

Instrumento de Criterios para la Implementación de IA en Evaluación de Políticas Públicas

Marco de Gobernanza y Capacidades Institucionales

Alfredo Domínguez Díaz, Diana Valadez — ACEVAL | Academia Nacional de Evaluadores de México — 2026

Presentación

Este instrumento define los criterios mínimos para la implementación de Inteligencia Artificial (IA) en procesos de evaluación de políticas públicas. Está diseñado para uso institucional por equipos de evaluación en dependencias, organismos y firmas especializadas.

El marco no prescribe herramientas específicas. Define capacidades que cada institución debe construir según su contexto, recursos y nivel de madurez. Su propósito es hacer posible un uso de IA **auditable, seguro, estratégicamente justificado y críticamente supervisado**.

Destinatarios: Coordinadores de evaluación, responsables de gobernanza de datos, directores de área y cualquier persona que tome decisiones sobre el uso de IA en proyectos en el ciclo de la política pública.

Cómo usar este instrumento: Cada ámbito incluye una matriz de madurez con indicadores observables. Use la matriz para diagnosticar el estado actual de su equipo y priorizar acciones. Avance de Inicial hacia Consolidado con acciones concretas, responsables y fechas.

1. Giro conceptual: del error a la capacidad

Enfoque tradicional

El análisis organizacional suele partir del **error**: se identifica qué salió mal y se diseña un control para evitarlo. Este enfoque tiene una limitación estructural — los errores son síntomas, no causas. Tratar errores como punto de partida produce listas de verificación que no construyen capacidad real. El enfoque propuesto por ACEVAL se centra en desarrollo de capacidades de forma preventiva y en atención de los riesgos de la adopción de la IA de forma inadecuada.

Enfoque de capacidades

Este instrumento parte de un principio diferente: **los errores son señales de capacidades no construidas**. En lugar de preguntar “¿qué salió mal?”, pregunta “¿qué capacidad falta que habría detectado, prevenido o contenido el error?”

Un error no es una categoría de análisis — es una señal de que un ámbito no se construyó.

2. Los cuatro ámbitos institucionales

Cada ámbito representa una capacidad institucional que debe construirse. No son riesgos que se evitan, son capacidades que se desarrollan. Los cuatro ámbitos forman un sistema: son interdependientes y su debilidad en uno afecta a los demás.

#	Ámbito	Capacidad que propone
1	Privacidad y Seguridad	Clasificar, proteger y controlar la información que ingresa a las herramientas de IA.
2	Gobernanza	Institucionalizar prompts, normas y roles para que el uso de IA sea auditable y coherente.
3	Estrategia y Control	Incorporar IA con propósito, según una necesidad de la evaluación, no por moda o presión externa.
4	Capacidades y Uso Crítico	Pensar críticamente sobre la IA, dominar las herramientas metodológicas y documentar los procesos.

3. Arquitectura del proceso de IA

El uso de IA en evaluación no es un acto aislado — es un proceso sociotécnico con dos capas: una **capa operativa** (interacción con el modelo) y una **capa de gobernanza** (contexto institucional que habilita y controla esa interacción).

Las cuatro componentes

Componente	Tipo	Función	Contenido principal
Inputs	Operativa	Entrada al modelo	Prompt, instrucciones, alcance, formato esperado, criterios de calidad, versión.

Componente	Tipo	Función	Contenido principal
Ingresos de gobernanza	Institucional	Precondiciones de entrada	Datos depurados, contexto validado, restricciones institucionales, marco normativo, clasificación de información.
Outputs	Operativa	Salida del modelo	Respuesta estructurada, verificable, en formato útil y sujeta a control.
Resultados de gobernanza	Institucional	Trazabilidad y control	Logs, bitácoras, trazabilidad completa, evidencia de uso, responsable, decisión de aceptación o rechazo.

Los Inputs y Outputs pertenecen principalmente al intercambio entre analista y sistema. Los Ingresos y Resultados de gobernanza pertenecen al entorno institucional que habilita, limita, registra y valida ese intercambio.

La información de entrada no aparece solo como dato, sino también como evidencia clasificada, restricción, regla y contexto validado. La salida no se agota en la respuesta — se completa con los registros que permiten controlar, explicar y auditar su uso.

4. Matriz de madurez por ámbito

Cada ámbito tiene cuatro niveles de madurez. Use esta matriz para diagnosticar el estado actual de su equipo y planificar la transición hacia niveles superiores.

Ámbito 1 — Privacidad y Seguridad

Capacidad objetivo: Clasificar, proteger y controlar la información que ingresa a las herramientas de IA.

