

Generador Inteligente de Extractos del Boletín Oficial

Silvana Rica ¹[0009-0006-3570-1298], Mauro Escobar ¹[0009-0008-7607-9436] y
Gabriela Calabrese¹[0009-0000-6757-4437]

¹Dirección Provincial de Informática

Subsecretaría Técnica, Administrativa y Legal

Ministerio de Economía de la Provincia de Buenos Aires, Argentina

srica@ec.gba.gov.ar, mdescobar@ec.gba.gov.ar, gcalabrese@ec.gba.gov.ar

Resumen. La creciente incorporación de tecnologías emergentes en la gestión pública ha intensificado la demanda de información estratégica en tiempo real por parte de los responsables de la toma de decisiones. En particular, en el ámbito normativo, el acceso oportuno a la publicación de leyes, decretos y resoluciones resulta crítico para el diseño, seguimiento y ajuste de políticas públicas.

A fin de satisfacer esta necesidad, en el Ministerio de Economía de la Provincia de Buenos Aires, se conformó un equipo legal especializado encargado de la lectura diaria del Boletín Oficial, la interpretación de las normas y la elaboración de informes ejecutivos de alto nivel. Esta dinámica no solo implicaba una elevada carga operativa manual, sino también tensiones entre la exhaustividad del análisis y la disponibilidad temprana de los reportes.

Con el fin de optimizar la operatoria, se desarrolló el **Generador Inteligente de Extractos del Boletín Oficial**, un asistente digital cognitivo entrenado para el manejo de información desestructurada escrita en lenguaje natural, que combina procesamiento automático y revisión humana calificada.

La solución automatiza la identificación, clasificación y generación de informes sobre decretos, resoluciones y demás actos administrativos publicados, incluyendo resúmenes y enlaces a las fuentes originales. Incorpora un esquema de validación humana que hace posible la revisión y ajuste de los resultados, habilitando un proceso de aprendizaje continuo. A través de su uso sostenido y supervisado, se adapta progresivamente a criterios específicos de análisis y redacción, incrementando su precisión y desempeño.

Esta experiencia propuso repensar las formas de abordar los trabajos y evidenció mejoras sustantivas en la eficiencia operativa, facilitando un acceso más ágil y confiable a los contenidos, a la vez que habilitó la orientación del talento experto hacia tareas de mayor valor analítico.

Palabras Claves: Inteligencia Artificial; Automatización; Procesamiento de lenguaje natural; *Human-in-the-loop*.

Intelligent Extract Generator for the Official Gazette

Abstract. The increasing incorporation of emerging technologies in public management has intensified the demand for real-time strategic information by decision-makers. In particular, within the regulatory domain, timely access to the publication of laws, decrees, and resolutions is critical for the design, monitoring, and adjustment of public policies.

In order to meet this need, within the Ministry of Economy of the Province of Buenos Aires, a specialized legal team was established, responsible for the daily reading of the Official Gazette, the interpretation of regulations, and the preparation of high-level executive reports. This dynamic not only implied a high manual operational workload, but also tensions between the exhaustiveness of the analysis and the early availability of reports.

In order to optimize operations, the Intelligent **Extract Generator for the Official Gazette** was developed, a cognitive digital assistant trained to handle unstructured information written in natural language that combines automated processing and qualified human review.

The solution automates the identification, classification, and generation of reports on decrees, resolutions, and other published regulations, including summaries and links to the original sources. It incorporates a human validation scheme that allows the review and adjustment of the results, enabling a continuous learning process. Through sustained and supervised use, it progressively adapts to specific analysis and writing criteria, improving its accuracy and performance.

This experience proposed rethinking the ways of approaching work and demonstrated substantial improvements in operational efficiency, facilitating more agile and reliable access to regulatory content, while enabling expert talent to be oriented toward higher-value analytical tasks.

Keywords: Artificial Intelligence; Automation; Natural Language Processing; Human-in-the-loop.

1. Introducción

La transformación digital del sector público ha avanzado de manera sostenida en sistemas transaccionales, gestión de expedientes, portales de servicios y trámites. No obstante, persiste un conjunto de tareas institucionales cuya complejidad no reside en la captura de datos, sino en la necesidad de procesar oportunamente documentos extensos, heterogéneos y con redacción no estructurada. Entre ellas se encuentra la revisión diaria de publicaciones oficiales.

