# Diseño y desarrollo de una aplicación web con inteligencia artificial para la gestión documental y recomendación de respuestas en Mipymes de Bucaramanga.

Modalidad: Desarrollo Tecnológico

Mario Andrés Jácome Mantilla.
CC 1098753829

**UNIDADES TECNOLÓGICAS DE SANTANDER**
**Facultad de Ciencias Naturales e Ingeniería**
**Ingeniería de Sistemas**
**Bucaramanga, 28 de agosto de 2026**

# Diseño y desarrollo de una aplicación web con inteligencia artificial para la gestión documental y recomendación de respuestas en Mipymes de Bucaramanga.

Modalidad: Desarrollo Tecnológico

Mario Andrés Jácome Mantilla.
CC 1098753829

**Trabajo de Grado para optar al título de**
Ingeniero de Sistemas

**DIRECTOR**
Yuli Andrea Álvarez Pizarro

Grupo de investigación: GRIIS

**UNIDADES TECNOLÓGICAS DE SANTANDER**
**Facultad de Ciencias Naturales e Ingeniería**
**Ingeniería de Sistemas**
**Bucaramanga, 28 de agosto de 2026**

Nota de Aceptación

Este informe final de trabajo de grado, en modalidad Desarrollo de Software fue APROBADO en cumplimiento de uno de los requisitos exigidos por la Institución Universitaria Tecnológica de Santander para optar por el título de Ingeniero de Sistemas según acta No. 24 de octubre de 2026 del Comité de Trabajo de Grado.



Firma del Evaluador



Firma del Director

# DEDICATORIA

Dedico este trabajo a mi pareja, por estar a mi lado y por acompañarme en este proceso, y a mi mamá, por su cariño, paciencia y su apoyo durante mi formación. A las dos, por ser parte de mi vida y acompañarme en esta meta.

# AGRADECIMIENTOS

Agradezco a Yuli Andrea Álvarez Pizarro, directora del trabajo de grado, por su orientación y por las observaciones que me ayudaron a mejorar el proyecto y el informe.

A las Unidades Tecnológicas de Santander y a los docentes que hicieron parte de mi formación, por los conocimientos que pude aplicar en el desarrollo de MiDocIA.

A mis compañeros de la universidad, por el apoyo y los conocimientos compartidos durante la carrera.

# TABLA DE CONTENIDO

**RESUMEN EJECUTIVO** ..... **12**

**INTRODUCCIÓN** ..... **13**

**1. DESCRIPCIÓN DEL TRABAJO DE INVESTIGACIÓN** ..... **15**

**1.1. PLANTEAMIENTO DEL PROBLEMA 15**
**1.2. JUSTIFICACIÓN 18**
**1.3. OBJETIVOS 21**
1.3.1. OBJETIVO GENERAL 21
1.3.2. OBJETIVOS ESPECÍFICOS 21
**1.4. ESTADO DEL ARTE 22**
1.4.1. ANTECEDENTES INTERNACIONALES 23
1.4.2. ANTECEDENTES NACIONALES 24
1.4.3. RELACIÓN DE LOS ANTECEDENTES CON MIDOCIA 25

**2. MARCO REFERENCIAL** ..... **27**

**2.1. MARCO TEÓRICO 27**
2.1.1. RECONOCIMIENTO ÓPTICO DE CARACTERES Y PROCESAMIENTO DOCUMENTAL 27
2.1.2. MODELOS DE LENGUAJE Y ANÁLISIS DEL CONTENIDO 28
2.1.3. INSTRUCCIONES PARA ORIENTAR EL ANÁLISIS 29
2.1.4. SUPERVISIÓN HUMANA Y USO DE LOS RESULTADOS 30
2.1.5. PRIVACIDAD DESDE EL DISEÑO Y PROCESAMIENTO TEMPORAL 31
2.1.6. RELACIÓN DE LOS FUNDAMENTOS CON LA EVALUACIÓN 32
**2.2. MARCO CONCEPTUAL 33**
2.2.1. GESTIÓN DOCUMENTAL EN EL ALCANCE DEL PROYECTO 33
2.2.2. DOCUMENTO DE ENTRADA Y CONTENIDO DOCUMENTAL 33
2.2.3. CATEGORÍA Y CLASIFICACIÓN DOCUMENTAL 34
2.2.4. TEXTO EXTRAÍDO Y CAMPOS RELEVANTES 34
2.2.5. RESUMEN DOCUMENTAL 35
2.2.6. ALERTA DOCUMENTAL 35
2.2.7. SUGERENCIA DE RESPUESTA Y RECOMENDACIÓN DE ACCIÓN 36
2.2.8. PROCESAMIENTO TEMPORAL Y ZERO STORAGE 36
2.2.9. REVISIÓN HUMANA Y CONFIRMACIÓN DEL ANÁLISIS 37
2.2.10. METADATOS DE USO Y VALORACIÓN DEL USUARIO 37
2.2.11. DOCUMENTOS REALES, SINTÉTICOS Y VALORES ESPERADOS 38
2.2.12. TIEMPOS UTILIZADOS EN LA EVALUACIÓN 38

**2.3. MARCO LEGAL 39**
2.3.1. PROTECCIÓN DE DATOS PERSONALES: LEY 1581 DE 2012 39
2.3.2. AUTORIZACIÓN Y CONSERVACIÓN DE LA INFORMACIÓN: DECRETO 1074 DE 2015 40
2.3.3. DERECHO DE PETICIÓN: LEY 1755 DE 2015 41
2.3.4. ALCANCE DE LAS RECOMENDACIONES Y DE LA REVISIÓN HUMANA 42
**2.4. MARCO TECNOLÓGICO 43**
2.4.1. ARQUITECTURA WEB Y ORGANIZACIÓN DE LOS COMPONENTES 43
2.4.2. INTERFAZ DE USUARIO: REACT Y VITE 44
2.4.3. SERVIDOR DE APLICACIÓN: NODE.JS Y EXPRESS 45
2.4.4. RECEPCIÓN Y PROCESAMIENTO TEMPORAL DE ARCHIVOS 45
2.4.5. EXTRACCIÓN DE TEXTO: GOOGLE CLOUD DOCUMENT AI 46
2.4.6. ANÁLISIS DEL CONTENIDO: GOOGLE GEMINI 47
2.4.7. REGISTRO DE USO Y EVALUACIÓN: SUPABASE 47
2.4.8. PUBLICACIÓN Y MANTENIMIENTO DE LA APLICACIÓN 48
**2.5. MARCO AMBIENTAL 49**
2.5.1. REVISIÓN DE DOCUMENTOS EN FORMATO DIGITAL 49
2.5.2. RECURSOS NECESARIOS PARA EL PROCESAMIENTO 50
2.5.3. PROCESAMIENTO TEMPORAL Y ALMACENAMIENTO 51
2.5.4. ALCANCE DE LA VALORACIÓN AMBIENTAL 51

**3. DISEÑO DE LA INVESTIGACIÓN .....52**

**3.1. MODALIDAD DEL TRABAJO 52**
**3.2. ALCANCE DE LA INVESTIGACIÓN 53**
**3.3. ENFOQUE DE LA INVESTIGACIÓN 54**
**3.4. MÉTODO EMPLEADO 55**
**3.5. TÉCNICAS E INSTRUMENTOS UTILIZADOS 55**
**3.6. CONJUNTO DOCUMENTAL DE EVALUACIÓN 56**
**3.7. PROCEDIMIENTO DE EVALUACIÓN Y TRATAMIENTO DE LOS RESULTADOS 57**
**3.8. FASES DE DESARROLLO Y EVALUACIÓN 58**

**4. DESARROLLO DEL TRABAJO DE GRADO .....60**

**4.1. ANÁLISIS DE LA PROBLEMÁTICA Y REQUERIMIENTOS 60**
4.1.1. DELIMITACIÓN DE LA NECESIDAD 60
4.1.2. CATEGORÍAS DOCUMENTALES E INFORMACIÓN RELEVANTE 61
4.1.3. REQUERIMIENTOS FUNCIONALES 62
4.1.4. REQUERIMIENTOS NO FUNCIONALES 64
4.1.5. ALCANCE DE LAS RESPUESTAS Y PARTICIPACIÓN DEL USUARIO 65
4.1.6. RESULTADO DEL ANÁLISIS DE REQUERIMIENTOS 66
**4.2. DISEÑO DE LA ARQUITECTURA DEL SISTEMA 66**
4.2.1. ORGANIZACIÓN POR CAPAS 67
4.2.2. DISEÑO DE LA CAPA DE PRESENTACIÓN 67

4.2.3. ORGANIZACIÓN INTERNA DEL BACKEND 68
4.2.4. COMUNICACIÓN ENTRE LOS COMPONENTES 69
4.2.5. PROCESAMIENTO TEMPORAL DEL CONTENIDO DOCUMENTAL 70
4.2.6. DISEÑO DE LOS REGISTROS DE SEGUIMIENTO 70
4.2.7. INCORPORACIÓN DE LA REVISIÓN HUMANA 71
4.2.8. ORGANIZACIÓN DEL DESPLIEGUE Y LA CONFIGURACIÓN 72
**4.3. IMPLEMENTACIÓN Y DESARROLLO 74**
4.3.1. CONFIGURACIÓN DEL ENTORNO DE TRABAJO 74
4.3.2. CONSENTIMIENTO Y SELECCIÓN DEL DOCUMENTO 75
4.3.3. RECEPCIÓN Y VALIDACIÓN DE ARCHIVOS EN EL BACKEND 75
4.3.4. IMPLEMENTACIÓN DE LA EXTRACCIÓN DE TEXTO 76
4.3.5. CLASIFICACIÓN DOCUMENTAL CON GEMINI 77
4.3.6. ANÁLISIS ESPECIALIZADO Y ORGANIZACIÓN DE LA RESPUESTA 77
4.3.7. CONSTRUCCIÓN DE LA VISTA DE RESULTADOS 78
4.3.8. REGISTRO DEL ANÁLISIS, CONFIRMACIÓN Y ENCUESTA 79
4.3.9. IMPLEMENTACIÓN DEL PANEL ADMINISTRATIVO 80
4.3.10. DESPLIEGUE Y DOCUMENTACIÓN DE LA SOLUCIÓN 80
**4.4. PRUEBAS Y OPTIMIZACIONES 82**
4.4.1. AJUSTE DE LA COMUNICACIÓN CON EL SERVICIO OCR 82
4.4.2. ADECUACIÓN DEL PROCESAMIENTO DE PDF EXTENSOS 83
4.4.3. SEPARACIÓN DE LA CLASIFICACIÓN Y EL ANÁLISIS ESPECIALIZADO 83
4.4.4. AJUSTES EN LA GENERACIÓN Y PREPARACIÓN DE LAS RESPUESTAS 84
4.4.5. COMPROBACIÓN DEL FLUJO COMPLETO 85
4.4.6. ALCANCE DE LAS OPTIMIZACIONES REALIZADAS 85
4.4.7. VERSIÓN UTILIZADA PARA LA EVALUACIÓN 86

**5. RESULTADOS ..... 86**

**5.1. APLICACIÓN WEB OBTENIDA 87**
5.1.1. INTERFAZ DE CARGA Y ESTADO DEL PROCESAMIENTO 87
5.1.2. PRESENTACIÓN DEL DOCUMENTO Y DEL ANÁLISIS 88
5.1.3. ORGANIZACIÓN DE LA INFORMACIÓN GENERADA 89
5.1.4. REVISIÓN Y FINALIZACIÓN DEL ANÁLISIS 89
5.1.5. ALCANCE DEL PRODUCTO OBTENIDO 90
**5.2. CONJUNTO DOCUMENTAL UTILIZADO PARA LA EVALUACIÓN 92**
5.2.1. ORIGEN DE LOS DOCUMENTOS 92
5.2.2. DISTRIBUCIÓN POR CATEGORÍA DOCUMENTAL 93
5.2.3. FORMATO, EXTENSIÓN Y CALIDAD VISUAL 94
5.2.4. ORGANIZACIÓN DE LA INFORMACIÓN PARA LA EVALUACIÓN 95
5.2.5. CONFIDENCIALIDAD Y ALCANCE DEL CONJUNTO 95
**5.3. RESULTADOS DE LA EVALUACIÓN MANUAL 96**
5.3.1. CLASIFICACIÓN DOCUMENTAL 97
5.3.2. CALIDAD DEL TEXTO OBTENIDO MEDIANTE OCR 98
5.3.3. EXTRACCIÓN DE INFORMACIÓN 98
5.3.4. ALERTAS Y OBSERVACIONES SOBRE EL ANÁLISIS 99
5.3.5. UTILIDAD DE LAS RECOMENDACIONES Y CORRECCIÓN HUMANA 100

