

INNOVACIÓN PÚBLICA · MATERIAL FORMATIVO · SEPTIEMBRE DE 2026

De datos dispersos a decisiones públicas

Minería de datos aplicada a proyectos, políticas públicas y trabajo legislativo: una introducción crítica, práctica y prospectiva para América Latina y Argentina

Dr. Gustavo Martin

Politólogo · Docente universitario · Director de la Revista Nueva Ola · Río Cuarto, Argentina

Resumen. La minería de datos dejó de ser un recurso reservado a laboratorios técnicos. Hoy puede asistir a equipos de proyectos, áreas de gobierno y cuerpos legislativos para ordenar información, identificar patrones, formular hipótesis, monitorear resultados y construir explicaciones más trazables. Este artículo presenta sus fundamentos, límites y aplicaciones en tres campos: proyectos de distinta escala, políticas públicas municipales y nacionales, y minería legislativa. Propone un enfoque latinoamericano: útil ante restricciones de capacidad y calidad de datos, cuidadoso con derechos, y crítico frente a la promesa de automatizar decisiones complejas. El punto de partida es simple: una base no decide; una institución tampoco mejora por tener más tableros. La mejora aparece cuando una pregunta pública bien formulada se conecta con datos pertinentes, método explícito, deliberación y responsabilidad.

Palabras clave: minería de datos; proyectos; políticas públicas; minería legislativa; gobernanza de datos.

Sumario. 1. Conceptos y marcos de trabajo · 2. Diagnóstico y preguntas públicas · 3. Contexto latinoamericano y argentino · 4. Aplicación a proyectos · 5. Políticas públicas · 6. Minería legislativa · 7. Método responsable · 8. Herramientas, riesgos y prospectiva.

Una advertencia inicial

El entusiasmo tecnológico suele invertir el orden correcto. Primero se compra una herramienta, luego se busca una base y, recién al final, se pregunta qué problema resolver. Esa secuencia genera tableros vistosos, informes sin decisión y automatizaciones difíciles de explicar. La minería de datos debe empezar al revés: con un problema delimitado, una decisión concreta, una pregunta verificable y un criterio para juzgar si el resultado será útil.

En las tres áreas estudiadas, la técnica no sustituye política, derecho ni gestión. Puede mejorar el diagnóstico, reducir tiempos de exploración, mostrar relaciones antes invisibles y ordenar un universo documental inmanejable. Pero no determina por sí misma prioridades, distribución de recursos, competencia normativa ni justicia de una decisión. Allí reside su potencial y también su riesgo. Un algoritmo no vuelve pública una decisión: apenas puede volver más visible una parte de su lógica.

Situación editorial: documento de lectura y formación. Pendiente de catalogación.

NOTA DE LECTURA

1. Dato: registro de una observación; información: dato organizado y contextualizado. El paso decisivo es convertirlo en evidencia útil para una decisión concreta.

1. Qué significa minar datos - y qué no

La minería de datos es el conjunto de procedimientos que permite descubrir regularidades, agrupamientos, asociaciones, anomalías o patrones potencialmente significativos en conjuntos de datos. No se limita a usar inteligencia artificial y tampoco es sinónimo de hacer gráficos. Incluye técnicas descriptivas, predictivas y de clasificación, pero sólo adquiere sentido cuando esos resultados se interpretan frente a un problema real.

Su genealogía combina estadística, gestión de bases de datos, aprendizaje automático, investigación operativa y análisis de textos. Durante décadas, las organizaciones acumularon registros más rápido de lo que podían comprenderlos. La idea de descubrimiento de conocimiento en bases de datos apareció para nombrar el proceso más amplio: seleccionar información pertinente, depurarla, transformarla, analizarla, interpretar los hallazgos y decidir qué hacer con ellos. La minería es una etapa central de ese recorrido, pero no puede reemplazar las otras.

En el campo público, el término requiere una precisión adicional. No se trata de extraer datos de la ciudadanía como si fueran un recurso libre de contexto. Se trata de producir evidencia con finalidad pública, bajo reglas de legalidad, minimización, seguridad, transparencia y no discriminación. Un dato sensible o mal integrado puede aumentar la capacidad de clasificación de una institución y, a la vez, deteriorar derechos y confianza.

Tres marcos organizan el trabajo:

Marco	Pregunta que ordena	Aporte
KDD	¿Cómo se pasa de registros dispersos a conocimiento interpretable?	Obliga a mirar selección, limpieza, minería e interpretación como un proceso completo.
CRISP-DM adaptado	¿Cómo se conecta el problema con un procedimiento repetible?	Parte de la comprensión del problema, entiende los datos, prepara, modela, evalúa y comunica/implementa.
Ciclo de política/proyecto	¿Para qué decisión pública sirve el hallazgo?	Vincula diagnóstico, alternativas, diseño, ejecución, monitoreo, evaluación y aprendizaje.