El Boletín Oficial de la Provincia de Buenos Aires constituye una fuente primaria para el seguimiento jurídico-administrativo del accionar estatal. Su consulta permite identificar piezas de interés para altos funcionarios de organismos del Sector Público que requieren conocer con inmediatez actos administrativos publicados, evaluar su impacto y sintetizar la información para la estructuración de las políticas públicas. En la práctica, esta labor recae en equipos técnicos legales que deben leer grandes volúmenes de documentación, identificar cuáles son relevantes para la organización y elaborar informes ejecutivos en plazos muy reducidos.

En el caso del Ministerio de Economía de la Provincia de Buenos Aires, la confección diaria de extractos normativos implica localizar decretos y resoluciones vinculadas a las máximas autoridades identificando las firmas de los funcionarios, producir un breve resumen y distribuir el resultado a primera hora. Esta labor, central para la gestión, presentaba una marcada dependencia de procedimientos manuales.

En este contexto, se desarrolló el **Generador Inteligente de Extractos del Boletín Oficial**, orientado a automatizar el relevamiento, filtrado, segmentación, resumen, almacenamiento y presentación de las publicaciones oficiales. A diferencia de un esquema puramente experimental, la solución fue concebida desde el inicio como una aplicación utilizable en operación cotidiana, con módulos específicos para cubrir las distintas etapas y con una integración explícita entre automatización, inteligencia artificial, almacenamiento en la nube y revisión humana. El enfoque híbrido reposiciona el rol del trabajo experto, minimizando su esfuerzo al aseguramiento de calidad.

2. La potencia de la automatización y la IA combinada con la validación humana y el entrenamiento continuo

2.1. La generación manual de extractos

El informe diario se elaboraba a partir de una operatoria artesanal. Un equipo de abogados y profesionales especializados revisaba la edición del Boletín Oficial publicada durante la madrugada, identificaba las resoluciones y decretos relevantes para cada cartera ministerial y preparaba un informe con listado de normas, breve resumen y enlace a la publicación oficial.

Este procedimiento evidenciaba diversas limitaciones: exigía disponibilidad inmediata de recursos humanos calificados al inicio de la jornada, imponía tareas repetitivas sobre perfiles de alto valor analítico y generaba riesgos de inconsistencias de formato y diferencias de criterio en la redacción.

La identificación de las piezas de interés requería combinar lectura del texto, reconocimiento de firmantes, conocimiento de cargos y capacidad de asociar cada acto con un ministerio u organismo determinado. En consecuencia, no bastaba con descargar documentos: era necesario interpretar su estructura y aplicar criterios institucionales de clasificación. Desde el punto de vista organizacional, suponía un uso ineficiente del capital profesional disponible. Gran parte del tiempo de especialistas se absorbía en tareas mecánicas de localización, segmentación y síntesis preliminar, reduciendo el tiempo disponible para análisis de fondo, evaluación de impacto y asesoramiento estratégico.

2.2. Un nuevo enfoque automatizado e inteligente

Frente a este escenario, surge una propuesta cuyo objetivo principal consiste en automatizar la elaboración de informes sobre decretos, resoluciones y demás actos administrativos publicados en el Boletín Oficial, reduciendo tiempos de procesamiento y mejorando la calidad, uniformidad, trazabilidad y disponibilidad temprana de la información producida.

Como objetivos específicos se definen:

- automatizar la obtención de las publicaciones;
- identificar piezas relevantes mediante criterios basados en firmantes, cargos y acrónimos institucionales;
- segmentar en unidades individuales;
- emitir resúmenes preliminares utilizando IA generativa;
- ofrecer una interfaz de revisión humana con edición;
- incorporar un mecanismo de retroalimentación supervisada;
- almacenar productos y enlaces para vincularlos en la nube;
- producir un informe final consolidado apto para circulación;

La automatización aborda tareas estructuradas, como la obtención y procesamiento de archivos. Las reglas institucionales formalizan criterios de relevancia. Y la inteligencia artificial aporta capacidad de síntesis sobre contenido no estructurado.

El componente central es la validación humana *-human-in-the-loop-*, que garantiza el control sobre los resultados y habilita un proceso de evolución sostenida.

