

DIRECCIÓN DE COMPRAS Y CONTRATACIÓN PÚBLICA

**APRUEBA DIRECTIVA DE CONTRATACIÓN PÚBLICA N°44 SOBRE
RECOMENDACIONES PARA LA ADQUISICIÓN DE PROYECTOS QUE INCLUYEN
CIENCIA DE DATOS E INTELIGENCIA ARTIFICIAL (IA)**

SANTIAGO, 07 DE DICIEMBRE DEL 2023

VISTOS:

Lo dispuesto en la ley N°19.886, de Bases sobre Contratos Administrativos de Suministro y Prestación de Servicios; en el Decreto N°250, de 2004, del Ministerio de Hacienda, que aprueba su reglamento; en el Decreto N°792, de 2023, del Ministerio de Hacienda, que nombra a la Directora de la Dirección de Compras y Contratación Pública; y en las Resoluciones N°7, de 2019, y N°14, de 2022, de la Contraloría General de la República, que fijan normas sobre exención del trámite de toma de razón y montos en unidades tributarias mensuales.

CONSIDERANDO:

1. La función legal de la Dirección de Compras y Contratación Pública de asesorar a los organismos públicos en la planificación y gestión de sus procesos de compras, dispuesto en el artículo 30, letra a), de la ley N° 19.886.
2. La facultad de emitir orientaciones y recomendaciones generales, conducentes a difundir buenas prácticas y a fortalecer la probidad, la transparencia y la eficiencia en las compras públicas por parte de los compradores y los proveedores, contemplada en el artículo 104 bis del Reglamento de la Ley de Compras Públicas.
3. La necesidad de emitir una directiva con el objeto de identificar y prevenir riesgos en el tratamiento ético de datos en las contrataciones de proyectos que incluyan ciencia de datos e inteligencia artificial, considerando particularmente la existencia de las Bases Tipo para la “Adquisición de Proyectos de Ciencia de Datos e Inteligencia Artificial”.
4. Que, para efectos de aprobar la señalada directiva, debe dictarse el correspondiente acto administrativo.

RESOLUCIÓN

EXENTA N° 583-B

1. **APRUÉBASE** la Directiva de Contratación Pública N°44 sobre “RECOMENDACIONES PARA LA ADQUISICIÓN DE PROYECTOS QUE INCLUYEN CIENCIA DE DATOS E INTELIGENCIA ARTIFICIAL (IA)”, cuyo texto se transcribe a continuación:



DIRECTIVA DE CONTRATACIÓN PÚBLICA N°44

RECOMENDACIONES PARA LA ADQUISICIÓN DE PROYECTOS QUE INCLUYEN CIENCIA DE DATOS E INTELIGENCIA ARTIFICIAL (IA)

1. OBJETIVO DE LA DIRECTIVA

Esta Directiva entrega recomendaciones para la adquisición de proyectos de Ciencia de Datos e Inteligencia Artificial (IA) integrando los requisitos para la gestión responsable y ética de los datos, aplicable a todas las modalidades que dispone ChileCompra, y en particular, para facilitar la contratación a través de la modalidad de compra de Bases Tipo de proyectos de Ciencia de Datos e Inteligencia Artificial (IA).

La gestión ética de los datos se refiere al conjunto de prácticas que se implementan para garantizar un uso responsable de los datos que respete los derechos de las personas. Esto implica considerar y respetar ciertos principios a lo largo de todo el ciclo de vida de los datos: creación y captura, el almacenamiento, la transmisión y el análisis, hasta el archivo o eliminación. Entre los principios se encuentra la creación de valor público, el respeto a la privacidad, la identificación y mitigación de sesgos, la transparencia y la rendición de cuentas.

Si los proyectos se planifican con estas buenas prácticas desde un comienzo, podrán contribuir al valor público y enfrentar adecuadamente los riesgos a la privacidad, sesgos y falta de transparencia. Con un uso ético y responsable de los datos, los compradores públicos contribuirán a construir confianza con la ciudadanía y aprovechar el potencial que tienen los datos para mejorar el diseño y la gestión de las políticas públicas.

El uso de esta Directiva permitirá efectuar de manera más adecuada la contratación de bienes y servicios para estos rubros.

2. ALCANCE

Este documento permite facilitar la construcción de las especificaciones para establecer los antecedentes, requerimientos y condiciones de contratación para la adquisición y/o modificación de proyectos de Ciencia de Datos e Inteligencia Artificial.

¿Qué se entiende por ciencia de datos?	La ciencia de datos se define comúnmente como una metodología mediante la cual se pueden inferir ideas procesables a partir de los datos. La realización de la ciencia de datos es una tarea con un objetivo ambicioso: la producción de creencias informadas por datos y para ser utilizadas como base para la toma de decisiones (Igual y Seguí, 2017). En palabras más simples, es un conjunto de herramientas que permiten pasar de datos a información que sirva para la toma de decisiones.
¿Qué se entiende por Inteligencia Artificial (IA)?	Inteligencia Artificial (en adelante, IA) es un término general para una gama de tecnologías basadas en algoritmos que a menudo intentan imitar el pensamiento humano para resolver tareas complejas (The Alan Turing Institute).
¿Qué es un algoritmo?	Conjunto de instrucciones sobre cómo una computadora debe realizar una tarea en particular. Son utilizados por muchas organizaciones para tomar decisiones y asignar recursos basados en grandes conjuntos de datos. Se



	comparan con las recetas, que toman un conjunto específico de ingredientes y los transforman a través de una serie de pasos explicables en un resultado predecible (Caplan et al., 2018).
¿Qué es un modelo?	Representación matemática que se calibra a partir de un conjunto de datos para realizar tareas específicas. Esta capacidad de aprendizaje implica identificar patrones y regularidades en los datos, permitiendo al modelo hacer predicciones o tomar decisiones basadas en información nueva y no vista previamente. Los modelos se entrenan utilizando conjuntos de datos de entrada (características) y salida (objetivos), ajustando iterativamente sus parámetros para mejorar la precisión en sus predicciones. Con el tiempo, y a medida que se procesan más datos, el modelo se vuelve más eficaz en su tarea específica, ya sea la clasificación, la regresión, o cualquier otro tipo de análisis predictivo (Helm, JM, Swiergosz, AM, Haeberle, HS et al. Aprendizaje automático e inteligencia artificial: definiciones, aplicaciones y direcciones futuras. Curr Rev Musculoskelet Med 13, 69–76 (2020) https://doi.org/10.1007/s12178-020-09600-8).

3. CONSIDERACIONES AL ELABORAR UN PROYECTO CUYA SOLUCIÓN CONTENGA CIENCIA DE DATOS O INTELIGENCIA ARTIFICIAL

Si se adquieren bienes y servicios tecnológicos de manera estandarizada y frecuente, el proceso de compras suele ser simple y de bajo riesgo. Existen procesos más complejos, entre los cuales pueden existir situaciones que importen un mayor riesgo de impactar a las personas cuyos datos serán procesados por estos sistemas o cuyos datos serán el insumo para la toma de decisiones automatizadas.

Los riesgos de los proyectos de ciencia de datos e IA pueden ser de diverso tipo, a continuación, se enumeran algunos riesgos como los siguientes:

- Cuando en el mismo contrato concurren simultáneamente varias prestaciones diferentes (por ejemplo, soluciones integrales que incluyan desarrollos, softwares paquetizados, etc.).
- Cuando se transan montos altos (por ejemplo, superiores a 5.000 UTM).
- Cuando en la contratación se involucran tecnologías emergentes que hacen un uso intensivo de datos, como el uso de algoritmos o de IA, las cuales pueden generar dudas en relación a los aspectos de cumplimiento normativo y respeto a los derechos de las personas.
- Cuando se introduce el componente tecnológico en un proceso manual u operativo, que implica riesgos inherentes de ciberseguridad no considerados.
- Cuando el objeto de la contratación versa sobre bienes y/o servicios que son complejos de definir o de satisfacer en el requerimiento, por ejemplo, modelos no conocidos que se transforman en “cajas negras” informáticas o limitaciones de la propiedad intelectual, ya sea porque resultan más innovadores o porque están relacionados directamente con funciones críticas del órgano público, donde los organismos pierden visibilidad y control.