Fuente: elaboración propia.

La adaptación es necesaria: los modelos originales nacieron en contextos empresariales. En política pública, el éxito no se mide sólo por precisión o ahorro de tiempo. Debe considerar cobertura, equidad, explicabilidad, costo institucional, cumplimiento de derechos y posibilidad de corregir un error.

Conviene separar tres operaciones que suelen confundirse. El análisis exploratorio describe distribuciones, tendencias o valores atípicos; la predicción estima un resultado no observado a partir de patrones previos; la inferencia causal pregunta qué habría ocurrido sin una intervención y exige un diseño comparativo mucho más exigente. Un tablero puede describir una brecha. Un modelo puede anticipar una demanda. Ninguno, por sí solo, demuestra que una política produjo un efecto.

Regla básica: una correlación puede ser una pista útil; no demuestra, por sí sola, que una medida haya producido un resultado. Confundir asociación con causalidad es uno de los errores más costosos en gestión pública.

NOTA DE LECTURA

2. KDD (Knowledge Discovery in Databases): proceso completo de seleccionar, preparar, analizar e interpretar datos. La minería es una etapa, no el proceso entero.

2. El problema público antes que la base

La calidad de un análisis comienza antes de abrir una planilla. Una formulación útil debe expresar quién está afectado, qué ocurre, dónde, desde cuándo, con qué intensidad y qué decisión puede modificarse. “Analizar la inseguridad” es demasiado amplio. “Identificar franjas horarias, zonas y modalidades de reclamos que permitan reasignar preventivamente recursos municipales sin estigmatizar barrios ni personas” ya orienta variables, fuentes y límites.

En proyectos, políticas y normas, las preguntas analíticas cumplen funciones diferentes. En un proyecto ayudan a diagnosticar demanda, costos, tiempos y resultados. En una política permiten monitorear cobertura, brechas, acceso y efectos no previstos. En una legislatura permiten organizar antecedentes, trayectorias de expedientes, temas, modificaciones y relaciones entre normas. La técnica puede ser parecida; la unidad de análisis y la decisión cambian.

Una matriz mínima evita trabajar a ciegas:

Elemento	Pregunta de control	Ejemplo
Decisión	¿Qué se decidirá o revisará?	Priorizar la ampliación de un programa de apoyo escolar.
Unidad	¿Qué representa cada fila?	Estudiante, escuela, barrio, expediente o norma.
Variables	¿Qué atributos son relevantes?	Asistencia, distancia, beca, turno, período y escuela.
Criterio	¿Cómo se reconocerá un hallazgo útil?	Brecha persistente y accionable, no sólo diferencia estadística.
Salvaguarda	¿Qué no se hará con los datos?	No publicar identificadores ni automatizar sanciones.

Fuente: elaboración propia.

El diagnóstico no debe confundirse con un listado de carencias. Es una explicación provisoria de un problema y sus determinantes. Por eso necesita combinar datos cuantitativos, saber territorial, normativa, entrevistas y observación. Un patrón estadístico puede indicar dónde mirar; no reemplaza la experiencia de quienes viven o ejecutan una política.

En América Latina, una dificultad recurrente es la fragmentación institucional. Registros administrativos creados para tramitar prestaciones rara vez nacen con una arquitectura para evaluar políticas. Tienen faltantes, categorías cambiantes, duplicaciones y sesgos de cobertura. Eso no los vuelve inútiles: obliga a documentar sus límites, cruzarlos sólo cuando exista base legal y evitar afirmaciones más fuertes que la evidencia disponible.

Principio operativo: si una decisión no puede describirse antes del análisis, probablemente tampoco podrá justificarse después de él.

NOTA DE LECTURA

3. Variable: atributo que puede asumir valores distintos, por ejemplo barrio, monto ejecutado o edad. Una base sin diccionario de variables no es auditable.

3. Diagnóstico latinoamericano y argentino a septiembre de 2026

La región reúne capacidades avanzadas de estadística, datos abiertos y gobierno digital con brechas severas de infraestructura, interoperabilidad, financiamiento y formación. La expansión de herramientas generativas redujo barreras para consultar, resumir y visualizar información, pero también incrementó la circulación de respuestas no verificadas. El desafío ya no es sólo abrir conjuntos de datos: es volverlos comprensibles, comparables, actualizados y socialmente responsables.