Nivel	Estado	Indicadores observables
Inicial	Sin clasificación	No existe clasificación de información antes de usar IA. Datos personales, datos de evaluados o información institucional se tratan como inputs sin filtro.
En desarrollo	Clasificación informal	Existe conciencia sobre qué información no debe enviarse a herramientas de IA, pero no hay proceso documentado. Cada analista decide por su cuenta.
Avanzado	Checklist formal en uso	Existe un checklist de clasificación de información que se aplica antes de cada proyecto. El equipo conoce qué categorías de datos requieren restricción.

Nivel	Estado	Indicadores observables
Consolidado	Checklist actualizado periódicamente	El checklist se revisa y actualiza al menos una vez al año. Se documentan las excepciones y las decisiones tomadas. Existe registro de versiones.

Acciones de transición:

- **Inicial → En desarrollo:** Documentar los tipos de información que el equipo ha identificado como sensibles (datos personales, información fiscal, datos de salud, entrevistas no consentidas, etc.).
- **En desarrollo → Avanzado:** Crear un checklist de clasificación con al menos tres categorías: información pública, información restringida, información confidencial.
- **Avanzado → Consolidado:** Definir responsable de actualizar el checklist, calendarizar la revisión y documentar incidentes de clasificación fallida.

Ámbito 2 — Gobernanza

Capacidad objetivo: Institucionalizar prompts, normas y roles para que el uso de IA sea auditable y coherente.

Nivel	Estado	Indicadores observables
Inicial	Sin documentación	Cada quien construye sus propios prompts. No hay repositorio, versión o revisión. Los prompts se borran o se pierden.
En desarrollo	Conciencia sin documentos	El equipo reconoce que los prompts son importantes, pero no hay proceso formal. Puede haber documentos compartidos sin versión ni responsable.
Avanzado	Lineamientos aprobados	Existen lineamientos internos que definen cómo se construyen, almacenan y revisan los prompts. Hay un responsable por proyecto y un formato mínimo.
Consolidado	Revisión periódica y auditoría	Los lineamientos se revisan al menos semestralmente. Existe auditoría de uso de IA por proyecto. Los prompts archivados incluyen versión, fecha, alcance y resultado.

Acciones de transición:

- **Inicial → En desarrollo:** Crear un repositorio compartido (carpeta en drive, SharePoint, o similar) para almacenar prompts con nombre, fecha y alcance.

- **En desarrollo → Avanzado:** Definir un formato mínimo para cada prompt: nombre, versión, fecha, responsable, alcance (qué preguntas resuelve), restricciones (qué no debe responder). Aprobar formalmente los lineamientos.
- **Avanzado → Consolidado:** Establecer auditoría semestral de uso de IA: muestreo de prompts archivados, revisión de outputs aceptados y rechazados, actualización de lineamientos con base en hallazgos.

Ámbito 3 — Estrategia y Control

Capacidad objetivo: Incorporar IA con propósito, según una necesidad de la evaluación, no por moda o presión externa.

Nivel	Estado	Indicadores observables
Inicial	Sin criterios de adopción	La IA se usa cuando alguien la sugiere o cuando está de moda. No hay documentación de necesidad o beneficio esperado.
En desarrollo	Criterio informal	El equipo tiene idea de cuándo conviene usar IA y cuándo no, pero no hay criterios escritos. Las decisiones varían entre proyectos y personas.
Avanzado	Marco formal y escala aplicada	Existe un marco formal con criterios para decidir cuándo usar IA en un proyecto. Se aplica a todos los proyectos nuevos. Se documentan las decisiones de adopción.
Consolidado	Marco medido y actualizado	El marco se evalúa con indicadores de uso: ¿la IA está agregando valor? ¿se reduce el tiempo sin comprometer calidad? Se actualiza con base en evidencia.

Acciones de transición:

- **Inicial → En desarrollo:** Documentar los casos en que se usó IA en proyectos recientes. Para cada caso, registrar: qué problema se quería resolver, qué herramienta se usó, qué resultado se obtuvo.
- **En desarrollo → Avanzado:** Crear un marco con al menos tres criterios de adopción: necesidad específica (no moda), capacidad de verificar outputs, disponibilidad de gobernanza en el ámbito correspondiente.
- **Avanzado → Consolidado:** Medir el impacto de IA por proyecto: tiempo ahorrado vs. calidad percibida. Definir indicador de “valor agregado” y revisar trimestralmente.

