

ANEXO I

PROTOCOLO PARA LA CREACIÓN, IMPLEMENTACIÓN Y USO RESPONSABLE DE AGENTES PÚBLICOS DE INTELIGENCIA ARTIFICIAL Y HERRAMIENTAS DE INTELIGENCIA ARTIFICIAL GENERATIVA EN EL ÁMBITO DE LA ADMINISTRACIÓN PÚBLICA PROVINCIAL

1. Objeto

El presente protocolo tiene por objeto establecer los principios, reglas y procedimientos aplicables a la solicitud, evaluación, diseño, desarrollo, implementación, operación, supervisión, registro, monitoreo y eventual discontinuidad de agentes públicos de inteligencia artificial y de herramientas de inteligencia artificial generativa utilizadas en el ámbito de la Administración Pública Provincial.

2. Alcance

Las disposiciones del presente protocolo son obligatorias para todo el personal que preste servicios en jurisdicciones y entidades comprendidas en el artículo 2° de la resolución aprobatoria, cualquiera sea su modalidad de vinculación, cuando en el ejercicio de funciones públicas utilice, configure, supervise, solicite o promueva el uso de herramientas de inteligencia artificial generativa o agentes públicos de inteligencia artificial.

Quedan excluidos del presente protocolo, sin perjuicio de las pautas generales aquí previstas, los desarrollos institucionales de inteligencia artificial de alto riesgo que por su complejidad requieran marcos específicos de adquisición, entrenamiento, validación, auditoría y gobernanza técnica.

3. Definiciones

A los fines del presente protocolo, se entenderá por:

- a. **Inteligencia Artificial (IA):** sistemas basados en máquinas que, para objetivos definidos por seres humanos, pueden generar predicciones, recomendaciones, clasificaciones, contenido o decisiones que influyen en entornos físicos o virtuales.
- b. **Inteligencia Artificial Generativa (IAG):** sistemas de IA cuyo propósito principal es generar contenido nuevo, en formato de texto, imagen, audio, video, código u otros, a partir de patrones aprendidos en datos de entrenamiento.
- c. **Agente Público de Inteligencia Artificial:** sistema o configuración tecnológica basada en IAG o IA que, bajo parámetros definidos por el Estado provincial, ejecuta o asiste tareas específicas, acotadas y predefinidas, con intervención y supervisión humana.

///... RESOLUCIÓN N° 02

d. **Supervisión Humana Significativa y Competente:** intervención activa, crítica e informada de un agente o funcionario público con autoridad suficiente para comprender el alcance del sistema, evaluar sus resultados, corregirlos, rechazarlos o decidir su no utilización.

e. **Datos Personales y Datos Sensibles:** los definidos por la normativa vigente en la materia.

f. Trazabilidad: aptitud del sistema y de su operación para permitir la reconstrucción verificable de las entradas, instrucciones, fuentes consultadas, intervenciones humanas, salidas generadas y decisiones adoptadas durante todo su ciclo de uso.

g. Explicabilidad e Interpretabilidad: condiciones técnicas y organizacionales que permiten comprender, en un nivel adecuado a la criticidad del caso, el funcionamiento general del sistema, los factores relevantes que incidieron en sus resultados y el alcance de sus limitaciones.

h. Consentimiento informado: manifestación de voluntad libre, expresa e informada del titular de los datos, precedida de información clara y suficiente sobre la finalidad del tratamiento, sus destinatarios, el carácter obligatorio o facultativo de la entrega de datos y la posibilidad de ejercer los derechos de acceso, rectificación y supresión, cuando ello resulte exigible por la normativa vigente.

i. A los fines operativos del presente protocolo, se consideran especialmente datos sensibles los que revelen origen racial o étnico, opiniones políticas, convicciones religiosas, filosóficas o morales, afiliación sindical, e información referente a la salud o a la vida sexual.

j. Registro de interacción: conjunto de evidencias técnicas y funcionales que documentan las interacciones con el agente, incluyendo consultas, archivos o insumos aportados, respuestas generadas, fuentes invocadas, fecha, usuario interviniente y validaciones realizadas, con sujeción a las reglas de confidencialidad y seguridad aplicables.

4. Principios rectores

Toda utilización de IA en el ámbito alcanzado deberá ajustarse a los siguientes principios:

a. Legalidad y juridicidad.

b. Centralidad de la persona y supervisión humana.

c. Transparencia

d. Responsabilidad y rendición de cuentas.

e. Protección de datos personales y privacidad.

f. Seguridad de la información.

g. Protección y promoción del talento y del empleo público, fortalecimiento de capacidades y reconversión laboral.

h. Equidad y no discriminación.

i. Eficiencia, proporcionalidad y mejora continua.

j. Accesibilidad.

...///

5. Regla general de uso

La IA tendrá carácter de herramienta de asistencia y apoyo al criterio profesional y/o técnico del agente público. Ningún resultado generado por una herramienta de IA producirá por sí mismo efectos jurídicos ni sustituirá la voluntad administrativa legítima, el análisis técnico o jurídico ni las competencias decisorias reservadas por el ordenamiento vigente.