Argentina dispone de activos relevantes: estadísticas públicas, portales de datos, registros sectoriales, información presupuestaria, documentos normativos y repositorios académicos. En materia legislativa, la Base InfoLEG/SAIJ ofrece conjuntos descargables de normativa nacional y complementos sobre normas modificadas o modificatorias. El recurso es valioso para exploración y construcción de mapas normativos, pero no reemplaza la verificación del texto vigente, su reglamentación ni la competencia de cada jurisdicción.

Cifras de contexto verificadas. La edición 2026 del índice de gobierno digital de la OCDE registró una mejora del componente de sector público orientado por datos, desde 0,63 en 2023 hasta 0,74 en 2025; el mismo estudio advirtió que la medición y el monitoreo continúan rezagados. En la encuesta de 36 países de la OCDE, la IA se usaba en al menos un área de gobierno en 35 países, pero sólo 14 exigían evaluación de riesgos previa al despliegue y 11 realizaban auditorías posteriores. Estos valores describen ese universo institucional, no son un indicador de Argentina ni de América Latina. La base InfoLEG, por su parte, reportó actualización mensual y datos al 31 de agosto de 2026. Las fuentes, fecha de corte y enlaces de verificación se consignan al final.

A nivel municipal, el panorama es más desigual. Existen ciudades con portales, mapas y sistemas de reclamos; otras conservan información en planillas heterogéneas o expedientes no estructurados. La dificultad no es únicamente tecnológica. También intervienen rotación de equipos, ausencia de diccionarios de datos, escasa continuidad, contratos opacos y falta de una autoridad clara que responda por la calidad del dato.

Cuatro tensiones atraviesan el presente:

Tensión	Riesgo si se ignora	Respuesta institucional
Apertura vs. protección	Exponer personas, domicilios o trayectorias individuales.	Anonimización, minimización y revisión de reidentificación.
Integración vs. finalidad	Cruzar registros para usos incompatibles con su origen.	Base legal, propósito explícito, control de accesos y registro de usos.
Velocidad vs. calidad	Decidir con datos atrasados, incompletos o no comparables.	Metadatos, pruebas de calidad y fechas de actualización visibles.
Automatización vs. explicación	Delegar decisiones sensibles en sistemas opacos.	Supervisión humana, trazabilidad y mecanismos de revisión.

Fuente: elaboración propia.

La Ley 25.326 de protección de datos personales y la Ley 27.275 de acceso a la información pública ofrecen un piso relevante para pensar derechos, aunque la práctica exige además criterios institucionales específicos. El dato público no es automáticamente dato reutilizable sin condiciones; el dato disponible no es automáticamente dato apto para inferir sobre personas o territorios.

NOTA DE LECTURA

4. Indicador: medida que aproxima un fenómeno. No equivale al fenómeno mismo y debe leerse junto con su definición, fuente, período y limitaciones.

4. Minería de datos aplicada a proyectos

Un proyecto es una intervención organizada para transformar una situación mediante recursos, actividades, productos y resultados esperados. La minería de datos puede fortalecer cada etapa, siempre que no reemplace la teoría de cambio: esa explicación de por qué ciertas actividades deberían producir resultados en una población y contexto determinados.

En la formulación, permite estimar demanda, segmentar población objetivo, reconocer estacionalidad, georreferenciar brechas y detectar restricciones. En la ejecución, ayuda a seguir avances físicos y financieros, advertir desvíos y priorizar asistencia. En la evaluación, permite comparar trayectorias, explorar heterogeneidad de resultados y producir preguntas para una indagación más rigurosa.

Fase	Uso analítico	Decisión que puede mejorar
Diagnóstico	Cruces, series, mapas, agrupamientos.	Definir población, alcance, línea de base y prioridades.
Diseño	Escenarios de costo, demanda y capacidad.	Elegir alternativa, cronograma, presupuesto y riesgos.
Implementación	Alertas de retraso, cobertura y ejecución.	Reasignar recursos o corregir procesos.
Evaluación	Comparaciones, segmentación y análisis de resultados.	Continuar, rediseñar, escalar o cerrar la intervención.

Fuente: elaboración propia.

Ejercicio didáctico. Una universidad quiere implementar una especialización en geriatría y gerontología. En lugar de suponer demanda, puede combinar matrícula histórica de carreras afines, egresados potenciales, oferta regional, distancia de traslado, encuestas de intención, capacidades docentes y costos. La minería no decide si la carrera es pertinente; permite dejar visibles supuestos, segmentos y escenarios. Si la población potencial está dispersa, el hallazgo puede orientar una modalidad híbrida o convenios territoriales.