**5.4. RESULTADOS RELACIONADOS CON EL TIEMPO DE PROCESAMIENTO 101**
5.4.1. COMPONENTES DEL TIEMPO EVALUADO 101
5.4.2. AHORRO PROMEDIO OBTENIDO 102
5.4.3. CÁLCULO DE LA REDUCCIÓN PORCENTUAL 103
5.4.4. PARTICIPACIÓN DE LA REVISIÓN HUMANA 103
5.4.5. ALCANCE DE LOS RESULTADOS DE TIEMPO 104
**5.5. MÉTRICAS REGISTRADAS AUTOMÁTICAMENTE POR MIDOCIA 105**
5.5.1. INFORMACIÓN REGISTRADA DURANTE EL USO 106
5.5.2. CONFIRMACIÓN DE LA REVISIÓN HUMANA 107
5.5.3. CALIFICACIÓN DE UTILIDAD 107
5.5.4. ESTIMACIÓN DEL TIEMPO DE REVISIÓN MANUAL 108
5.5.5. TIEMPO AUTOMÁTICO DE PROCESAMIENTO 108
5.5.6. CONSULTA E INTERPRETACIÓN DEL PANEL ADMINISTRATIVO 109
**5.6. ANÁLISIS GENERAL DE LOS RESULTADOS 110**
5.6.1. RELACIÓN ENTRE CLASIFICACIÓN, EXTRACCIÓN Y PRESENTACIÓN 111
5.6.2. RELACIÓN ENTRE AHORRO DE TIEMPO Y CALIDAD DEL ANÁLISIS 111
5.6.3. UTILIDAD DE LAS RECOMENDACIONES Y PARTICIPACIÓN HUMANA 112
5.6.4. MANEJO DE LA INFORMACIÓN Y SEGUIMIENTO DEL USO 113
5.6.5. ALCANCE DE LA EVALUACIÓN Y APORTE DEL PROYECTO 113

**6. CONCLUSIONES ..... 114**

**7. RECOMENDACIONES ..... 117**

**8. REFERENCIAS BIBLIOGRÁFICAS ..... 120**

**9. APÉNDICES ..... 122**

**9.1. APÉNDICE A. INSTRUMENTO DE EVALUACIÓN DE MIDOCIA 122**
**9.2. APÉNDICE B. ENCUESTA DE UTILIDAD 123**

**10. ANEXOS ..... 123**

**10.1. ANEXO A. MANUAL DE USUARIO DE MIDOCIA 123**
**10.2. ANEXO B. MANUAL DE INSTALACIÓN Y MANTENIMIENTO DE MIDOCIA 123**
**10.3. ANEXO C. MATRIZ COMPLETA DE VALIDACIÓN Y EVALUACIÓN DE MIDOCIA 124**

# LISTA DE ILUSTRACIONES

**ILUSTRACIÓN 1 ARQUITECTURA GENERAL DE MIDOCIA..... 73**
**ILUSTRACIÓN 2 FLUJO DE PROCESAMIENTO DE MIDOCIA ..... 81**
**ILUSTRACIÓN 3 INTERFAZ PRINCIPAL DE MIDOCIA..... 91**
**ILUSTRACIÓN 4 VISUALIZACIÓN DEL DOCUMENTO Y RESULTADO DEL ANÁLISIS ..... 91**
**ILUSTRACIÓN 5 COMPARACIÓN DEL TIEMPO PROMEDIO DE REVISIÓN..... 105**
**ILUSTRACIÓN 6 PANEL DE ADMINISTRACIÓN CON MÉTRICAS DE EVALUACIÓN DE MIDOCIA.....110**

# LISTA DE TABLAS

**TABLA 1 FASES DE DESARROLLO Y EVALUACIÓN ..... 59**
**TABLA 2 REQUERIMIENTOS FUNCIONALES DE MIDOCIA ..... 62**
**TABLA 3 COMPOSICIÓN DEL CONJUNTO DE EVALUACIÓN ..... 93**
**TABLA 4 RESULTADOS GENERALES DE LA EVALUACIÓN MANUAL ..... 96**
**TABLA 5 MÉTRICAS REGISTRADAS POR EL APLICATIVO ..... 105**
**TABLA 6 C1 FICHA DE VALIDACIÓN ..... 125**
**TABLA 7 C2 BANCO DOCUMENTAL: IDENTIFICACIÓN Y ORIGEN ..... 126**
**TABLA 8 C3 BANCO DOCUMENTAL: CARACTERÍSTICAS Y DATOS DE EVALUACIÓN ..... 130**
**TABLA 9 C4 BANCO DOCUMENTAL: OBSERVACIONES ..... 132**
**TABLA 10 C5 EVALUACIÓN: RESULTADOS DE CLASIFICACIÓN DOCUMENTAL Y CALIDAD DEL OCR ..... 134**
**TABLA 11 C6 EVALUACIÓN: RESULTADOS DE EXTRACCIÓN DE CAMPOS Y DETECCIÓN DE ALERTAS ..... 135**
**TABLA 12 C7 EVALUACIÓN: UTILIDAD DE LAS RECOMENDACIONES Y CORRECCIÓN HUMANA REQUERIDA ..... 136**
**TABLA 13 C8 EVALUACIÓN: TIEMPOS DE PROCESAMIENTO, REVISIÓN Y AHORRO OBTENIDO ..... 137**
**TABLA 14 C9 EVALUACIÓN: OBSERVACIONES CUALITATIVAS DE LA EVALUACIÓN ..... 139**
**TABLA 15 C10 CAMPOS EVALUADOS: VALIDACIÓN DETALLADA DE LOS CAMPOS ESPERADOS POR DOCUMENTO ..... 144**
**TABLA 16 C11 RESUMEN GENERAL DE LOS RESULTADOS DE LA EVALUACIÓN DE MIDOCIA ..... 191**

## RESUMEN EJECUTIVO

El presente proyecto tuvo como objetivo desarrollar una aplicación web orientada al apoyo de la gestión documental en Mipymes, mediante la integración de reconocimiento óptico de caracteres e inteligencia artificial, con el propósito de reducir los tiempos de revisión y apoyar la toma de decisiones. Como resultado se desarrolló **MiDocIA**, una aplicación capaz de recibir documentos digitalizados, extraer su contenido, clasificarlos, identificar información relevante y alertas, y generar sugerencias de respuesta o recomendaciones de acción.

El desarrollo se realizó mediante fases de análisis, diseño, implementación, integración y evaluación. La arquitectura incorporó los principios de **Zero Storage**, evitando el almacenamiento persistente del contenido documental, y **Human-in-the-Loop**, manteniendo al usuario como responsable de revisar y confirmar los resultados generados.

La evaluación se realizó con **32 documentos**, de los cuales 20 correspondieron a documentos reales de origen empresarial tratados de manera confidencial y 12 a documentos sintéticos controlados. La aplicación alcanzó una precisión de clasificación del **93,75 %**, una precisión promedio de extracción de información del **82,33 %** y una valoración promedio del OCR de **4,97 sobre 5**. Además, el uso de MiDocIA permitió reducir en promedio **50,65 %** el tiempo empleado frente al procedimiento manual evaluado. Los resultados permiten concluir que MiDocIA funciona como una herramienta de apoyo para la gestión documental, facilitando la organización inicial de la información y reduciendo el tiempo de revisión, sin sustituir el criterio humano en la decisión final.

**PALABRAS CLAVE.** Inteligencia artificial, gestión documental, OCR, Zero Storage, Human-in-the-Loop.

# INTRODUCCIÓN

La gestión documental comprende actividades que van más allá de recibir y guardar archivos. En las áreas administrativas también es necesario leer su contenido, reconocer el tipo de documento, identificar los datos relevantes y determinar qué acción debe realizarse. Una factura, un contrato o una solicitud pueden requerir revisiones diferentes, aunque formen parte del trabajo cotidiano de una misma organización. Por esta razón, el proyecto se centró en apoyar la comprensión inicial de los documentos y la preparación de posibles respuestas.

El presente trabajo surgió de una necesidad planteada en el banco de proyectos de las Unidades Tecnológicas de Santander, relacionada con el uso de inteligencia artificial para apoyar la gestión documental en las Mipymes de Bucaramanga. A partir de esta problemática se propuso desarrollar una aplicación web que facilitara la clasificación y el análisis de documentos, con el propósito de reducir el tiempo de revisión y aportar información útil para la toma de decisiones.

Como resultado se desarrolló MiDocIA, una aplicación que recibe documentos digitalizados y presenta al usuario un análisis de su contenido. El sistema permite trabajar con siete categorías documentales: factura, contrato, PQRS, documento judicial, carta, cotización y otro. Esta última categoría permite analizar documentos que no corresponden directamente a las seis anteriores, conservando un tratamiento general de la información.

El procesamiento integra reconocimiento óptico de caracteres, conocido como OCR, e inteligencia artificial. Mediante Google Cloud Document AI se obtiene el texto del archivo cargado. Posteriormente, Google Gemini utiliza ese contenido para clasificar el documento e identificar información de interés según su categoría. El resultado incluye un resumen, datos relevantes, alertas y una sugerencia de respuesta o recomendación de acción. La interfaz permite consultar el documento original junto con el análisis para facilitar su revisión.

Durante el desarrollo se tuvieron en cuenta la privacidad del contenido y la participación del usuario. Para el tratamiento de los archivos se incorporó el principio Zero Storage, mediante el cual MiDocIA procesa temporalmente los documentos en memoria y evita almacenar de forma persistente el archivo original, el texto extraído y el análisis generado dentro de su infraestructura. De manera separada, la aplicación registra metadatos de uso y datos autorizados para el seguimiento y la evaluación del sistema.

La participación del usuario se mantuvo mediante el principio Human-in-the-Loop. Los resultados producidos por la inteligencia artificial se presentan para que la persona los contraste con el documento y determine si pueden utilizarse o requieren ajustes. La confirmación forma parte de este proceso de revisión. Las sugerencias generadas sirven como apoyo, mientras que la responsabilidad de decidir cómo actuar permanece en el usuario.

El trabajo se organizó en fases de análisis, diseño, implementación, integración y evaluación. En estas etapas se definieron las funciones de la aplicación, se estructuró su arquitectura y se desarrollaron los componentes necesarios para procesar los documentos y presentar los resultados. También se realizaron ajustes relacionados con la extracción de texto, el análisis mediante inteligencia artificial y el funcionamiento del flujo completo.

La evaluación se realizó con 32 documentos: 20 reales de origen empresarial, tratados de manera confidencial, y 12 sintéticos elaborados para las pruebas. El instrumento fue diseñado por el autor del proyecto y diligenciado por la persona designada por la empresa colaboradora. Se valoraron la clasificación, la calidad del texto extraído, la identificación de campos, las alertas, la utilidad de las recomendaciones y los tiempos empleados. Los registros automáticos de la aplicación complementaron esta información.

El informe presenta la problemática, los objetivos y los fundamentos del proyecto. Luego describe la metodología y el desarrollo de MiDocIA, seguido de los resultados

de la evaluación y su análisis. Finalmente, expone las conclusiones y recomendaciones, e identifica los manuales y la matriz de validación que acompañan el trabajo.

# 1. DESCRIPCIÓN DEL TRABAJO DE INVESTIGACIÓN

## 1.1. PLANTEAMIENTO DEL PROBLEMA

El proyecto parte de una necesidad identificada en el banco de proyectos de las Unidades Tecnológicas de Santander y planteada por la directora del trabajo de grado, relacionada con el procesamiento de documentos en las áreas administrativas de las Mipymes de Bucaramanga. La problemática considera el tiempo y la atención que requieren actividades como leer un documento, reconocer su propósito, identificar información relevante y establecer una posible respuesta o acción.

Los documentos que puede recibir una organización tienen características diferentes. Una factura contiene información sobre conceptos, valores y condiciones de pago; un contrato establece acuerdos y obligaciones entre las partes; una cotización presenta una oferta comercial; y una PQRS comunica una petición, queja, reclamo o solicitud que debe ser revisada. De igual manera, las cartas y los documentos judiciales pueden incluir instrucciones, requerimientos o fechas que necesitan atención. Esta diversidad hace necesario adaptar la revisión al contenido y al propósito de cada documento.

La dificultad no se limita a identificar la categoría documental. Después de reconocer el tipo de archivo, la persona encargada debe localizar los datos que permiten comprenderlo y decidir qué hacer. Según el caso, puede necesitar verificar quién lo envía, a quién está dirigido, qué se solicita, cuáles son los valores relacionados o qué compromisos se mencionan. Parte de esta información puede encontrarse

distribuida entre párrafos, tablas o distintas páginas, lo que exige revisar el contenido y relacionar sus elementos.

Cuando estas actividades se realizan de manera manual, el tiempo empleado depende de factores como la extensión del documento, la claridad de su presentación y la cantidad de información que debe verificarse. Además, una revisión incompleta puede llevar a pasar por alto una condición, interpretar de manera equivocada una solicitud o preparar una respuesta sin considerar todos los datos disponibles. Para el proyecto, estas situaciones representan dificultades que una herramienta de apoyo documental puede ayudar a atender.

