

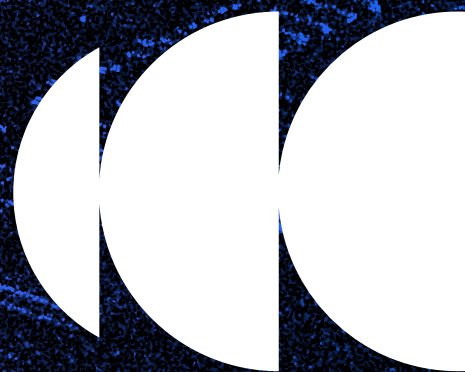


Procuración General
de la Ciudad

Marco de Gobernanza de Inteligencia Artificial

Principios rectores y arquitectura institucional

CABA 2026





1. Objetivos y alcance

El presente documento establece los principios, define los roles y determina procesos de gobernanza de la inteligencia artificial para la Procuración General de la Ciudad Autónoma de Buenos Aires.

Constituye la capa normativa de primer orden del marco de gobernanza de IA del organismo y se guía por la siguiente pregunta rectora: qué debe garantizar la Procuración General para asegurar que el uso de inteligencia artificial sea seguro, responsable, trazable y alineado con sus objetivos institucionales.

El presente marco se dicta en consonancia con la Ley 6.684, el Decreto 97/26 y la Resolución N.º 58/SECITD/26, que aprobó el Plan Estratégico de Inteligencia Artificial en el Gobierno de la Ciudad de Buenos Aires.

En ese carácter, operativiza los lineamientos de dicho Plan en el ámbito de la Procuración General y constituye el instrumento guía para que el organismo los observe en todo proyecto que involucre sistemas de inteligencia artificial, conforme al artículo 2º de la citada resolución.

Sus disposiciones se interpretan en armonía con la Guía de recomendaciones para el uso ético, responsable y adecuado de IA generativa en el ámbito de la Procuración General de la Ciudad Autónoma de Buenos Aires (RESOL-2025-50-GCABA-PG y Anexo IF-2025-06873959-GCABA-PG), la Guía para el diseño, desarrollo y uso ético, responsable y adecuado de sistemas de inteligencia artificial no generativos en el ámbito de la Procuración General de la Ciudad de Buenos Aires (RESOL-2025-397-GCABA-PG y Anexo IF-2025-50388218-GCABA-PG) y las normas complementarias que dicte la Secretaría de Innovación y Transformación Digital en el marco del artículo 4º del Decreto 97/26; ante el dictado de nuevas normas, la DGIA propondrá las adecuaciones correspondientes en la revisión periódica del presente.

El documento define el propósito institucional de la adopción de IA para el organismo, los principios que la rigen, los roles y responsabilidades, el ciclo de vida de las soluciones y la gestión del riesgo.

Sus disposiciones son de aplicación obligatoria para toda solución basada en inteligencia artificial que se diseñe, desarrolle, integre, despliegue, use, monitoree o dé de baja en cualquier dependencia de la Procuración General, independientemente del área que origine la iniciativa, del equipo que lleve adelante el desarrollo y de la tecnología empleada.

A los efectos del presente marco, el término solución de inteligencia artificial comprende los sistemas de inteligencia artificial a los que refieren el Decreto 97/26 y la Resolución N.º 58/SECITD/26, cualquiera sea la tipología por la que se incorporen conforme al Plan Estratégico de IA del GCBA

—IA colaborativa, IA transaccional o IA embebida—. La tipología opera como descriptor que informa la evaluación de riesgos pero no determina, por sí misma, la clasificación de riesgo ni las obligaciones aplicables a la solución.

2. Propósito institucional y objetivos

Para la Procuración General la inteligencia artificial no constituye una agenda tecnológica aislada, sino una capacidad institucional y una fuerza transversal que atraviesa y potencia los tres verticales de su Plan Estratégico —Gestión Estratégica del Litigio, Abogacía Pública de Excelencia y PG Inteligente—, que debe guiarse por el valor institucional de innovar con propósito.

Su desarrollo planificado, gobernado y medible sienta las bases para la evolución progresiva del organismo hacia un modelo de organismo inteligente y agéntico, capaz de servir con mayor eficacia a la ciudadanía y de generar valor público, respetando los derechos de las personas.

Esta evolución se inscribe, asimismo, en la visión y los objetivos del Plan Estratégico de Inteligencia Artificial del Gobierno de la Ciudad de Buenos Aires, a cuya realización el organismo contribuye desde su función específica.

El presente documento surge a partir de una premisa: la IA se gobierna mucho antes de implementarse.

En ese contexto, la gobernanza no opera como obstáculo a la innovación: es la condición que la vuelve sostenible, confiable y legítima en un organismo jurídico-público.

Este marco se concibe, en consecuencia, como un instrumento vivo, sujeto a revisión y actualización periódica conforme evolucionan las tecnologías, los usos institucionales, los riesgos asociados y los marcos normativos y éticos aplicables.

Son objetivos del presente:

- Garantizar que toda solución de IA del organismo opere bajo supervisión humana significativa, con trazabilidad de su funcionamiento y con una línea de responsabilidades identificable.
- Asegurar la protección de los datos personales y de la información institucional en todo el ciclo de vida de las soluciones.
- Graduar las exigencias de gobernanza en función del riesgo asociado a cada solución, evitando tanto la desprotección de los derechos de ciudadanos y usuarios, como la burocratización de la innovación.
- Establecer una arquitectura documental estable, que distinga la norma —principios y resultados exigidos— de los instrumentos de trabajo que la operativizan.

