calidad, integridad, temporalidad, consistencia y posibles sesgos del conjunto de datos que será utilizado para entrenar un modelo de aprendizaje automático (Gebru et al., 2018). Para que el equipo técnico pueda generar un perfil de datos adecuado, el director del proyecto deberá garantizar el acceso a los conjuntos de datos (cuando son internos) o coordinar su obtención con organismos externos, junto con la documentación sobre cada variable dentro de ellos. Asimismo podrá colaborar en la identificación de dimensiones importantes para las cuales la muestra de datos pueda ser diferente a la población en general, así como en la identificación de estados indeseados que puedan incluir sesgos, desigualdades perjudiciales para ciertos subgrupos y/o cualquier otro patrón que se considere subóptimo o no deseable desde un punto de vista de la política social. Por último, puede ayudar a identificar las razones por las cuales falta información y decidir, junto con el equipo técnico, la manera de llenar este vacío.

Rol del director del proyecto

Descripción general de datos y motivación
Nombre de conjunto de datos usado
¿Qué institución creó la base de datos?
¿Con qué objetivo?
¿Qué mecanismos o procedimientos se utilizaron para recopilar los datos (por ejemplo encuesta en vivienda, sensor, software, API)?
Tamaño de la muestra
Frecuencia de captura (semanal, mensual, diaria) o número de observaciones promedio por individuo. ¿Se actualizará el conjunto de datos (por ejemplo, añadiendo nuevas instancias o eliminando otras)?

Conectar al equipo técnico con las personas a cargo de los datos (tanto internos como externos) para conseguir información detallada acerca de las bases, la forma de recolección y las posibles debilidades.

Discutir con el equipo legal las restricciones de uso de los datos, dada la forma en que se capturaron. Por ejemplo: ¿Es necesario obtener un consentimiento informado? ¿Se trata de información pública o privada?

Discutir, junto con el equipo técnico, si se considera que la muestra es representativa. Si no lo es, ¿cómo se afectará el entrenamiento de la herramienta?

Rol del director del proyecto

Controles esenciales
Obtener documentación para cada variable dentro del conjunto de datos. Breve descripción que incluya su nombre y tipo, lo que representa, cómo se mide su valor, etc.
Realizar un análisis exploratorio de los datos calculando estadísticas descriptivas, identificando el porcentaje de los valores faltantes y la distribución de cada variable dentro de la base de datos.
Analizar la cobertura espacial y temporal de los datos.
Analizar la cobertura de los grupos protegidos (sexo, raza, edad, etc.)
Describir las dimensiones importantes en las cuales la muestra de datos puede ser diferente a la población, y en particular los sesgos de selección no medidos. Apoyarse en la literatura especializada y en información de expertos.
Identificar posibles 'estados indeseables' en los datos, trátase de sesgos o inequidades perjudiciales para subgrupos o de cualquier otro patrón que se considere subóptimo o no deseable desde el punto de vista de la política social.
¿Existen valores faltantes? En caso afirmativo, explique las razones por las que no se tiene esa información. (Esto incluye información eliminada intencionalmente) Identificar razones por las que existen datos faltantes y piense si la falta de datos está asociada con la variable a predecir.

Analizar los resultados del análisis exploratorio de datos. Validar que la madurez de la base de datos sea la adecuada para solucionar el problema de política pública seleccionado.

Comunicar al equipo técnico qué grupos serán considerados protegidos.

Explicar al equipo técnico, con base en conocimiento experto, las diferencias encontradas.

Explicar, con base en conocimiento experto, los posibles sesgos de los datos que puedan surgir por un estado indeseado en la recolección. Definir lo que se considera un estado subóptimo. Discutir con el equipo técnico las medidas de mitigación de sesgos en los grupos protegidos.

Explicar, con base en conocimiento experto, los valores faltantes (por ejemplo, una circunstancia adversa que no permitió el relevamiento correcto de datos).

Anexo 5. Perfil del modelo

El perfil del modelo es una herramienta que resume las características principales de un sistema de toma y/o soporte de decisiones basado en aprendizaje automático, al tiempo que resalta sus principales supuestos y características, así como las medidas de mitigación implementadas (Mitchell et al., 2019). El director del proyecto podrá corroborar que la conceptualización y diseño de la política pública, así como los casos de uso, corresponden al problema que se esté abordando, para lo cual deberá verificar los antecedentes, la población destinataria, el horizonte de tiempo, y los actores y sistemas involucrados. Asimismo deberá asegurar que los casos de uso se ajusten a las situaciones donde se espera que se utilice el sistema de soporte de decisiones, incluso teniendo en cuenta los usos no considerados y las advertencias relacionadas. Por último deberá validar los grupos protegidos y los posibles sesgos. El perfil del modelo deberá ser completado conjuntamente por el director del proyecto y el equipo técnico.