3. La herramienta por dentro

3.1. Arquitectura en capas

La solución se estructura como una arquitectura que puede describirse como una cadena de procesamiento de datos *-pipeline-*, compuesta por seis capas funcionales, de ejecución secuencial tal como se diagrama en la siguiente figura:



Fig.1 Arquitectura modular organizada en capas funcionales

Se asigna una funcionalidad individual a cada una:

- **Capa de adquisición:** obtiene información desde la publicación oficial del Boletín Oficial, descarga y releva el contenido necesario para su análisis.
- **Capa de extracción y normalización:** transforma documentos no estructurados en texto utilizable por los módulos de clasificación y resumen.
- **Capa de análisis y filtrado:** identifica piezas relevantes según reglas institucionales configurables.
- **Capa de síntesis y retroalimentación:** emite resúmenes preliminares mediante IA y capitaliza memoria supervisada derivada de correcciones previas.
- **Capa de edición y validación:** habilita la revisión humana de resultados antes de la emisión final.
- **Capa de almacenamiento y publicación:** produce el extracto final consolidado y lo almacena localmente o en la nube.

El diseño en capas propicia el reparto de responsabilidades y facilita la parametrización. Por ejemplo, cambios en la lógica de identificación de firmantes no afectan la generación documental, ajustes sobre el flujo de resúmenes se realizan sin reescribir el módulo de adquisición. El almacenamiento remoto se incorporó sin alterar las funcionalidades de las capas previas.

3.2. Flujo operativo

El circuito de trabajo puede resumirse en las siguientes etapas:

- Inicio: contenedor de capas;

Fig.2 Pantalla inicial

- Parametrización de organismos, firmantes, cargos y acrónimos; definición de ajustes operativos y destino de salida;

Organismo	Firmante	Cargo	Acrónimo 2	Acrónimo 3
Agencia De Recaudación	Cristian Alexis Girard	Director Ejecutivo	NORMATIVA	DEARBA
Asesoría General De Gobierno	Santiago Carlos Pérez Tenel	Asesor General	AGG	
Comisión de Investigaciones	Roberto Carlos Salvarezza	Presidente	CIC	
CORFO Río Colorado	Ramiro Vergara	Administrador General	CORFO	
Dirección General De Cultura	Alberto Estanislao Sileoni	Director	DGCYE	
Dirección Provincial de Hosp	Juan Sebastian Riera	Director	DPHMSALGP	
Instituto Cultural	Florencia Juana Saintout	Presidenta	ICULGP	
Instituto Provincial De Asoci	Gildo Ezequiel Onorato	Presidente	IPAYCMPCEITGP	
Jefatura De Asesores Del Gol	María Cristina Álvarez Rodri	Responsable	JAGGP	
Ministerio De Ambiente	Daniela Marina Vilar	Ministra	MAMGP	
Ministerio De Comunicaci	Jesica Laura Rey	Ministra	MCPGP	
Ministerio De Desarrollo Agr	Javier Leonel Rodriguez	Ministro	MDAGP	
Ministerio De Desarrollo De l	Andrés Larroque	Ministro	MDCGP	
Ministerio De Economía	Pablo Julio López	Ministro	MECONGP	MHYFGP
Ministerio De Gobierno	Carlos Alberto Bianco	Ministro	MGGP	
Ministerio De Hábitat Y Desz	Silvina Batakis	Ministra	MHYDUGP	
Ministerio De Infraestructur	Gabriel Nicolás Katopodis	Ministro	MYSPGP	
Ministerio De Justicia Y Dere	Juan Martín Mena	Ministro	MYDHP	
Ministerio De Las Mujeres Pi	Estela Olivia Díaz	Ministra	MBYDGP	
Ministerio De Producción Ci	Augusto Eduardo Costa	Ministro	MPCEITGP	

Fig.3 Editor de organismos y firmantes

- Consulta y descarga automática de publicaciones del BO; (con activación de estrategia de “fallback” en caso de publicación no visible por el circuito principal);

```

Buscando normas firmadas por: Ministro - Pablo Julio
López
Se encontraron 1 documentos entre el 11/07/2025 y 11/07/
2025.
Descargando Documento desde: https://boletinoficial.gba.
gob.ar/secciones/13583/ver
✓ Firma encontrada en línea 1867: Pablo Julio López, Mi
nistro.
✓ Firma encontrada en línea 1971: Pablo Julio López, Mi

```

Fig.4 Registro de eventos de la descarga del Boletín Oficial

- Extracción de texto desde los archivos obtenidos; selección según reglas institucionales; separación en extractos individuales;
- Creación automática de resumen preliminar por cada norma seleccionada (IA generativa);

