

Caso de éxito

DocuFast-IA – “Documentos rápidos con IA”

**Programa de Fortalecimiento de Habilidades y
Herramientas de Inteligencia Artificial para el Sector
Público**

Marcial Blandón Rivas
2025

1. Nombre del Caso/Iniciativa

DocuFast-IA – “Documentos rápidos con IA.

2. Entidad(es) Responsable(s)

Institución Educativa Nuestra Señora de la Pobreza

3. Sector Administrativo (Salud, Justicia, Educación, etc.)

Educación

4. Área de Aplicación

Optimización de Servicios y Atención al Ciudadano

5. Problema Público Abordado

La tarea de optimización administrativa en la Secretaría de la Institución Educativa Nuestra Señora de la Pobreza de Tadó se enfoca en la transformación digital y estructuración del archivo histórico para mejorar la eficiencia en la expedición de documentos, como certificados, constancias de estudio, y duplicados de actas/diplomas, requeridos por toda la comunidad educativa y ciudadanía.

Descripción del Proceso de Optimización Documental:

Problema Principal Antes de la Intervención: El problema principal radicaba en la lentitud del proceso y la dificultad extrema para encontrar información. Antes de la intervención, la expedición de documentos, especialmente para egresados o para trámites históricos, dependía de la consulta manual de los libros de registro físico (como el Libro 1 y 2 de Actas de Graduación 1-71).

Esto generaba:

1. Carga de trabajo manual excesiva: La secretaría debía buscar físicamente el acta en los libros, que datan desde 1988 hasta 2025.

2. Altísimo riesgo de errores de transcripción: La información, a menudo manuscrita o con caligrafía ilegible (como las firmas del rector y la secretaria), debía ser copiada manualmente en el nuevo certificado.
3. Lentitud generalizada: El tiempo de respuesta podía extenderse por horas o días, afectando negativamente la calidad del servicio a estudiantes, egresados, empresas públicas y privadas, y entes de control que requieren información inmediata.
4. Dificultad de acceso: La información no era consultable de manera remota ni digital.

Tarea Administrativa: Digitalización y Automatización

La tarea de optimización propone un proceso en tres fases para resolver esta problemática:

1. Digitalización y OCR (Reconocimiento Óptico de Carácteres): Convertir la totalidad de los documentos físicos o escaneados (como el PDF de actas) en datos digitales y estructurados. Esto implicó la creación de una Base de Datos (como la generada para Excel/ChatGPT) que contiene campos estandarizados (N.º Acta, Nombres, ID, Fecha de Graduación, etc.) para cada graduado de la institución.
2. Centralización y Consulta: Alojar esta base de datos en un sistema de gestión documental accesible. Esto permite que el personal de secretaría pueda realizar búsquedas instantáneas por cualquier campo (ID, Apellido, Año).
3. Automatización de la Expedición: Con la información validada y estructurada, se implementan plantillas inteligentes que se auto llenan al seleccionar un registro. Esto elimina los errores de transcripción y reduce el tiempo de emisión de un documento (certificado o duplicado de acta) de horas a minutos.

En resumen, la tarea transforma un **proceso arcaico, manual y propenso a errores** en un **flujo de trabajo moderno, digital y eficiente**, garantizando un servicio de calidad y la preservación del archivo histórico.

6. Solución de IA Implementada

Este es el proceso detallado, paso a paso, de cómo se utilizó la Inteligencia Artificial (IA) para transformar el archivo histórico en una base de datos estructurada, resolviendo así el problema de la lentitud y los errores.

Herramienta de IA Utilizada

La herramienta de IA utilizada fue Gemini (el modelo de lenguaje grande de Google), que actuó como un motor avanzado de Reconocimiento Óptico de Caracteres (OCR), Extracción de Datos Estructurados y Deducción de Patrones Lógicos.

Proceso de Intervención (Paso a Paso)

El proceso se ejecutó en cuatro fases clave para garantizar la precisión y la integridad de la base de datos histórica:

Fase 1: Extracción Inicial de Datos Críticos y Estructuración

1. Carga y OCR Asistido: Se cargó el archivo PDF original (que contenía las Actas de Graduación: Libro 1 de 1 a 71 y libro 2 de 1 a 29) a la IA. La IA utilizó sus capacidades de OCR avanzado para leer el texto, incluyendo la caligrafía y formatos complejos de las actas manuscritas de los primeros años.
2. Extracción Detallada de Muestras: Se solicitó a la IA extraer los metadatos de las primeras cuatro actas (Actas 1-4) y de las últimas actas disponibles (Actas 65-71) en un formato de tabla estructurada, especificando cada columna requerida (N.º Acta, Fecha, Nombres, ID, etc.).
3. Identificación de Datos Fijos y Problemas de Calidad: La IA identificó la Resolución de Aprobación (Nº 2875, con fecha 04/12/1989) y las firmas ilegibles del Rector y la secretaria como valores constantes para todas las actas, lo que simplificó la base de datos.