Para estos y otros casos más complejos y riesgosos, esta Directiva entrega algunas recomendaciones que permiten mitigar o disminuir dichos riesgos. Sin embargo, cada proyecto y la realidad de cada contrato puede requerir medidas complementarias o presentar particularidades, como, por ejemplo, la adquisición de tecnología en el sector



Defensa puede requerir estándares de seguridad mucho más exigentes que en el caso del área de la Educación.

Por tanto, **será responsabilidad de cada órgano contratante** velar por la implementación y cumplimiento de las medidas más adecuadas a la situación de su compra en específico.

4. FLUJO PARA LA CONTRATACIÓN DE SOLUCIONES QUE INCORPOREN CIENCIA DE DATOS E INTELIGENCIA ARTIFICIAL

4.1. Flujo para determinar el objeto de la contratación

A fin de entender el contexto en el cual las recomendaciones de esta directiva son aplicables, es importante comprender cómo se genera el flujo para identificar y desarrollar el requerimiento de compra de un proyecto que incorpora la ciencia de datos y/o la inteligencia artificial. Para ello, a continuación, se presenta un diagrama habitual para determinar el objeto de un proyecto de este tipo y su posterior adquisición y gestión contractual:



Imagen N°1 extractada y adaptada de *Permitido Innovar: Guías para transformar el Estado chileno ¿Cómo podemos desarrollar proyectos de ciencia de datos para innovar en el sector público?*

4.2. Etapas del proceso y recomendaciones

El proceso estructurado en el diagrama anterior abarca etapas consecutivas con interrogantes a resolver para avanzar a la solución final, y así determinar finalmente cómo realizar el requerimiento de compra y posteriormente elaborar el contrato. En dicha estructura es posible encontrar una serie de pasos para cumplir las fases del flujo referido, los que se explican a continuación. Como es una adaptación de las fases contenidas en la “Guía Permitido Innovar: ¿Cómo podemos desarrollar proyectos de ciencia de datos para innovar en el sector público?”¹, desarrollada por el Laboratorio de Gobierno del Ministerio de Hacienda y la Universidad Adolfo Ibáñez, se puede encontrar un detalle de la mayoría de los pasos y herramientas para implementarlos en dicho documento, exceptuando aspectos propios de los procesos de compra pública.

FASE 1: INVESTIGACIÓN DEL PROBLEMA

Paso 1: Conformar el equipo del proyecto

¹ El documento *Permitido Innovar: Guías para transformar el Estado chileno ¿Cómo podemos desarrollar proyectos de ciencia de datos para innovar en el sector público?* Laboratorio de Gobierno, Gobierno de Chile y Universidad Adolfo Ibáñez (2022) se encuentra disponible en <https://www.lab.gob.cl/permitido-innovar>.



Este documento ha sido firmado electrónicamente de acuerdo con la ley N° 19.799

Para verificar la integridad y autenticidad de este documento ingrese al siguiente link:

<https://doc.digital.gob.cl/validador/AIYTLU-419>

Es común pensar que los proyectos de ciencia de datos son de exclusiva responsabilidad del equipo técnico de la institución, es decir, de un conjunto de analistas y responsables de los datos. Esto es un error. Estos proyectos son llevados a cabo por equipos multidisciplinarios con competencias que se complementan. Más aún, la ciencia de datos no es posible sin un acabado conocimiento del dominio y entorno en el que se implementará el modelo. Los roles mínimos a considerar en esta fase son: jefatura del proyecto, representantes del área a intervenir (referente del proceso y funcionario/a ejecutor del proceso), analista de datos, responsable de datos y responsable legal. Además, se debe considerar el apoyo de la línea directiva de la institución. También se podría requerir el apoyo del área de comunicaciones o participación ciudadana, y de estudios y evaluación.²

Paso 2: Describir el problema

El objetivo de esta fase es obtener un conocimiento profundo del problema que afecta a las personas usuarias, y poner a prueba la idea de que la ciencia de datos es la forma de encarar las soluciones.

Una tentación es definir el problema como la ausencia de un modelo de ciencia de datos, o creer que la mera existencia de datos en una institución exige la existencia de este modelo, pero no se debe olvidar que su valor yace en los problemas que resuelve y el impacto final para las personas usuarias. La descripción será un paso necesario, pero no suficiente para que quienes toman las decisiones consideren necesario abordar la problemática.

Paso 3: Analizar la prefactibilidad del proyecto

La prefactibilidad del proyecto se puede analizar considerando los ámbitos políticos, económicos, sociales, tecnológicos y legales. Se debe determinar si existe una solución empaquetada que sea pertinente para resolver el problema o si es necesario un desarrollo. También debiera evaluarse si hay otras soluciones mejores que la implementación de ciencia de datos e IA, de manera de asegurar la proporcionalidad en el uso de la tecnología. Se plantean preguntas como:

- ¿Existe claridad sobre quiénes son las autoridades internas de la institución cuyas decisiones puedan influenciar el desarrollo del proyecto?
- ¿Existen los recursos financieros para el desarrollo de un modelo de ciencia de datos y su sostenibilidad en el tiempo?
- ¿Existen beneficios sociales o externalidades positivas asociadas al proyecto?
- ¿Existen los datos necesarios para el proyecto?
- ¿Está dentro de las funciones y competencias de la institución actuar sobre el problema?

Paso 4: Identificar actores y sus actividades clave

Es relevante identificar los actores involucrados en el problema y aquellos que posiblemente se involucrarían en la solución. A nivel micro, se identifica a las personas usuarias, interacciones y puntos de contacto donde el servicio es entregado. A nivel meso, las organizaciones o grupos de personas usuarias. Y a nivel macro, los sistemas que rigen y definen a los anteriores. Por otro lado, es necesario comprender detalladamente cómo se conforma el proceso detrás del problema descrito, cuáles son sus etapas y qué actividades las componen. Las actividades son un conjunto de acciones o tareas específicas dentro de un proceso que ejecuta una persona responsable de un cargo. La importancia de detallar

² Ver pg. 50-51 de Denis, Gabriela; Hermosilla, María Paz; Aracena, Claudio; Sánchez Ávalos, Roberto; González Iarcón, Natalia; Pombo, Cristina. Uso responsable de IA para política pública: manual de formulación de proyectos (2021). Banco Interamericano del Desarrollo. <http://dx.doi.org/10.18235/0003631>.



las actividades clave yace en que es necesario contar con un entendimiento acabado de cómo se realizan actualmente y cómo el modelo de ciencia de datos las podría modificar, enriquecer, facilitar o reemplazar.

Paso 5: Mapear datos

La materia prima de un proyecto de ciencia de datos o inteligencia artificial son, precisamente, los datos. Sin ellos no existe proyecto ni hay posibilidad de ofrecer una solución de estas características a las personas usuarias del servicio. Se deben identificar todas las fuentes de datos disponibles para modelar, y evaluar si contienen lo necesario para solucionar el problema. Es relevante plantear preguntas como:

- ¿Cómo se almacena la información?
- ¿Qué información se recolecta?
- ¿Cómo se accede a los datos?

Paso 6: Definir objetivos del proyecto

En el momento de definir los objetivos, particularmente en proyectos de ciencia de datos, recomendamos basarse en la metodología de objetivos SMART, para avanzar con más certeza y con mayor control del proceso, definiendo objetivos específicos, medibles, alcanzables, realistas y que se ajusten a la temporalidad del proyecto.

FASE 2: PROPUESTA DE SOLUCIÓN

Paso 1: Realizar estudio de mercado

Realizar un estudio de mercado implica recopilar y analizar información detallada sobre las opciones de modelos de IA o ciencia de datos que existen para solucionar el problema, incluyendo demanda, competidores, tendencias y factores externos, con el propósito de comprender mejor las oportunidades y desafíos a los que se enfrentan. Esta investigación permite tomar mejores decisiones y adaptar la solución a las necesidades de los usuarios del proyecto.

En tal sentido, considérese que el artículo 13 bis del reglamento de la Ley N°19.886 regula la posibilidad de que las instituciones públicas efectúen, antes de la elaboración de las bases de licitación, procesos formales de consultas o reuniones con proveedores, mediante llamados públicos y abiertos, convocados a través del Sistema de Información, con el objeto de obtener información acerca de los precios, características de los bienes o servicios requeridos, tiempos de preparación de las ofertas, o cualquier otra que requiera para la confección de las bases.

