

**Informe de Proyecto: Solución de Estandarización Documental**  
**con IA "On-Demand"**

PARA: Dirección General / Gerencia de Sistemas / Socios / Gonzalo Recalde

DE: Diego Cutignola

FECHA: 30 de Noviembre 2025

**ASUNTO:** *Propuesta de implementación de  
herramienta complementaria de IA para organización de archivos.*

---

## 1. Resumen Ejecutivo

El objetivo de este proyecto es desplegar una herramienta interna ("Antigravity App") que permita a los colaboradores estandarizar la nomenclatura de documentos en Google Drive mediante Inteligencia Artificial Generativa.

A diferencia de nuestros procesos *batch* actuales (que operan automáticamente sobre carpetas estructurales predefinidas), esta solución está diseñada para ser **ejecutada manualmente por el usuario** en carpetas de trabajo temporales, auditorías específicas o repositorios personales, cerrando la brecha de desorden en la "última milla" de la gestión documental.

## 2. El Problema y la Oportunidad

Actualmente, el estudio cuenta con una **estrategia híbrida** de gestión documental que tiene un punto ciego:

1. **La Capa Automática (Existente):** Disponemos de procesos desatendidos que ordenan las carpetas estructurales críticas ("*Doc de Respaldo*", "*Doc Clientes*"). Esto garantiza el orden en el "archivo central".
2. **El Desafío Pendiente:** El trabajo diario de los colaboradores ocurre fuera de esas carpetas maestras. Los contadores y auditores crean carpetas temporales, reciben Zips de clientes o trabajan en borradores. En estos espacios, la nomenclatura sigue siendo caótica (IMG-2024.jpg, scan001.pdf).

**Solución Propuesta:** Desarrollar una interfaz web interna simple y segura donde el colaborador tiene el control para ordenar *sus* carpetas de trabajo bajo los mismos estándares del estudio.

### Sinergia Operativa: Dos Herramientas, Un Objetivo

La nueva aplicación funcionará en conjunto con la solución actual sin solaparse:

Característica	Solución Actual (Automática)	Nueva App (Propuesta)
Tipo de Ejecución	Autónoma / Programada (Background).	<b>Manual / A demanda del usuario.</b>
Alcance	Carpetas Rígidass ("Doc Clientes").	<b>Cualquier carpeta</b> donde el usuario tenga acceso.
Objetivo	Mantenimiento de la estructura central.	Herramienta de productividad diaria.
Control	Definido por Sistemas.	Definido por el Colaborador.

## 3. Especificaciones Técnicas y Arquitectura

La solución se desplegará íntegramente en **Google Cloud Platform (GCP)** bajo una arquitectura *Serverless* (sin servidor), garantizando escalabilidad automática y costos reducidos.

### 3.1. Stack Tecnológico & Alojamiento

- **Alojamiento (Frontend & Backend): Google Cloud Run.** Contenedores Docker que escalan a cero cuando no se usan (costo de infraestructura nulo fuera de horario laboral).
- **Gestión de Tareas: Google Cloud Tasks.** Gestiona el procesamiento asíncrono ("Fire and Forget") para que el usuario no deba esperar con la ventana abierta.
- **Motor de Inteligencia (Cerebro): Vertex AI (Gemini Pro).** Modelo LLM empresarial configurado para respuestas determinísticas.

- **Motor de Lectura (OCR Híbrido):**

- Nivel 1: pypdf para extracción de texto nativo (rápido y sin costo).
- Nivel 2: **Google Cloud Document AI** para documentos escaneados/imágenes (alta precisión).

### 3.2. Modelo de Seguridad

- **Autenticación:** Vía Google OAuth restringido al dominio del estudio (@estudio.com).
- **Alcance (Scopes):** La aplicación opera utilizando el token del usuario. **Solo puede procesar carpetas que el empleado ya tiene permiso de ver.**
- **Integridad:** No se modifica el contenido del archivo, solo sus metadatos (nombre).