Disponer de un archivo digital tampoco resuelve por sí solo la necesidad de comprender su contenido. Un documento escaneado o una imagen pueden consultarse en pantalla, pero la persona todavía debe leerlos, ubicar los datos importantes y determinar su relación con la actividad que está realizando. Por esta razón, la problemática abordada comprende tanto la obtención del texto como su organización y análisis para facilitar la revisión.

Otra parte del proceso corresponde a la preparación de respuestas o recomendaciones. Después de leer un documento, el usuario puede necesitar redactar una comunicación, solicitar información adicional o establecer qué aspectos requieren verificación. Esta actividad exige interpretar el contenido y seleccionar la información que resulta pertinente. Una sugerencia inicial puede apoyar esa tarea, siempre que la persona tenga la posibilidad de contrastarla con el documento y ajustar lo que considere necesario.

El uso de inteligencia artificial para apoyar estas actividades también plantea una dificultad: los resultados generados pueden contener omisiones o interpretaciones que no correspondan completamente al contenido original. Por ello, la solución debía permitir que el usuario consultara el documento junto con el análisis y mantuviera la responsabilidad de revisar la información antes de utilizarla. Esta

necesidad se incorporó al proyecto mediante el principio Human-in-the-Loop, que conserva la participación humana en el proceso.

La privacidad fue otro criterio considerado desde el planteamiento de la solución. Los documentos empresariales pueden contener información personal, comercial o administrativa cuyo tratamiento requiere cuidado. En consecuencia, se definió que la aplicación debía realizar el análisis sin conservar de forma persistente el archivo original ni su contenido dentro de la infraestructura de MiDocIA. Este criterio dio lugar a la incorporación del principio Zero Storage y a la separación entre el contenido documental y los metadatos autorizados para seguimiento.

Los beneficiarios potenciales de la solución son las Mipymes y, en particular, las personas que desarrollan actividades administrativas relacionadas con la revisión de documentos. Sin embargo, el proyecto no realizó una caracterización directa de las empresas de Bucaramanga ni aplicó entrevistas o encuestas durante su etapa inicial. El punto de partida fue la problemática definida en el banco de proyectos. La participación posterior de una empresa colaboradora permitió evaluar el funcionamiento de la herramienta con documentos reales, dentro del alcance establecido para las pruebas.

A partir de esta problemática, se planteó el desarrollo de una aplicación web que apoyara la clasificación documental, la extracción de información, la identificación de alertas y la preparación de posibles respuestas. Su alcance se concentró en facilitar el análisis inicial y la revisión por parte del usuario. Para valorar el aporte de la solución, se estableció la necesidad de evaluar tanto la calidad de los resultados como el tiempo empleado en el proceso.

Con base en lo anterior, se formuló la siguiente pregunta de investigación: ¿Cómo podría una aplicación web con inteligencia artificial optimizar la gestión documental y apoyar la toma de decisiones en las Mipymes de Bucaramanga, garantizando la privacidad de la información?

## 1.2. JUSTIFICACIÓN

El desarrollo de MiDocIA se justifica en la posibilidad de utilizar reconocimiento óptico de caracteres e inteligencia artificial para apoyar la revisión de documentos en las Mipymes. A partir de la necesidad planteada en el banco de proyectos de las Unidades Tecnológicas de Santander, se propuso una aplicación que facilitara la identificación de información relevante y la preparación de posibles respuestas. Su aporte consiste en ofrecer al usuario un análisis inicial que pueda revisar y utilizar como apoyo en sus actividades administrativas.

Desde el punto de vista práctico, la aplicación reúne en un mismo proceso la extracción del texto, la clasificación documental, la identificación de datos, la presentación de alertas y la generación de recomendaciones. Esta integración permite que el usuario consulte información organizada según el documento que está revisando. Por ejemplo, en una cotización pueden resultar relevantes los productos, los valores y las condiciones comerciales, mientras que en un contrato la atención puede centrarse en las partes, las obligaciones y la vigencia. El tratamiento por categoría permite orientar el análisis hacia estas diferencias.

La presentación del documento original junto con los resultados también tiene una utilidad dentro del proceso. El usuario puede contrastar los datos extraídos con su fuente, reconocer información incompleta y determinar qué aspectos necesitan una revisión más detallada. De esta manera, la aplicación proporciona un punto de partida para el análisis, conservando la posibilidad de verificar el contenido antes de emplearlo en una respuesta o decisión.

Desde el componente tecnológico, el proyecto permitió integrar servicios de OCR y modelos de lenguaje dentro de una aplicación web funcional. El trabajo comprendió la recepción de archivos, el procesamiento temporal de su contenido, la comunicación con servicios externos y la presentación de los resultados en una interfaz. El aporte técnico se encuentra en la articulación de estos componentes

dentro de un flujo orientado a una necesidad documental concreta, incluyendo el manejo de las condiciones y dificultades encontradas durante su implementación.

La elección de una aplicación web facilita su acceso desde un navegador y evita que el usuario deba instalar localmente las herramientas utilizadas para el análisis. El procesamiento se realiza mediante la infraestructura y los servicios configurados para MiDocIA. Esta característica resulta pertinente para el alcance del proyecto, porque permite ofrecer las funciones de análisis mediante una interfaz de uso directo. Su utilización requiere conexión a Internet y depende de la disponibilidad de los servicios externos que intervienen en el proceso.

El tratamiento de la información constituye otra razón que orientó el diseño. La aplicación incorpora el principio Zero Storage para evitar la conservación persistente del archivo original, el texto extraído y el análisis generado dentro de la infraestructura de MiDocIA. Esta decisión permite delimitar qué información se procesa temporalmente y qué datos se registran para seguimiento. Los metadatos autorizados de uso se gestionan de manera separada del contenido documental, manteniendo una distinción entre el análisis del archivo y la evaluación del funcionamiento del sistema.

La revisión humana es igualmente necesaria para que los resultados puedan utilizarse de forma responsable. Una recomendación generada por inteligencia artificial puede requerir correcciones, información adicional o ajustes según el contexto del documento. Por ello, MiDocIA incorpora el principio Human-in-the-Loop y mantiene al usuario como responsable de contrastar la información y decidir cómo utilizarla. Esta participación permite aprovechar el apoyo tecnológico conservando el criterio de la persona encargada de la actividad.

En el ámbito económico y organizacional, el beneficio esperado se relaciona con la reducción del tiempo dedicado a la lectura inicial, la búsqueda de datos y la preparación de respuestas. Un menor tiempo en estas actividades puede permitir que el usuario dedique mayor atención a verificar condiciones, resolver dudas y

atender las situaciones identificadas. En este proyecto, dicho aporte se evaluó mediante la comparación de los tiempos registrados durante las pruebas. El alcance de esa medición corresponde al conjunto documental utilizado y no constituye una estimación de ahorro económico para todas las Mipymes.

La evaluación también aporta información sobre las posibilidades y limitaciones de la solución. El uso de documentos reales de origen empresarial permitió observar su comportamiento frente a contenido suministrado por una empresa colaboradora. Los documentos sintéticos complementaron las pruebas con escenarios controlados y valores esperados previamente definidos. La combinación de ambas fuentes permitió revisar la clasificación, la extracción de información y la utilidad de las recomendaciones, además de identificar aspectos que requieren mejoras.

Para las Unidades Tecnológicas de Santander y el programa de Ingeniería de Sistemas, el proyecto representa una aplicación de conocimientos de análisis de requerimientos, arquitectura de software, desarrollo web e integración de servicios. La elaboración de los manuales y del instrumento de evaluación complementa el producto desarrollado, al documentar su funcionamiento y los criterios utilizados para valorar sus resultados. Estos materiales pueden servir como referencia para otros trabajos académicos relacionados con procesamiento documental e inteligencia artificial.

En cuanto al componente ambiental, MiDocIA permite realizar la revisión de los documentos desde una interfaz digital y puede complementar prácticas de trabajo que ya utilizan archivos electrónicos. Sin embargo, el proyecto no midió una reducción del consumo de papel ni el impacto energético de los servicios empleados. Por esta razón, su justificación ambiental se limita a la posibilidad de apoyar la consulta y el análisis digital, sin atribuir beneficios cuantitativos que no fueron evaluados.

En conjunto, el proyecto se justifica por el desarrollo y la evaluación de una herramienta que apoya tareas documentales mediante una integración tecnológica

concreta. Su utilidad se relaciona con facilitar la revisión, organizar la información relevante y proponer posibles acciones, manteniendo la supervisión del usuario y un tratamiento temporal del contenido procesado.

## 1.3. OBJETIVOS

### *1.3.1. OBJETIVO GENERAL*

Desarrollar una aplicación web orientada a la recomendación de acciones para la gestión documental en Mipymes, mediante la integración de visión artificial y modelos de lenguaje natural, para optimizar los tiempos de respuesta y apoyar la toma de decisiones.

### *1.3.2. OBJETIVOS ESPECÍFICOS*

**1.3.2.1** Analizar las necesidades de gestión documental definidas a partir de la problemática planteada y los tipos de documentos contemplados en el proyecto, con el propósito de establecer los requerimientos funcionales y de información de la aplicación.

**1.3.2.2** Diseñar la arquitectura de software y la interfaz de usuario de la aplicación, incorporando un flujo de procesamiento temporal basado en Zero Storage y mecanismos Human-in-the-Loop orientados a preservar la privacidad de la información y mantener la supervisión humana.

**1.3.2.3** Implementar un módulo de procesamiento documental capaz de recibir archivos digitalizados, extraer su contenido mediante

reconocimiento óptico de caracteres y gestionar diferentes formatos documentales sin almacenamiento persistente del contenido.

**1.3.2.4** Integrar un motor de inteligencia artificial capaz de clasificar documentos, extraer información relevante, identificar alertas y generar sugerencias de respuesta o recomendaciones de acción para apoyar la toma de decisiones del usuario.

**1.3.2.5** Evaluar el desempeño y la utilidad de la aplicación mediante pruebas con documentos reales provenientes de un entorno empresarial y un banco controlado de documentos sintéticos representativos de las categorías soportadas, valorando la clasificación documental, la extracción de información, la identificación de alertas, la utilidad de las recomendaciones y el tiempo de procesamiento.

## 1.4. ESTADO DEL ARTE

La revisión de antecedentes se centró en trabajos relacionados con automatización documental, extracción de información y generación de respuestas mediante inteligencia artificial. Se consideraron investigaciones de Perú, Ecuador y Colombia que permiten conocer diferentes formas de abordar estas actividades. Para establecer su relación con MiDocIA se tuvieron en cuenta el problema atendido, el enfoque tecnológico descrito y el alcance de cada propuesta.
Los antecedentes seleccionados comprenden soluciones aplicadas a compras públicas, gestión documental empresarial y atención de peticiones, quejas y reclamos, además de una monografía sobre ciencia de datos y Big Data. Esta variedad permite comparar aplicaciones concretas con una revisión conceptual del

tema. La información presentada se apoya en las descripciones y los resúmenes disponibles en los registros académicos consultados.

### *1.4.1. ANTECEDENTES INTERNACIONALES*

Peña Del Aguila (2024), de la Universidad Tecnológica del Perú, desarrolló el trabajo denominado Automatización inteligente en la gestión documental de compras públicas: un enfoque basado en inteligencia artificial (IA). La investigación aborda las dificultades para extraer información de documentos PDF con estructuras diferentes utilizados en el Sistema Perú Compras. La propuesta integra reconocimiento óptico de caracteres y procesamiento de lenguaje natural para analizar órdenes de compra y organizar sus datos.
El resumen describe una metodología cuantitativa, descriptiva y transversal, con una muestra censal de 30 trabajadores. Se aplicó una encuesta antes y después del uso del software para valorar aspectos de precisión, eficiencia y satisfacción. El autor reporta mejoras en la extracción y los tiempos de procesamiento, aunque el registro consultado no presenta porcentajes que permitan una comparación numérica con MiDocIA.Este antecedente se relaciona con el proyecto por la integración de extracción de texto y análisis del contenido. La comparación permite reconocer una diferencia de alcance: MiDocIA contempla siete categorías documentales y evalúa sus resultados mediante un instrumento aplicado a documentos reales y sintéticos. Por su parte, Ruales Aldás (2024), de la Universidad Politécnica Salesiana de Ecuador, presentó el trabajo Análisis, diseño y desarrollo de un sistema de automatización de la gestión documental en empresas a través de una interfaz gráfica intuitiva. Su objetivo se orienta a facilitar el acceso, la organización, la conservación y la difusión de documentos empresariales.
El trabajo describe una metodología aplicada con enfoque mixto. Incluye modelamiento de procesos mediante BPMN 2.0 y herramientas como Power Apps, Power Automate, SharePoint, Azure SQL Database y Power BI. La propuesta

relaciona la automatización con la interacción del usuario y la consulta de información.