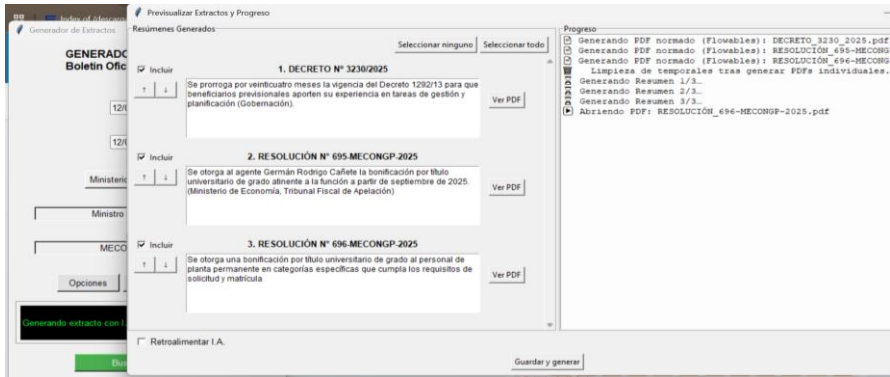


Fig.5 Previsualización de los resúmenes individuales

- Presentación de vista previa editable; corrección humana y validación; incorporación opcional de memoria supervisada si el usuario habilita retroalimentación de IA;
- Carga de documentos a la nube; inserción de enlaces a cada uno en el informe resultante;
- Elaboración del extracto final consolidado; distribución o almacenamiento del informe final.



PUBLICADO EN EL BOLETÍN OFICIAL
DE LA PROVINCIA DE BUENOS AIRES

RESUMEN DEL 14/11/2025

Creado el 17 de noviembre de 2025

Resolución / Decreto	Extracto	Link
DECRETO N° 2640/2025	Se crea el "Programa de Fortalecimiento para la Salida de las Violencias de Género" para mujeres y LGBTQ+ en riesgo, brindando acompañamiento especializado y asistencia económica directa por hasta se	Ver Documento
RESOLUCIÓN N° 614-MECONGP-2025	Se realiza una transferencia de créditos y otra de importes del Primer Orden para solventar valores de las Elecciones Legislativas e incrementar partidas de Bienes de Uso (Ministerio de Economía)	Ver Documento
RESOLUCIÓN N° 623-MECONGP-2025	Se refuerza el presupuesto vigente mediante una transferencia de créditos para atender erogaciones en servicios no personales y transferencias (Comisión de Investigaciones Científicas)	Ver Documento
RESOLUCIÓN N° 626-MECONGP-2025	El Ministro de Economía resuelve ampliar el Cálculo de Recursos y adecuar el Presupuesto General prorrogado 2025 para la Cuenta Especial debido a una mayor recaudación y la necesidad de atender gastos	Ver Documento

RESOLUCIÓN N° 729-MECONGP-2025

VISTO el EX-2025-40388877-GDEBA-DAEPYPSMECONGP, General Ejercicio 2023 - Ley N° 15.394 prorrogado para el Ejercicio 2025 en los términos del DECRETO 2024-3681-GDEBA-

CONSIDERANDO:

Que resulta necesario ampliar el Cálculo de Recursos de la Administración Central del Presupuesto General Ejercicio 2023 - Ley N° 15.394 prorrogado para el Ejercicio 2025 en los términos del DECRETO 2024-3681-GDEBA-GPBA, dentro del presupuesto asignado en la imputación de Ingresos Tributarios, durante el período enero-setiembre del corriente Ejercicio Fiscal, respecto a los niveles determinados por la Ley para algunos Rubros de Recursos, con el fin de incrementar el Cálculo de Recursos, con el fin de adecuar el Presupuesto de Erogaciones de Obligaciones a cargo del Tesoro, con amplitud a los porcentajes de afectación que son las normas vigentes. Que la totalidad incrementada la Dirección Provincial de Presupuesto Público y en los expedientes fuertemente Controlada General de la Provincia. Que la presente medida se dicta en uso de las atribuciones conferidas por los artículos 16, 17 y 20 de la Ley N° 15394 - Presupuesto General Ejercicio 2023 prorrogado para el Ejercicio 2025 en los términos del DECRETO 2024-3681-GDEBA-

Por ello,

EN EJERCICIO DE LAS FACULTADES QUE LE CONFIERE EL DECRETO N° 3681/24 EL MINISTRO DE ECONOMÍA DE LA PROVINCIA DE BUENOS AIRES

RESUELVE

ARTÍCULO 1°: Ampliar el Cálculo de Recursos de la Administración Central del Presupuesto General Ejercicio 2023 - Ley N° 15394 prorrogado para el Ejercicio 2025 en los términos del DECRETO 2024-3681-GDEBA-GPBA - Parilla N° 4, por la suma de pesos cuatro millones cuatrocientos veinte mil ochocientos noventa y nueve mil ochocientos ochenta mil (\$4.420.899.280.000.), conforme al detalle consignado en ANEXO 1 (F-2025-44011427-GDEBA-DAEPYPSMECONGP), que forma parte integrante de la presente resolución.