Por otro lado, debe recordarse que para el caso de las contrataciones en las cuales la evaluación de las ofertas revista gran complejidad y en aquellas superiores a 5.000 UTM, de acuerdo al artículo 13 ter del mismo reglamento citado es obligatorio que las entidades obtengan y analicen la información acerca de las características de los bienes o servicios requeridos, de sus precios, de los costos asociados o de cualquier otra característica necesaria, pudiendo para ello utilizar el mecanismo de consultas al mercado ya referido

Paso 2: Seleccionar tipo de análisis de datos

En un proyecto de ciencia de datos, parte importante de la ideación consiste en seleccionar el o los tipos de análisis de datos que se llevarán a cabo. Entre los más destacados se encuentran: descriptivo, exploratorio, inferencial, predictivo, causal y mecánico.



Paso 3: Analizar factores éticos que afecten la propuesta

Este paso busca resolver la pregunta respecto a ¿qué elementos de ética y seguridad se deben considerar en la compra? Para eso sugerimos el uso de la Guía “Formulación ética de proyectos de ciencia de datos”³, que divide las consideraciones éticas en 4 dimensiones:

- 1) Uso y protección de datos personales
- 2) Transparencia y rendición de cuentas
- 3) Sesgos y discriminación
- 4) Otras consideraciones éticas: proporcionalidad y licencia social

Además, la guía identifica riesgos éticos según casos de usos habituales, como la predicción, personalización, o detección de eventos, lo que podría facilitar el análisis de los riesgos éticos para el proyecto que se comprará.

Se agrega a estas consideraciones, lo relativo a la ciberseguridad, un componente indispensable al trabajo con datos y sistemas digitales hoy en día.

A continuación, se explica de manera somera cada uno de los ámbitos de riesgos éticos. Para más detalles se sugiere consultar la guía antes referida.

1) Uso y protección de datos personales

La protección de datos personales está reconocida como derecho fundamental en la Constitución Política de la República de Chile actualmente vigente, en su artículo 19 N°4, y se ha regulado a nivel legal principalmente a través de la Ley N°19.628 sobre Protección de la Vida Privada, la que dispone expresamente que las instituciones públicas pueden utilizar datos personales siempre y cuando su uso esté dentro de materias de su competencia, resguardando la confidencialidad de la información y solicitando el consentimiento previo, expreso e informado del titular cuando corresponda. El tratamiento de datos personales debe respetar los derechos de las y los titulares de los datos, resguardando su confidencialidad.

Para identificar los riesgos en esta materia, se debe realizar lo siguiente:

- 1) Identificar si se trata de datos personales y/o sensibles.
- 2) Identificar si se cuenta con las competencias para tratar los datos personales.
- 3) Si no se cuenta con las competencias, definir cómo se obtendrán los datos.
- 4) Identificar cuáles deberían ser los principales resguardos por adoptarse durante la ejecución del proyecto. Por ejemplo, se podría requerir una evaluación de impacto de protección de datos.

En la medida que sea posible, es recomendable entregar al proveedor de un contrato los datos de manera anonimizada o seudo anonimizada, de manera de adoptar un enfoque de preservación de la privacidad de las personas.

2) Transparencia y rendición de cuentas

La transparencia y rendición de cuentas de los proyectos de ciencia de datos se enmarcan en diferentes normativas chilenas, tales como:

- Ley N° 20.285 sobre Acceso a la Información Pública.
- Ley N° 20.500 sobre Asociaciones y Participación Ciudadana en la Gestión Pública.

³ Formulación ética de proyectos de ciencia de datos (2022), División de Gobierno Digital y Universidad Adolfo Ibáñez. <https://digital.gob.cl/transformacion-digital/estandares-y-guias/guia-formulacion-etica-de-proyectos-de-ciencia-de-datos/>.



- Ley Orgánica Constitucional de Bases Generales de la Administración del Estado (LOCBGAE).

La transparencia busca disminuir la opacidad que caracteriza a los modelos de ciencia de datos e IA, de manera que las decisiones que se tomen mediante o con el apoyo de estos sistemas, sea comprensible por las personas y comprendida por quienes operan estos sistemas.

En esta etapa, se requiere identificar en qué grado será necesario explicar los mecanismos de toma de decisión a los actores identificados en la fase de investigación del problema, ya sea que se trate de usuarios internos o externos. En la medida que el problema a resolver afecte significativamente los derechos de las personas, será necesario un mayor grado de transparencia en el proyecto.

Es deseable que se creen mecanismos estables de comunicación hacia las personas y de manera interna sobre los aspectos claves del sistema de ciencia de datos o IA, de manera de promover la confianza entre las personas y la institución.

La rendición de cuentas es un deber legal y ético que tiene todo funcionario o persona de responder e informar por la administración, manejo y rendimiento de fondos, bienes o recursos y los resultados, la verificación o validación en el cumplimiento del mandato o actividades que realiza.

Las auditorías son procesos que están diseñados para la verificación de hechos, para poder verificar el cumplimiento de una actividad según lo planeado y las directrices estipuladas. La finalidad de una auditoría es diagnosticar e identificar qué actividades se desarrollan según lo esperado, cuáles no y aquellas que son susceptibles de mejora. En materia de sistemas de IA, es muy importante poder verificar el cumplimiento de su funcionamiento, tanto respecto de medidas de seguridad, análisis de sesgo y protección de datos comprometidas en su oferta y exigidas en las bases. Una auditoría algorítmica debería permitirnos comprender el funcionamiento técnico del algoritmo, sus especificaciones técnicas de desarrollo, los datos utilizados, método de desarrollo, procesos, justificación del modelo de análisis elegido, parámetros y rendimiento.

3) Sesgos y discriminación

La Constitución Chilena garantiza a todas las personas la igualdad ante la ley, estableciendo que en Chile no hay persona ni grupo privilegiado. Ni la ley ni autoridad alguna podrán establecer diferencias arbitrarias. También se garantiza la igual protección de la ley en el ejercicio de los derechos de las personas y un debido proceso, esto es, un procedimiento justo.

En la ciencia de datos, en general, es fundamental reconocer que pueden existir sesgos, es decir, que puede haber errores sistemáticos en los datos, por diversos motivos, que condicionan cierta información en alguna dirección. Esto podría implicar que los resultados sean injustos y, en algunos casos, la situación de grupos vulnerados se replique y se agrave. Identificar estos sesgos y saber cómo abordarlos para que el modelo se implemente sin consecuencias negativas en la ciudadanía es una tarea especialmente compleja en política pública.

Existen distintos tipos de sesgos, como el sesgo muestral, el sesgo de medición y el sesgo por modelamiento, y siempre se deben tomar las precauciones y acciones necesarias para prevenirlos o corregirlos en la medida de lo posible.



4) Otras consideraciones éticas: proporcionalidad y licencia social

Es importante al formular un proyecto de ciencia de datos o IA, identificar si existen otras soluciones no basadas en tecnología que pueden dar respuesta al problema identificado. El desarrollo de modelos de ciencia de datos o IA puede ser costoso en recursos y tiempo, por lo cual es necesario realizar un análisis costo-beneficio de la posible solución. Además, puede tener efectos adversos en la ciudadanía, por lo que es de especial importancia comunicar el por qué la solución tomada es la mejor para solucionar el problema, identificando claramente los beneficios y riesgos potenciales.

El principio de proporcionalidad, y su debida comunicación a las personas es clave para lograr la aceptación general del proyecto de ciencia de datos o IA, lo que se denomina licencia social.

Para que el modelo sea efectivamente utilizado debe poseer licencia social, es decir, que la ciudadanía acepte la implementación del modelo. La licencia social es clave para la viabilidad del proyecto, ya que, si no se logra, se corre el riesgo de que el proyecto nunca se implemente. Para eso, no bastará con simplemente cumplir con las disposiciones legales, sino que se deberá cumplir mayores estándares de comunicación. Una buena práctica es que las organizaciones posean una plataforma de información con un lenguaje claro que entregue los conceptos más relevantes del modelo, y que dicha información se vaya actualizando con la retroalimentación de los usuarios.

5) Ciberseguridad

El uso de tecnologías de información, comunicaciones y el uso intensivo de datos, conlleva riesgos que pueden afectar los derechos de las personas, la seguridad pública, las infraestructuras críticas, los intereses esenciales y la seguridad exterior de Chile. Estos riesgos pueden provenir de múltiples fuentes y se pueden manifestar mediante actividades de espionaje, sabotaje, fraudes o ciberataques realizados por otros países, por grupos organizados o por particulares, entre otros.