Para MiDocIA, este antecedente resulta pertinente porque permite considerar la interfaz como parte del proceso documental. Como comparación de alcance, el proyecto de Ruales Aldás contempla la conservación y organización de documentos, mientras que MiDocIA se concentra en su análisis temporal y en la presentación de resultados para revisión. Esta diferencia ayuda a precisar que MiDocIA funciona como asistente de análisis y no como repositorio de archivos empresariales.

### *1.4.2. ANTECEDENTES NACIONALES*

En Colombia, Fernández Villalobos (2025), de la Universidad Nacional Abierta y a Distancia, elaboró la monografía Evolución de la ciencia de datos y Big Data en la gestión documental. El trabajo estudia la transición desde métodos tradicionales hacia prácticas apoyadas en tecnologías digitales y revisa conceptos, técnicas y herramientas relacionadas con el tratamiento de la información. Según el resumen institucional, la monografía incluye casos de organizaciones que han incorporado estas tecnologías y examina beneficios, dificultades y oportunidades asociados con el acceso, la seguridad y la eficiencia documental. Su naturaleza corresponde a una revisión académica, por lo que se utiliza como antecedente conceptual. La relación con MiDocIA se establece en el aprovechamiento del contenido documental para apoyar la consulta y la toma de decisiones. Esta relación no implica que MiDocIA implemente una arquitectura de Big Data. El proyecto desarrollado procesa los documentos cargados por el usuario y devuelve un análisis para su revisión, dentro de un alcance distinto al estudio general de grandes volúmenes de información.

En el contexto local, Guerrero Urango (2025), de la Universidad Autónoma de Bucaramanga, desarrolló el trabajo Sistema automatizado de gestión y respuesta

de PQR en propiedades horizontales con inteligencia artificial y automatización de procesos. La investigación aborda dificultades relacionadas con los tiempos de atención y la trazabilidad de las solicitudes.

El resumen institucional señala que se analizaron más de 24.000 registros históricos y se integraron clasificación de texto, automatización de procesos y generación de respuestas apoyada en recuperación de información, conocida como RAG. También describe mecanismos de asignación de responsables, actualización de estados e integración con Odoo para seguimiento.

Este antecedente tiene relación con MiDocIA por la clasificación de información y el apoyo en la preparación de respuestas. La comparación muestra alcances diferentes: MiDocIA analiza varias categorías documentales y presenta recomendaciones al usuario; el trabajo de Guerrero Urango comprende además la gestión del recorrido de las PQR dentro de una organización. Por ello, las métricas de ambos proyectos deben interpretarse según las actividades que cada uno evalúa.

### *1.4.3. RELACIÓN DE LOS ANTECEDENTES CON MIDOCIA*

A partir de esta revisión se identifican distintas necesidades que pueden atenderse mediante automatización documental. Para el presente proyecto, la comparación permite distinguir entre obtener información de un archivo, organizar documentos para su conservación y gestionar una solicitud hasta su atención. Estas actividades pueden relacionarse dentro de una organización, pero tienen propósitos y criterios de evaluación diferentes.

MiDocIA se ubica principalmente en el análisis inicial del documento y en el apoyo a la preparación de una respuesta o acción. Su funcionamiento comprende la extracción del texto, la clasificación, la identificación de información relevante y la

presentación de recomendaciones. La revisión humana forma parte de ese alcance, porque el resultado debe contrastarse con el contenido original antes de utilizarse. La comparación también orienta la interpretación de los resultados. En MiDocIA, medir el tiempo de procesamiento automático permite conocer cuánto tarda el sistema en producir el análisis. Sin embargo, para valorar el tiempo total de la actividad es necesario considerar también la revisión realizada por el usuario. De manera semejante, una clasificación correcta no demuestra por sí sola que todos los campos hayan sido extraídos adecuadamente. Por esta razón, el instrumento del proyecto contempla dimensiones de evaluación separadas.
En cuanto al tratamiento de la información, MiDocIA adopta como decisión de diseño el procesamiento temporal del contenido y el registro separado de metadatos autorizados. Esta característica delimita su funcionamiento y debe explicarse mediante la arquitectura implementada. La información consultada sobre los antecedentes no permite establecer que sus soluciones carezcan de mecanismos de privacidad o supervisión humana, por lo que la comparación se limita a los alcances descritos por sus autores.

El aporte de MiDocIA se plantea, entonces, en la integración de estas funciones dentro de una aplicación web orientada a documentos de uso administrativo en Mipymes. La evaluación con documentos reales y sintéticos permite examinar su desempeño, reconocer los ajustes que requieren las recomendaciones y documentar las limitaciones encontradas. Los antecedentes sirven como referencia para situar ese trabajo dentro de otras experiencias de automatización y análisis documental.

# 2. MARCO REFERENCIAL

## 2.1. MARCO TEÓRICO

El marco teórico presenta los fundamentos que permiten comprender el procesamiento documental realizado por MiDocIA. Se abordan el reconocimiento óptico de caracteres, los modelos de lenguaje, las instrucciones utilizadas para orientar el análisis, la supervisión humana y la privacidad desde el diseño. Estos elementos se relacionan con las funciones de la aplicación y con los criterios empleados para evaluar sus resultados.

### *2.1.1.* RECONOCIMIENTO ÓPTICO DE CARACTERES Y PROCESAMIENTO DOCUMENTAL

El reconocimiento óptico de caracteres, conocido como OCR, permite obtener texto procesable a partir de documentos digitalizados e imágenes. Esta operación constituye una etapa inicial del análisis documental, porque transforma el contenido visual del archivo en información que puede ser utilizada por otros componentes del sistema.

La documentación de Google Cloud (2026) distingue entre la obtención del texto y la extracción de campos específicos. La primera permite recuperar el contenido escrito y aspectos de su distribución; la segunda busca identificar datos concretos, como nombres, valores o fechas. Esta distinción resulta útil para comprender que disponer del texto de un documento no equivale a haber identificado correctamente toda su información relevante.

En MiDocIA, Google Cloud Document AI se utiliza para obtener el texto que posteriormente recibe el componente de inteligencia artificial. La aplicación separa estas tareas para que el análisis se realice sobre el contenido extraído. Por ejemplo,

recuperar una cifra mediante OCR constituye una operación diferente de determinar si corresponde a un subtotal, un impuesto o un valor total.

Esta separación también se conserva en la evaluación del proyecto. La calidad del OCR se valora de manera independiente de la extracción de campos. Así, un documento puede presentar un texto legible y, al mismo tiempo, contener datos que el análisis posterior identifique de forma parcial o no incluya en el resultado.

### *2.1.2.* MODELOS DE LENGUAJE Y ANÁLISIS DEL CONTENIDO

Los modelos de lenguaje trabajan con unidades de texto denominadas tokens, que pueden representar palabras, fragmentos de palabras o caracteres. A partir del contexto disponible, estiman posibles secuencias y generan una salida. Este funcionamiento permite realizar tareas como producir texto y elaborar resúmenes de documentos (Google for Developers, 2026a).

El contexto tiene importancia porque una expresión puede adquirir distintos significados según la información que la acompaña. En el análisis de documentos, una fecha puede corresponder a la expedición del archivo, al vencimiento de una obligación o al inicio de una actividad. Para interpretar su función es necesario relacionarla con el contenido cercano.

Los modelos de lenguaje de gran tamaño utilizan arquitecturas capaces de representar relaciones entre elementos del texto. La documentación de Google for Developers (2026b) explica que los mecanismos de atención permiten considerar la relevancia de unos elementos respecto de otros. También advierte que estos modelos pueden generar información incorrecta o no sustentada, incluso cuando la respuesta presenta una redacción coherente.

En MiDocIA, el modelo de lenguaje se utiliza para clasificar el documento, identificar datos relevantes y generar un resumen, alertas y recomendaciones. El proyecto integra un modelo existente mediante un servicio externo; no comprende el entrenamiento de un modelo propio desde cero.

Para interpretar sus resultados es necesario distinguir entre información extraída y contenido generado. Los nombres, valores o condiciones identificados deben poder contrastarse con el documento. Una recomendación, en cambio, constituye una propuesta elaborada a partir del análisis. Esta diferencia orienta la revisión del usuario y evita presentar una sugerencia como si fuera una afirmación incluida expresamente en el archivo original.

### *2.1.3.* INSTRUCCIONES PARA ORIENTAR EL ANÁLISIS

Las instrucciones proporcionadas a un modelo de lenguaje se conocen como prompts. Su función es comunicar la tarea que debe realizarse, el contexto que debe considerar y la forma esperada de la respuesta. Google for Developers (2025) explica que el diseño de estas instrucciones permite orientar la salida del modelo sin modificar sus parámetros, a diferencia de un proceso de entrenamiento o ajuste del modelo.

En el procesamiento documental, las instrucciones permiten precisar qué información debe buscarse según el tipo de documento. Una factura requiere identificar elementos distintos de los que resultan relevantes en una carta o un contrato. Por esta razón, el proyecto utiliza una clasificación inicial y un análisis posterior orientado a la categoría detectada.

La clasificación permite seleccionar las instrucciones que se aplicarán al contenido. Después, el análisis especializado solicita los campos y resultados definidos para esa categoría. Esta organización corresponde a una decisión de implementación de MiDocIA, mediante la cual se distribuyen las tareas entre dos etapas del procesamiento.

La revisión de los resultados permite valorar si esas instrucciones son adecuadas para los documentos evaluados. Cuando se presenta una omisión o una interpretación que no corresponde al contenido, es necesario examinar el

documento original, el texto obtenido mediante OCR y el resultado del análisis antes de atribuir el error a una sola causa.

### *2.1.4.* SUPERVISIÓN HUMANA Y USO DE LOS RESULTADOS

La participación humana es un componente relevante en los sistemas de inteligencia artificial que apoyan decisiones. El marco de gestión de riesgos del National Institute of Standards and Technology propone definir las responsabilidades de las personas que intervienen en estos sistemas y establecer procesos de supervisión acordes con su contexto de uso. También reconoce que las formas de interacción entre personas e inteligencia artificial pueden variar según las tareas y los riesgos involucrados (Tabassi, 2023).

En el proyecto, el principio Human-in-the-Loop se aplica a la revisión de los resultados antes de finalizar el análisis. El usuario tiene acceso al documento original y a la información presentada por MiDocIA, lo que le permite contrastar datos, identificar omisiones y decidir si la recomendación resulta pertinente.

Esta participación tiene una función concreta: conservar el criterio de la persona frente a una salida que puede requerir ajustes. Por ejemplo, una sugerencia de respuesta puede estar redactada de manera comprensible, pero omitir una condición del documento o asumir una intención que el remitente no expresó. En ese caso, corresponde al usuario corregirla antes de utilizarla. La confirmación incluida en la aplicación registra que se realizó la acción prevista para cerrar la revisión. Sin embargo, no constituye una comprobación automática de que todos los datos sean correctos. Por esta razón, en la evaluación de MiDocIA se distinguen la confirmación del proceso y la valoración de la calidad del análisis. La intervención del usuario tampoco se plantea como un mecanismo de entrenamiento automático. En el alcance documentado del proyecto, la revisión sirve para valorar y utilizar los resultados. No se implementó un proceso mediante el cual cada confirmación modificara los parámetros del modelo de inteligencia artificial.

### *2.1.5.* PRIVACIDAD DESDE EL DISEÑO Y PROCESAMIENTO TEMPORAL

El enfoque Privacy by Design plantea que la protección de la privacidad debe incorporarse desde la definición y construcción de una solución. Cavoukian (2012) desarrolla este enfoque mediante principios que incluyen la prevención, la privacidad como configuración predeterminada, su integración en el diseño, la protección durante el ciclo de vida de la información y la transparencia frente al usuario.

La autora también relaciona la privacidad con la definición de los propósitos del tratamiento y la reducción de la información recolectada a lo necesario. Para el diseño de software, estas orientaciones permiten examinar qué datos requiere una función, cómo se utilizan y qué conservación resulta necesaria para cumplir su propósito (Cavoukian, 2012).

En MiDocIA, este enfoque se relaciona con la decisión de procesar temporalmente el contenido documental. El principio denominado Zero Storage se utiliza en el proyecto para describir una arquitectura que evita guardar de forma persistente el archivo original, el texto obtenido mediante OCR y el análisis generado dentro de la infraestructura de la aplicación.

El alcance de esta decisión debe distinguirse del registro de metadatos de uso. MiDocIA conserva información autorizada para seguimiento y evaluación, como tiempos de procesamiento, categoría documental y confirmación de la revisión. Estos registros cumplen una función diferente de la conservación del contenido del documento.