ARTÍCULO 2°: Incrementar en la suma de pesos cuatro millones cuatrocientos veinte mil ochocientos noventa y nueve mil ochocientos ochenta mil (\$4.420.899.280.000.), el Presupuesto General Ejercicio 2023 - Ley N° 15394 prorrogado para el Ejercicio 2025 en los términos del DECRETO 2024-3681-GDEBA-GPBA - Sector Público Provincial en Finanzas - Administración Provincial - Administración Central, conforme al detalle consignado en ANEXO 2 (F-2025-44012027-GDEBA-DAEPYPSMECONGP), que forma parte integrante de la presente resolución.

ARTÍCULO 3°: Ampliar dentro del Presupuesto General Ejercicio 2023 - Ley N° 15394 prorrogado para el Ejercicio 2025 en los términos del DECRETO 2024-3681-GDEBA-GPBA, las Contribuciones Fijas de Administración Central a Organismos Descentralizados (Parilla N° 15), en la suma de pesos veinte mil doscientos cuarenta y cuatro mil ochocientos mil (\$20.244.400.000.), conforme al detalle consignado en ANEXO 3 (F-2025-44012112-GDEBA-DAEPYPSMECONGP), que forma parte integrante de la presente resolución.

ARTÍCULO 4°: Incrementar en la suma de pesos veinte mil doscientos cuarenta y cuatro mil ochocientos mil (\$20.244.400.000.), el Presupuesto General Ejercicio 2023 - Ley N° 15394 prorrogado para el Ejercicio 2025 en los términos del DECRETO 2024-3681-GDEBA-GPBA - Sector Público Provincial en Finanzas - Administración Provincial - Organismos Descentralizados, conforme al detalle consignado en ANEXO 4 (F-2025-44012113-GDEBA-DAEPYPSMECONGP), que forma parte integrante de la presente resolución.

Fig.6 Visualización del extracto final consolidado

3.3. Estructura modular habilita

Uno de los rasgos centrales del proyecto es su organización en módulos especializados, cada uno orientado a una etapa del proceso. Esta estructura desacopla funciones, facilita

la evolución e incrementa la trazabilidad operativa. A diferencia de diseños monolíticos, permite intervenir sobre componentes específicos sin afectar el funcionamiento integral.

- **Módulo de búsqueda y filtrado de normas**

Releva la publicación oficial e identifica resoluciones o decretos firmados por autoridades definidas por el usuario. Opera con nombres, cargos y acrónimos institucionales como criterios de coincidencia. Constituye la primera instancia de reducción de complejidad, al focalizar el procesamiento sobre lo relevante para la organización.

- **Módulo de segmentación y generación de extractos individuales**

Separa las piezas detectadas en unidades documentales independientes. Esto simplifica etapas posteriores, como la elaboración de resúmenes, enlaces y almacenamiento.

- **Módulo de resumen asistido por IA**

Crea una síntesis preliminar del contenido de cada texto previamente segmentado y depurado. El componente generativo es invocado mediante un cliente automatizado y opera como apoyo inicial para la revisión posterior.

- **Módulo de retroalimentación supervisada de IA**

Reutiliza correcciones humanas como contexto de referencia para futuras ediciones. No implica reentrenamiento del modelo base, sino una estrategia de mejora incremental basada en memoria supervisada y ajuste de *prompting*.

- **Módulo de vista previa y corrección**

Materializa el enfoque *human-in-the-loop*. Ofrece una interfaz editable para revisar resultados, ajustar resúmenes, incluir o excluir actos y validar la versión final antes de su emisión.

- **Módulo de carga local o en la nube**

Almacena las unidades creadas en destinos previamente definidos y conserva sus referencias de acceso. Los enlaces se incorporan al informe final, fortaleciendo accesibilidad y trazabilidad.

- **Módulo de generación DOCX/PDF**

Produce el extracto consolidado con títulos, resúmenes y enlaces asociados, en formatos aptos para circulación institucional.