La ciberseguridad es clave en todo el proceso de adaptabilidad a la sociedad digital, para la aplicación y desarrollo de tecnologías como la inteligencia artificial, en los diversos procesos socio-relacionales, en la generación de servicios y los procesos productivos. Existen actualmente directrices gubernamentales que obligan a las instituciones públicas a implementar procesos que permitan implementar las medidas técnicas y organizativas que permitan garantizar la protección, disponibilidad, integridad, confidencialidad y resiliencia de los sistemas digitales.

Paso 4: Elaborar el requerimiento de compra

La elaboración del requerimiento de compra es un proceso crucial en la adquisición de bienes o servicios para una organización. En este requerimiento deben detallarse las especificaciones técnicas del bien o servicio que se requiere contratar, así como las necesidades o problemas que se desean cubrir o solucionar con la compra. En cuanto al procedimiento de compra correspondiente, el organismo público debe regirse por lo establecido en la Ley N°19.886, de Compras Públicas, y su reglamento.

Para el caso de la licitación pública, la Dirección ChileCompra aprobó un formato de bases tipo para la adquisición de “Proyectos de Ciencia de Datos e Inteligencia Artificial”, mediante Resolución N°60, de 28 de diciembre de 2022, tomadas razón por la Contraloría General de la República el 18 de enero de 2023. Este documento contiene todos los aspectos necesarios para la gestión ética de los datos y, tal como lo ordenó la Contraloría



General en su dictamen N° E370752 de 2023, existe la obligatoriedad para los organismos públicos del uso de las Bases Tipo para los rubros en las que se encuentran disponibles.

Paso 5: Realizar el proceso de compra

Luego de la elaboración del requerimiento de compra, el organismo público procederá a efectuar la adquisición según el procedimiento de compra que corresponda. Así, si se realizará mediante licitación pública, es necesario que se elaboren las bases de licitación para convocar a los oferentes del rubro a presentar la mejor propuesta frente al requerimiento de compra, determinándose requisitos de admisibilidad (administrativos, técnicos y/o económicos) y criterios de evaluación, los cuales no deben ser arbitrarios ni vulnerar los principios de igualdad de los oferentes y de libre concurrencia. Se recuerda que el proceso de contratación pública es un proceso reglado, regido por la Ley N°19.886 y su reglamento. Supletoriamente se le aplicarán las normas de Derecho Público, y en defecto de estas, las normas de Derecho Privado.

FASE 3: GESTIÓN DE UN CONTRATO DE IA O CIENCIA DE DATOS

Las etapas de esta fase dependerán si la solución es un desarrollo a medida o la compra de una solución existente.

En ambos casos, se recomienda considerar lo siguiente:

- **Revisión de las actividades de la carta Gantt respecto de los entregables y requisitos de gestión ética de datos:** Un aspecto relevante en la gestión del contrato es la carta Gantt entregada por el proveedor, que habitualmente se ajusta al inicio del proyecto. Para poder tener adecuada visibilidad de las exigencias respecto de gestión de riesgos éticos, se sugiere que estén consideradas en las actividades de la carta Gantt antes del entregable respectivo, de manera de asegurar su ejecución oportuna.
- **Verificación de cumplimiento de las especificaciones:** Esto tiene que ver con el control sobre el producto entregado, es el visto bueno a que todo lo que se solicitó en las bases, se cumpla. En este sentido, los planes de pruebas y las simulaciones en ambientes no productivos serán cruciales para determinar qué se espera, qué se recibe y que, tanto el organismo, como los desarrolladores tengan coincidencia en que ese está en presencia de un producto terminado.
- **Documentación:** Las organizaciones cambian, y suele existir rotación de personas, por lo cual siempre es recomendable documentar la solución y las decisiones asociadas a cada parte de las herramientas; así, el conocimiento, los detalles y reglas de negocio, quedan en la organización para su correcta operación. Se sugiere que el proveedor elabore minutas de las reuniones donde se detallan decisiones y su justificación y que luego estas decisiones y sus justificaciones se reflejen en los informes.
- **Capacitación:** Si es necesario, capacitar a los usuarios finales y eventualmente ajustar la perspectiva en que se está realizando la capacitación, ya que luego del comienzo del uso de la solución aparecerán posibles sugerencias o mejoras de los mismos usuarios.

A continuación, se distinguirá, según si el desarrollo es a medida o se compra una solución existente, cuáles son las actividades de gestión contractual:

I. Solución se debe buscar mediante un desarrollo a la medida:

as fases en este caso son las siguientes:



Este documento ha sido firmado electrónicamente de acuerdo con la ley N° 19.799

Para verificar la integridad y autenticidad de este documento ingrese al siguiente link:

<https://doc.digital.gob.cl/validador/AIYTLU-419>

Paso 1: Obtener / entregar y cargar los datos

En este punto desde la gestión ética de los datos valen las recomendaciones respecto a la gestión adecuada de datos personales, si aplica al proyecto, respecto a documentos de transferencia de datos, tratamientos por encargo o la necesidad de realizar una evaluación de impacto en protección de datos.

Paso 2: Transformar los datos

Este paso consiste en ejecutar distintos procedimientos que permitan convertir los datos en un formato listo para el análisis, es decir, que permiten el desarrollo del modelo propiamente tal. Para ello, es necesario considerar conjuntos de técnicas y transformaciones cuyo uso depende de los objetivos de cada proyecto. Contar con los datos en un formato adecuado es indispensable para un proyecto exitoso y habitualmente toma tiempo. A nivel técnico la transformación de datos en un flujo de un sistema de ciencia de datos o IA se refiere al proceso de alterar, modificar o convertir los datos de entrada en un formato que sea más adecuado para las tareas subsiguientes, como el análisis, la visualización o el modelado. Estas transformaciones pueden incluir una amplia variedad de operaciones, como la normalización, la limpieza, la codificación de variables categóricas, el manejo de datos faltantes, la extracción de características y la reducción de dimensionalidad, entre otros.

El objetivo es preparar los datos de una forma que maximice la eficacia y la precisión del modelo de ciencia de datos o IA en su tarea específica. Este es un paso crucial en cualquier proyecto, ya que la calidad y la preparación de los datos tienen un impacto directo en el rendimiento del modelo.

Paso 3: Realizar análisis exploratorio de datos

Tal como se indicó los sistemas de ciencia de datos o inteligencia artificial pueden resultar sesgados y una de sus causas puede ser que la base de datos que se utiliza para su entrenamiento posee sesgos. El análisis exploratorio de los datos es una descripción de cada uno de los conjuntos de datos utilizados en la creación del modelo de ciencia de datos o IA. Este análisis podría detectar anomalías en los datos y proporcionar información para evaluar la calidad, integridad, temporalidad, consistencia y posibles sesgos del conjunto de datos que será utilizado para entrenar un modelo.

Paso 4: Desarrollar y ajustar el modelo

En lo relativo a la gestión ética, en el desarrollo del modelo el proveedor deberá comparar el desempeño de las distintas opciones con las respectivas métricas de evaluación de equidad estadística que hayan sido seleccionadas por la entidad compradora, como aspectos relevantes a considerar para el desempeño del sistema, a objeto de verificar en los modelos los posibles sesgos.

Paso 5: Validar el modelo

Desde el punto de vista de la gestión ética de datos, el proveedor deberá entregar la documentación que permita a la entidad compradora comprender la justificación y el funcionamiento del modelo seleccionado para, la solución de Inteligencia Artificial y/o Ciencia de Datos. Esto puede hacerse mediante una *ficha de modelo*, como también incorporando mecanismos de explicabilidad del sistema, que le permitan comprender por qué una solución de Inteligencia Artificial y/o Ciencia de Datos conduce a una decisión o



resultado en particular, indicando claramente los factores clave que han llevado al sistema a un resultado determinado.

Paso 6: Generar conclusiones

Esto corresponde a la validación de los resultados del modelo de ciencia de datos por todo el equipo desde el punto de vista técnico y sustantivo, revisando si este se alinea con los objetivos del proyecto.

Todos estos pasos se encuentran descritos en detalle y existen herramientas para implementarlos en la “Guía Permitido Innovar: ¿Cómo podemos desarrollar proyectos de ciencia de datos para innovar en el sector público?”⁴ entre las páginas 50 y la 76.