- Sostener el desarrollo y la preservación de las capacidades humanas que la adopción de IA requiere, en conjunto con la Escuela de Formación en Abogacía Pública como plataforma institucional de formación.
- Contribuir, mediante indicadores verificables, a la medición del aporte de la IA a los objetivos del Plan Estratégico y a las métricas de impacto y valor público definidas en el Plan Estratégico de IA del GCBA.

3. Principios de gobernanza

Los principios siguientes rigen el diseño, el desarrollo, la implementación, el uso, la operación y la baja de toda solución de IA del organismo, cualquiera sea su arquitectura tecnológica. Operan como criterios de decisión en cada etapa del ciclo de vida y como parámetros de interpretación del marco en su conjunto.

3.1. Supervisión humana significativa

Ninguna solución de IA del organismo decide por sí misma. La última palabra sobre todo producto que genere efectos jurídicos o institucionales corresponde siempre a un agente humano competente, conforme al esquema de supervisión que se defina como más adecuado para cada solución.

Las tareas de decisión y de validación no se delegan en soluciones de IA en ningún caso: la solución puede preparar los elementos para la decisión o asistir la producción de su borrador pero la decisión es siempre humana. La validación debe estar garantizada por el propio diseño de la solución, no solamente por una instrucción de uso.

La supervisión se define por su sustancia cognitiva y no por su forma: el organismo no la considera cumplida por la sola designación formal de un validador, sino por la concurrencia de las condiciones que la hacen posible —tiempo suficiente, competencia preservada y capacidades humanas específicas—. El deber de supervisar va asociado a la capacidad de ejercerlo.

La intensidad y la modalidad de la supervisión se establecen en función de la autonomía y del riesgo de cada solución.

3.2. Responsabilidad distribuida e indelegable a lo largo del ciclo de vida

La utilización de IA no exime al firmante ni a la autoridad competente de verificar la suficiencia, exactitud, legalidad, pertinencia y razonabilidad del contenido que adopta.

Hacia el interior de la Procuración la responsabilidad siempre permanece atribuible a personas

físicas. La responsabilidad profesional, funcional o jurídica correspondiente a la aprobación y firma del producto final asistido o generado por inteligencia artificial corresponde siempre a quien lo aprueba y/o firma.

No obstante, la responsabilidad por el funcionamiento de la solución no recae exclusivamente sobre quien aprueba y/o firma el resultado. Comprende a todos los actores que intervienen en su ciclo de vida, de acuerdo con sus respectivas funciones, decisiones, conocimientos, capacidades de intervención y grado de control sobre los riesgos.

Cada interviniente será responsable por las decisiones, acciones u omisiones comprendidas dentro de su ámbito de competencia y control. El usuario o firmante no asume riesgos originados en defectos de diseño, fallas técnicas, configuraciones inadecuadas que no estuvieran razonablemente bajo su conocimiento o capacidad de intervención.

Cuando el daño o riesgo se origine en la conducta de un tercero, incluida la manipulación deliberada de datos, documentos, instrucciones o materiales procesados por la solución, la responsabilidad deberá determinarse considerando tanto la conducta de dicho tercero como la existencia, suficiencia y correcta aplicación de los controles destinados a prevenir, detectar, contener y responder a esa manipulación.

La actuación dolosa o negligente de un tercero no excluirá automáticamente la responsabilidad de quienes, conforme a sus funciones, hubieran debido implementar o aplicar controles razonables.

3.3. Transparencia

El organismo documentará y mantendrá actualizado el inventario de las soluciones de inteligencia artificial que utilice, indicando, como mínimo, su finalidad, ámbito de aplicación, nivel de exigencia, responsable institucional, estado de implementación y principales características funcionales. El inventario estará disponible para el público general.

La información que se publique o comunique deberá ser compatible con las necesidades de seguridad, protección de datos personales, confidencialidad, secreto profesional y protección de la infraestructura crítica.

3.4. Explicabilidad

Toda solución de inteligencia artificial deberá diseñarse y operar de manera que permita comprender, en un nivel adecuado a su contexto de utilización, a los riesgos asociados y al público destinatario, cómo interviene en la producción de sus resultados y cuál es su función dentro del proceso

de trabajo o de toma de decisiones.

La información deberá presentarse de forma adecuada a las necesidades, competencias y derechos de usuarios y partes interesadas, de modo que resulte útil para el ejercicio de sus funciones o para la protección de sus derechos o intereses.

Cuando la solución incorpore modelos, servicios o componentes desarrollados por terceros, el organismo deberá poder explicar, como mínimo, el propósito de su utilización, la forma en que dichos componentes se integran a la solución, sus limitaciones conocidas y las medidas de supervisión y control implementadas.

3.5. Trazabilidad

Toda solución de inteligencia artificial deberá generar y conservar, en la medida del control que el organismo ejerza sobre su operación, registros suficientes para reconstruir las actuaciones relevantes realizadas durante su utilización.

La trazabilidad deberá permitir identificar, según corresponda, la versión de la solución utilizada, los principales datos o fuentes de información empleados, los resultados generados, las intervenciones humanas efectuadas y las decisiones adoptadas durante el proceso.

El organismo mantiene, asimismo, un catálogo institucional de sus sistemas de IA con el registro de sus ciclos de vida, y un repositorio único de su documentación técnica en sus versiones vigentes e históricas.

3.6. Uso proporcional de la inteligencia artificial

La utilización de inteligencia artificial deberá encontrarse debidamente justificada. El organismo sólo incorporará soluciones de IA cuando éstas resulten adecuadas para alcanzar una finalidad pública legítima y aporten un valor agregado verificable respecto de alternativas razonablemente disponibles.