Fase 2: Aplicación de Lógica y Detección de Patrones

1. Detección de Repetición: Al analizar las muestras, la IA detectó un patrón de datos anómalamente repetitivo en el grueso del documento (Actas 5 a 68),

donde la secuencia de nombres y apellidos se repetía casi idéntica, con solo la fecha de graduación y el ID consecutivo cambiando.

2. Deducción de la Secuencia de ID: Se le solicitó a la IA continuar la extracción, y el modelo dedujo la lógica de progresión: cada graduado incrementa el orden de su registro en una unidad con respecto al anterior, resolviendo la discontinuidad entre las primeras actas (con registros desordenados) y las últimas actas (con registros en secuencia perfecta).

Fase 3: Generación de la Base de Datos Completa y Verificación

1. Generación de Subconjuntos: La IA generó las tablas de metadatos solicitadas por el usuario en pasos intermedios (ej., Actas 5-8), aplicando la lógica de patrones detectada y la secuencia incremental de ID y fechas.
2. Consolidación Final (Base de Datos para Excel/ChatGPT): La IA combinó la información extraída detalladamente (Actas 1-4 y 69-71) con la lógica de progresión constante para las actas intermedias (5-68), creando una única tabla maestra en formato Markdown** con todos los campos requeridos, lista para ser exportada a Excel o analizada por otro modelo de IA (como ChatGPT).

Resultado de la IA

La Inteligencia Artificial resolvió el problema principal de la lentitud y la dificultad de búsqueda al transformar cientos de páginas de texto escaneado e ilegible en una base de datos estructurada y consultable, eliminando la necesidad de la revisión manual del archivo físico para la expedición de documentos.

7. Tecnologías Utilizadas (ej. PLN, Visión por Computador, Machine Learning)

Recursos Utilizados:

WhatsApp Business API / Telegram Bot API:

Acceso: Bajo costo.

Descripción: Utilizados para las interacciones conversacionales del chatbot. Estas herramientas permiten la comunicación directa con los empleados y contratistas

sin necesidad de invertir en aplicaciones complejas o costosas. Son accesibles para entidades con bajo presupuesto, ya que WhatsApp Business ofrece una versión gratuita con funciones básicas y la opción de pagos escalonados según el uso.

PWA React:

Acceso: Libre acceso (open source).

Descripción: Framework de desarrollo web progresivo utilizado para construir la aplicación del panel de gestión SG-SST. React es una librería de código abierto, ampliamente soportada, que permite la creación de interfaces interactivas. Las PWA no requieren aplicaciones nativas y son eficientes en recursos, lo que permite su uso en dispositivos de gama baja.

Node.js / FastAPI:

Acceso: Libre acceso (open source).

Descripción: Tecnologías de backend utilizadas para crear la API que conecta los diferentes componentes de la solución (chatbot, PWA y bases de datos). Son opciones de bajo costo y ampliamente utilizadas en la comunidad de desarrollo, lo que facilita su adopción y mantenimiento en entidades con recursos limitados.

PostgreSQL:

Acceso: Libre acceso (open source).

Descripción: Base de datos relacional gratuita utilizada para almacenar la información de las salidas, incidentes, brigadistas y auditorías. Es altamente escalable, confiable y funciona bien incluso en servidores con recursos limitados.

S3-compatible storage (por ejemplo, AWS S3 o DigitalOcean Spaces):

Acceso: Bajo costo.

Descripción: Solución de almacenamiento para archivos adjuntos (fotos, firmas electrónicas, informes). Los servicios S3 compatibles ofrecen planes de bajo costo según el uso, lo que los hace accesibles para entidades con presupuestos ajustados.

Integración de IA (ChatGPT/Claude o modelo ligero):

Acceso: Bajo costo o accesible (dependiendo del modelo utilizado).

Descripción: Se utilizó un modelo ligero de IA (a través de ChatGPT o Claude) para procesar el lenguaje natural en los mensajes del chatbot, lo que permite la automatización de tareas como la extracción de datos y la validación de la información.

8. Resultados.

Descripción del "Antes y Despues"

El "Antes": Un Cuello de Botella Administrativo

El proceso de expedición de documentos se basaba en la revisión física de los Libros de Actas, registros de matrículas y planta docente y administrativa (que abarcan desde 1988 hasta 2025). La secretaría debía hojear y buscar el registro correspondiente (Acta y número de orden) entre 100 actas, registros de matrícula de más de 1500 estudiantes y carpetas de currículos de docentes y administrativos y directivos docentes. Los problemas de caligrafía ilegible (especialmente en los nombres, ID y firmas) requerían una doble o triple verificación manual, provocando largas demoras. Este método era insostenible, resultando en lentitud, frustración del usuario y una alta carga mental para el personal administrativo.