Si el presupuesto para la compra es acotado, los pasos 1 y 2 podrían realizarse con personal interno y entregar los datos lo más procesados posible al proveedor adjudicado.

Además, existen dos subfases que podrían o no realizarse en el marco de la gestión del contrato, dependiendo del tipo de problema a resolver, la extensión del contrato y el presupuesto disponible. Estas son:

- **Desarrollo del piloto**
 - Paso 1: Diseñar evaluación de impacto
 - Paso 2: Implementar el piloto de la solución
 - Paso 3: Evaluar resultados del piloto

- **Implementación de la Solución:**
 - Paso 1: Planificar el despliegue
 - Paso 2: Monitorear el desempeño
 - Paso 3: Robustecer el modelo

El detalle de los pasos para el desarrollo del piloto y la implementación de la solución está en la guía referida previamente, entre las páginas 82 y 100.

II. Compra de una solución existente:

En el caso de una compra de una solución existente, las actividades de gestión del contrato deberían estar orientadas a determinar la validez que tiene sobre la problemática específica de la institución pública. Por ejemplo, a realizar un piloto que permita testear el nivel de desempeño sobre los datos de la institución y que este esté en línea con el desempeño prometido de la solución empaquetada.

Se sugieren los siguientes pasos para evaluar una solución existente de ciencia de datos o IA con el fin de realizar una correcta implementación en la institución:

- **Definir métricas claras:** Identificar las métricas clave de rendimiento (KPI, por sus siglas en inglés) que son relevantes para el proceso y la aplicación en cuestión. Puede ser precisión, tasa de error, retorno de la inversión (ROI), tiempo de respuesta, etc.
- **Comparación con modelos anteriores/línea base:** Si ya hay una solución existente, ya sea humana o automatizada, se debe comparar el rendimiento del nuevo modelo con el anterior.

⁴ Para mayor detalle de estos pasos revisar *Permitido Innovar: Guías para transformar el Estado chileno ¿Cómo podemos desarrollar proyectos de ciencia de datos para innovar en el sector público?* Laboratorio de Gobierno, Gobierno de Chile y Universidad Adolfo Ibáñez (2022). <https://www.lab.gob.cl/guias-permitido-innovar>



- **Validación cruzada y conjuntos de prueba:** Utilizar técnicas de validación como la validación cruzada para evaluar el rendimiento del modelo en diferentes subconjuntos del conjunto de datos. Esto ayudará a garantizar que el modelo es robusto y generaliza bien a nuevos datos.
- **Pruebas piloto:** Además de la validación estadística, realizar pruebas en entornos en vivo o semi-en vivo para evaluar cómo el sistema se desempeña en condiciones reales.
- **Revisión por expertos:** Invitar a expertos en el dominio del negocio y técnicos a revisar la solución. Pueden proporcionar una evaluación cualitativa de su utilidad y fiabilidad.
- **Aspectos éticos:** Asegurarse de que la solución no tiene efectos secundarios indeseables, como sesgos no deseados, y que cumple con todas las directrices éticas y regulaciones relevantes. Puede exigirse al proveedor informes de auditoría algorítmica, las fichas de datos, la ficha del modelo y mediciones de equidad estadística.

5. CLÁUSULAS SUGERIDAS EN LA ELABORACIÓN DEL REQUERIMIENTO DE COMPRA, BASES DE LICITACIÓN Y/O EL CONTRATO

Dada la complejidad de este tipo de contratos, que incluyen Ciencia de Datos o Inteligencia Artificial, se sugiere la inclusión de cláusulas que permitan precaver los riesgos que hemos señalado anteriormente, tales como las que se insertan a continuación, sin perjuicio que cada organismo público debiese adaptarlas a su realidad institucional y contractual, según sea el caso.

5.1. Seguridad de la Información y Ciberseguridad

El objetivo de una cláusula de seguridad de información es establecer disposiciones y responsabilidades específicas relacionadas con la seguridad y ciberseguridad, para el cumplimiento de la confidencialidad, disponibilidad e integridad de la información, y la resiliencia de los sistemas informáticos. Se propone la siguiente:

“El proveedor deberá adoptar todas las medidas técnicas y organizativas de seguridad que sean efectivas para efectos de evitar que la información del órgano contratante sea accedida por terceros no autorizados.

Lo anterior se extiende, además, a las comunicaciones electrónicas de dicha información entre el proveedor y el órgano comprador.

En tal caso, el proveedor deberá emplear las medidas de seguridad que sean necesarias y adecuadas para que estas comunicaciones no sean interceptadas.

Por el sólo hecho de participar en el presente procedimiento de compras, el oferente debe dar cumplimiento a las Políticas y Procedimientos vigentes de Seguridad y Políticas tecnológicas de la entidad compradora las cuales se presumen conocidas por el oferente, para todos los efectos legales. Asimismo, el trabajo o proyectos realizados por el proveedor para el contratante deben sujetarse estrictamente a los estándares de seguridad de la información establecidos por el comprador.

Para lo anterior, seguirá a lo menos, los estándares de seguridad establecidos en el decreto N°83, de 2004, del Ministerio Secretaría General de la Presidencia, sobre seguridad y confidencialidad de los documentos electrónicos, o aquella norma que lo reemplace. Cada institución deberá hacer el análisis del tipo de seguridad que requiere implementar según la naturaleza de los datos tratados”.



5.2. Notificación de incidentes de seguridad

En materia de Seguridad de la Información y Ciberseguridad, el Decreto N°273, de 2022, del Ministerio del Interior y Seguridad Pública, establece la obligación de notificación de incidentes de ciberseguridad, por parte de los jefes de servicio de los Ministerios y demás organismos de la Administración centralizada y descentralizada del Estado, al Centro de Respuesta ante Incidentes de Seguridad Informática (CSIRT) del Ministerio del Interior y Seguridad Pública. El artículo 3 del mencionado Decreto obliga a solicitar a los proveedores de servicios de tecnologías de la información, que compartan la información sobre las amenazas y vulnerabilidades que puedan afectar a las redes, plataformas y sistemas informáticos de los órganos de la administración del Estado. En tal sentido, se plantea a continuación una cláusula tipo a modo de recomendación:

“Si el Proveedor tuviese conocimiento de cualquier acceso ilícito a los datos de la entidad pública compradora y sus datos de soporte, almacenados en sus equipos o instalaciones, o tuviese conocimiento de un acceso no autorizado a dichos equipos o instalaciones y, tuviera como resultado la pérdida, revelación o alteración de los datos de la entidad (en adelante cada uno un “Incidente de Seguridad”), el proveedor deberá sin demora (1) notificar el Incidente de Seguridad al órgano comprador en un plazo máximo de 2 horas; (2) investigar el Incidente de Seguridad y proporcionar a la entidad información detallada sobre el Incidente de Seguridad; y (3) tomar medidas razonables para mitigar los efectos y minimizar los daños resultantes del Incidente de Seguridad.

Las notificaciones de Incidentes de Seguridad se remitirán a uno o más administradores de la entidad a través de cualquier medio que el órgano público seleccione, incluyendo correo electrónico. Es responsabilidad exclusiva del proveedor asegurarse de que sus administradores mantengan en todo momento datos de contacto exactos y actualizados. La obligación del proveedor de notificar o responder a un Incidente de Seguridad según lo previsto en esta sección no constituye reconocimiento por parte del proveedor en cuanto a incumplimiento o responsabilidad alguna con respecto al Incidente de Seguridad.

La entidad deberá notificar al comprador, sin demora, acerca de cualquier posible uso indebido que se haya producido en sus cuentas o credenciales de autenticación, o acerca de cualquier incidente de seguridad relacionado con la prestación de los servicios contractuales”.

5.3. Evaluación de Impacto en Protección de Datos

Las evaluaciones de impacto en protección de datos personales son un análisis de riesgos, previo al tratamiento de datos personales, que buscan prevenir que, en ciertas hipótesis de tratamiento de datos, se produzcan efectos que puedan afectar la privacidad o los derechos de las personas cuyos datos serán insumos del sistema del tratamiento de información de ciencia de datos e IA. Las evaluaciones de impacto en Protección de Datos se realizan mediante metodologías validadas a nivel internacional para identificar los riesgos del tratamiento, medidas a implementar, responsables y plazos.