El procesamiento temporal tampoco significa que la información permanezca exclusivamente en el dispositivo del usuario. El flujo implementado requiere enviarla a los servicios externos utilizados para OCR y análisis. Por tanto, la descripción de privacidad debe considerar tanto las decisiones de almacenamiento de MiDocIA como la participación de esos servicios.

De esta manera, Zero Storage se presenta como una medida arquitectónica del proyecto, vinculada con la reducción de la conservación del contenido. Su implementación debe explicarse mediante el funcionamiento real del sistema y no interpretarse como una garantía absoluta frente a cualquier riesgo de privacidad

### *2.1.6.* RELACIÓN DE LOS FUNDAMENTOS CON LA EVALUACIÓN

Los fundamentos anteriores permiten organizar la evaluación de MiDocIA según las funciones que intervienen en el procesamiento. La obtención del texto, la clasificación documental, la extracción de campos y la generación de recomendaciones producen resultados distintos, por lo que requieren criterios de revisión específicos.

En este proyecto, la evaluación del OCR permite valorar la legibilidad del contenido recuperado. La clasificación se contrasta con la categoría esperada del documento, mientras que la extracción se revisa mediante los campos definidos en el instrumento. La recomendación se valora considerando su utilidad y los ajustes necesarios antes de emplearla.

Esta organización permite interpretar los resultados con mayor precisión. Una clasificación correcta puede acompañarse de campos incompletos, y una recomendación útil puede necesitar modificaciones. Por ello, los indicadores se analizan de forma separada antes de establecer una valoración general del desempeño.

La medición del tiempo también incorpora la participación humana. El tiempo de procesamiento automático corresponde a una parte de la actividad, mientras que el tiempo total con MiDocIA incluye la revisión posterior del usuario. Esta distinción mantiene la evaluación relacionada con el propósito del proyecto: apoyar el trabajo documental conservando la intervención de la persona responsable.

## 2.2. MARCO CONCEPTUAL

El marco conceptual establece el significado de los términos utilizados para describir el funcionamiento y la evaluación de MiDocIA. Las definiciones se presentan de acuerdo con el alcance del proyecto, con el propósito de mantener una interpretación consistente de sus funciones, resultados y límites.

### 2.2.1. GESTIÓN DOCUMENTAL EN EL ALCANCE DEL PROYECTO

En MiDocIA, el apoyo a la gestión documental comprende actividades de lectura, clasificación, identificación de información relevante y preparación de posibles respuestas o acciones. La aplicación interviene en el análisis inicial del documento y presenta información organizada para facilitar su revisión.
Este alcance no comprende todas las actividades que puede desarrollar un sistema de gestión documental. MiDocIA no funciona como un archivo empresarial ni incorpora la conservación de expedientes, el control de versiones de documentos o el seguimiento completo de trámites. Su función se concentra en analizar el archivo cargado y entregar resultados que el usuario pueda contrastar con su contenido.

### 2.2.2. DOCUMENTO DE ENTRADA Y CONTENIDO DOCUMENTAL

El documento de entrada es el archivo que el usuario selecciona para su procesamiento. Puede corresponder a un PDF o a una imagen en alguno de los formatos admitidos por la aplicación. Este archivo constituye la fuente a partir de la cual se realiza el análisis.

El contenido documental comprende la información presente en sus páginas, incluyendo texto, cifras, nombres, fechas, condiciones y demás elementos relevantes. Para interpretar los resultados de MiDocIA es necesario conservar la

distinción entre el archivo original, el texto recuperado mediante OCR y el análisis generado por la inteligencia artificial. Cada uno corresponde a una etapa diferente del procesamiento.

### 2.2.3. CATEGORÍA Y CLASIFICACIÓN DOCUMENTAL

La categoría documental es el grupo utilizado para orientar el análisis de un documento según su propósito y contenido. En el proyecto se establecieron siete categorías: factura, contrato, PQRS, documento judicial, carta, cotización y otro. La categoría PQRS comprende peticiones, quejas, reclamos y solicitudes.
La clasificación documental corresponde a la identificación de la categoría que mejor representa el archivo analizado. En MiDocIA, esta identificación permite seleccionar las instrucciones utilizadas en la etapa posterior de extracción y generación de resultados.

La categoría otro permite tratar documentos que no corresponden directamente a las seis categorías específicas. Su utilización no significa que el archivo carezca de utilidad o que no pueda analizarse, sino que requiere un tratamiento general dentro del alcance previsto por la aplicación.

### 2.2.4. TEXTO EXTRAÍDO Y CAMPOS RELEVANTES

El texto extraído es el contenido recuperado por el componente OCR y utilizado como entrada para el análisis posterior. Su revisión permite establecer si la información escrita del documento fue obtenida de manera legible y suficientemente completa para continuar el procesamiento.
Los campos relevantes son los datos que se busca identificar dentro de ese contenido. Su selección depende de la categoría documental. Por ejemplo, pueden corresponder a las partes de un contrato, el valor de una factura, la vigencia de una cotización o el asunto de una carta. En la evaluación, estos campos se comparan

con la información esperada del documento. Se distingue entre datos identificados correctamente, datos recuperados de manera parcial y datos no detectados. Esta revisión permite valorar la extracción con mayor detalle que una apreciación general del análisis.

### 2.2.5. RESUMEN DOCUMENTAL

El resumen documental es una presentación breve del asunto principal y de los elementos necesarios para comprender el propósito del archivo. Dentro de MiDocIA, su función es proporcionar una orientación inicial que facilite al usuario reconocer de qué trata el documento.

El resumen debe interpretarse como una síntesis del contenido. Su lectura no sustituye la consulta de los detalles que puedan ser relevantes para una decisión. Por ejemplo, una descripción general de un contrato puede mencionar su objeto, pero la verificación de obligaciones específicas requiere revisar las condiciones correspondientes en el documento original.

### 2.2.6. ALERTA DOCUMENTAL

Una alerta documental es una observación presentada por la aplicación sobre un elemento que puede requerir atención. En el proyecto, puede estar relacionada con fechas, plazos, condiciones, información faltante o aspectos que conviene verificar antes de utilizar el resultado.La alerta funciona como una señal para orientar la revisión. Debe contrastarse con el documento y con el contexto de la actividad que realiza el usuario. Su aparición no demuestra por sí sola la existencia de un incumplimiento, ni su ausencia permite concluir que el documento no contiene aspectos relevantes que deban examinarse.

### 2.2.7. SUGERENCIA DE RESPUESTA Y RECOMENDACIÓN DE ACCIÓN

La sugerencia de respuesta es un texto generado a partir del análisis que puede servir como base para preparar una comunicación. La recomendación de acción es una orientación sobre una posible actividad posterior, como verificar un dato, solicitar información adicional o revisar una condición.Ambos resultados tienen una finalidad de apoyo. El usuario debe determinar si corresponden a su situación y realizar los ajustes necesarios antes de utilizarlos. En el alcance de MiDocIA, la generación de una sugerencia no implica su envío automático ni la ejecución de la acción propuesta.La utilidad de la recomendación se relaciona con su pertinencia para el documento evaluado y con el apoyo que ofrece al usuario. Una recomendación puede resultar útil y, al mismo tiempo, requerir modificaciones antes de emplearse.

### 2.2.8. PROCESAMIENTO TEMPORAL Y ZERO STORAGE

El procesamiento temporal corresponde al uso del archivo y de su contenido durante la ejecución del análisis, sin incorporarlos a un repositorio documental permanente de MiDocIA. El sistema recibe el archivo en memoria y utiliza los servicios externos configurados para realizar la extracción del texto y su análisis. Zero Storage es el nombre empleado en el proyecto para identificar el principio arquitectónico que evita el almacenamiento persistente del archivo original, el texto OCR y el análisis generado dentro de la infraestructura de la aplicación. Esta definición se refiere al contenido documental y debe diferenciarse del registro de información autorizada para seguimiento. Por ello, no se utiliza para afirmar que la aplicación carece de cualquier almacenamiento ni que el procesamiento se realiza exclusivamente en el equipo del usuario.

### 2.2.9. REVISIÓN HUMANA Y CONFIRMACIÓN DEL ANÁLISIS

La revisión humana es la actividad mediante la cual el usuario compara los resultados de MiDocIA con el documento original y determina cómo utilizarlos. Comprende verificar los datos presentados, examinar las alertas y valorar si la sugerencia de respuesta requiere ajustes.
La confirmación es la acción disponible en la aplicación para registrar la finalización de esa revisión dentro del flujo previsto. El principio Human-in-the-Loop se concreta en la participación del usuario antes de cerrar el proceso.
Para efectos de la evaluación, una confirmación no se interpreta como una calificación de exactitud. La finalización del flujo y la calidad del resultado son aspectos distintos y se registran mediante mecanismos diferentes.

### 2.2.10. METADATOS DE USO Y VALORACIÓN DEL USUARIO

Los metadatos de uso son registros asociados con el funcionamiento de la aplicación, como la categoría documental identificada, el tiempo de procesamiento y el estado de confirmación. Su propósito es apoyar el seguimiento del sistema y la evaluación académica.

La aplicación también permite registrar respuestas voluntarias sobre la utilidad percibida del análisis. Estas valoraciones complementan la evaluación realizada con el instrumento del proyecto, pero no reemplazan la comparación entre los resultados y el documento original.

En consecuencia, se distingue entre el comportamiento registrado por el sistema, la percepción expresada por el usuario y la revisión de la calidad del análisis. Esta separación permite interpretar cada dato según lo que efectivamente representa.

### 2.2.11. DOCUMENTOS REALES, SINTÉTICOS Y VALORES ESPERADOS

Los documentos reales corresponden a archivos de origen empresarial suministrados de manera confidencial para las pruebas. Su utilización permite examinar el comportamiento de MiDocIA frente a contenido procedente de una operación empresarial. Los documentos sintéticos son archivos elaborados con información ficticia pero realista para representar situaciones de prueba. En este proyecto se utilizaron para complementar las categorías documentales evaluadas y establecer escenarios controlados. Los valores esperados son los datos de referencia con los que se comparan los resultados del sistema. En los documentos sintéticos se definieron previamente a partir de su contenido. Esta referencia permite identificar qué información fue recuperada correctamente y qué elementos requieren una revisión adicional.

### 2.2.12. TIEMPOS UTILIZADOS EN LA EVALUACIÓN

El tiempo de procesamiento corresponde al intervalo empleado por la aplicación para ejecutar el análisis, según el mecanismo de medición utilizado. El tiempo de revisión humana corresponde al período que necesita el usuario para examinar los resultados.

En el instrumento de evaluación, el tiempo total con MiDocIA combina el procesamiento y la revisión humana. Este valor se compara con el tiempo registrado para realizar la actividad de manera manual.

El tiempo ahorrado expresa la diferencia entre ambos procedimientos. La reducción porcentual representa esa diferencia en relación con el tiempo manual. Estas medidas permiten valorar el aporte de la aplicación considerando que la participación del usuario continúa formando parte de la actividad.

## 2.3. MARCO LEGAL

El desarrollo de MiDocIA toma como referentes las normas colombianas relacionadas con la protección de datos personales y el derecho de petición. Estas disposiciones son pertinentes porque los documentos administrativos pueden contener información de personas naturales y solicitudes que requieren atención dentro de determinados términos. Por ello, además de revisar el funcionamiento de la aplicación, es necesario considerar las condiciones en las que se procesa la información y el alcance de las recomendaciones generadas.
Este marco legal permite explicar la relación entre la normativa y algunas decisiones del proyecto, como solicitar consentimiento antes del procesamiento, evitar la conservación permanente de los documentos y mantener la revisión humana de los resultados. Las normas se presentan como referentes del diseño y del uso de la herramienta, sin que esto constituya una certificación de cumplimiento jurídico integral.

### *2.3.1.* PROTECCIÓN DE DATOS PERSONALES: LEY 1581 DE 2012

La Ley 1581 de 2012 establece disposiciones generales para la protección de datos personales. Su artículo 3 comprende como tratamiento operaciones como recolectar, almacenar, utilizar, circular o suprimir información vinculada con personas naturales. En consecuencia, la protección de los datos debe considerarse durante su uso, incluso cuando no se conserven permanentemente.
El artículo 4 establece, entre otros, los principios de finalidad, libertad, seguridad y confidencialidad. Estos orientan el uso de la información hacia propósitos legítimos e informados, con autorización cuando corresponda y medidas que prevengan accesos indebidos. Asimismo, los artículos 9 y 12 regulan la autorización y la información que debe recibir el titular sobre el tratamiento, su finalidad, sus derechos y la identificación del responsable (Congreso de la República de Colombia, 2012).