Sólo podrán utilizarse herramientas, modelos o configuraciones de IA previamente autorizados por la Autoridad de Aplicación, seleccionados en función de la naturaleza de la tarea, la criticidad del proceso, la sensibilidad de los datos, la modalidad de despliegue y la disponibilidad de garantías de trazabilidad, explicabilidad e interpretabilidad requeridas para el caso.

La verificación y validación humana de los resultados no excluyen, por sí solas, las responsabilidades institucionales, funcionales o estatales que pudieren derivarse del diseño, implementación, supervisión o utilización de estas herramientas conforme el ordenamiento vigente.

6. Inicio del procedimiento

Las áreas interesadas en la creación de un agente público de inteligencia artificial deberán presentar solicitud formal firmada por la autoridad del área, ante la Autoridad de Aplicación, con indicación mínima de:

- a. problema, oportunidad de mejora y/o necesidad a resolver;
- b. tarea específica, acotada y claramente definida;
- c. descripción del proceso actual;
- d. volumen estimado de trabajo;
- e. tiempos actuales de ejecución;
- f. beneficios esperados;
- g. información o bases de conocimiento involucradas;
- h. riesgos previsibles;
- i. responsables funcionales y técnicos;
- j. propuesta inicial de indicadores de evaluación de desempeño.
- k. tipo de herramienta o modelo de IA generativa propuesto, proveedor, modalidad de despliegue y justificación estratégica de su elección;
- l. categorías de datos personales y sensibles eventualmente involucradas, base jurídica de tratamiento y, cuando corresponda, mecanismo previsto para obtención y resguardo del consentimiento informado;

No serán admitidas solicitudes genéricas, ambiguas o dirigidas a sustituir integralmente funciones humanas, decisorias o de control en puntos críticos de los procesos.

7. Evaluación de viabilidad

Recibida la solicitud, la Autoridad de Aplicación realizará la evaluación de viabilidad técnica, funcional, operativa, jurídica y de seguridad, contemplando:

- a. factibilidad del desarrollo y/o implementación;
- b. calidad y disponibilidad de la información;

...///

///... RESOLUCIÓN N° 02

- c.- riesgos en materia de sesgos, errores o alucinaciones;



*Ministerio de Economía
y Servicios Públicos*

- d.- impacto sobre procedimientos, derechos o servicios;
- e.- requerimientos de seguridad y confidencialidad;
- f.- necesidad de advertencias de uso y límites operativos;
- g.- prioridad institucional de la iniciativa.
- h.- adecuación estratégica del tipo de herramienta o modelo propuesto a la criticidad del caso y a la sensibilidad de los datos involucrados;
- i.- mecanismos de trazabilidad, auditabilidad, explicabilidad e interpretabilidad de corresponder;
- j.- determinación de las categorías de datos personales y sensibles comprometidas, de la base jurídica del tratamiento y, cuando corresponda, de las condiciones para el consentimiento informado;

8. Fase de co-diseño

De considerarse viable la propuesta, se desarrollará una instancia de co-diseño coordinada por la Autoridad de Aplicación con participación del área solicitante y de los agentes que ejecutan la tarea, la cual comprenderá:

- a. relevamiento del proceso actual;
- b. identificación de oportunidades de mejora;
- c. diseño del proceso objetivo;
- d. definición de casos de uso;
- e. delimitación del alcance funcional;
- f. determinación de reglas de negocio;
- g. definición de fuentes de información autorizadas;
- h. establecimiento de límites de actuación.
- i. definición de la base de datos de alimentación o conocimiento, sus criterios de selección, curación, actualización y depuración;
- j. determinación del conjunto de instrucciones, ejemplos, parámetros y demás componentes del kit de entrenamiento, configuración o ajuste operativo autorizado para el agente;
- k. diseño de controles para la detección, mitigación, corrección y reporte de sesgos, errores sistemáticos y desvíos de comportamiento.

9. Desarrollo, configuración y prueba

La Autoridad de Aplicación, con intervención del área solicitante, tendrá a su cargo la configuración, desarrollo o parametrización del agente, la preparación de la base de conocimiento, la definición de instrucciones, integraciones y controles, y la realización de pruebas controladas, interactivas y documentadas.

Durante esta etapa deberán registrarse los resultados, desvíos detectados, medidas correctivas y condiciones de uso.

...///

///... RESOLUCIÓN Nº 02

Durante esta etapa, los responsables de sistemas y los responsables funcionales

deberán efectuar controles permanentes sobre sesgos, deriva del modelo, calidad de la base de conocimiento y adecuación del kit de entrenamiento o configuración utilizado, dejando constancia documentada de las revisiones practicadas y de las medidas de corrección adoptadas.