La selección del tipo de solución, modelo o técnica de inteligencia artificial deberá ser proporcional a la finalidad perseguida, adecuada al contexto de utilización, sustentarse en fundamentos técnicos razonables y ser compatible con los principios y derechos protegidos por el presente marco. No deberán incorporarse componentes de inteligencia artificial cuando una solución de menor complejidad permita alcanzar razonablemente los mismos resultados con menores riesgos o costos.

3.7. Gobernanza proporcional al riesgo

Las medidas de gobernanza previstas en el presente marco se aplicarán de manera proporcional al nivel de riesgo y al contexto de utilización de cada solución de inteligencia artificial, conforme al modelo de gestión del riesgo y a los criterios del presente.

Ningún control se impone por la mera disponibilidad de la tecnología que lo hace posible, ni se omite por la dificultad de implementarlo.

3.8. Protección de datos personales

El tratamiento de datos personales mediante soluciones de inteligencia artificial deberá ajustarse a la normativa aplicable en materia de protección de datos personales. La privacidad constituye un derecho fundamental que deberá ser respetado y protegido durante todo el ciclo de vida de la solución.

Los datos personales deberán recopilarse, utilizarse, compartirse, conservarse y eliminarse conforme a las finalidades autorizadas y a los marcos jurídicos aplicables, adoptando medidas técnicas y organizativas adecuadas para garantizar su seguridad y confidencialidad. La naturaleza y sensibilidad de los datos tratados constituirán un criterio central en la evaluación de riesgos y en la definición de las medidas de gobernanza aplicables a cada solución de inteligencia artificial.

3.9. Seguridad

La seguridad de las soluciones de IA es una condición de su despliegue y de su permanencia en operación. Comprende la protección de la información que procesan, la gestión de accesos y credenciales, el registro de eventos, la evaluación de la superficie de ataque y la capacidad institucional de prevenir, detectar y responder a incidentes de seguridad, incluidas las técnicas de manipulación específicas de los sistemas basados en modelos de lenguaje.

La dimensión de ciberseguridad se integra en la evaluación de riesgos, en los estándares mínimos y en la documentación técnica de cada solución, y se desarrolla en trabajo conjunto con la Dirección de Tecnología e Informatización, de acuerdo a los estándares del Gobierno de la Ciudad de Buenos Aires.

3.10. Mejora continua

El marco de gobernanza y las soluciones que gobierna se revisan periódicamente a la luz de la evidencia de su funcionamiento. El monitoreo no es un control terminal sino una fuente de aprendizaje institucional: sus resultados alimentan ajustes, medidas de mitigación y, cuando corresponde, la suspensión o baja de soluciones.

3.11. Equidad y no discriminación

Las soluciones de IA del organismo no deben generar ni reproducir tratos injustificadamente desiguales entre personas o grupos. La identificación y mitigación de sesgos —en los datos, en los modelos y en los usos— integra la evaluación de riesgos y el monitoreo de toda solución, con intensidad proporcional a su impacto sobre las personas, en particular sobre los administrados. Este principio comprende, además, la accesibilidad de las soluciones para todos sus usuarios.

4. Roles y responsabilidades

4.1. Dirección General de Inteligencia Artificial (DGIA)

La DGIA diseña e implementa la estrategia institucional de adopción de inteligencia artificial. Elabora y mantiene el presente marco; asiste a las áreas en la identificación y el análisis funcional de casos de uso; conduce o acompaña el desarrollo de las soluciones según corresponda; verifica el cumplimiento de los estándares mínimos como condición del despliegue; administra el catálogo institucional de sistemas de IA y el repositorio del marco; y monitorea las soluciones en producción, proponiendo ajustes, mitigaciones o la baja cuando la evidencia lo justifique.

4.2. Áreas usuarias

Las áreas de la Procuración General que emplean soluciones de IA aportan el conocimiento del proceso durante el análisis funcional, cumplen la capacitación exigida como condición del uso, operan las soluciones conforme a su diseño y a las guías aplicables, ejercen la validación humana que corresponda a su rol y reportan los errores e incidentes que detecten.

4.3. Dirección de Tecnología e Informatización

La Dirección de Tecnología e Informatización, dependiente de la Dirección General Técnica Administrativa y Legal (DGTAL), interviene en el desarrollo, la integración y el mantenimiento de las soluciones que involucran los sistemas de gestión institucional o infraestructura a su cargo, y participa con la DGIA, de manera interdisciplinaria, en la definición y verificación de los controles de ciberseguridad de las soluciones de IA. La distribución de responsabilidades entre la DGIA y la Dirección de Tecnología e Informatización se documenta para cada solución en su documentación técnica.

4.4. Responsables por solución

Toda solución de IA cuenta, como mínimo, desde su aprobación y durante toda su vida operativa, con dos responsables designados de manera nominal:

 **Responsable funcional:** responde por la adecuación de la solución al proceso que asiste, por la vigencia de su necesidad institucional, por la supervisión de su uso en el área y por el reporte de los problemas funcionales.

 **Responsable técnico:** responde por la operación técnica de la solución, por la aplicación de los controles exigidos según su nivel de riesgo, por su mantenimiento y por la respuesta técnica ante incidentes.

La designación de ambos responsables es un estándar mínimo obligatorio. La asignación de los roles de gestión de la información asociados a cada solución —propietario, custodio y usuario de la información— se instrumenta mediante la ficha prevista en las guías operativas.