En MiDocIA, esta norma se relaciona con el procesamiento de documentos que pueden incluir nombres, números de identificación, correos electrónicos o datos de contacto. Aunque el propósito de la aplicación es apoyar la lectura y el análisis documental, estos datos pueden formar parte del contenido enviado por el usuario. Por esta razón, la aplicación incorpora una aceptación previa al procesamiento. Este mecanismo forma parte del flujo de uso y busca que la persona conozca las condiciones bajo las cuales se analizará el archivo. Sin embargo, la aceptación de quien carga un documento no demuestra, por sí sola, que exista autorización para tratar los datos de todas las personas mencionadas en él. Esta distinción resulta relevante cuando se utilizan documentos empresariales que contienen información de clientes, trabajadores o proveedores.

### *2.3.2.* AUTORIZACIÓN Y CONSERVACIÓN DE LA INFORMACIÓN: DECRETO 1074 DE 2015

El Decreto 1074 de 2015 reúne disposiciones reglamentarias del sector Comercio, Industria y Turismo. Dentro de estas, el capítulo 25 desarrolla aspectos de la protección de datos personales. El artículo 2.2.2.25.2.1 limita la recolección a los datos pertinentes para la finalidad prevista, mientras que los artículos 2.2.2.25.2.2 y 2.2.2.25.2.5 se refieren a la solicitud de autorización y a la conservación de su prueba.

Por su parte, el artículo 2.2.2.25.2.8 relaciona la duración del tratamiento con el tiempo necesario para cumplir sus finalidades, considerando también las obligaciones legales o contractuales de conservación. Estas disposiciones muestran que la protección de la información involucra tanto su recolección como las condiciones de uso y permanencia (Presidencia de la República de Colombia, 2015).

En el proyecto, estos referentes se relacionan con la decisión de procesar temporalmente los archivos. MiDocIA utiliza el documento para extraer su contenido y generar el análisis, sin crear un repositorio permanente de archivos originales,

textos extraídos o respuestas completas dentro de su propia infraestructura. Esta decisión corresponde al alcance del enfoque Zero Storage descrito en el proyecto. Su propósito es reducir la información documental que permanece almacenada después del procesamiento. No significa que la aplicación funcione sin circulación de datos, pues utiliza servicios externos para el reconocimiento de texto y el análisis mediante inteligencia artificial. También debe distinguirse el contenido documental de los registros de uso. La aplicación conserva determinados metadatos y puede recibir información voluntaria del usuario. Por ello, el alcance del almacenamiento debe explicarse según el tipo de información involucrada, evitando presentar todos los registros como anónimos o afirmar que no existe ningún dato persistente.

### *2.3.3.* DERECHO DE PETICIÓN: LEY 1755 DE 2015

La Ley 1755 de 2015 regula el derecho fundamental de petición. El artículo 14 del Código de Procedimiento Administrativo y de lo Contencioso Administrativo, sustituido por esta ley, establece un término general de quince días para resolver peticiones, salvo disposición legal especial. También contempla diez días para solicitudes de documentos e información y treinta días para consultas dirigidas a las autoridades sobre materias a su cargo.

El artículo 32 contempla el derecho de petición ante organizaciones privadas para garantizar derechos fundamentales y remite, salvo norma especial, a las reglas del capítulo correspondiente. Por tanto, la aplicación de estos términos requiere identificar la solicitud y su contexto, sin asignar automáticamente un mismo plazo a todos los documentos clasificados como PQRS (Congreso de la República de Colombia, 2015).

Para el cómputo de los términos, el artículo 62 de la Ley 4 de 1913 establece que en los plazos expresados en días se excluyen los feriados y de vacancia, salvo disposición contraria. Esta regla sustenta la distinción entre días hábiles y días

calendario (Congreso de la República de Colombia, 1913). En MiDocIA, este referente es pertinente para el análisis de peticiones, quejas, reclamos y solicitudes. La herramienta puede apoyar la identificación del asunto y presentar recomendaciones para su revisión. Sin embargo, clasificar un archivo como PQRS no equivale a establecer de manera definitiva el procedimiento jurídico aplicable.

Por ejemplo, la fecha escrita en una carta puede ser diferente de la fecha en que fue recibida por su destinatario. Si el documento no contiene información suficiente para identificar la recepción, el análisis no permite determinar con certeza el vencimiento. Por ello, cualquier plazo sugerido debe contrastarse con los soportes disponibles y con las condiciones del trámite.

La utilidad de la aplicación consiste en facilitar la revisión inicial y llamar la atención sobre aspectos que requieren seguimiento. La verificación de las fechas y la preparación de la respuesta corresponden a la persona encargada del documento.

### *2.3.4.* ALCANCE DE LAS RECOMENDACIONES Y DE LA REVISIÓN HUMANA

MiDocIA genera resúmenes, identifica información relevante y propone acciones o respuestas según el contenido analizado. Estas funciones apoyan el trabajo administrativo, pero sus resultados pueden contener omisiones o interpretaciones que requieren corrección. Esto adquiere especial importancia cuando un documento incluye obligaciones contractuales, solicitudes formales o actuaciones judiciales. Por esta razón, el proyecto incorpora una revisión humana en la que el usuario puede comparar el documento con el análisis obtenido. Esta revisión permite comprobar si los nombres, valores, fechas y asuntos identificados corresponden al contenido original antes de utilizar las recomendaciones.

La confirmación del usuario forma parte del funcionamiento de la herramienta, pero no constituye una validación jurídica del resultado. De igual manera, una respuesta sugerida debe revisarse y ajustarse a las circunstancias del caso antes de utilizarse en una comunicación empresarial.

El alcance de este marco legal consiste en identificar referentes normativos y explicar su relación con el proyecto. La evaluación realizada sobre MiDocIA se concentra en su funcionamiento y utilidad para apoyar el análisis documental; no corresponde a una auditoría integral de cumplimiento de la normativa sobre protección de datos ni a una evaluación jurídica de cada recomendación generada.

## 2.4. MARCO TECNOLÓGICO

MiDocIA integra herramientas de desarrollo web, reconocimiento óptico de caracteres, inteligencia artificial y almacenamiento de registros de uso. Cada tecnología cumple una función dentro del recorrido que realiza el documento, desde su carga hasta la presentación del análisis y la confirmación del usuario. La solución se desarrolló mediante una arquitectura cliente-servidor. La interfaz permite interactuar con la aplicación, mientras que el servidor coordina el procesamiento y la comunicación con los servicios externos. Esta organización permite diferenciar la presentación de los resultados, la lógica de la aplicación y las operaciones realizadas por los proveedores de OCR e inteligencia artificial.

### *2.4.1.* ARQUITECTURA WEB Y ORGANIZACIÓN DE LOS COMPONENTES

La arquitectura de MiDocIA comprende tres grupos de componentes. El primero corresponde a la interfaz de usuario, desde la cual se acepta el consentimiento, se selecciona el archivo y se consulta el análisis. El segundo corresponde al servidor, encargado de recibir las solicitudes, validar los archivos y coordinar las etapas del procesamiento. El tercero reúne los servicios externos utilizados para extraer texto, analizar el contenido y registrar los datos autorizados de seguimiento. En esta organización, el navegador se comunica con el servidor de MiDocIA y este realiza las solicitudes a los servicios externos. De esta manera, la integración con

los proveedores y el manejo de sus credenciales se concentran en el backend. El código del frontend y del backend se encuentra organizado en un mismo repositorio. Aunque ambos componentes tienen responsabilidades diferentes, esta estructura permite mantener juntos los archivos del proyecto y los procedimientos necesarios para su ejecución. En producción, la interfaz compilada y el servidor se publican como un único servicio en Railway.

### *2.4.2.* INTERFAZ DE USUARIO: REACT Y VITE

React permite construir interfaces mediante componentes que reúnen elementos visuales y comportamientos. Su documentación explica cómo presentar información, responder a eventos y actualizar la pantalla según las acciones del usuario (React, s. f.).

En MiDocIA, React se utiliza para construir las partes de la aplicación con las que interactúa el usuario. Estas incluyen el consentimiento inicial, la selección del documento, los estados del procesamiento, la consulta del resultado y la confirmación de la revisión humana.

La interfaz también permite presentar el documento junto con la información generada. Esta disposición apoya la comparación entre el contenido original y los datos identificados por la inteligencia artificial. Así, la presentación del resultado forma parte del proceso de revisión previsto en el proyecto.

Como herramienta complementaria se utiliza Vite, que proporciona un entorno de desarrollo y permite generar los archivos del frontend destinados a producción (Vite, s. f.). En MiDocIA, estos archivos compilados son entregados al navegador por el servidor Express.

### *2.4.3.* SERVIDOR DE APLICACIÓN: NODE.JS Y EXPRESS

Node.js es un entorno que permite ejecutar JavaScript fuera del navegador. Cuenta con mecanismos para realizar operaciones de entrada y salida de forma asíncrona, lo que resulta pertinente para aplicaciones que necesitan comunicarse con servicios externos (Node.js, s. f.).

En el proyecto, Node.js ejecuta la lógica del backend. Desde esta parte de la aplicación se coordina la recepción del archivo, la extracción del texto, el envío del contenido al modelo de inteligencia artificial y la devolución de los resultados. Express complementa este entorno mediante herramientas para organizar las rutas y el tratamiento de las solicitudes web. También permite incorporar funciones intermedias, conocidas como middleware, que intervienen durante el procesamiento de una solicitud (Express, s. f.-a).

En MiDocIA, estas herramientas se utilizan para atender las operaciones de análisis documental, confirmación humana y registro de la valoración del usuario. El servidor también entrega los archivos de la interfaz en el entorno de producción. Esta distribución permite que la interfaz se concentre en la interacción con el usuario y que el backend reúna las operaciones necesarias para completar el análisis. Por ejemplo, cuando se solicita procesar un documento, el servidor coordina las etapas y devuelve una respuesta que la interfaz presenta de forma organizada.

### *2.4.4.* RECEPCIÓN Y PROCESAMIENTO TEMPORAL DE ARCHIVOS

La recepción de archivos se implementa mediante Multer. Esta herramienta dispone de una opción denominada memoryStorage, que mantiene el archivo recibido como un búfer en memoria. Un búfer permite representar temporalmente el contenido binario del documento para utilizarlo durante el procesamiento (Express, s. f.-b). MiDocIA utiliza esta configuración para evitar que la carga genere una copia persistente en el sistema de archivos del servidor. La aplicación también verifica los

formatos admitidos y establece un límite de tamaño de 10 MB por archivo. Para el manejo de documentos PDF extensos se utiliza la biblioteca pdf-lib. Según la configuración documentada del proyecto, los archivos que lo requieren se dividen en bloques de hasta catorce páginas. Esta división se realiza en memoria y permite enviar las partes al servicio de reconocimiento de texto. Después, el texto obtenido se reúne temporalmente para continuar con el análisis.
Estas decisiones forman parte de la implementación del enfoque Zero Storage. Su alcance corresponde a evitar la conservación persistente del contenido documental en la infraestructura de MiDocIA. El procesamiento sigue requiriendo memoria y comunicación con los servicios externos que participan en el análisis.

### *2.4.5.* EXTRACCIÓN DE TEXTO: GOOGLE CLOUD DOCUMENT AI

Google Cloud Document AI ofrece herramientas para procesar documentos y extraer texto e información sobre su estructura. Dentro de sus capacidades se encuentra el reconocimiento óptico de caracteres, que permite obtener texto procesable a partir del contenido documental (Google Cloud, 2026).
En MiDocIA, Document AI realiza la etapa de extracción de texto. El backend envía el contenido mediante la API REST del servicio y recibe la información extraída para utilizarla en la siguiente etapa.
Esta función resulta necesaria para trabajar con documentos digitalizados y archivos cuyo contenido se presenta como imagen. El resultado del OCR permite que el modelo de lenguaje reciba una representación textual del documento y pueda analizarla según las instrucciones establecidas.
Dentro de la aplicación, la extracción de texto y el análisis con inteligencia artificial tienen funciones diferenciadas. Document AI obtiene el texto, mientras que Gemini lo utiliza para clasificar el documento y generar el resultado. Por ello, la calidad del texto extraído influye en la información disponible para el análisis posterior.

### *2.4.6.* ANÁLISIS DEL CONTENIDO: GOOGLE GEMINI

La API de Gemini permite integrar modelos de inteligencia artificial capaces de generar texto a partir de instrucciones y del contenido proporcionado en una solicitud (Google, 2026).

En la versión documentada de MiDocIA se utiliza el modelo gemini-2.5-flash. Su función consiste en clasificar el documento, identificar datos relevantes, elaborar un resumen, señalar posibles alertas y proponer una respuesta o recomendación según el caso.

El procesamiento se organiza en dos llamadas. En la primera se identifica la categoría documental. En la segunda se utilizan instrucciones específicas para analizar el contenido de acuerdo con la categoría obtenida. Esta organización permite solicitar información diferente según se trate, por ejemplo, de una factura, un contrato o una PQRS.