“El proveedor deberá entregar un informe de riesgos en protección de datos, identificando particularmente la probabilidad de impactos en los derechos de las personas cuyos datos serán tratados, la cual puede abordarse mediante una evaluación de impacto en datos personales, de acuerdo a alguna metodología



propuesta por autoridades públicas reguladoras en la materia (Reino Unido, España, Uruguay, Argentina)⁵. Este informe deberá contener a lo menos:

- Un flujo y descripción de las operaciones de tratamiento de datos relacionadas con el proyecto.
- Un análisis normativo frente a la legislación de protección de datos identificando⁶ brechas de cumplimiento de las operaciones de tratamiento de datos descritas.
- Probabilidades de amenazas o vulneraciones a la seguridad de los datos.
- Descripción de escenarios de uso no adecuados.
- Descripción de los procesos de concientización de su equipo ejecutor, respecto al impacto de un manejo erróneo de los datos, su manejo legal y ético.
- Un análisis de riesgos de las brechas detectadas y recomendaciones para mitigarlos”.

5.4. Tratamiento de datos personales por mandato

El contrato de tratamiento por mandato o encargo es aquel en que los datos personales son entregados a un tercero denominado encargado del tratamiento de datos o mandatario, el cual le brinda un servicio al responsable que conlleva el tratamiento de datos personales. El encargado o mandatario sólo puede realizar lo que específicamente se le ha contratado y esta cláusula busca resguardar los términos de este tipo de tratamiento.

“Por la presente cláusula se encarga al proveedor el efectuar tratamiento de datos personales, por cuenta del órgano contratante.

Dicho mandato tiene por objeto [indicar la finalidad del mandato] y recae sobre los siguientes tipos de datos personales: [indicar los tipos de datos personales que se transferirán].

El mecanismo de transferencia será el siguiente [indicar lo que corresponda si habrá una o varias o permanentes transferencias].

El tratamiento durará exclusivamente durante la vigencia del presente contrato o hasta cumplir la finalidad del encargo, si ello ocurriera antes. Ocurrido cualquiera de los casos mencionados precedentemente, el proveedor deberá realizar la devolución de los datos y su eliminación efectiva, lo que podrá ser verificado por el órgano contratante.

El mandatario está obligado a garantizar el resguardo de la confidencialidad de los datos por parte del personal a su cargo, asegurando que estos conozcan las obligaciones de protección de datos a las que están sujetos. También deberá implementar las medidas de seguridad exigidas por el órgano contratante, las que también podrán ser auditadas por este último. Si el mandatario tuviese conocimiento o sufriera un incidente o violación a los datos personales, cualquier acceso ilícito a los datos de la entidad y sus datos de soporte, almacenados en sus equipos o instalaciones, o tuviese conocimiento de un acceso no autorizado a dichos equipos o instalaciones y, tuviera como resultado la pérdida, revelación o alteración de los datos de la entidad el proveedor deberá en el plazo de 2 horas (1) notificar el Incidente de Seguridad al órgano comprador; (2) investigar el Incidente de Seguridad y proporcionar a la entidad información detallada sobre el Incidente de Seguridad; y (3)

⁵ Las bases tipo proponen estas metodologías en el anexo 3, página 57, pero el proveedor puede utilizar otras. Ej: Ley N°19.628 sobre Protección de la Vida Privada.



tomar medidas razonables para mitigar los efectos y minimizar los daños resultantes del Incidente de Seguridad.

Las notificaciones de Incidentes de Seguridad se remitirán a [.....] a través de cualquier medio que el órgano público seleccione, incluyendo correo electrónico. Queda expresamente prohibido: 1) la subcontratación de los servicios sobre los datos objeto del mandato a menos que se cuente con la autorización expresa del órgano público, 2) el uso de dichos datos personales para fines distintos a los indicados en esta cláusula, quedando además expresamente prohibida su comunicación a terceros”.

5.5. Análisis exploratorio de los datos

Una de las razones de por qué los sistemas de ciencia de datos o inteligencia artificial pueden resultar sesgados se debe a que la base de datos que se utiliza para su entrenamiento posee sesgos. El sesgo se puede entender como el error sistematizado que tiende a favorecer constantemente en una dirección. Una herramienta sesgada puede llevar a discriminar a usuarios y beneficiarios del sistema de ciencia de datos e IA, de lo cual existen múltiples casos documentados en el mundo. Por eso, es importante comprender los sesgos existentes en los datos durante la etapa de análisis exploratorio, para poder adoptar las medidas de mitigación correspondientes. Por ello, se sugiere la siguiente cláusula:

*“El proveedor deberá realizar un análisis exploratorio de los datos utilizados en el desarrollo del modelo de IA, indicando su fuente, utilización, y una descripción de sus posibles sesgos. Este análisis deberá ocurrir en la etapa de **Preparación o preprocesamiento de los datos**, luego del mapeo de las distintas bases de datos.*

El análisis exploratorio de los datos es una descripción de cada uno de los conjuntos de datos utilizados en la creación del modelo de IA. Este análisis podría detectar anomalías en los datos y proporcionar 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.

El comprador dará acceso al proveedor a los conjuntos de datos, junto con la documentación sobre cada variable dentro de ellos. Asimismo, el proveedor 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. En caso de detectar falta de datos, el proveedor propondrá al comprador estrategias para llenar ese vacío”.

5.6. Cláusula de análisis de sesgo en los modelos

Los modelos de IA a menudo se entrenan en conjuntos de datos que reflejan las posibles desigualdades y prejuicios presentes en la sociedad. Si no se abordan estos sesgos, los modelos pueden perpetuar y amplificar las desigualdades existentes, lo que resulta en decisiones injustas o discriminatorias. Se propone analizar los modelos no solamente considerando sus efectos globales en la población o considerando las métricas tradicionales de desempeño, sino también el impacto en grupos protegidos definidos por la entidad compradora. Un ejemplo de cláusula en este sentido es:



“Previo a la elección del modelo propuesto, el proveedor deberá comparar el desempeño de las opciones de modelos con las respectivas métricas de evaluación de equidad estadística que hayan sido seleccionadas por la entidad compradora, como aspectos relevantes a considerar para el desempeño del sistema, a objeto de verificar en los modelos los posibles sesgos. Para esto podrá utilizar herramientas como Aequitas (<http://aequitas.dssg.io/>), fairlearn (<https://fairlearn.org/>), AI Fairness 360 (<https://ai-fairness-360.org/>) u otras a su elección”.

5.7. Exigencia de transparencia y razonamiento de la inteligencia artificial o IA explicable.

El sector público se encuentra sometido al principio de transparencia a nivel constitucional y legal. Esto implica la apertura, accesibilidad y rendición de cuentas de las acciones, decisiones y procesos gubernamentales a la ciudadanía, lo que incluye los procesos en los que participan decisiones automatizadas o algorítmicas. El órgano público que adquiere este tipo de tecnologías debe comprender cómo se toman las decisiones, para poder responder a la demanda ciudadana de transparencia y en ciertos casos ser capaz de explicar a los ciudadanos la toma de decisiones automatizada.

Para estos efectos, en los casos en los que se desarrolle un sistema a la medida, el organismo comprador deberá establecer la obligación del proveedor de hacer entrega del código fuente asociado al sistema de inteligencia artificial a la entidad compradora. En tal sentido, se propone la siguiente cláusula:

“Entrega del Código fuente:

El adjudicatario deberá entregar los códigos fuente de los modelos mediante el uso de un repositorio privado de alguna plataforma de versionamiento (ej: github, bitbucket. Etc.) convenida con el comprador.

Resulta relevante que el organismo público pueda comprender cómo los sistemas de ciencia de datos e IA llegan a sus resultados, a fin de poder dar explicaciones internamente o a la ciudadanía cuando corresponda. Para ello el contrato deberá incorporar reglas que permitan el traspaso de esta información desde el proveedor al comprador, en términos comprensibles para la entidad compradora. Para ello, se sugiere la cláusula siguiente:

Justificación del modelo:

El proveedor deberá entregar la documentación que permita a la entidad compradora comprender la justificación de la elección de la solución de Inteligencia Artificial y/o Ciencia de Datos seleccionada. Dentro de esta justificación se deberá hacer referencia a la complejidad del modelo, su explicabilidad así como el cumplimiento de los resultados exigidos, considerando las métricas tanto de rendimiento como de equidad y los métodos de evaluación para la selección. Lo anterior debe hacerse considerando, no obstante, que hay requerimientos que solamente pueden satisfacerse con algoritmos de alta complejidad (ej: reconocimiento o segmentación de imágenes, procesamiento de lenguaje natural, entre otros).