5. Modelo de clasificación por nivel de riesgo

Toda solución de inteligencia artificial será clasificada de acuerdo con su nivel de riesgo, de acuerdo con el Modelo de Gestión de Riesgo contenido en este documento. La clasificación se realizará sobre la base de las capacidades, el contexto de utilización y el impacto potencial de la solución, y no de la tecnología empleada. La arquitectura técnica subyacente —el modelo utilizado, sea propietario u open source; el framework; el orquestador; la programación clásica o las combinaciones híbridas de reglas deterministas y componentes de IA— no determina por sí misma el nivel de riesgo, sino que constituye uno de los elementos a considerar en su evaluación.

El riesgo se evalúa conforme al modelo de gestión del riesgo, sobre la base de las capacidades y el impacto de la solución, considerando al menos: el impacto en los derechos de las personas, en particular sobre los ciudadanos y ciudadanas; la sensibilidad de los datos tratados; el grado de autonomía de la solución, desde la asistencia puntual con intervención humana en cada paso hasta la operación por objetivos; su capacidad de ejecutar acciones con efectos, y no solamente de leer o sugerir; y la criticidad institucional del proceso que asiste.

Del nivel de riesgo resultante se sigue el nivel de exigencia de las medidas de gobernanza aplicables: los estándares mínimos obligatorios se aplican con la intensidad que corresponda a ese nivel. Las soluciones de bajo riesgo transitan una vía liviana, con requisitos reducidos y aprobación ágil; las soluciones de mayor riesgo requieren evaluaciones, controles y supervisión reforzados. En ningún caso el nivel de exigencia puede ser tal que carezca de relación razonable con el riesgo que mitiga.

Las soluciones de terceros que incorporen capacidades de IA —la IA embebida, en los términos del Plan Estratégico de IA del GCBA— quedan comprendidas en el presente marco.

Su adquisición o contratación requiere la evaluación de riesgos previa, con la profundidad que corresponda a su nivel de exigencia, e incorpora la debida diligencia sobre el proveedor: información suficiente sobre los datos tratados, los modelos empleados y los controles aplicados, y las previsiones contractuales de seguridad y confidencialidad que el nivel de riesgo exija. La circunstancia de que la operación técnica corresponda al proveedor no exime de la designación del responsable funcional ni de la aplicación de los estándares mínimos en la capa de control institucional.

El nivel de riesgo de toda solución de inteligencia artificial deberá determinarse antes de su diseño, desarrollo, adquisición o implementación. Esta clasificación define el nivel de exigencia de las medidas de gobernanza aplicables y deberá revisarse cuando se produzcan modificaciones relevantes en la solución o en su contexto de utilización.

6. Ciclo de vida de las soluciones

Toda solución de IA del organismo transita un ciclo de vida cuyas etapas se enuncian a continuación como garantías: lo que debe estar asegurado en cada una para poder avanzar a la siguiente. El modo concreto de satisfacer cada garantía se desarrolla en guías operativas y se gradúa según el nivel de riesgo de cada solución.



Idea. Toda iniciativa identifica el proceso institucional al que se propone asistir, la necesidad que la origina y su alineación con los objetivos del Plan Estratégico. Ninguna solución se desarrolla sin propósito institucional identificado.



Evaluación. Antes del desarrollo, la solución cuenta con un análisis funcional del proceso proporcional a su complejidad, con la evaluación de riesgos y con la determinación del nivel de exigencia aplicable. La evaluación define, además, los puntos de intervención humana y los indicadores con que se medirá la solución. Las pruebas de concepto y los pilotos que precedan al desarrollo se documentan y sus aprendizajes se sistematizan como conocimiento institucional reutilizable, en línea con el enfoque de experimentación responsable del Plan Estratégico de IA del GCBA.



Diseño y desarrollo. La solución se diseña y construye aplicando los controles que su nivel de riesgo exige y produce la documentación necesaria para su operación, su auditoría y la eventual transferencia de conocimiento entre equipos.



Despliegue. Ninguna solución entra en producción sin la verificación de calidad correspondiente a su nivel de riesgo, sin sus responsables funcional y técnico designados, y sin que sus usuarios hayan completado la capacitación exigida. El desarrollo de capacidades es condición del despliegue, al mismo nivel que la aprobación técnica.



Monitoreo. Toda solución en producción se monitorea mediante indicadores definidos en su evaluación, sobre el núcleo común institucional. Los resultados del monitoreo alimentan ajustes, mitigaciones y decisiones sobre la continuidad de la solución.



Baja. La desafectación es una etapa plena del ciclo, no un abandono de hecho. Comprende la decisión documentada de baja o suspensión, la preservación de los registros que la trazabilidad exige, el tratamiento de los datos conforme a la normativa aplicable y la comunicación a las áreas usuarias.

El pasaje de una etapa a la siguiente queda documentado en la documentación técnica de gobernanza de la solución. Las soluciones que se encuentren en operación al momento de la entrada en vigencia del presente marco se adecuan conforme al cronograma priorizado que establezca la DGIA.

Modelo de Gestión del Riesgo

1. Objeto y regla general

El riesgo de una solución se evalúa por sus capacidades y su impacto. La arquitectura técnica subyacente —por ejemplo, el modelo empleado, sea propietario u open source; el framework; el orquestador; la programación clásica o determinista, y las combinaciones híbridas de reglas fijas con componentes de IA— opera como descriptor que informa la evaluación. Del mismo modo opera la tipología de incorporación de la solución conforme al Plan Estratégico de IA del GCBA —colaborativa, transaccional o embebida—.