- **Módulo de parametrizaciones**

Administra organismos, titulares, cargos y acrónimos usados en el filtrado. Su actualización se adapta ante cambios institucionales sin modificar la lógica central de la aplicación.

- **Módulo de ajustes y configuración operativa**

Centraliza definiciones operativas, como formato de salida y ubicaciones de almacenamiento. Esto otorga flexibilidad de uso sin necesidad de alterar el código fuente.

3.4. Tecnologías utilizadas

Dentro del stack tecnológico se destacan los siguientes componentes:

- **Lenguaje y entorno de desarrollo:** el proyecto fue desarrollado en **Python**, por su versatilidad para integrar automatización, procesamiento documental, interfaces de usuario y creación de archivos de salida.
- **Interfaz del usuario:** se empleó **Tkinter**, biblioteca estándar de Python, para ventanas de selección, cuadros de progreso, vistas previas editables, formularios auxiliares y secciones de configuración. Se logró una interfaz liviana, embebida en el propio ecosistema y sencilla de distribuir junto con el ejecutable final.
- **Automatización de navegador:** la interacción automatizada con componentes externos de IA se resolvió con **Selenium**, con control de navegador Chrome/Chromium y perfiles persistentes, preservando la sesión de usuario cuando resultó necesario.
- **Procesamiento de documentos PDF:** la obtención de texto desde PDF se apoyó en bibliotecas de lectura y extracción textual. Se aplicaron expresiones regulares, heurísticas de segmentación y reglas específicas para detectar títulos, considerandos, firmas y artículos.
- **Generación documental:** se incluyeron bibliotecas específicas de **DOCX** y de **PDF** orientadas a la maquetación.
- **Seguridad y protección de archivos de configuración:** para el resguardo de contenido crítico se incorporó cifrado a través de la biblioteca **Fernet** del paquete *cryptography*, evitando la modificación por fuera del entorno y preservando la integridad del universo de datos usado en el filtrado.
- **Despliegue y distribución:** se empaqueta para Windows mediante **PyInstaller**, obteniendo ejecutables autónomos bajo esquema *onefile*, y manteniendo compatibilidad con ejecución directa de scripts en entornos Linux. Esta decisión permitió sostener compatibilidad multiplataforma.

4. Fortalezas del diseño

4.1. Relevamiento automatizado, extracción textual y filtrado por reglas

La obtención de la información del Boletín Oficial se diseñó sobre la base de un patrón de **relevamiento automatizado + extracción textual + filtrado por reglas**.

El proceso comienza con la consulta a la publicación oficial con un rango de fechas determinadas. Se recuperan ediciones asociadas, se identifican las secciones pertinentes y se descarga el material necesario para su análisis posterior.

Una vez obtenidos los archivos fuente, se ejecuta la extracción textual. Dado que el Boletín Oficial no se presenta como una base de datos estructurada lista para consumo analítico, se trabaja sobre texto extraído, y a partir de allí reconstruye unidades lógicas como títulos de resoluciones, referencias a firmantes, artículos y cuerpos normativos.

La metodología puede describirse como una combinación de:

- automatización de captura de documentos;
- extracción de texto desde PDF;
- *parsing* heurístico basado en patrones;
- clasificación por reglas institucionales.

4.2. Estrategia de “*fallback*” para la detección temprana

Un aspecto relevante de la ingeniería implementada consiste en la estrategia de “*fallback*” de publicación temprana. Como el Boletín Oficial no siempre aparece indexado o accesible de inmediato al flujo principal de consulta, se incorpora una lógica alternativa para detectar la versión más reciente a través del acceso directo a la última edición disponible.

Ante esta situación, se dispara un circuito secundario de contingencia para recuperar, analizar y verificar en su contenido la fecha efectiva de emisión. De corresponder a la jornada buscada, el procesamiento continúa aun cuando la publicación todavía no se encuentre plenamente visible en el circuito primario.

Este mecanismo reduce demoras operativas y permite anticipar el inicio del procesamiento diario, promoviendo la disponibilidad temprana de los informes.

4.3. Inteligencia artificial y técnica aplicada

La inteligencia artificial interviene en la etapa de síntesis textual. El modelo de lenguaje generativo externo es invocado mediante un cliente automatizado, utilizando un perfil persistente de navegador para asegurar continuidad de sesión y estabilidad de uso.

La técnica aplicada puede caracterizarse como **resumen abstractivo** o **generación de resúmenes abstractivos**, es decir, producción de una síntesis nueva a partir del contenido textual de entrada, en lugar de mera extracción literal de frases del documento. Este enfoque resulta particularmente útil cuando el texto fuente tiene estilo administrativo y requiere una reformulación breve y legible.