El límite metodológico es claro: los datos de interés no sustituyen una evaluación de factibilidad. Un proyecto viable necesita consistencia entre problema, objetivos, actividades, presupuesto, recursos, normativa, capacidades, riesgos e indicadores. La minería aporta evidencia; no reemplaza el estudio técnico, legal, organizacional, ambiental, económico-financiero o social que corresponda.

Un error frecuente consiste en perseguir predicción sofisticada con registros pobres. Es preferible un tablero sencillo, con indicador definido, fuente conocida, responsable y frecuencia de actualización, antes que un modelo imposible de explicar o mantener.

NOTA DE LECTURA

5. Interoperabilidad: capacidad de sistemas y organizaciones para intercambiar y reutilizar datos con reglas técnicas, semánticas, jurídicas y de gobernanza compartidas.

5. Políticas públicas municipales y nacionales

Una política pública no es una planilla de indicadores: es una respuesta estatal frente a un problema que involucra valores, recursos, instituciones y actores. La minería de datos sirve cuando se integra al ciclo de la política y no cuando opera como un atajo tecnocrático. Su mayor aporte es reducir incertidumbre de manera explícita, no suprimir la deliberación democrática.

En municipios, los usos más inmediatos aparecen en gestión de reclamos, movilidad, residuos, mantenimiento urbano, salud primaria, cultura, seguridad vial, habilitaciones, desarrollo local y ejecución presupuestaria. La proximidad ofrece ventaja: los equipos conocen el territorio y pueden contrastar el patrón con experiencia cotidiana. También aumenta riesgo de reidentificación, porque pequeños conjuntos de datos permiten reconocer personas o familias con facilidad.

A escala nacional o provincial, el valor se amplía hacia programas con gran cobertura, registros administrativos y series temporales. Allí son frecuentes los desafíos de coordinación federal, compatibilidad de definiciones y heterogeneidad territorial. Un promedio nacional puede ocultar que una medida funciona en una provincia, es irrelevante en otra y deja sin cobertura a un grupo específico.

Ejemplo práctico de gestión local. Un municipio observa reclamos reiterados por microbasurales. El análisis puede reunir fecha, ubicación aproximada, tipo de residuo, frecuencia de recolección, cercanía a puntos limpios, denuncias, costos de limpieza y campañas realizadas. Un mapa de calor puede orientar recorridos y comunicación. Pero no debe derivar en vigilancia indebida ni en imputar responsabilidades individuales a partir de patrones territoriales.

Ejemplo práctico de política nacional. Un programa de becas puede analizar tiempos de tramitación, renovación, abandono, modalidad educativa, región y condiciones de acceso. La pregunta responsable no es “qué estudiante fallará”, sino “qué barreras institucionales y territoriales se asocian con trayectorias interrumpidas y qué ajustes son factibles sin discriminar ni automatizar exclusiones”. Cambiar la pregunta cambia la ética y el diseño.

Para ambos niveles, la secuencia aconsejable es: problema y decisión; inventario de fuentes; evaluación de calidad; análisis exploratorio; contraste con actores; hipótesis y alternativas; decisión documentada; monitoreo de efectos y corrección. La última etapa importa: una política que no aprende reproduce sus propios errores con mejores gráficos.

NOTA DE LECTURA

6. Sesgo: distorsión sistemática que aparece por cómo se recogen, clasifican, procesan o interpretan los datos. Más datos no corrigen por sí solos un sesgo de origen.

6. Minería legislativa: textos, trayectorias y calidad normativa

La minería legislativa aplica métodos de análisis de datos y textos al trabajo parlamentario, normativo y documental. Puede analizar proyectos, leyes, ordenanzas, dictámenes, sesiones, votaciones, expedientes, presupuestos y relaciones entre normas. No sustituye la interpretación jurídica ni la decisión política. Su función es organizar un campo documental complejo, mostrar patrones, recuperar antecedentes y hacer más trazable la revisión.

La unidad de análisis no siempre es la norma completa. Puede ser un artículo, una modificación, una remisión, un expediente, una intervención en sesión, una comisión, un autor o un período. La elección es decisiva: comparar leyes por título ofrece una imagen pobre; comparar artículos, objetos, sujetos alcanzados, autoridades de aplicación y remisiones abre un análisis más útil, aunque exige mejor preparación de texto.