La evaluación de riesgos se realiza para cada solución en particular, desde la etapa de diseño de su ciclo de vida, y su análisis completo —dimensión por dimensión— se documenta en la documentación técnica de la solución, que integra su expediente de gobernanza.

La evaluación se vuelve a realizar frente a cambios sustantivos en la solución como, por ejemplo: la ampliación de la autonomía o de la capacidad de acción de la solución, la incorporación de nuevos datos o fuentes, el cambio del proceso asistido o la ocurrencia de un incidente relevante.

Ningún perfil de riesgo se modifica de hecho, siempre requerirá una nueva evaluación formal y la

actualización, cuando corresponda, del nivel de exigencia aplicable y de las medidas de gobernanza asociadas.

2. Dimensiones de evaluación

Toda solución se evalúa en las cinco dimensiones siguientes. Para cada una, la documentación técnica de la solución consigna el nivel identificado —bajo, medio o alto— y su fundamento. Los descriptores que siguen orientan la asignación de nivel; la evaluación pondera las particularidades del caso.

2.1. Impacto sobre personas

-  **Bajo:** la solución asiste tareas de uso interno sin efecto sobre terceros; sus outputs no integran actos ni comunicaciones dirigidas a personas externas al organismo.
-  **Medio:** los outputs de la solución integran, previa validación humana, insumos de actos o presentaciones que pueden afectar a terceros, sin intervenir en la determinación de su contenido decisorio.
-  **Alto:** la solución interviene en la elaboración de actos administrativos, dictámenes o presentaciones judiciales que afectan situaciones jurídicas de administrados, o procesa información cuya inexactitud puede perjudicarlos directamente.

La evaluación de esta dimensión comprende el análisis del riesgo de sesgo: la posibilidad de que la solución genere o reproduzca, en sus outputs o en sus usos, tratos injustificadamente desiguales entre personas o grupos. Las medidas de prevención, identificación, evaluación y mitigación adoptadas se consignan en la documentación técnica y se verifican en el monitoreo, con intensidad proporcional al nivel identificado.

2.2. Naturaleza y sensibilidad de la información tratada

-  **Bajo:** información pública o de gestión interna sin datos personales.
-  **Medio:** datos personales en los términos de la Ley 25.326, sin alcanzar las categorías del nivel alto.
-  **Alto:** datos sensibles en los términos del artículo 2 de la Ley 25.326, información reservada del organismo, o datos de expedientes cuya difusión indebida comprometa la posición de la Procuración General en juicio o los derechos de terceros.

2.3. Arquitectura funcional de la solución

-  **Bajo** — Asistente de IA: la solución brinda asistencia puntual a una persona, quien conduce la interacción, determina los pasos y utiliza o descarta cada resultado.
-  **Medio** — Flujo agéntico: la solución ejecuta una secuencia de tareas previamente diseñada, que puede combinar reglas deterministas, componentes de IA, herramientas y puntos de validación humana definidos.
-  **Alto** — Agente de IA: la solución opera a partir de objetivos y determina, dentro de los límites establecidos, los pasos, herramientas y acciones necesarios para alcanzar la finalidad encomendada.

2.4. Autonomía operativa y capacidad de acción

Bajo: la solución consulta, lee, clasifica, extrae, resume o formula recomendaciones. Sus resultados no producen efectos sin una decisión humana posterior.

Medio: la solución genera documentos, registros, comunicaciones o propuestas de actuación, pero requiere validación humana antes de incorporarlos a sistemas institucionales o producir efectos.

Alto: la solución ejecuta acciones sobre sistemas institucionales o frente a terceros —como registrar, modificar, remitir o notificar— sin validación humana previa para cada acción, aun cuando opere dentro de límites predefinidos o esté sujeta a supervisión posterior.

2.5. Criticidad institucional del proceso

-  **Bajo:** proceso de apoyo cuya interrupción o falla no compromete funciones sustantivas ni plazos.
-  **Medio:** proceso relevante cuya falla degrada la operación de un área o la calidad de sus productos, con alternativas manuales disponibles.
-  **Alto:** proceso sustantivo cuya falla compromete plazos procesales, la función dictaminante o la posición del organismo en juicio, o carece de alternativa manual viable en tiempo útil.

2.6. La ciberseguridad como factor transversal

La ciberseguridad no constituye una dimensión separada: atraviesa la evaluación de las cinco dimensiones. Al analizar cada una, la documentación técnica identifica los factores de exposición pertinentes: el procesamiento de contenido producido por terceros —vector del prompt injection indirecto—, la superficie de ataque de la solución, la gestión de accesos y credenciales, la conexión con sistemas institucionales críticos y las consecuencias de un

compromiso de seguridad sobre los datos tratados y sobre el proceso asistido. Este análisis se realiza en trabajo conjunto con la Dirección de Tecnología e Informatización, en consonancia con los estándares estipulados por la Agencia de la Información (ASI), cuando la solución involucre sistemas de gestión institucional o infraestructura a su cargo, y sus resultados alimentan los controles de ciberseguridad exigibles.

3. Determinación del nivel riesgo de la solución

Sobre la base del análisis de las distintas dimensiones, la DGIA determinará, mediante decisión fundada, el nivel de riesgo de la solución —bajo, medio o alto—. Este nivel impacta en la intensidad mínima de los requisitos de evaluación, documentación, autorización, supervisión, monitoreo y control que deberán cumplirse durante su ciclo de vida.

La determinación se realizará de manera fundada, caso por caso. Para ello, se ponderarán los niveles identificados en cada dimensión, los factores de exposición derivados del análisis transversal de ciberseguridad, y las particularidades del contexto de aplicación y del proceso asistido.