10. Validación y puesta en funcionamiento

Finalizada la etapa de prueba, la puesta en funcionamiento requerirá:

- a. validación técnica de la Autoridad de Aplicación;
- b. conformidad funcional del área requirente;
- c. definición expresa de alcance, responsables y condiciones de uso;
- d. incorporación al Registro Provincial de Agentes Públicos de Inteligencia Artificial;
- e. capacitación básica del personal usuario.
- f. aprobación del esquema de trazabilidad, conservación y acceso a los registros de interacción de corresponder;
- g. validación de las medidas de anonimización, seudonimización y demás resguardos aplicables al tratamiento de datos;
- h. aprobación del plan de monitoreo permanente de sesgos, desempeño, actualización de la base de conocimiento y revisión del kit de entrenamiento o configuración.

11. Registro y publicación interna

Todo agente aprobado deberá inscribirse en el Registro creado por el artículo 6° de la resolución, consignándose como mínimo:

- a. nombre institucional;
- b. área responsable;
- c. finalidad;
- d. alcance funcional;
- e. modelo de lenguaje extenso utilizado;
- f. link y/o forma de acceso al agente;
- g. responsable de supervisión y actualización;
- h. fecha de habilitación;
- i. estado operativo;
- j. indicadores de evaluación de desempeño;
- k. restricciones de uso.
- l. proveedor, tipo de modelo, versión y modalidad de despliegue autorizada;
- m. base de datos de alimentación o conocimiento autorizada y descripción del kit de entrenamiento, configuración o ajuste operativo aprobado;
- n. categorías de datos involucradas, base jurídica del tratamiento y, cuando corresponda, referencia al mecanismo de consentimiento aplicable;
- o. política de trazabilidad, conservación y acceso a registros de interacción de corresponder;
- p. criterios mínimos de auditabilidad, explicabilidad e interpretabilidad exigibles para su operación.

La Autoridad de Aplicación podrá disponer, cuando corresponda, su publicación en catálogos o repositorios.

...///

12. Protección de datos y seguridad

En toda interacción con herramientas de IA deberán observarse estrictamente las reglas de protección de datos personales, confidencialidad, minimización de datos y seguridad de la información.

Antes de utilizar información institucional en cualquier herramienta de IA, el agente responsable deberá verificar la naturaleza de los datos involucrados y aplicar, cuando corresponda, técnicas de anonimización, seudonimización u otros resguardos adecuados.

La anonimización o disociación de datos deberá adecuarse al caso concreto y podrá comprender, entre otras medidas, la supresión o enmascaramiento de identificadores directos, la seudonimización, la generalización, la agregación, la tokenización u otras técnicas equivalentes, debiendo dejarse constancia de la modalidad aplicada cuando ello resulte relevante para la auditoría o revisión posterior.

Se consideran especialmente protegidos los datos sensibles relativos al origen racial o étnico, opiniones políticas, convicciones religiosas, filosóficas o morales, afiliación sindical, salud o vida sexual. Su tratamiento mediante herramientas de IA sólo podrá realizarse en los supuestos legalmente habilitados, con estricta necesidad, proporcionalidad, medidas reforzadas de seguridad y, cuando corresponda, consentimiento libre, expreso e informado.

13. Responsabilidades

Son responsabilidades del área solicitante:

- a. supervisar el uso operativo del agente;
- b. mantener actualizada la información que alimenta su funcionamiento;
- c. reportar incidentes, errores o desvíos;
- d. garantizar la intervención humana en los puntos *críticos* del proceso.
- e. custodiar, conforme a los perfiles de acceso autorizados, los registros de interacción y antecedentes funcionales necesarios para la revisión posterior del uso del agente, en caso de corresponder.
- f. supervisar de manera permanente la presencia de sesgos, la calidad de las bases de datos de alimentación y la adecuación del kit de entrenamiento, configuración o ajuste operativo autorizado para que se encuentre actualizado.

Son responsabilidades de la Autoridad de Aplicación:

- a. aprobar o rechazar solicitudes;
- b. coordinar el diseño y desarrollo;
- c. definir estándares técnicos mínimos;
- d. administrar el Registro;
- e. capacitar y asistir a las áreas usuarias;
- f. monitorear el desempeño y cumplimiento del protocolo;
- g. disponer revisiones, suspensiones o discontinuidades.

...///



14. Monitoreo, evaluación y mejora continua

Los agentes públicos de inteligencia artificial podrán ser sometidos a evaluaciones periódicas sobre uso, desempeño, reducción de tiempos, calidad de resultados, nivel de errores, satisfacción de usuarios, riesgos emergentes y adecuación normativa. Dichas evaluaciones deberán contemplar, como mínimo, las revisiones periódicas que correspondan según el nivel de complejidad, sobre trazabilidad, sesgos, deriva del modelo, calidad y actualización de la base de datos de alimentación o conocimiento, vigencia del kit de entrenamiento o configuración y suficiencia de los mecanismos de explicabilidad, interpretabilidad y acceso a registros.

La Autoridad de Aplicación podrá ordenar ajustes, reentrenamientos, limitaciones funcionales, suspensión preventiva o discontinuidad definitiva.

15. Incumplimientos

El incumplimiento de las obligaciones previstas en el presente protocolo constituirá falta administrativa, sin perjuicio de las responsabilidades civiles o penales que pudieran corresponder, y dará lugar a la adopción de las medidas preventivas, correctivas o disciplinarias pertinentes conforme la normativa vigente.