Para lo anterior, el proveedor deberá entregar una ficha de modelo de aprendizaje automático según la propuesta de Margaret Mitchell, Simone Wu, Andrew Zaldivar, Parker Barnes, Lucy Vasserman, Ben Hutchinson, Elena Spitzer, Inioluwa Deborah Raji, Timnit Gebru en Model Cards for Model Reporting (<https://arxiv.org/pdf/1810.03993.pdf>) o la señalada en las páginas 75 y 76 del



documento *Uso responsable de IA para política pública: Manual de formulación de proyectos del Banco Interamericano del Desarrollo*⁷, u otra propuesta afín”.

Dependiendo del proyecto, otra cláusula posible es la siguiente:

Mecanismos de explicabilidad:

El comprador deberá exigir al proveedor mecanismos de explicabilidad del sistema, que le permitan comprender por qué una solución de Inteligencia Artificial y/o Ciencia de Datos conduce a una decisión o resultado en particular indicando claramente los factores clave que han llevado al sistema a un determinado resultado, y los cambios en la entrada que deben modificarse para llegar a un resultado diferente.

Puede sugerir a los proveedores el uso de herramientas de explicación contrafactual que funcionan con modelos de cajas negras, aunque existen otras herramientas que podrían utilizarse como Local Interpretable Model-Agnostic Explanations (LIME), What if tool (<https://pair-code.github.io/what-if-tool/>), u otro que el proveedor proponga según el modelo a utilizar.

Lo anterior responde a la necesidad de poder contribuir a los procesos de auditoría tanto interna como externa que buscan evaluar la calidad del sistema y permitir a los afectados por dicho sistema, comprender los resultados y discernir en base a ello cuando sea necesario”.

5.8. Auditoría algorítmica de los sistemas

La auditoría algorítmica es un proceso posterior al desarrollo de un modelo de ciencia de datos e IA, que busca corroborar las pautas de funcionamiento del modelo con el objetivo de detectar sesgos posteriores, problemas de desempeño, que permitan proporcionar recomendaciones para mejorar los sistemas. La auditoría busca establecer la responsabilidad de las partes involucradas en la creación, implementación y mantenimiento de los modelos. Esto puede incluir a desarrolladores y a las organizaciones que los ejecutan. Una cláusula de este estilo podría ser la siguiente:

“El órgano contratante podrá solicitar, por sí o a través de un auditor independiente, la auditoría específica del funcionamiento algorítmico, debiendo el adjudicatario colaborar para permitir la verificación de los compromisos contractuales, particularmente, y sin que signifique exclusión de otro tipo de casos, como en el caso de algoritmos que incorporan técnicas de aprendizaje automático, que pueden integrar y tratar cantidades masivas de datos, incluyendo datos de carácter personal y sensible.

La auditoría algorítmica podrá realizarse sobre diseño, desarrollo e implementación de un modelo de ciencia de datos e IA cuando estos puedan producir un impacto desventajoso en categorías de grupos protegidos de acuerdo con la Ley 20.609 que establece medidas contra la discriminación y a la normativa de protección de la vida privada, Ley 19.628, de cara a corregir dichos problemas detectados.

El adjudicatario deberá cooperar con dicha auditoría brindando a) Transparencia Técnica, esto es, aquella que permita comprender el funcionamiento técnico del modelo, sus especificaciones técnicas de desarrollo, los datos utilizados, método de desarrollo, procesos, justificación del modelo de análisis elegido, parámetros y



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

Este documento ha sido firmado electrónicamente de acuerdo con la ley N° 19.799

Para verificar la integridad y autenticidad de este documento ingrese al siguiente link:

<https://doc.digital.gob.cl/validador/AIYTLU-419>

rendimiento y b) Transparencia de los procesos, esto es la provisión de información sobre el propósito del Sistema algorítmico y el proceso seguido en su desarrollo, los datos o categorías de datos utilizados en ese contexto, una comprensión de las opciones y supuestos realizados, la forma en que se prevé la intervención humana, el método utilizado para identificar los riesgos, los riesgos identificados y las medidas adoptadas para mitigar los riesgos, así como las partes que estuvieron involucradas en el desarrollo del Sistema Algorítmico y sus roles.

Un informe recogerá las conclusiones de la auditoría, el cual deberá indicar que el adjudicatario cumple con las obligaciones previstas en el Contrato. El adjudicatario estará obligado a subsanar los defectos identificados dentro de un plazo fijado de común acuerdo. Si el adjudicatario no subsana los defectos identificados dentro del plazo para subsanar tales defectos fijados en el informe, se entenderá este incumplimiento como una infracción grave a las obligaciones del contrato.

Las auditorías algorítmicas serán complementarias e independientes a las auditorías internas o externas que puede solicitar el organismo comprador, con el objeto de evitar duplicar esfuerzos”.

5.9. Confidencialidad

La cláusula de confidencialidad cumple un doble objetivo, resguardar la reserva de la información que será conocida por el adjudicatario y asegurar que la información se mantenga dentro del contexto de tratamiento definido en el desarrollo, atendida la finalidad y personas autorizadas para acceder a la información definida como confidencial.

“El adjudicatario no podrá utilizar para ninguna finalidad ajena a la ejecución del contrato, la documentación, los antecedentes y, en general, cualquier información, que haya conocido o a la que haya accedido, en virtud de cualquier actividad relacionada con el contrato.

El adjudicatario, así como su personal dependiente que se haya vinculado a la ejecución del contrato, en cualquiera de sus etapas, deben guardar confidencialidad sobre los antecedentes relacionados con el desarrollo de los servicios.

El adjudicatario debe adoptar medidas para el resguardo de la confidencialidad de la información, reservándose el órgano comprador el derecho de ejercer las acciones legales que correspondan, de acuerdo con las normas legales vigentes, en caso de divulgación no autorizada, por cualquier medio, de la totalidad o parte de la información referida.

La divulgación, por cualquier medio, de la totalidad o parte de la información referida en los párrafos anteriores, por parte del proveedor, durante la vigencia del contrato o dentro de los 5 años siguientes después de finalizado éste, podrá dar pie a que la Entidad entable en su contra las acciones judiciales que correspondan. Con todo, tratándose de bases de datos de carácter personal, la obligación de confidencialidad dura indefinidamente, de acuerdo con la Ley N°19.628, sobre Protección de la Vida Privada”.

6. GLOSARIO TÉCNICO

Para un mejor entendimiento de los conceptos utilizados en la presente directiva, se incorpora a este documento el siguiente glosario técnico:



- ❖ **Algoritmo:** Conjunto de instrucciones sobre cómo una computadora debe realizar una tarea en particular. Son utilizados por muchas organizaciones para tomar decisiones y asignar recursos basados en grandes conjuntos de datos. Se comparan con las recetas, que toman un conjunto específico de ingredientes y los transforman a través de una serie de pasos explicables en un resultado predecible (Caplan, Donovan, Hanson, & Matthews, 2018).
- ❖ **Anonimización de datos:** Consiste en la eliminación de los datos de identificación individual realizado para evitar de forma irreversible su identificación, de tal manera que ya no se pueden relacionar con un individuo determinado.
- ❖ **Base de datos:** Una base de datos (DB) es una recopilación de información, que puede estar conformada por datos estructurados o no estructurados. Es común que una DB esté controlada por un sistema de gestión de bases de datos (DBMS, por sus siglas en inglés). Gran parte de las DBs hacen uso del lenguaje SQL para consultar y escribir información. Las DBs permiten acceder, actualizar, organizar, modificar, y gestionar la información que contienen. Además, cuando una DB es relacional (común en ciencia de datos e inteligencia artificial) por lo general está conformada por datasets únicos que forman juntos una DB.
- ❖ **Ciencia de datos:** Se define comúnmente como una metodología mediante la cual se pueden inferir ideas procesables a partir de los datos. La realización de la ciencia de datos es una tarea con un objetivo ambicioso: la producción de creencias informadas por datos y para ser utilizadas como base para la toma de decisiones (Iguar & Seguí, 2017).
- ❖ **Datos personales o datos de carácter personal:** Los relativos a cualquier información concerniente a personas naturales, identificadas o identificables (artículo 2° literal f) de la Ley N°19.628).
- ❖ **Datos personales sensibles:** Aquellos datos personales que se refieren a las características físicas o morales de las personas o a hechos o circunstancias de su vida privada o intimidad, tales como los hábitos personales, el origen racial, las ideologías y opiniones políticas, las creencias o convicciones religiosas, los estados de salud físicos o psíquicos y la vida sexual (artículo 2° literal g) de la Ley N°19.628).
- ❖ **Datos estructurados:** Datos altamente organizados y fáciles de descifrar por algoritmos. Por ejemplo: fechas, nombres, direcciones, entre otras. Es similar a un archivo entre filas y columnas diferenciados (IBM, 2021).
- ❖ **Datos no estructurados:** Datos que no pueden ser procesados ni analizados con herramientas de análisis tradicionales. Por ejemplo: texto, imágenes, actividad, redes sociales, entre otras (IBM, 2021).
- ❖ **Entrenamiento de un modelo:** Es la fase de la ciencia de datos donde se intenta ajustar de la mejor manera el algoritmo en base a los datos disponibles (C3.ai).
- ❖ **Licencia social:** Describe cómo las expectativas de la sociedad con respecto a algunas actividades pueden ir más allá del cumplimiento de los requisitos de la regulación formal; aquellos que no cumplan las condiciones para la licencia social (incluso si cumplen formalmente) pueden experimentar desafíos y cuestionamientos continuos. Son las expectativas de la sociedad con respecto a la conducta y actividades de las corporaciones que van más allá de los requisitos de la regulación formal (Carter, Laurie & Dixon-Woods, 2015).