Las soluciones de menor riesgo estarán sujetas a requisitos simplificados, mientras que las de mayor riesgo requerirán medidas reforzadas. La clasificación y sus fundamentos deberán consignarse en la documentación técnica de la solución.

Estándares Mínimos Obligatorios

1. Objeto y reglas de aplicación

La presente sección establece los estándares mínimos obligatorios aplicables a toda solución de inteligencia artificial de la Procuración General de la Ciudad Autónoma de Buenos Aires.

Estos estándares integran el marco de gobernanza de IA del organismo, y traducen sus principios en requisitos mínimos, concretos y verificables aplicables a cada solución.

Los estándares se formulan como resultados verificables y no como implementaciones técnicas determinadas. Cada estándar establece el resultado que debe garantizarse, mientras que la documentación técnica y las guías operativas de cada solución definen los mecanismos adoptados para alcanzarlo y las evidencias que permiten verificar su cumplimiento.

Los estándares son tecnológicamente neutrales y no podrán interpretarse de manera que su cumplimiento quede condicionado al uso de una tecnología, arquitectura o proveedor específico.

Los estándares son aplicables a todas las soluciones. Sin embargo, la profundidad de su implementación, la documentación requerida, los controles aplicables y las evidencias de cumplimiento se adaptarán al nivel de riesgo asignado a cada solución conforme las pautas que se brindan seguidamente.

El cumplimiento de los estándares constituye una condición para el despliegue de la solución y para su continuidad en operación. Será verificado por la DGIA y quedará documentado en el documento de gobernanza de cada solución.

2. Estándares mínimos obligatorios

Regla general de aplicación y graduación

Los estándares previstos en esta sección son obligatorios para todas las soluciones de inteligencia artificial, cualquiera sea su nivel de riesgo. La graduación no determina si el estándar debe cumplirse, sino la profundidad de su implementación, la documentación y las evidencias requeridas, la frecuencia de los controles y el grado de intervención de las áreas competentes.

Las soluciones de riesgo bajo podrán cumplir los estándares mediante mecanismos simplificados. Las de riesgo medio requerirán controles y evidencias adicionales. Las de riesgo alto estarán sujetas a medidas reforzadas, mayor formalización documental y seguimiento periódico.

La clasificación general de una solución no impedirá aplicar medidas reforzadas respecto de un riesgo específico cuando sus características, los datos tratados, su contexto de utilización o la ocurrencia de un incidente así lo requieran.

2.1. Responsable funcional designado

Toda solución deberá contar, desde la aprobación del proyecto y durante todo su ciclo de vida, con un responsable funcional designado de manera nominal. La designación, sus funciones y sus modificaciones deberán constar en el expediente de gobernanza y mantenerse actualizadas.

El responsable funcional deberá velar por que la solución responda a la finalidad institucional definida, se utilice conforme a las condiciones autorizadas y mantenga su adecuación al proceso que asiste.



Qué se gradúa. Se gradúan la dedicación requerida, la frecuencia del seguimiento, las obligaciones de reporte y la participación del responsable en las revisiones de la solución.

En todos los niveles deberá existir una designación nominal; en los niveles medio y alto deberán documentarse, además, sus funciones de seguimiento, reporte y participación en las evaluaciones periódicas.

2.2. Responsable técnico designado

Toda solución deberá contar con una referencia técnica identificable durante su desarrollo, implementación, operación y mantenimiento. Cuando involucre sistemas de gestión institucional o infraestructura a cargo de la Dirección de Tecnología e Informatización, deberá identificarse también el referente de dicha área.

El responsable técnico deberá intervenir, dentro de sus competencias, en la definición de la arquitectura, las integraciones, los controles técnicos, la gestión de cambios y la respuesta ante incidentes.



Qué se gradúa. Se gradúan el grado de formalización de la designación, la dedicación requerida y la amplitud de sus responsabilidades. En el nivel bajo, la función podrá ser asumida por el área que administre o habilite la herramienta. En los niveles medio y alto deberá designarse un responsable técnico específico y documentarse su intervención en las etapas relevantes del ciclo de vida.

2.3. Evaluación de riesgos realizada y actualizada

Toda solución deberá contar, antes de iniciar su diseño, desarrollo, adquisición o implementación, con una evaluación inicial de riesgos. Esta comprenderá el análisis preliminar de las dimensiones previstas en el modelo de gestión del riesgo, el factor transversal de ciberseguridad y la determinación fundada del nivel de riesgo aplicable.

La evaluación deberá actualizarse durante el desarrollo y antes del despliegue, a medida que se definan y prueben la arquitectura, los datos, las integraciones, los controles y las capacidades efectivas de la solución. Una vez en operación, deberá revisarse sobre la base del monitoreo, los incidentes, los usos no previstos y los cambios en la solución o en su contexto.

Las evaluaciones y sus actualizaciones deberán documentarse en el expediente de gobernanza.



Qué se gradúa. Se gradúan la profundidad del análisis, la extensión de la documentación, la intervención de áreas especializadas y la frecuencia de las revisiones. En el nivel bajo podrá utilizarse una evaluación simplificada. En el nivel medio deberá documentarse el análisis completo y actualizarse antes del despliegue. En el nivel alto se requerirá una evaluación reforzada y revisiones periódicas durante la operación.

2.4. Trazabilidad

Toda solución deberá contar con mecanismos de trazabilidad adecuados a su modalidad de uso, su nivel de riesgo y el grado de control que el organismo ejerza sobre su operación.