Ámbito 4 — Capacidades y Uso Crítico

Capacidad objetivo: Pensar críticamente sobre la IA, dominar las herramientas metodológicas y documentar los procesos.

Nivel	Estado	Indicadores observables
Inicial	Confianza sin revisión	El equipo acepta las respuestas de IA sin cuestionarlas. No hay protocolo de verificación. Se confunde usar IA con saber usar IA.
En desarrollo	Conciencia con sensibilización	El equipo reconoce las limitaciones de la IA, pero no ha recibido formación estructurada. Puede haber intuición sin herramientas concretas.
Avanzado	Programa de formación activo	Existe un programa de formación con al menos: pensamiento crítico sobre IA, limitaciones y sesgos conocidos, protocolo de verificación de outputs, uso ético.
Consolidado	Formación revisada y compartida	La formación se actualiza con base en nuevos riesgos, herramientas o casos. Existe material documentado que se comparte con nuevos integrantes. Se evalúa la comprensión del equipo.

Acciones de transición:

- **Inicial → En desarrollo:** Realizar al menos una sesión de sensibilización sobre limitaciones de IA en evaluación: alucinaciones, sesgos de entrenamiento, falta de contexto mexicano, riesgos de dependencia.
- **En desarrollo → Avanzado:** Diseñar programa de formación con objetivos de aprendizaje, contenidos mínimos y evaluación. Incluir al menos: qué puede y qué no puede hacer la IA en evaluación, cómo verificar un output, cómo documentar el proceso.
- **Avanzado → Consolidado:** Revisar el programa al menos una vez al año. Documentar las actualizaciones. Crear un manual de referencia rápida para el equipo.

5. Matriz de decisión rápida

Si su equipo necesita priorizar por dónde empezar, use esta matriz de urgencia-impacto:

Ámbito	Urgencia (qué pasa si no se construye)	Impacto en la evaluación
Privacidad y Seguridad	Alta si maneja datos sensibles	Confidencialidad institucional

Ámbito	Urgencia (qué pasa si no se construye)	Impacto en la evaluación
Gobernanza	Alta si usa IA de forma recurrente	Activos institucionales y trazabilidad
Capacidades y Uso Crítico	Alta si el equipo toma decisiones basadas en outputs	Rigor metodológico y validez
Estrategia y Control	Media — importante pero no urgente al inicio	Uso eficiente de recursos

Recomendación general: Si el equipo está en nivel Inicial en todos los ámbitos, comenzar por **Privacidad y Seguridad** (protección de información) y **Gobernanza** (registro y trazabilidad). Estos dos ámbitos habilitan los otros dos.

6. Principios transversales

Estos principios atraviesan los cuatro ámbitos y deben orientar cualquier decisión de implementación.

Principio 1 — Gobernanza primero, tecnología después

No se adopta una herramienta porque está disponible o está de moda. Se adopta porque responde a un problema diagnóstico en un ámbito donde la gobernanza ya está en su lugar o está en proceso de construirse.

Principio 2 — Formar para pensar, no solo para usar

La capacitación no se trata de enseñar a hacer clicks en una herramienta. Se trata de desarrollar pensamiento crítico sobre las limitaciones, los sesgos y las implicaciones éticas del uso de IA en la evaluación.

Principio 3 — La responsabilidad es institucional

Los errores no ocurren porque alguien no supo — ocurren porque el sistema no protegió a esa persona. La función de este instrumento es crear sistemas que reduzcan la probabilidad de error, no culpar a individuos.

Principio 4 — El contexto precede al modelo

Antes de seleccionar qué IA usar, debe existir el contexto gobernado: la información clasificada, las restricciones definidas, los roles asignados. La herramienta viene después.

7. Errores comunes por ámbito

Conocer los errores más frecuentes ayuda a prevenirlos.

Ámbito	Error común	Cómo evitarlo
Privacidad	Enviar información sensible a IA sin revisar políticas de privacidad de la herramienta	Aplicar checklist de clasificación antes del primer uso
Gobernanza	Guardar prompts en notas personales o conversaciones	Usar repositorio institucional desde el primer prompt
Estrategia	Adoptar IA porque “todos la usan” sin necesidad específica	Documentar la necesidad antes de la adopción
Capacidades	Confundir “usar IA” con “saber usar IA”	Formar en verificación antes de formar en uso

Instrumento desarrollado por: Alfredo Domínguez Díaz, Diana Valadez — ACEVAL | Academia Nacional de Evaluadores de México

Versión: 1.0 — 2026

Contacto: Para sugerencias o actualizaciones, contacte a la Academia Nacional de Evaluadores de México.