No obstante, para brindar mayor precisión y relevancia, la IA opera sobre un texto previamente procesado. Ese pre-procesamiento incluye segmentación por norma, limpieza de encabezados innecesarios, detección de considerandos y selección de secciones de mayor valor semántico. En consecuencia, el resumen preliminar combina dos capas: una capa determinística basada en reglas que delimita correctamente el contenido a resumir, y una capa generativa que produce la síntesis en lenguaje natural.

A ello se suma el mecanismo de retroalimentación supervisada, mediante el cual los resúmenes corregidos por los expertos se reutilizan como memoria contextual. La combinación del pre-procesamiento heurístico, el *prompting* dirigido y la memoria supervisada beneficia la pertinencia de la salida sin alterar el modelo generativo subyacente. El uso de IA acelera la confección del extracto preliminar y el punto de partida para la revisión experta.

4.4. Supervisión humana sobre la IA y sobre la mejora continua

Un aspecto central del diseño es la decisión de mantener una instancia explícita de validación humana. Los especialistas ya no realizan el trabajo desde cero, sino que revisan y ajustan una salida ya estructurada. Se produce un cambio significativo en la distribución del esfuerzo y tiempos.

Desde el punto de vista metodológico, este enfoque responde a un patrón de **supervisión humana sobre automatización asistida por IA**.

La evolución del sistema no depende exclusivamente de entrenar el modelo generativo, sino de varios mecanismos complementarios como el ajuste de *prompts*, el refinamiento de reglas de filtrado, la actualización de criterios institucionales, el aprovechamiento de memoria supervisada y el aprendizaje operativo derivado del uso cotidiano.

Por ello, resulta más adecuado hablar de **mejora continua supervisada** que de aprendizaje autónomo pleno, ya que la herramienta evoluciona en estrecha interacción con el criterio experto.

5. Reflexiones finales

5.1. Resultados y beneficios

La implementación consigue avances concretos en la confección de extractos normativos. Reduce significativamente el tiempo de elaboración e incrementa la uniformidad del producto final tanto en el orden de presentación como en la forma de redactar las síntesis. Por último y para destacar, la automatización libera tiempo de profesionales especializados, cuya intervención se concentra ahora en validación y ajuste.

El encadenamiento de las etapas convierte un conjunto de tareas manuales en un *pipeline* reproducible, verificable y reutilizable. Cada etapa puede trazarse, ajustarse y depurarse de manera independiente:



Fig.7 Comparativa entre el proceso manual y el flujo automatizado inteligente

La incorporación de almacenamiento remoto y enlaces directos optimizó la accesibilidad y trazabilidad. El informe ya no funciona únicamente como resumen ejecutivo, sino también como índice de acceso a los documentos normativos procesados y almacenados en un repositorio compartido.

Desde el punto de vista técnico, demuestra la viabilidad de articular *scraping* documental, *parsing* heurístico, reglas institucionales, IA generativa, memoria supervisada, cifrado de configuración, carga en la nube y edición humana en un único flujo de trabajo aplicado al sector público. En cuanto a la visión organizacional, brinda disponibilidad temprana de información procesada tanto para autoridades como para equipos técnicos.

En promedio se examinan diariamente de **80 a 120 normas**, observándose una notable reducción en los tiempos de procesamiento. Mientras que la elaboración manual de extractos requería de **2 a 4 horas de dedicación especializada por jornada**, ahora el informe se obtiene **en menos de 15 minutos**, concentrando la intervención humana en instancias de validación y ajuste.

5.2. Conclusiones

El **Generador Inteligente de Extractos del Boletín Oficial** constituye una experiencia exitosa aplicada de automatización e inteligencia artificial en el ámbito de la gestión pública. Su principal aporte no reside únicamente en el componente generativo, sino en una arquitectura integral que articula adquisición de archivos, extracción textual, clasificación por reglas institucionales, síntesis asistida por IA, retroalimentación supervisada, revisión humana, generación documental y almacenamiento en la nube.

Desde una perspectiva técnica, el proyecto muestra que es posible intervenir sobre procedimientos apoyados en documentos no estructurados mediante una arquitectura híbrida, donde conviven heurísticas determinísticas, automatización de navegador, módulos de procesamiento textual, memoria supervisada, cifrado de archivos sensibles y componentes generativos.