Las instrucciones enviadas al modelo se conocen como prompts. En el proyecto, estos orientan la información que debe buscarse y la forma en que debe presentarse la respuesta. El trabajo realizado corresponde a la integración y configuración del modelo dentro de la aplicación, sin desarrollar ni entrenar un modelo de lenguaje propio.

Los resultados generados se presentan al usuario para su revisión. La aplicación mantiene esta participación humana porque la generación de una respuesta no demuestra, por sí misma, que todos los datos o interpretaciones sean correctos.

### *2.4.7.* REGISTRO DE USO Y EVALUACIÓN: SUPABASE

Supabase proporciona una base de datos PostgreSQL y herramientas para administrar y consultar la información almacenada (Supabase, s. f.).

En MiDocIA, se utiliza para conservar los registros autorizados de seguimiento y evaluación académica. Entre estos se encuentran la fecha del análisis, la categoría

documental, el tiempo de procesamiento y el estado de confirmación humana. También se registran los datos correspondientes a la valoración voluntaria cuando el usuario los proporciona.

Esta información permite consultar aspectos del uso de la aplicación sin almacenar el archivo original, el texto extraído ni el análisis completo. Por ejemplo, es posible conocer si se confirmó una revisión o qué calificación recibió el resultado sin conservar el contenido del documento procesado.

La base de datos tiene, por tanto, una función de seguimiento. La existencia de estos registros debe distinguirse del almacenamiento documental, ya que algunos campos, como el correo electrónico o un comentario voluntario, pueden contener información personal.

### *2.4.8.* PUBLICACIÓN Y MANTENIMIENTO DE LA APLICACIÓN

Railway permite desplegar aplicaciones Express mediante diferentes mecanismos, entre ellos la conexión con un repositorio de GitHub. Su documentación contempla la configuración del servicio y de las variables de entorno necesarias para ejecutar la aplicación (Railway, s. f.).

En MiDocIA, Railway aloja el backend y la interfaz compilada como un único servicio. Ambos componentes funcionan bajo el mismo dominio, desde el cual el usuario accede a la aplicación y realiza las solicitudes de procesamiento.

El proyecto utiliza Git y GitHub para mantener el código fuente y registrar sus modificaciones. Por su parte, npm permite administrar las dependencias y ejecutar las tareas definidas para el frontend y el backend. Estas herramientas apoyan la instalación, la actualización y el mantenimiento de la solución.

La configuración también diferencia los archivos del proyecto de las credenciales requeridas para acceder a los servicios externos. Estas credenciales se gestionan mediante la configuración del entorno y deben mantenerse fuera del código publicado.

El funcionamiento de MiDocIA depende de la conexión a Internet, de la disponibilidad del servicio desplegado y de los proveedores utilizados para OCR, inteligencia artificial y registro de datos. Por ello, el manual de instalación y mantenimiento documenta la preparación del entorno, la configuración de las integraciones y los procedimientos de verificación necesarios para poner en funcionamiento la aplicación.

## 2.5. MARCO AMBIENTAL

MiDocIA corresponde al desarrollo de una aplicación web para apoyar el análisis de documentos administrativos. Las actividades del proyecto se concentraron en el diseño, la programación, la integración de servicios y la evaluación de la herramienta. Dentro de este alcance no se realizaron intervenciones directas sobre el suelo, las fuentes hídricas, la vegetación o la biodiversidad.
Su relación con el ambiente se encuentra principalmente en el uso de equipos de cómputo, conexión a Internet y servicios de procesamiento en la nube. También se relaciona con la posibilidad de revisar documentos en formato digital. Estos aspectos permiten reconocer consideraciones ambientales del funcionamiento de la aplicación, aunque no fueron objeto de medición durante el proyecto.

### *2.5.1.* REVISIÓN DE DOCUMENTOS EN FORMATO DIGITAL

MiDocIA permite cargar documentos digitales y consultar su contenido junto con el análisis generado. Este funcionamiento ofrece la posibilidad de realizar la revisión desde el navegador, sin que la impresión sea un paso requerido por la aplicación. En una actividad administrativa, esta característica puede facilitar la consulta de archivos que ya se reciben por medios electrónicos. Por ejemplo, una factura o una comunicación disponible en PDF puede analizarse directamente, sin imprimirla para

utilizar la herramienta. Sin embargo, el usuario o la organización pueden mantener procedimientos que requieran copias físicas.

Por esta razón, la posibilidad de trabajar digitalmente se considera una característica del sistema y un beneficio ambiental potencial. Durante la evaluación no se registró la cantidad de impresiones realizadas antes y después de utilizar MiDocIA, ni se midió el consumo de papel o tinta. En consecuencia, los resultados del proyecto no permiten establecer una reducción comprobada de estos materiales.

### *2.5.2.* RECURSOS NECESARIOS PARA EL PROCESAMIENTO

El funcionamiento de MiDocIA requiere recursos tanto en el equipo del usuario como en los servicios externos utilizados. El navegador presenta el documento y los resultados, mientras que el servidor coordina las solicitudes y los servicios de OCR e inteligencia artificial procesan la información.

La Agencia Internacional de Energía señala que el funcionamiento de la inteligencia artificial depende de la electricidad utilizada en los centros de datos. Su informe Energy and AI estudia esta relación y sus implicaciones para la demanda energética y las emisiones (International Energy Agency, 2025). Este referente permite reconocer que el procesamiento digital también depende de infraestructura física y consumo de energía.

En MiDocIA no se midió la energía consumida por cada análisis ni se dispuso de una medición del consumo atribuible al proyecto dentro de los servicios contratados. Por ello, los tiempos de procesamiento registrados durante la evaluación se utilizan como indicadores de funcionamiento, sin convertirlos en estimaciones de consumo energético o emisiones.

### *2.5.3.* PROCESAMIENTO TEMPORAL Y ALMACENAMIENTO

La arquitectura de MiDocIA evita conservar permanentemente el archivo original, el texto extraído y el análisis completo en su propia infraestructura. Los documentos se procesan temporalmente en memoria y la base de datos conserva los registros autorizados de seguimiento.

Esta decisión se adoptó principalmente para atender el alcance de privacidad del proyecto. Desde la perspectiva del almacenamiento, significa que la aplicación no construye un repositorio creciente de documentos analizados. Sin embargo, el uso de memoria, la transmisión de información y las operaciones de los servicios externos siguen formando parte de su funcionamiento.

En consecuencia, el enfoque Zero Storage no se presenta como evidencia de una reducción de la huella de carbono. Para establecer una relación de ese tipo sería necesario medir los recursos utilizados y compararlos con otra alternativa bajo condiciones equivalentes. Esa comparación no formó parte de la evaluación realizada.

### *2.5.4.* ALCANCE DE LA VALORACIÓN AMBIENTAL

La consideración ambiental del proyecto tiene un alcance descriptivo. Se identifican los recursos tecnológicos que intervienen en el funcionamiento de la aplicación y la posibilidad de revisar documentos sin requerir su impresión, pero no se determina un balance ambiental del sistema.

La evaluación de MiDocIA se concentró en la clasificación documental, la extracción de información, la utilidad de los resultados, los tiempos de revisión y la participación humana. No incluyó indicadores de consumo de papel, electricidad, agua o emisiones asociadas al procesamiento.

Una evaluación posterior podría incorporar el registro de impresiones evitadas, la cantidad de análisis realizados y los datos de consumo disponibles en la

infraestructura utilizada. Estos registros permitirían estudiar con mayor precisión la relación entre el apoyo al trabajo administrativo y los recursos necesarios para ofrecerlo.

# 3. DISEÑO DE LA INVESTIGACIÓN

El diseño de la investigación reúne las decisiones que orientaron el desarrollo de MiDocIA y la evaluación de su funcionamiento. Comprende la modalidad del trabajo, el alcance, el enfoque, el método, las técnicas utilizadas y el procedimiento seguido para analizar los resultados.

El proyecto partió de una problemática planteada en el banco de proyectos de las Unidades Tecnológicas de Santander, relacionada con el tiempo requerido para leer, clasificar e interpretar documentos administrativos en las Mipymes. A partir de esta necesidad se construyó una aplicación web que integra reconocimiento óptico de caracteres e inteligencia artificial, manteniendo la revisión humana dentro del proceso.

La evaluación se realizó mediante documentos reales de origen empresarial y documentos sintéticos preparados para las pruebas académicas. Esta combinación permitió observar el comportamiento de la aplicación frente a diferentes contenidos y comparar sus resultados con la información de los documentos utilizados.

## 3.1. MODALIDAD DEL TRABAJO

El trabajo se desarrolló bajo la modalidad de desarrollo tecnológico, debido a que estuvo orientado al diseño, construcción e integración de una solución de software. El producto obtenido fue una aplicación web funcional que permite cargar un documento, extraer su texto, identificar su categoría y presentar información relevante para apoyar su revisión.

El desarrollo comprendió la definición de requerimientos, el diseño de la arquitectura, la construcción de la interfaz, la integración de servicios externos y la ejecución de pruebas. También incluyó la elaboración de los manuales necesarios para explicar el uso, la instalación y el mantenimiento de la aplicación.
La contribución del proyecto se encuentra en la integración de estos componentes dentro de un flujo de análisis documental. Para ello se utilizaron servicios existentes de OCR e inteligencia artificial, junto con funciones desarrolladas para recibir archivos, organizar los resultados, solicitar la confirmación del usuario y registrar información de seguimiento.
La modalidad adoptada permitió relacionar el desarrollo del producto con su evaluación práctica. Por esta razón, además de comprobar que la aplicación ejecutara el proceso completo, se revisó la correspondencia entre la información generada y el contenido de los documentos.

## 3.2. ALCANCE DE LA INVESTIGACIÓN

El alcance de la investigación es descriptivo. La evaluación se orientó a observar, registrar y analizar el comportamiento de MiDocIA durante el procesamiento de los documentos seleccionados. Se consideraron aspectos como la clasificación documental, la calidad del texto extraído, la identificación de campos, las alertas, la utilidad de las recomendaciones y los tiempos empleados.
Este alcance permitió describir los aciertos y las dificultades encontrados durante las pruebas. Por ejemplo, se revisó si la categoría asignada correspondía al documento, si los datos relevantes habían sido identificados y si la recomendación podía utilizarse o requería ajustes.
La investigación no buscó establecer relaciones causales entre variables ni demostrar que los resultados obtenidos fueran aplicables a todas las Mipymes de Bucaramanga. Los hallazgos corresponden al conjunto documental evaluado y a las

condiciones bajo las cuales se realizaron las pruebas. Tampoco se realizó una caracterización, encuesta o selección estadística de una población de empresas durante la etapa inicial. La problemática provino del banco de proyectos institucional, mientras que la participación empresarial correspondió a la validación práctica de la herramienta mediante documentos reales.

## 3.3. ENFOQUE DE LA INVESTIGACIÓN

Se adoptó un enfoque mixto con predominio cuantitativo. El componente cuantitativo estuvo representado por los registros que permitieron contar resultados, calcular proporciones y comparar tiempos. Entre estos se encuentran la cantidad de documentos clasificados correctamente, los campos identificados y los tiempos de revisión manual y de revisión apoyada por MiDocIA. También se utilizaron las valoraciones numéricas registradas en el instrumento de evaluación y los datos de seguimiento generados por la aplicación. Estos permitieron consolidar indicadores sobre el funcionamiento del sistema y la percepción de utilidad de sus resultados.

El componente cualitativo correspondió a las observaciones escritas sobre el contenido del análisis y la pertinencia de las recomendaciones. Estas observaciones permitieron explicar situaciones que una calificación numérica no describe por completo, como la necesidad de completar información, corregir una interpretación o ajustar la redacción de una respuesta.

Ambos componentes se utilizaron de manera complementaria. Las mediciones permitieron resumir el desempeño observado, mientras que los comentarios ayudaron a comprender las dificultades encontradas en casos concretos. De esta forma, la evaluación consideró tanto los resultados numéricos como las razones que justificaban algunas de las valoraciones.

## 3.4. MÉTODO EMPLEADO

El método empleado fue principalmente analítico. Se partió de la problemática definida para identificar las funciones que debía cumplir la solución y organizar el procesamiento documental en componentes con responsabilidades diferenciadas. Durante el desarrollo se abordaron por separado la recepción de archivos, la extracción de texto, la clasificación documental, el análisis del contenido y la presentación de resultados. Posteriormente, estos componentes se integraron en un mismo flujo de funcionamiento.

En la evaluación se mantuvo esta separación para revisar cada aspecto del resultado. La identificación correcta de la categoría, por ejemplo, se valoró de manera independiente de la extracción de los campos. Asimismo, la legibilidad del texto obtenido mediante OCR se distinguió de la utilidad de la recomendación. Esta forma de revisión permitió reconocer que un mismo análisis podía presentar aciertos en algunos componentes y dificultades en otros. Por ello, la valoración de MiDocIA se apoyó en varios criterios, evitando reducir todo su funcionamiento a una única calificación.