Cuando el organismo controle el flujo de la solución, los registros deberán permitir reconstruir, en la medida necesaria, la información procesada, la solución, el modelo de IA y versión utilizadas, los resultados generados, las acciones ejecutadas y las intervenciones humanas realizadas.

Cuando se utilicen plataformas o herramientas cuya operación interna no sea controlada por el organismo, la trazabilidad se garantizará en la capa institucional mediante el registro de las herramientas y versiones autorizadas, las personas habilitadas, la formación recibida, los prompts institucionales validados y los incidentes reportados.

Toda solución deberá incorporarse al catálogo institucional de sistemas de inteligencia artificial, con los registros correspondientes a su ciclo de vida.



Qué se gradúa. Se gradúan, el detalle de los registros, su plazo de conservación y la capacidad de reconstrucción exigida. En el nivel bajo podrá bastar el registro institucional de la solución, sus usuarios autorizados y los incidentes. En el nivel medio deberán conservarse, además, las operaciones y validaciones relevantes. En el nivel alto deberá poder reconstruirse el proceso completo en los aspectos bajo control institucional.

2.5. Supervisión humana proporcional al riesgo

Toda solución deberá contar con mecanismos de supervisión humana definidos desde su diseño y adecuados a su autonomía, capacidad de acción y nivel de riesgo. Los puntos de intervención y validación humana obligatoria deberán documentarse y, cuando corresponda, quedar garantizados por la propia arquitectura de la solución.

Las funciones de decisión, aprobación y validación final no podrán delegarse en la solución cuando el resultado produzca efectos jurídicos, institucionales o materiales relevantes.



Qué se gradúa. Se gradúan la frecuencia, el momento y la profundidad de la intervención humana. En el nivel bajo podrá admitirse revisión por muestreo o por excepción cuando los resultados no produzcan efectos por sí mismos. En el nivel medio deberán establecerse puntos de validación antes de las actuaciones relevantes. En el nivel alto se requerirá validación previa a toda acción con efectos sobre las personas y revisión periódica de la efectividad sustantiva de la supervisión.

2.6. Gestión de incidentes

Toda solución deberá contar con un procedimiento para identificar, reportar, analizar y responder a incidentes relacionados con su funcionamiento, seguridad o uso.

Se incluyen, entre otros, los errores que afecten la calidad o confiabilidad de sus resultados, los accesos o acciones no autorizados y los intentos de alterar su comportamiento mediante instrucciones maliciosas incorporadas en la información procesada, como el prompt injection.

El procedimiento deberá establecer canales de reporte accesibles, responsables de intervención, criterios de gravedad, medidas de contención y corrección, y un registro de los incidentes, sus causas, efectos y tratamiento.

Todo incidente que evidencie riesgos no identificados, fallas relevantes en los controles o cambios en el perfil de riesgo deberá dar lugar a la revisión de la evaluación y de las medidas de gobernanza aplicables.



Qué se gradúa. Se gradúan los plazos de respuesta, el nivel de escalamiento, las obligaciones de comunicación y la profundidad del análisis posterior. En el nivel bajo será suficiente un canal de reporte, un responsable y un registro. En el nivel medio deberán definirse criterios de gravedad y plazos de respuesta. En el nivel alto se requerirán procedimientos formales de contención, análisis de causa, comunicación y reevaluación.

La gravedad del incidente podrá justificar la aplicación de medidas reforzadas con independencia del nivel general de riesgo de la solución.

Los plazos de respuesta, las obligaciones de comunicación y la profundidad del análisis posterior se definen según la gravedad del incidente, el nivel de riesgo de la solución y sus efectos potenciales o producidos.

Todo incidente que evidencie riesgos no identificados, fallas relevantes en los controles o una modificación del perfil de riesgo deberá dar lugar a la revisión de la evaluación de riesgos y de las medidas de gobernanza aplicables.

2.7. Monitoreo con indicadores definidos por proyecto

Toda solución en producción deberá contar con indicadores definidos durante su evaluación y validados antes del despliegue. Los indicadores deberán ser pertinentes para su finalidad, archi-

tectura, nivel de riesgo y contexto de utilización.

La selección podrá comprender, según corresponda, indicadores de desempeño técnico, calidad de los resultados, eficiencia operativa, intervención humana, impacto sobre el proceso, seguridad, cumplimiento y experiencia de uso.

Cada indicador deberá contar con una definición, una fuente de información, una frecuencia de medición, un responsable y, cuando corresponda, valores de referencia o umbrales para la adopción de medidas correctivas.



Qué se gradúa. Se gradúan la cantidad de indicadores, la frecuencia de medición, el nivel de automatización del monitoreo y las obligaciones de actuación ante desvíos. En el nivel bajo podrán utilizarse pocos indicadores esenciales, revisados periódicamente. En el nivel medio deberán incorporarse indicadores de calidad, funcionamiento, uso e incidentes. En el nivel alto se requerirá monitoreo reforzado, umbrales de alerta y procedimientos de actuación.

2.8. Controles de ciberseguridad proporcionales

Toda solución deberá aplicar controles de ciberseguridad adecuados a su exposición y a las consecuencias de un eventual compromiso. Cuando involucre sistemas o infraestructura a cargo de la Dirección de Tecnología e Informatización, los controles deberán definirse con la intervención de dicha área y de conformidad con los estándares de la Agencia de Sistemas de Información.

Los controles deberán contemplar, según corresponda, la gestión de accesos y credenciales, la protección de los datos en tránsito y en reposo, el registro de eventos de seguridad, la seguridad de las integraciones y la evaluación de la superficie de ataque, incluida la exposición a contenido producido por terceros.