Rol del director del proyecto

Conceptualización y diseño de la política pública		
1. Información básica	Personas que desarrollaron el modelo, fecha, versión, tipo	Está a cargo de toda la etapa de conceptualización y diseño del proyecto.
2. Casos de uso	Antecedentes	Definir correctamente el problema de política pública que busca resolver la institución.
	Población destinataria y horizonte de tiempo de las predicciones	Identificar correctamente a la población afectada. Definir de manera concisa y clara los objetivos que deben cumplirse para solucionar el problema descrito.
	Actores y componentes que interactuarán con los resultados	Contactar a las personas que actualmente dan respuesta a la situación para determinar cómo han procedido y cómo se podrían hacer mejoras con una herramienta de IA. Tener en cuenta las particularidades propias de la acción que puedan poner en riesgo la implementación de la herramienta.
	Casos de uso considerados durante el desarrollo	Junto con el equipo técnico, identificar a aquellos directores de proyecto que ya hayan resuelto —o que están intentando resolver— problemas similares con IA y contactarlos para familiarizarse con sus experiencias.
	Usos no considerados y advertencias relacionadas	Junto con el equipo técnico, identificar escenarios de uso incorrecto de la herramienta de IA, si es que existen, y diseñar medidas de mitigación para evitar su ocurrencia.
	Definición de grupos protegidos	Definir, con base en conocimiento experto, qué grupos serán considerados protegidos y a cuáles se quiere garantizar equidad en los resultados.

Recolección y procesamiento de datos		
3. Datos de entrenamiento	Conjunto de datos usados y su etiquetado	
	Pasos de preprocesamiento o preparación de datos.	Revisar, junto con el equipo técnico, los resultados del análisis exploratorio de los datos; explicar las anomalías encontradas (si las hay) y los resultados del análisis con base en conocimiento experto; y contextualizar los principales hallazgos.
	Sesgos y deficiencias potenciales según el caso de uso (véase el punto 2 aquí)	Solicitar al equipo técnico un análisis exploratorio de los datos por subgrupos de la población. Adoptar medidas de mitigación de sesgos y de equidad algorítmica.
Desarrollo del modelo		
4. Modelamiento	Algoritmos usados para entrenar, parámetros supuestos o restricciones	
5. Métricas de desempeño	Métricas técnicas usadas para seleccionar y evaluar los modelos	
	Análisis costo-beneficio del modelo para su caso de uso (véase el punto 2 aquí)	Realizar el análisis de costo-beneficio de la implementación de la herramienta y determinar si el problema de política pública puede solucionarse sin utilizar IA.
	Definición de grupos protegidos y medidas de equidad seleccionadas	Definir medidas de equidad en los resultados (subgrupos de población a los cuales se quiere garantizar equidad).
6. Datos de validación	Conjuntos de datos usados y su etiquetado	
	Pasos de preprocesamiento	
	Evaluación de adaptación de datos de validación según el caso de uso (véase el punto 2 aquí)	
	Sesgos y deficiencias potenciales según el caso de uso (véase el punto 2 aquí)	Establecer umbrales aceptables de sesgos junto con el equipo técnico, con base en conocimiento experto. Elaborar medidas de mitigación basadas en política pública, en caso de que se requiera.

7. Resumen del análisis cuantitativo	Error de validación reportado
	Resumen del análisis costo-beneficio
	Informe sobre medidas de equidad para grupos protegidos
Uso y monitoreo	
8. Recomendaciones de monitoreo	Estrategia de monitoreo y mejora en producción
	Estrategias de monitoreo humano de predicciones (si aplica)
Rendición de cuentas	
9. (Opcional) Explicabilidad de predicciones	Estrategia para explicar predicciones particulares (si es necesario)
	Estrategia para entender la importancia de distintos atributos
10. Otras consideraciones éticas, recomendaciones y advertencias	

Decidir, junto con el equipo técnico, qué modelo se implementará.

Definir, junto con el equipo técnico, un programa piloto o ensayo aleatorio antes de implementar el modelo en toda la población.

Programar monitoreos periódicos de la herramienta (lo cual dependerá de la naturaleza de esta última) con el equipo técnico y también con el funcionariado que interactúa con ella.

Formular una política de respuesta o de revisión frente a posibles solicitudes de la ciudadanía sobre el resultado de la herramienta.

Discutir con el equipo legal los riesgos que puedan surgir durante la implementación del modelo (protección de datos personales, sesgos u otros), y elaborar medidas de mitigación y de explicabilidad, en caso de que se requiera.

