

P O N E N C I A

Del archivo al activo

Evaluación pública en la era de la inteligencia artificial

Salma Leticia Jalife Villalón

Centro México Digital · CMD

*Seminario Querétaro Avanza Contigo · Semana gLOCAL 2026
2 de junio de 2026*

APERTURA

Un sistema que produce bien, *y circula mal.*

En veinte años, México ha producido uno de los acervos evaluativos más densos de América Latina. Vive mayoritariamente como PDFs en repositorios institucionales: cien páginas por documento, sin metadatos consistentes, sin indexación cruzada.

Cuando un evaluador diseña hoy una evaluación, rara vez consulta sistemáticamente qué se encontró en programas análogos, qué recomendaciones se repitieron, cuáles se atendieron. No por falta de voluntad: por la fricción que impone el formato.

TESIS

Los datos de evaluación dejan de ser un *archivo* para convertirse en un *activo*.

ARCHIVO MUERTO

Algo que se guarda.

Documento terminado, cerrado, esperando a ser consultado uno por uno.

INFORMACION ESTÁTICA, AISLADA Y PASIVA

ACTIVO VIVO

Algo que produce valor.

Capa consultable de inteligencia pública: lectura, síntesis y comparación a escala.

DATOS DINÁMICOS, INTERCONECTADOS Y PREDICTIVOS

No se puede extraer inteligencia de archivos de mala calidad.

CURADURIA Y GOBERNANZA DE DATOS

Estandarización de metadatos

Limpieza de información

Seguridad y privacidad

EL PROCESO

El reto principal es convertir la información no estructurada en un formato que el modelo pueda “razonar”.

INFRAESTRUCTURA TÉCNICA: DE PDF A CONOCIMIENTO

RAG

Retrieval Augmented Generation

Knowledge Graphs

Extracción de entidades y relaciones

EL FACTOR HUMANO

La tecnología por sí sola no genera valor:
Los procesos institucionales si.

CAMBIO CULTURAL Y ORGANIZACIONAL

- **Del "Cumplimiento" a la "Utilidad"**: El funcionario suele ver la evaluación como un requisito de transparencia o fiscalización. Es necesario rediseñar el proceso para que la herramienta de IA sea útil para el diseñador de políticas públicas (ej. un asistente que resuma lecciones aprendidas de 10 años de evaluaciones en segundos).
- **Integridad y "Human-in-the-loop"**: La IA debe ser un copiloto, no un tomador de decisiones autónomo. La capacidad de verificar la fuente (que el sistema indique exactamente de qué página del PDF obtuvo la información) es crucial para generar confianza institucional.

LA ESTRATEGIA DE VALOR

Los archivos cambian de evidencia de cómo se hizo algo a verlos como insumos para diseñar el futuro.

¿QUÉ ES EL ACTIVO?

- **Accesibilidad:** que cualquier funcionario público sin ser programador pueda consultar el archivo histórico.
- **Interoperabilidad:** que el activo de evaluación pueda hablar con los activos de presupuesto o indicadores de desempeño.
- **Capacidad predictiva:** que el sistema no sólo describa qué pasó (archivo), sino que sugiera cómo mitigar riesgos en nuevas intervenciones basándose en hallazgos históricos (activo).



P A R T E

Cuatro usos hoy

Vectorización (La “comprensión” matemática)

Cuatro cosas que un evaluador puede hacer este año con la tecnología disponible. No son escenarios futuros: son aplicaciones en operación.

- 01 Ingesta y pre-procesamiento** (La limpieza)
- 02 Vectorización** (La comprensión matemática)
- 03 Recuperación inteligente**
- 04 Generación y Trazabilidad**

Los archivos institucionales (PDFs) suelen ser complejos (tablas, gráficos, formatos escaneados).

CÓMO FUNCIONA

Segmentación Inteligente (*Chunking*): No puedes alimentar al modelo con un documento de 200 páginas completo. Debes dividir el texto en fragmentos coherentes (párrafos, secciones de hallazgos o recomendaciones).

OCR Avanzado: El uso de herramientas de reconocimiento óptico de caracteres para rescatar información de archivos antiguos que fueron guardados como imágenes.

Enriquecimiento de Metadatos: Cada "fragmento" de texto debe etiquetarse con metadatos (año, dependencia, tipo de programa, presupuesto asignado).

BENEFICIO

Se pre-procesa y clasifica la información, se producen patrones y primeras síntesis

El evaluador revisa y corrige

Se hace un filtrado de la información antes de procesarla



USO 02 · TRES USOS HOY - VECTORIZACION

Para que una IA entienda el contexto de una evaluación debemos convertir el texto en vectores (significado semántico)

MODELOS EMBEBIDOS

Se utilizan modelos de lenguaje para convertir los párrafos e las evaluaciones en vectores.

En lugar de buscar solamente por palabras clave se utilizan conceptos.