- ❖ **Sesgo (estadística):** Posible consecuencia de negar a determinados miembros de la población la oportunidad de ser seleccionados para la muestra. Como resultado, la muestra puede no ser representativa de la población (Lind, Marchall & Wathen, 2005).
- ❖ **Sesgo (sistema computacional):** Sistemas informáticos que discriminan de forma sistemática e injusta a determinadas personas o grupos de individuos a favor de otros. Un sistema discrimina injustamente si niega una oportunidad o un bien o si asigna un resultado indeseable a un individuo o grupo de individuos por motivos que no son razonables o apropiados (Friedman & Nissenbaum, 1996).
- ❖ **Pseudonimización:** Consiste en la sustitución de un atributo, normalmente un atributo único, por otro atributo en un registro. Este es un proceso de seguridad que reduce la vinculabilidad de un conjunto de datos con la identidad del titular de los datos. La pseudonimización es una técnica de desidentificación que sustituye un identificador (o identificadores) de un titular de datos por un seudónimo para ocultar la identidad de ese titular. El dato pseudononimizado continúa siendo tratado bajo las reglas y procedimientos establecidos para los datos personales.
- ❖ **Métricas en Aprendizaje de máquina:** Las métricas en aprendizaje automático son medidas cuantitativas utilizadas para evaluar y comparar el rendimiento de los modelos. En problemas de clasificación, algunas métricas comunes son la exactitud, que mide la proporción de predicciones correctas, la precisión y el recall, que se centran en la calidad de las predicciones positivas, y el F1-Score, que es una media armónica de precisión y recall. También se usa la curva ROC y el área bajo esta (AUC-ROC) para evaluar la capacidad discriminativa del modelo. En problemas de regresión, el Error Cuadrático Medio (MSE) y el Error Absoluto Medio (MAE) son métricas populares que cuantifican la diferencia entre las predicciones y los verdaderos valores. El coeficiente de determinación R^2 mide cuánta variabilidad del resultado puede explicarse con las variables independientes. Para clustering, índices como Davies-Bouldin o la Silueta evalúan la calidad de las agrupaciones. Cada métrica tiene su propia relevancia dependiendo del contexto y del objetivo específico del modelo.
- ❖ **Equidad:** En algoritmos de aprendizaje automático, la equidad se refiere a la capacidad del modelo para hacer predicciones y tomar decisiones que son justas y no discriminan arbitrariamente contra ciertos grupos o individuos basados en características sensibles como género, raza, orientación sexual, religión, entre otras que pueden estar reflejadas en los datos o en el modelo que utilizará un sistema de IA. La equidad es crucial para la ética en inteligencia artificial y se evalúa mediante análisis de métricas específicas y técnicas de auditoría, que los compradores deberán exigir a los proveedores a fin de reducir los riesgos de sesgos que puedan producir los resultados de los sistemas que serán desplegados por el Estado. El objetivo de los análisis de equidad es minimizar los impactos dispares de los modelos considerando las características protegidas que sean relevantes en el sistema específico.
- ❖ **Transparencia:** En algoritmos de aprendizaje automático implica que tanto el proceso de toma de decisiones del modelo como su funcionamiento interno sean comprensibles y explicables para los usuarios, expertos y partes interesadas. Esta claridad ayuda a generar confianza en el sistema, facilita la identificación y corrección de errores o sesgos, y permite un escrutinio ético y legal. La transparencia es especialmente importante en aplicaciones críticas, como la



atención médica, la justicia y la toma de decisiones gubernamentales, donde la falta de entendimiento puede tener consecuencias significativas.

Referencias del glosario:

- C3.ai. Glossary. Model Training. disponible en: <https://c3.ai/glossary/data-science/model-training/>
- Cady, F. (2017). The data science handbook. John Wiley & Sons.
- Caplan, R., Donovan, J., Hanson, L., & Matthews, J. (2018). Algorithmic accountability: A primer. Data & Society, 18.
- Carter, P., Laurie, G. T., & Dixon-Woods, M. (2015). The social licence for research: why care. data ran into trouble. Journal of medical ethics, 41(5), 404-409.
- Friedman, B., & Nissenbaum, H. (1996). Bias in computer systems. ACM Transactions on Information Systems (TOIS), 14(3), 330-347.
- IBM Cloud Education (2021). Structured vs. Unstructured Data: What's the Difference?
- Igual, L., & Seguí, S. (2017). Introduction to Data Science. In Introduction to Data Science (pp. 1-4). Springer, Cham.
- Stock, J. H., Watson, M. W., & Larrión, R. S. (2012). Introducción a la Econometría (No. 330.1543 S8). Pearson.

7. DOCUMENTOS RELACIONADOS

Se hace presente que los lineamientos de este documento deben ser analizados de forma armónica y complementaria con la normativa actualmente vigente, especialmente en materia de compras públicas (Ley N°19.886 y su reglamento), la jurisprudencia administrativa de la Contraloría General de la República, como asimismo, con las guías, Bases Tipo y las Directivas de Compra de la Dirección ChileCompra –estas últimas, si bien actualmente no son vinculantes, ofrecen recomendaciones técnicas basadas en la experiencia y conocimiento del sistema que vale la pena considerar–, las que deben ser aplicadas de manera coherente e integral en concordancia con los manuales de compra de cada institución pública, y que se singularizan a continuación:

- Directiva de Compra N°26: Recomendaciones para una mayor eficiencia en la contratación de bienes y servicios.
- Directiva de Compra N°27: Recomendaciones para favorecer la generación de datos abiertos en la contratación pública.
- Directiva de Compra N°38, de recomendaciones para la contratación de bienes y servicios relacionados con tecnologías de información.
- Directiva de Compra N° 33: Recomendaciones para la innovación en las compras públicas.
- Guía ¿Cómo podemos desarrollar proyectos de ciencia de datos para innovar en el sector público? Laboratorio de Gobierno, Ministerio de Hacienda y Universidad Adolfo Ibáñez.
- Guía Formulación ética de proyectos de ciencia de datos. División de Gobierno Digital, SEGPRES y Universidad Adolfo Ibáñez. Uso responsable de IA para política pública: manual de formulación de proyectos del Banco Interamericano de Desarrollo.
- Base Tipo para la adquisición de “Proyectos de Ciencia de Datos e Inteligencia Artificial”, disponible en el sitio www.chilecompra.cl, sección bases tipo.



2. **PUBLÍQUESE** la presente resolución en la página web www.chilecompra.cl.

Anótese, Regístrese y Comuníquese,

**VERÓNICA VALLE SARÁH
DIRECTORA
DIRECCIÓN DE COMPRAS Y CONTRATACIÓN PÚBLICA**

RMZ/ECG/CPC/TMS/GLM/DEF

Distribución:

- Fiscalía
- División de Estudios e Inteligencia de Negocios
- División de Usuarios
- Dirección