## 4. Lógica de Negocio: Reglas de Nomenclatura

Para garantizar coherencia con los sistemas actuales, la IA seguirá reglas estrictas ("System Prompt").

**Formato Objetivo:** [FECHA]\_[CATEGORÍA]\_[EMISOR]\_[DETALLE\_BREVE].[EXT]

### 4.1. Taxonomía y Categorización

La IA clasificará cada documento en una de las siguientes categorías exclusivas:

Categoría	Descripción / Uso	Ejemplos
<b>CONTABLE</b>	Documentación formal contable.	Balances, Libro Diario, Sumas y Saldos.
<b>FACTURA</b>	Comprobantes fiscales.	Facturas A/B/C, Tickets, Notas de Crédito.
<b>SUELDO</b>	Documentación laboral.	Recibos de Haberes, F931, Liquidaciones.

Categoría	Descripción / Uso	Ejemplos
<b>RESUMEN</b>	Extractos financieros.	Resúmenes Bancarios, Tarjetas, Brokers.
<b>IMPUESTO</b>	Obligaciones tributarias.	VEPs, DDJJ IIBB/Ganancias, Tasas.
<b>LEGAL</b>	Documentos jurídicos.	Contratos, Estatutos, Actas.
<b>DOC-INTERNA</b>	Papeles de trabajo.	Excel auxiliares, borradores.
<b>CONSTANCIA</b>	Identificación fiscal.	Inscripciones, CUIT.

#### 4.2. Reglas de Inferencia de Fecha

- **Documentos Puntuales:** AAAA-MM-DD (Ej: Facturas).
- **Documentos Mensuales:** AAAA-MM (Ej: Resúmenes, DDJJ).
- **Documentos Anuales:** AAAA (Ej: Balances).

### 5. Funcionamiento (Experiencia de Usuario)

1. **Log-in Seguro:** El usuario ingresa con su cuenta corporativa.
2. **Input:** Pega el enlace de la carpeta de Drive a organizar.
3. **Procesamiento Inteligente:**
  - La App detecta automáticamente si un archivo ya tiene el formato correcto (Ej: 2025-01\_FACTURA...). Si es así, lo omite para ahorrar costos.
  - Si el nombre es incorrecto, lee el contenido y lo renombra.
4. **Reporte:** Al finalizar, el usuario recibe un email detallando qué archivos fueron renombrados.

## 6. Beneficios y ROI (Retorno de Inversión)

Beneficio	Impacto Directo
Ahorro de Tiempo	Procesar 50 archivos manualmente toma ~60 minutos. <b>La App lo hace en ~3 minutos</b> de tiempo de máquina, liberando al empleado para tareas de valor.
Búsqueda Instantánea	Al tener nombres estandarizados, encontrar un comprobante de hace 3 años toma segundos, no horas.
Consistencia Total	Aseguramos que un archivo se llame igual tanto si está en la carpeta automática de clientes como si está en la carpeta personal de un auditor.
Costo Eficiente	Al usar arquitectura <i>Serverless</i> , el costo de infraestructura es cercano a cero si no se usa.
Profesionalismo	Evita compartir con clientes archivos con nombres como WhatsApp_Image_2023... o Scan0056.pdf.

## 7. Plan de Implementación

### 1. Fase 1: Prototipo (2 Semanas)

- Desarrollo del *Worker* en Python.
- Configuración de reglas de nomenclatura en Vertex AI.

### 2. Fase 2: Despliegue Beta (1 Semana)

- Prueba con usuarios clave (Auditores/Seniors) en carpetas de "Sandbox".

### 3. Fase 3: Lanzamiento General

- Comunicación interna y puesta a disposición de la herramienta para todo el estudio.

## Conclusión

Esta propuesta cierra el círculo de la gestión documental del estudio. Mientras la solución actual cuida la estructura central, esta nueva aplicación **extiende esa capacidad de orden y estandarización a las manos de cada colaborador**, eliminando el caos documental en todas las instancias de trabajo de forma segura, auditada y eficiente.

**Solicito su aprobación para avanzar con la Fase 1 (Prototipo).**

**Atentamente,**