BASE DE DATOS VECTORIAL

Se almacenan los vectores en una base de datos especializada (Pinecone, Milvus, Weaviate, etc.).

Esta es la “memoria” del active institucional.

El funcionario comienza a hacer preguntas

Búsqueda semántica

El Sistema en la base vectorial los fragmentos que tienen un significado cercano a esa pregunta.

Re-ranking

Paso intermedio que clasifica los resultados obtenidos para asegurar que los fragmentos más relevantes y recientes ocupen los primeros lugares.

BENEFICIO *Pasar de evaluaciones aisladas a aprendizaje sistémico. La diferencia entre tener mil informes y tener una memoria institucional viva.*

Integridad de la fuente, el sistema debe restringirse a la base de datos oficiales

CONTEXTUALIZACIÓN

El Sistema envía al LLM la pregunta del funcionario + los fragmentos recuperados. El LLM redacta una respuesta coherente basada únicamente en esa información.

CITAS Y REFERENCIAS

Para que se a un acrtivo confinable el Sistema debe citar la Fuente (documento y página donde se encontró)

CONTROL DE ACCESO

EO TODA LA INFORMACIÓN ES PÚBLICA.El Sistema debe integrarse con el Directorio Activo Institucional.

DASHBOARD DE HALLAZGOS

El Sistema debe ser capaz de tabular automáticamente las recomendaciones pendientes de todas las evaluaciones históricas.

BENEFICIO *Pasar de evaluaciones aisladas a aprendizaje sistémico. La diferencia entre tener mil informes y tener una memoria institucional viva.*

Qué no hace la IA

La credibilidad metodológica de esta agenda depende de no sobrevenderla.

— **No verifica fuentes por sí sola**

Puede inventar citas con seguridad aparente. Requiere arquitectura con verificación.

— **No sustituye el juicio metodológico**

Aumenta al evaluador; no reemplaza diseño, muestreo ni interpretación causal.

— **No interpreta**

Sintetiza bien, pero borra matices. El criterio del evaluador es insustituible.

— **No protege datos sensibles por defecto**

Padrones, identificadores y datos personales requieren gobernanza explícita.

II

P A R T E

Tres condiciones para mañana

La diferencia entre un piloto que sorprende y uno que decepciona casi nunca está en el modelo. Está en los datos.

- 01 Datos
- 02 Infraestructura
- 03 Talento

CONDICIÓN 01 · TRES CONDICIONES PARA MAÑANA

Datos

Lo mínimo y lo aspiracional.

MÍNIMO VIABLE

- Formatos legibles por máquina (no PDFs escaneados)
- Metadatos consistentes: programa, año, tipo de evaluación, dependencia, indicadores
- Recomendaciones tipificadas y trazables

ASPIRACIONAL

- Vocabularios controlados compartidos entre dependencias
- Identificadores únicos de programas
- Exposición vía API y gobernanza explícita

La calidad del entregable que se produce hoy determina si será un activo aprovechable mañana.

Infraestructura

Conviene desmitificar.

STACK MÍNIMO VIABLE

Modelos	APIs de modelos comerciales o de código abierto.
Base vectorial	Indexación semántica del acervo para recuperación.
Almacenamiento	Documentos estructurados con metadatos —nube o servidor estándar.

Miles

de pesos al mes, no millones.

El costo de operar este stack para una unidad estatal de evaluación es del orden de miles de pesos mensuales.

La pregunta no es “¿podemos costearlo?” sino “¿bajo qué régimen de gobernanza opera y dónde residen los datos sensibles?”

Talento

El perfil que rinde no es el científico de datos puro.

El evaluador aumentado: analista de políticas públicas

Un evaluador con criterio metodológico que entiende qué le puede pedir a un modelo y qué no, y que sabe leer una respuesta para detectar cuándo el modelo se equivocó con confianza.

ERROR COMÚN Contratar especialistas externos como sustitutos del equipo.

TRANSICIÓN SOSTENIBLE Formar al equipo interno —que ya tiene el conocimiento sustantivo— en alfabetización de IA aplicada a su función.

C I E R R E

Una posición ventajosa, *antes de que se cierre.*

Los gobiernos subnacionales son lo suficientemente grandes para tener acervos significativos y lo suficientemente ágiles para reorganizarlos sin procesos de varios años. Querétaro, con la institucionalidad de evaluación que ya ha construido, puede ser uno de los primeros estados donde el acervo evaluativo opera como infraestructura de inteligencia pública —no como repositorio documental.

*Durante años, el gobierno ha invertido millones de pesos en evaluaciones que terminaron siendo documentos de archivo. Hoy, con la IA, esas evaluaciones dejan de ser un registro del pasado para convertirse en el mapa de ruta del futuro. **No estan haciendo más archivos, estan construyendo el cerebro institucional de Querétaro.***

El acervo siempre estuvo ahí. La pregunta es si lo seguimos tratando como archivo o decidimos convertirlo en activo.