Objeto de análisis	Preguntas posibles	Producto útil
Normas vigentes	¿Qué regula, modifica, deroga o remite?	Mapa temático y de relaciones normativas.
Proyectos/trámite	¿Qué temas, autores, comisiones y tiempos aparecen?	Tablero de actividad y cuellos de botella.
Textos normativos	¿Qué conceptos, obligaciones y ambigüedades se repiten?	Matriz comparada y alertas de revisión.
Presupuestos/planes	¿Qué autorizaciones, montos y metas se conectan?	Seguimiento de coherencia entre norma y ejecución.

Fuente: elaboración propia.

Argentina cuenta con fuentes oficiales para normativa nacional y diversos portales legislativos. No se identificó, en cambio, una base oficial nacional única, exhaustiva y homogénea de ordenanzas municipales. Por eso, una minería seria debe declarar cobertura, jurisdicción, período, criterios de inclusión y límites. Una muestra de ciudades grandes no representa a todos los municipios; una búsqueda web no garantiza vigencia; una norma publicada no equivale a una política implementada.

La minería de texto puede extraer términos, temas, similitud entre documentos, entidades, fechas y referencias. La minería de relaciones puede reconstruir modificaciones, derogaciones y remisiones. Estas operaciones reducen horas de lectura inicial, pero generan falsos positivos: una coincidencia textual no prueba identidad jurídica y una omisión automática puede ocultar una cláusula decisiva.

Por eso la revisión legislativa debe seguir una cadena: localizar; verificar fuente y versión; contextualizar competencia y jerarquía; analizar técnica, coherencia y efectos; y recién entonces formular una recomendación. La herramienta señala; el equipo responsable interpreta y decide.

NOTA DE LECTURA

7. Modelo: representación simplificada de una relación. Un modelo predictivo estima probabilidades; no prueba causalidad ni autoriza decisiones automáticas sobre derechos.

7. Método reproducible para una minería legislativa responsable

Una práctica robusta no comienza con una consulta en lenguaje natural ni termina con un resumen automático. Necesita un procedimiento que otra persona pueda reconstruir. La metodología VETEL aporta una orientación útil para esa exigencia: doble lectura, localización de hallazgos, distinción entre error formal y problema sustantivo, contraste con marco aplicable y advertencia de que la asistencia técnica nunca reemplaza revisión humana competente.

Paso	Acción	Control de calidad
1. Delimitar	Definir pregunta, jurisdicción, período y corpus.	Evitar mezclar proyectos, normas vigentes y documentos no comparables.
2. Validar	Comprobar legibilidad, integridad y tipo documental.	Registrar archivos excluidos y motivo.
3. Estructurar	Extraer metadatos y texto; normalizar fechas, tipos y campos.	Mantener original y versión procesada.
4. Analizar	Buscar patrones, temas, similitudes, relaciones y anomalías.	No confundir resultados automáticos con conclusiones.
5. Verificar	Realizar segunda lectura de hallazgos relevantes.	Ubicar artículo, página, cláusula o fuente.
6. Comunicar	Separar datos observados, inferencias y recomendaciones.	Declarar limitaciones, cobertura y necesidad de revisión.

Fuente: elaboración propia.

Ejemplo basado en fuente real. La Base InfoLEG incluye documentos normativos nacionales y conjuntos complementarios de normas modificadas y modificatorias. Una práctica inicial puede seleccionar un tema, descargar el conjunto pertinente, filtrar por tipo de instrumento y período, y construir una tabla que relacione norma, organismo, fecha, palabras clave y estado de modificación. Luego debe volver al texto oficial para revisar vigencia y alcance. Ese retorno al original no es un detalle: es la frontera entre exploración documental y afirmación jurídica.

Las alertas más útiles no son veredictos tajantes sino preguntas técnicas: ¿la autoridad de aplicación tiene competencia? ¿se asignan funciones sin recursos? ¿las remisiones existen y son coherentes? ¿hay definiciones contradictorias? ¿la norma distingue objeto, sujetos, procedimiento, plazos y consecuencias? ¿se preservan igualdad, debido proceso, protección de datos y acceso a la información?

El análisis de calidad normativa debe ser prudente. Un proyecto técnicamente débil puede responder a una demanda social legítima; su corrección no exige descartarlo, sino mejorar su arquitectura, competencia, implementación y garantías.

NOTA DE LECTURA

8. Dato personal: información de cualquier tipo referida a personas determinadas o determinables. Su tratamiento exige finalidad legítima, seguridad y proporcionalidad.

8. Herramientas, técnicas y decisiones de uso

No existe una única herramienta correcta. La elección depende del volumen, el tipo de dato, la sensibilidad, las capacidades del equipo y el resultado esperado. Para una primera cohorte de trabajo, una planilla bien diseñada y un tablero simple suelen ser superiores a un proyecto de inteligencia artificial sin mantenimiento. La sofisticación debe ser proporcional al problema.