Qué se gradúa. Se gradúan la profundidad de las evaluaciones, las pruebas requeridas, la intervención de áreas especializadas y la frecuencia de las revisiones. Toda solución deberá aplicar controles básicos de acceso, credenciales y protección de la información. En el nivel medio deberán evaluarse las integraciones, la superficie de ataque y la exposición a contenido externo. En el nivel alto se requerirán pruebas reforzadas, tratamiento formal de vulnerabilidades y revisión técnica antes del despliegue y ante cambios sustantivos.

2.9. Protección de datos personales

Toda solución que trate datos personales deberá documentar las categorías de datos involucradas, la finalidad y base jurídica del tratamiento, las personas o grupos afectados, los plazos de conservación y las medidas técnicas y organizativas adoptadas para su protección.

El cumplimiento deberá verificarse durante todo el ciclo de vida, incluyendo la recolección, utilización, acceso, transferencia, almacenamiento y supresión de los datos.



Qué se gradúa. Las obligaciones legales no se gradúan ni dejan de aplicarse por el nivel general de riesgo de la solución. Se gradúan la profundidad del análisis, la documentación requerida y las medidas técnicas y organizativas, de acuerdo con la naturaleza, sensibilidad y volumen de los datos, las finalidades del tratamiento y los riesgos para las personas.

2.10. Capacitación previa a la implementación

Toda persona habilitada para utilizar, supervisar, administrar o mantener una solución deberá recibir formación suficiente para comprender su finalidad, funcionamiento general, alcances, limitaciones, riesgos, reglas de uso y mecanismos de reporte de incidentes.

La capacitación deberá incluir, según el rol, la preparación adecuada de las entradas, la validación de los resultados, los errores previsibles y los criterios para interrumpir, escalar o cuestionar el funcionamiento de la solución.

La formación completada deberá quedar registrada y constituirá una condición para la habilitación de cada usuario.



Qué se gradúa. Se gradúan la extensión, especificidad, modalidad y frecuencia de actualización de la formación, de acuerdo con el rol y el nivel de riesgo. En el nivel bajo podrá utilizarse una capacitación breve y estandarizada. En el nivel medio deberá impartirse formación específica por función. En el nivel alto se requerirán ejercicios prácticos, evaluación de competencias y actualización periódica.

2.11. Soluciones de terceros con capacidades de IA incorporadas

Toda solución provista por terceros que incorpore capacidades de inteligencia artificial, incluida la IA embebida en los términos del Plan Estratégico de IA del GCBA, quedará sujeta a los

estándares del presente régimen en los aspectos bajo control institucional.

Antes de su adquisición, contratación o habilitación, deberá realizarse la evaluación de riesgos correspondiente, designarse un responsable funcional y, cuando proceda, un responsable técnico, y efectuarse una debida diligencia documentada sobre el proveedor y la solución.

La debida diligencia comprenderá, en la medida en que la información se encuentre disponible, los datos tratados, las capacidades y limitaciones de los componentes utilizados, las condiciones de seguridad y confidencialidad, los mecanismos de gestión de incidentes y los controles aplicados por el proveedor.

La contratación deberá incorporar previsiones adecuadas al riesgo, y la solución deberá registrarse en el catálogo institucional de sistemas de inteligencia artificial.



Qué se gradúa. Se gradúan la profundidad de la debida diligencia, las evidencias requeridas, el contenido contractual y la frecuencia de supervisión del proveedor. En el nivel bajo serán suficientes una evaluación simplificada, la designación del responsable funcional, la verificación de las condiciones esenciales del servicio y el registro en el catálogo. En el nivel medio deberán incorporarse previsiones contractuales sobre seguridad, confidencialidad, protección de datos, incidentes y cambios relevantes. En el nivel alto se requerirán, además, evidencias reforzadas de cumplimiento, mecanismos periódicos de supervisión y, cuando corresponda, derechos de auditoría o acceso a información equivalente.

La actuación del proveedor no eximirá al organismo de gestionar los riesgos comprendidos dentro de su capacidad de decisión y control.

3. Verificación

La DGIA verificará el cumplimiento de los estándares como condición previa a la implementación de cada solución y durante su operación, con la periodicidad y profundidad que correspondan a su nivel de riesgo, a los resultados del monitoreo y a los cambios producidos durante su ciclo de vida. La verificación deberá dejar constancia de los mecanismos, controles, documentos y evidencias mediante los cuales se acredita el cumplimiento de cada estándar. Sus resultados, las observaciones formuladas y las medidas correctivas requeridas integrarán el expediente de gobernanza de la solución.



Procuración General
de la Ciudad

www.procuracionba.com.ar



G O B I E R N O D E L A C I U D A D D E B U E N O S A I R E S

"2026 - Año del 30° Aniversario de la sanción de la Constitución de la Ciudad Autónoma de Buenos Aires"

Hoja Adicional de Firmas
Informe gráfico

Número: IF-2026-38274357-GCABA-PG

Buenos Aires, Martes 25 de Agosto de 2026

Referencia: Marco de Gobernanza de Inteligencia Artificial - Procuración General de CABA

El documento fue importado por el sistema GEDO con un total de 22 pagina/s.

Digitally signed by Comunicaciones Oficiales
Date: 2026.08.25 12:32:45 -03:00

Fernando Martin Ocampo
Procurador General
PROCURACION GENERAL
AREA JEFE DE GOBIERNO

Digitally signed by Comunicaciones Oficiales
Date: 2026.08.25 12:32:49 -03:00