Desde una perspectiva organizacional, revela que la automatización reduce tareas repetitivas sin perder control sobre el resultado final.

La capacidad de almacenar en la nube tanto los extractos individuales como los informes vinculados al procesamiento diario, brinda trazabilidad, facilita el acceso distribuido y reduce la dependencia de rutas locales. En contextos institucionales, resulta especialmente útil cuando los informes deben circular entre múltiples usuarios o áreas, o cuando se requiere conservar una copia centralizada de la documentación.

La experiencia confirma además que el valor de la IA en contextos públicos aumenta cuando se la utiliza como asistencia dentro de flujos bien definidos, con objetivos concretos y validación humana explícita. Mejora las condiciones de ejecución, permitiendo disponer de información normativa oportuna, homogénea y de mayor calidad para la toma de decisiones.

5.3. Desafíos y líneas futuras

Entre los principales desafíos se identifican la variabilidad del formato de las publicaciones, la dependencia de patrones textuales para identificar normas y la necesidad de conservar sesiones válidas para el componente generativo externo cuando es invocado mediante un navegador automatizado. El filtrado por firmantes y cargos requiere mantenimiento continuo, dado que los cambios de autoridades o variantes de firma deben reflejarse oportunamente en la parametrización. También resulta necesario controlar la calidad de la extracción textual, especialmente cuando la fuente presenta estructuras complejas o encabezados repetitivos.

El mecanismo de retroalimentación, por su parte, requiere cuidado en la selección de ejemplos reutilizados, a fin de no introducir sesgos ni arrastrar correcciones contextuales que no deban estandarizarse. Del mismo modo, la carga en la nube exige gestionar permisos, rutas de destino y consistencia entre los enlaces y documentos efectivamente almacenados.

Por último, la naturaleza institucional del proceso impone restricciones de estabilidad y seguridad, exigiendo un equilibrio entre flexibilidad técnica y robustez operativa, especialmente cuando la aplicación se distribuye como ejecutable en puestos de trabajo diversos o se integra con perfiles específicos de navegador.

Como evolución futura del proyecto se prevé:

- ampliar la base de criterios institucionales para clasificación automática;
- agilizar la detección de firmantes con títulos complejos;
- robustecer la obtención de extractos individuales;
- fortalecer la trazabilidad entre norma, resumen y extracto final;
- perfeccionar la interacción entre el cliente automatizado y el modelo generativo;
- optimizar el uso de memoria supervisada en la retroalimentación;
- ampliar las capacidades de configuración de destinos;
- extender la compatibilidad multiplataforma, migración hacia entorno web;
- evaluar alternativas de IA local o de integración más directa, reduciendo dependencia del navegador cuando el contexto técnico lo permita.

6. Referencias

Individual Identity & Security Team. (2025). *Cryptography: Python cryptography library (Version 45.0.1)*. Python Software Foundation. <https://pypi.org/project/cryptography/>

Selenium Project (s.f.) (2025). *Selenium IDE: Open source record and playback test automation for the web*. <https://www.selenium.dev/selenium-ide/>

Canny, S. (s.f.). (2025) *python-docx: Create and update Microsoft Word (.docx) files*. <https://python-docx.readthedocs.io/>

Google LLC. (2025). *Google APIs Client Library for Python (Version 2.158.0)*. Google Developers. <https://developers.google.com/api-client-library/python/>

Python Software Foundation. (2025). *Graphical User Interfaces with Tk*. Python.org. <https://docs.python.org/3/library/tkinter.html>

Padilla, J. (2025). *PyJWT: JSON Web Token implementation in Python (Version 2.10.0)*. PyPI. <https://pyjwt.readthedocs.io/>

McKie, J., & Liu, R. (2025). *PyMuPDF (Version 1.25.2)*. Artifex Software, Inc. <https://pymupdf.readthedocs.io/en/latest/>

Individual Identity & Security Team (s.f.) (2024). *Fernet (symmetric encryption)*. Cryptography.io. <https://cryptography.io/en/latest/fernet/>

JSON Web Token (RFC 7519): Jones, M., Bradley, J., & Sakimura, N. (2015). *JSON Web Token (JWT)*. IETF RFC Editor. <https://www.rfc-editor.org/rfc/rfc7519>

OAuth 2.0 Framework (RFC 6749): Hardt, D. (2012). *The OAuth 2.0 Authorization Framework*. IETF RFC Editor. <https://www.rfc-editor.org/rfc/rfc6749>