Nivel	Herramientas típicas	Uso adecuado	Precaución
Inicial	Excel/Sheets, filtros, tablas dinámicas, mapas básicos.	Ordenar, depurar, cruzar y visualizar conjuntos acotados.	Controlar fórmulas, versiones y categorías.
Intermedio	SQL, R/Python, BI, georreferenciación.	Integrar fuentes, reproducir análisis, automatizar reportes.	Documentar código, permisos y actualización.
Texto	OCR, PLN, búsqueda semántica, grafos.	Explorar corpus extensos y recuperar relaciones.	Verificar contexto, idioma, citas y vigencia.
IA generativa	Asistentes de consulta, resumen y clasificación preliminar.	Acelerar exploración y producir borradores de trabajo.	No cargar datos sensibles ni aceptar resultados sin contraste.

Fuente: elaboración propia.

El aprendizaje automático puede clasificar expedientes, detectar grupos o estimar probabilidades. Sin embargo, su desempeño depende de etiquetas consistentes, datos suficientes y criterios de validación. Un modelo entrenado con decisiones históricas puede reproducir inequidades pasadas. En cuestiones de derechos, la pregunta no debe ser sólo “¿funciona?”, sino “¿qué hace, sobre quién, con qué error, quién puede revisarlo y qué daño produce si falla?”.

Las herramientas de IA generativa merecen una regla adicional: sirven para explorar y redactar hipótesis, no para inventar fuentes, afirmar vigencia normativa ni evaluar personas. La interfaz conversacional reduce fricción, pero también oculta transformaciones, parámetros y errores. Por eso conviene guardar la consulta, la fecha, la fuente de los datos, el resultado inicial, las correcciones y la decisión final.

En instituciones con recursos limitados, el salto más importante no es técnico: es crear un diccionario de variables, nomenclatura compartida, permisos de acceso, responsable de actualización y protocolo para corregir datos. Sin esa infraestructura mínima, cualquier tablero se vuelve una fotografía desactualizada.

NOTA DE LECTURA

9. Procesamiento de lenguaje natural (PLN): técnicas para estructurar, buscar, clasificar o resumir textos. Sus resultados son hipótesis de lectura, no dictámenes jurídicos.

9. Riesgos, derechos y gobernanza

Una política basada en datos puede ser más justa o más discriminatoria. Todo depende de su diseño. Clasificar no es neutral cuando condiciona acceso a prestaciones, controles, sanciones o prioridades. Una variable aparentemente inocua - domicilio, horario de conexión, tipo de dispositivo - puede operar como sustituto de condición socioeconómica, género, origen o discapacidad. Por eso la ética no es una etapa final de cumplimiento: es un criterio de diseño.

La gobernanza de datos define quién puede acceder, para qué finalidad, con qué reglas de seguridad, bajo qué plazo de conservación, cómo se registran cambios y ante quién se responde. Es la parte menos visible y más importante. Sin ella, la institución depende de archivos personales, automatizaciones sin responsable y decisiones imposibles de auditar.

Un protocolo mínimo debe incluir:

Regla	Aplicación concreta
Finalidad	Describir el problema y prohibir reutilizaciones incompatibles.
Necesidad	Pedir y conservar sólo variables necesarias para el objetivo.
Calidad	Registrar origen, cobertura, fecha, errores conocidos y actualización.
Seguridad	Definir roles, accesos, copias, cifrado y respuesta ante incidentes.
Explicación	Poder comunicar en lenguaje claro cómo se produjo un hallazgo.
Revisión	Ofrecer supervisión humana y canales de corrección/impugnación cuando corresponda.
Evaluación	Medir errores, impactos distributivos y efectos no previstos después de implementar.

Fuente: elaboración propia.

En materia legislativa, el riesgo adicional es la falsa certeza jurídica. Los sistemas pueden hallar textos similares, proponer redacción o detectar inconsistencias aparentes. No pueden sustituir la verificación de competencia, jerarquía, vigencia, procedimiento parlamentario, interpretación constitucional o deliberación democrática. Presentar una salida automática como “dictamen” o “certificación” sería metodológicamente impropio.

El criterio más sólido es el de asistencia responsable: automatizar tareas repetitivas, elevar la calidad de la primera revisión y reservar la decisión para personas identificables, con competencia y obligación de fundamentar. La automatización debe aumentar la responsabilidad, no diluirla.