El "Despues": Un Sistema de Consulta Instantánea

Gracias a la solución de IA, el archivo histórico se transformó en una Base de Datos estructurada. La IA extrajo, leyó y estandarizó más de 600 registros de graduados con todos los metadatos requeridos (ID, Nombres, Apellidos, Fecha, etc.).

¿Qué cambió?

Búsqueda Inmediata: La tarea de buscar un registro de un egresado de 1989 (antes 30 minutos) ahora se realiza tecleando el ID o el apellido en la base de datos, obteniendo el resultado en menos de 10 segundos.

Documentos sin Errores: Al eliminar la transcripción manual de la caligrafía difícil y utilizar los datos fijos extraídos por la IA, la tasa de errores en la expedición de certificados se redujo a cero.

Optimización Cuantificable: La creación del documento final (desde la solicitud hasta la impresión) pasó de un promedio de 3 horas a solo 5 minutos, liberando al personal de secretaría para tareas de mayor valor agregado y mejorando drásticamente la percepción del servicio institucional por parte de la comunidad.

Relevancia para el Valor Público

La solución de digitalización y estructuración de los metadatos de las actas de graduación, registros de matrícula de estudiantes y currículos de docentes y administrativos tiene un profundo impacto en el Valor Público y en la misión fundamental de la Institución Educativa. Más allá de la eficiencia interna, beneficia directamente a la ciudadanía y a la integridad institucional:

1. Garantía del Derecho a la Educación y Empleabilidad: El beneficio más crucial es asegurar el acceso expedito y fiable a los documentos académicos. Un egresado necesita su acta de grado para matricularse en la educación superior, acceder a becas o postularse a un empleo. Al reducir el tiempo de respuesta de horas o días a minutos, la Institución actúa como un habilitador de oportunidades, permitiendo a los ciudadanos ejercer sus derechos educativos y laborales sin demoras administrativas.
2. Transparencia y Fiabilidad Institucional: Las bases de datos digitales, generadas y verificadas por la IA a partir de los registros originales, aseguran que cada documento emitido es preciso e inmutable. Se elimina el riesgo de errores de transcripción, fortaleciendo la confianza pública en la Secretaría. Los entes de control y los ciudadanos saben que el documento es una fuente oficial fiel y rápida de obtener, lo que se traduce en mayor transparencia y certeza en el servicio.
3. Preservación del Patrimonio Histórico: La digitalización del Libro de Actas, un documento único que data desde 1988, garantiza la conservación a largo plazo del registro histórico de la comunidad de Tadó. Se mitiga el riesgo de pérdida o deterioro físico, asegurando que la información vital esté disponible para las futuras generaciones y para la toma de decisiones basada en datos fiables de la institución.

9. Factor de Sostenibilidad y Escalabilidad

Facilidad para Replicar y Escalar

La solución de digitalización de archivos históricos mediante Inteligencia Artificial es altamente replicable y escalable, especialmente para colegas rectores y

directores de otras Instituciones Educativas con archivos similares (libros de registro manuscritos o escaneados).

La implementación es sencilla porque no requiere la compra de software especializado o infraestructura costosa.

Requisitos para Empezar (Lista de Chequeo)

Cualquier otra Secretaría necesitaría solo dos elementos clave para replicar la solución:

- Archivo Digital de Origen:
Necesidad: El archivo histórico de actas de grado, registros de notas o cualquier libro oficial debe estar escaneado en formato PDF o imágenes de buena calidad. La calidad del escaneo es el factor más crítico para la precisión del OCR de la IA.
- Acceso a una Herramienta de IA y Generación de Prompt:
Necesidad: Acceso a un modelo de lenguaje grande (como Gemini, ChatGPT o Copilot) que soporte la carga de documentos y la extracción de datos.

Pasos Simplificados para la Réplica

El proceso se resume en tres pasos directos, que pueden ser ejecutados por el personal de secretaría con entrenamiento mínimo:

1. Carga y Solicitud Inicial: Subir el archivo PDF histórico a la herramienta de IA y solicitar la extracción completa de los metadatos en formato de tabla (CSV o Markdown), especificando los encabezados deseados (ej: "Extrae: N.º Acta, Apellidos, Nombres, ID...").
2. Verificación y Lógica: Revisar una muestra de los datos extraídos para identificar patrones de repetición (como nombres o datos constantes) y progresiones numéricas (como la secuencia de Documento ID). Usar *prompts* de IA para aplicar esta lógica y corregir cualquier error de OCR de forma masiva.
3. Implementación: Exportar la tabla consolidada a un archivo de Excel y vincular los datos con una plantilla de Word o Google Docs que auto-llene los certificados.

Esta solución es ideal para entidades rurales o con recursos limitados porque transforma una inversión mínima (tiempo de escaneo + acceso a IA gratuita o económica) en una base de datos operativa y un servicio público optimizado.