NOTA DE LECTURA

10. Remisión normativa: referencia de una norma a otra disposición. Debe comprobarse su existencia, vigencia, numeración y sentido jurídico, no sólo su coincidencia de palabras.

10. Tendencias actuales y escenarios prospectivos

A septiembre de 2026, cuatro tendencias ya modifican el trabajo con datos públicos: la expansión de IA generativa en tareas de oficina; el uso de búsqueda semántica y recuperación aumentada para consultar corpus extensos; el avance de grafos de conocimiento que representan relaciones entre instituciones, normas y programas; y la presión por reglas de gobernanza, transparencia y evaluación de riesgos. La promesa no es una administración sin personas, sino capacidades públicas más informadas y auditables.

La prospectiva no permite adivinar el futuro. Permite ensayar futuros plausibles y tomar mejores decisiones ahora. Para 2027-2030 se pueden pensar tres escenarios:

Escenario	Rasgos	Decisión presente
Capacidad pública gobernada	Datos interoperables, equipos formados, reglas claras y auditorías periódicas.	Invertir en calidad, formación, repositorios y protocolos antes que en soluciones aisladas.
Adopción fragmentada	Muchas herramientas, poca coordinación; datos duplicados y dependencia de proveedores.	Definir estándares comunes, responsables y cláusulas de portabilidad.
Automatización opaca	Sistemas que clasifican o recomiendan sin explicación ni reparación efectiva.	Establecer supervisión humana, registros de uso, pruebas de sesgo y límites de decisión.

Fuente: elaboración propia.

La tendencia más importante no es técnica sino institucional: pasar de experimentar con herramientas a diseñar capacidades permanentes. Eso supone carreras y formación para equipos públicos, repositorios de casos, compras públicas que exijan transparencia, protocolos de prueba y una cultura que permita corregir un modelo sin ocultar el error.

Para la minería legislativa, es razonable esperar más corpus digitalizados, mejoras en OCR, comparaciones entre versiones, grafos de remisiones y asistentes que ayuden a explorar antecedentes. El riesgo será confundir accesibilidad con confiabilidad. Cuanto más fácil sea producir un resumen normativo, más valioso será el trabajo de verificar fuente, versión, contexto y consecuencia jurídica.

La pregunta prospectiva decisiva no es “¿qué podrá hacer la IA?”, sino “¿qué instituciones queremos construir para usarla sin perder legalidad, igualdad, privacidad y capacidad de decisión pública?”.

NOTA DE LECTURA

11. Trazabilidad: posibilidad de reconstruir con claridad fuentes, transformaciones, criterios, responsables y límites de un resultado analítico.

11. Ruta de implementación en noventa días

Para equipos que quieran comenzar sin sobredimensionar el proyecto, conviene trabajar con un caso acotado y una decisión real. La secuencia siguiente es aplicable a una organización, municipio, programa o área legislativa.

Período	Objetivo	Producto verificable
Días 1-30	Delimitar problema, decisión, fuentes y responsabilidades.	Ficha de caso, inventario de datos, matriz de riesgos y diccionario inicial.
Días 31-60	Depurar una base piloto y realizar análisis exploratorio.	Base documentada, tablero básico, hallazgos y lista de dudas a validar.
Días 61-90	Contrastar hallazgos, decidir acción y diseñar monitoreo.	Informe ejecutivo, decisión fundada, indicadores y protocolo de actualización.

Fuente: elaboración propia.

En proyectos, la prueba puede concentrarse en seguimiento de actividades y presupuesto. En políticas públicas, en una brecha concreta de cobertura o tiempos de acceso. En minería legislativa, en un corpus delimitado de proyectos o normas sobre una materia. La escala piloto permite detectar campos faltantes, definiciones incompatibles y decisiones que todavía no están suficientemente claras.

Un buen informe final distingue cuatro planos: dato observado (qué aparece en la base), inferencia (qué interpretación razonable se propone), recomendación (qué acción se sugiere) y límite (qué no puede concluirse). Esta separación evita que un gráfico se convierta, sin justificación, en una decisión.

También conviene institucionalizar un comité pequeño de revisión: persona responsable del área, perfil técnico, referente jurídico o de datos personales cuando corresponda, y una voz territorial o de usuarios. No es burocracia innecesaria. Es una forma de impedir que la interpretación quede capturada por quien conoce la herramienta pero desconoce el problema.

Checklist final: ¿la pregunta sigue vigente?, ¿la base representa lo que se dice que representa?, ¿se documentaron transformaciones?, ¿el resultado puede explicarse?, ¿hay riesgo de daño o discriminación?, ¿la decisión conserva responsable humano?, ¿se definió cómo medir efectos y corregir? Si la respuesta es no, todavía no corresponde automatizar.

NOTA DE LECTURA

12. Prospectiva: exploración sistemática de futuros plausibles para mejorar decisiones presentes. No es adivinación ni predicción lineal.

Conclusión

La minería de datos puede mejorar el Estado, los proyectos y el trabajo legislativo cuando se la trata como una práctica de conocimiento situado. Su valor no reside en convertir toda decisión en cálculo, sino en hacer visibles patrones, incertidumbres, brechas y alternativas que de otro modo quedarían ocultas. El mejor análisis no es el que produce más cifras, sino el que ayuda a formular una decisión más justa, fundada y revisable.

Para América Latina y Argentina, el desafío no es copiar modelos de alta capacidad institucional. Es construir una trayectoria propia: con datos abiertos cuando sea posible, protección efectiva cuando sea necesario, herramientas proporcionales, formación interdisciplinaria, perspectivas territoriales y control democrático. La técnica legislativa, la evaluación de proyectos y el análisis de políticas públicas tienen una exigencia común: no presentar como neutral lo que siempre contiene opciones sobre fines, recursos, derechos y poder.

La consigna final es menos espectacular que la promesa de una máquina que decide, pero mucho más útil: mejor pregunta, mejor dato, mejor método, mejor explicación y mejor responsabilidad.

Nota de integridad editorial. El texto fue revisado en dos lecturas: la primera controló coherencia conceptual, alcance de los ejemplos y consistencia entre proyectos, políticas y normas; la segunda contrastó cifras, fuentes institucionales y formulaciones que podían inducir una certeza excesiva. Los casos didácticos ilustran procedimientos posibles; no son evaluaciones de impacto, dictámenes jurídicos ni recomendaciones automáticas para personas o territorios.

Fuentes y materiales de consulta

No se incluyen citas intercaladas en el cuerpo para favorecer la lectura formativa. Las fuentes siguientes orientan la verificación, actualización y ampliación del contenido.

1. Argentina. Ley 25.326, Protección de los Datos Personales. Texto actualizado disponible en Argentina.gob.ar / SAIJ.
2. Argentina. Ley 27.275, Derecho de Acceso a la Información Pública. Texto actualizado disponible en Argentina.gob.ar / SAIJ.
3. Ministerio de Justicia de la Nación - SAIJ. Base InfoLEG de Normativa Nacional y bases complementarias de normas modificadas y modificaciones. Actualización mensual consultada en septiembre de 2026.
<https://datos.jus.gob.ar/es/dataset/base-de-datos-legislativos-infoleg>
4. CEPAL. Agenda Digital para América Latina y el Caribe (eLAC), observatorios y materiales sobre desarrollo digital, datos abiertos e inteligencia artificial. <https://desarrollodigital.cepal.org/>
5. OCDE. Digital Government Outlook 2026, capítulo sobre adopción y gobernanza de IA en gobierno. Cifras de 35/36 países, 14/36 evaluaciones previas y 11/36 auditorías posteriores verificadas contra el informe y su resumen oficial.
https://www.oecd.org/en/publications/digital-government-outlook_0496b2bc-en.html
6. OCDE. Governing with Artificial Intelligence: The State of Play and Way Forward in Core Government Functions, 2025. Referencia para adopción de IA en procesos y servicios públicos. <https://doi.org/10.1787/795de142-en>
7. BID. Curso y materiales: Datos para la efectividad de las políticas públicas.
<https://knowledge.iadb.org/es/curso/datos-para-la-efectividad-de-las-politicas-publicas>
8. INAP Argentina. Materiales de capacitación en analítica de datos para la gestión pública y minería de texto.
<https://capacitacion.inap.gob.ar/>
9. OECD. Recomendación del Consejo sobre Inteligencia Artificial (OECD/LEGAL/0449), para principios de IA confiable.
<https://legalinstruments.oecd.org/en/instruments/OECD-LEGAL-0449>
10. Naciones Unidas - UNESCO. Recomendación sobre la Ética de la Inteligencia Artificial, 2021.
<https://unesdoc.unesco.org/ark:/48223/pf0000381137>

Contacto: martingustavo@hum.unrc.edu.ar · revistanuevaola@rinconpolitico.com · WhatsApp +54 358 4332541 · revistanuevaola.ar · rinconpoliticoconsultora.ar

NOTA DE LECTURA

13. Reproducibilidad: posibilidad de que otra persona repita el procedimiento con los mismos insumos y obtenga resultados equivalentes, o explique por qué difieren.