## 3.5. TÉCNICAS E INSTRUMENTOS UTILIZADOS

Se utilizaron la revisión documental, las pruebas funcionales del software y las pruebas de procesamiento con documentos reales y sintéticos. La revisión documental permitió consultar antecedentes, fundamentos técnicos y referentes relacionados con el problema abordado. Esta actividad apoyó la definición de las funciones de MiDocIA y la comprensión de las tecnologías utilizadas.

Las pruebas funcionales se realizaron durante la construcción e integración de la aplicación. Se revisaron operaciones como la carga de archivos, la comunicación con los servicios externos, la presentación del análisis, la confirmación humana y el

registro de métricas. Estas pruebas permitieron identificar dificultades de funcionamiento y realizar ajustes antes de la evaluación final.
Para registrar la valoración de los resultados se utilizó un instrumento de evaluación diseñado por el autor del proyecto. En este se consolidaron los datos de cada documento y la revisión de la clasificación, el OCR, los campos extraídos, las alertas, la utilidad de las recomendaciones y los tiempos empleados. También se incluyeron observaciones sobre los aspectos que requerían corrección.
Como fuente complementaria se utilizaron los registros de uso de MiDocIA. Estos reúnen metadatos del procesamiento, la confirmación de la revisión humana y las respuestas voluntarias de la encuesta incorporada en la aplicación.
El instrumento y los registros automáticos cumplen funciones diferentes. El primero contiene la comparación entre el documento y el resultado generado; los segundos permiten consultar eventos y valoraciones registrados durante el uso del sistema. Por esta razón, sus indicadores se presentan por separado cuando corresponden a mediciones distintas.

## 3.6. CONJUNTO DOCUMENTAL DE EVALUACIÓN

La evaluación se realizó con 32 documentos: 20 reales de origen empresarial y 12 sintéticos. En conjunto se incluyeron las siete categorías soportadas por MiDocIA: factura, contrato, PQRS, documento judicial, carta, cotización y otro.
Los documentos reales fueron suministrados de manera confidencial por una empresa colaboradora. Su utilización permitió revisar el comportamiento de la aplicación frente a contenidos provenientes de una actividad empresarial, con sus propias estructuras, formas de redacción y datos administrativos.
Los documentos sintéticos se elaboraron con información ficticia pero realista y con valores esperados previamente definidos. Su función fue ampliar los escenarios de

prueba y facilitar la comparación entre la información incluida en cada archivo y los resultados generados por la aplicación.

Los dos grupos se utilizaron como fuentes complementarias de evaluación. Su distribución no corresponde a una muestra estadísticamente representativa de los documentos utilizados por las Mipymes ni a una división entre grupo experimental y grupo de control.

La unidad de evaluación fue cada documento procesado y su resultado correspondiente. Para la extracción de información también se revisaron los campos individuales definidos en el instrumento, lo que permitió identificar aciertos, resultados parciales y datos no encontrados.

## 3.7. PROCEDIMIENTO DE EVALUACIÓN Y TRATAMIENTO DE LOS RESULTADOS

La validación práctica fue realizada por la persona designada por la empresa colaboradora. El autor del proyecto diseñó y estructuró el instrumento de evaluación, mientras que las valoraciones registradas correspondieron a la revisión de la persona evaluadora. Para cada caso se utilizó el documento original como referencia y se comparó su contenido con el resultado presentado por MiDocIA. La revisión consideró la categoría identificada, la calidad del texto extraído, los datos relevantes, las alertas y la recomendación generada. Las diferencias encontradas se registraron mediante las calificaciones y observaciones del instrumento.

En la comparación de tiempos se distinguió el procedimiento manual del proceso apoyado por la aplicación. Este último incluyó el tiempo necesario para obtener el análisis y realizar la revisión humana. Así, la comparación tuvo en cuenta que el uso de MiDocIA requiere que una persona compruebe la información antes de utilizarla.

También se diferenciaron los tiempos registrados en el instrumento de los tiempos automáticos del servidor. Estos datos provienen de mecanismos de medición

distintos y se analizaron conservando esa distinción. De igual manera, las respuestas voluntarias de la encuesta se interpretaron como estimaciones del tiempo que el usuario habría necesitado para realizar la actividad sin MiDocIA. Estas respuestas se analizaron por separado del ahorro calculado mediante la comparación de tiempos registrada en la matriz.

Los datos obtenidos se consolidaron mediante conteos, porcentajes y promedios, según el indicador correspondiente. Las observaciones escritas se utilizaron para explicar dificultades y necesidades de ajuste identificadas durante la revisión. La confirmación Human-in-the-Loop se interpretó como el registro de la acción de confirmación realizada por el usuario. Este indicador permite verificar la finalización de ese paso del flujo, pero no demuestra que todos los datos del análisis sean correctos.

Los documentos empresariales utilizados no se reproducen ni se presentan como anexos del informe. La matriz de evaluación se incorpora con autorización de la empresa colaboradora, sin incluir copias de los documentos procesados. Durante el funcionamiento de la aplicación se mantienen los principios de procesamiento temporal y revisión humana definidos para MiDocIA.

## 3.8. FASES DE DESARROLLO Y EVALUACIÓN

El procedimiento se organizó en cinco fases relacionadas con los objetivos específicos del proyecto. Estas fases permitieron ordenar las actividades desde la revisión de la problemática hasta la evaluación de la aplicación. Durante la construcción se realizaron ajustes a partir de las pruebas funcionales. Por tanto, la organización por fases describe las actividades principales del trabajo y permite identificar su relación con los objetivos, sin desconocer las revisiones efectuadas durante el desarrollo.

**TABLA 1 FASES DE DESARROLLO Y EVALUACIÓN**

| Fase | Actividad principal | Objetivo relacionado |
|---|---|---|
| 1. Análisis | Revisión de la problemática planteada y definición de los tipos documentales, información relevante y requerimientos de la aplicación. | Objetivo 1 |
| 2. Diseño | Definición de la arquitectura de software, interfaz de usuario y principios Zero Storage y Human-in-the-Loop. | Objetivo 2 |
| 3. Implementación | Desarrollo del procesamiento de archivos y extracción de texto mediante reconocimiento óptico de caracteres. | Objetivo 3 |
| 4. Integración | Incorporación del componente de inteligencia artificial para clasificación, extracción de información, identificación de alertas y generación de recomendaciones. | Objetivo 4 |
| 5. Evaluación | Ejecución de pruebas con documentos reales de origen empresarial y documentos sintéticos controlados, registrando y comparando los resultados obtenidos. | Objetivo 5 |

Fuente: elaboración propia.

# 4. DESARROLLO DEL TRABAJO DE GRADO

El desarrollo de MiDocIA se realizó de manera progresiva, partiendo del análisis de la problemática y los requerimientos de la aplicación, seguido por el diseño de la arquitectura, la implementación de los diferentes componentes y las pruebas necesarias para obtener una versión funcional desplegada en producción. Estas etapas se relacionan directamente con los objetivos específicos planteados para el proyecto.

## 4.1. ANÁLISIS DE LA PROBLEMÁTICA Y REQUERIMIENTOS

El proyecto inició a partir de la problemática planteada en el banco de proyectos de las Unidades Tecnológicas de Santander, relacionada con el tiempo empleado en la lectura, clasificación y análisis de documentos administrativos dentro de las Mipymes. A partir de esta necesidad se definieron las funciones que debía cumplir MiDocIA y las condiciones bajo las cuales se procesaría la información. El análisis se concentró en identificar qué actividades podían recibir apoyo mediante una aplicación web. Se consideraron la extracción del contenido, la identificación del tipo de documento, la organización de los datos relevantes y la generación de orientaciones para su revisión. También se estableció que el usuario debía conservar la responsabilidad de comprobar los resultados y decidir cómo utilizarlos.

### 4.1.1. DELIMITACIÓN DE LA NECESIDAD

La necesidad abordada comprende varias tareas relacionadas. Para revisar un documento, la persona debe reconocer su propósito, ubicar los datos principales e

identificar si contiene solicitudes, fechas, valores u obligaciones que requieren atención. Según el contenido, también puede necesitar preparar una respuesta o establecer una acción posterior.

Estas tareas orientaron la definición del flujo de MiDocIA. La aplicación debía recibir el archivo, obtener su texto y presentar un análisis organizado que sirviera como punto de partida para la revisión. El resultado debía permitir consultar la información identificada y compararla con el documento original.
La definición inicial se realizó a partir de la problemática institucional y del alcance del proyecto. No se llevó a cabo una encuesta o caracterización de empresas para establecer estos requerimientos. La participación de la empresa colaboradora correspondió posteriormente a la evaluación práctica de la aplicación.
Esta delimitación permitió concentrar el desarrollo en el apoyo al análisis documental. La construcción de un archivo empresarial permanente, la gestión de expedientes y la ejecución automática de decisiones quedaron fuera de las funciones contempladas.

### 4.1.2. CATEGORÍAS DOCUMENTALES E INFORMACIÓN RELEVANTE

Se establecieron siete categorías: factura, contrato, PQRS, documento judicial, carta, cotización y otro. Esta clasificación permitió organizar las instrucciones de análisis según las características de cada documento y definir qué información debía buscar la aplicación. En las facturas se consideraron datos como el número del documento, el emisor, el receptor, las fechas, los conceptos y los valores. En las cotizaciones se incluyeron el proveedor, el cliente, los productos o servicios ofrecidos, la vigencia de la propuesta y las condiciones de pago y entrega. Aunque ambas categorías pueden contener información comercial, su propósito es diferente y requiere una revisión acorde con su contenido.
Para los contratos se contemplaron las partes, el objeto, el valor, la vigencia y las obligaciones. También se incluyó la identificación de condiciones que pudieran

requerir atención del usuario. Estos elementos sirven para organizar la lectura inicial, sin que el resultado constituya una revisión jurídica completa del contrato. En las PQRS se consideraron el remitente, el motivo, la solicitud y las fechas disponibles. En los documentos judiciales se incluyeron la autoridad, la identificación del proceso, las partes y la orden o actuación principal. Los plazos señalados por el análisis requieren una revisión particular, porque su interpretación depende del contenido y de las circunstancias del trámite. Para las cartas se contemplaron el remitente, el destinatario, la fecha, el asunto y la acción solicitada. Finalmente, la categoría otro permitió realizar un análisis general de documentos que no correspondieran a las categorías anteriores, buscando reconocer su propósito y la información relevante. Los campos definidos orientan la búsqueda de información. Su inclusión en las instrucciones no significa que todos estén presentes en cada archivo o que siempre puedan extraerse correctamente. Esta diferencia se tuvo en cuenta al evaluar los resultados frente a los documentos originales.

### 4.1.3. REQUERIMIENTOS FUNCIONALES

Los requerimientos funcionales describen las operaciones previstas para la aplicación. Se organizaron considerando la recepción del archivo, el procesamiento de su contenido, la presentación del resultado y la participación del usuario.

**TABLA 2 REQUERIMIENTOS FUNCIONALES DE MIDOCIA**

| **Código** | **Requerimiento** |
|---|---|
| RF-01 | Permitir la carga de documentos en formato PDF, PNG, JPEG y TIFF, con un tamaño máximo de 10 MB por archivo. |
| RF-02 | Extraer el texto del documento mediante reconocimiento óptico de caracteres para utilizarlo en el análisis. |
| **Código** | **Requerimiento** |

| | |
|---|---|
| RF-03 | Clasificar el documento en una de las siete categorías definidas para el proyecto. |
| RF-04 | Identificar los campos relevantes de acuerdo con la categoría documental detectada. |
| RF-05 | Presentar alertas sobre elementos que puedan requerir atención, como fechas, plazos y condiciones relevantes del documento. |
| RF-06 | Generar una sugerencia de respuesta o una recomendación de acción según el contenido analizado. |
| RF-07 | Mostrar el documento original junto con el análisis para facilitar su comparación. |
| RF-08 | Solicitar la confirmación explícita del usuario para finalizar el ciclo de revisión humana. |
| RF-09 | Ofrecer una encuesta voluntaria después de la confirmación para registrar la valoración del usuario. |
| RF-10 | Disponer de un panel administrativo para consultar las métricas de uso registradas por la aplicación. |
| RF-11 | Permitir el procesamiento de PDF extensos mediante su división temporal en bloques de hasta catorce páginas, respetando el tamaño máximo del archivo. |
| RF-12 | Solicitar y registrar la aceptación del consentimiento para el tratamiento de datos antes del procesamiento. |

Fuente: elaboración propia.

Estos requerimientos permiten seguir el recorrido del documento dentro de la aplicación. La carga y extracción de texto preparan el contenido; la clasificación y el análisis organizan la información; y la visualización y confirmación mantienen al usuario dentro del proceso.

La encuesta y el panel administrativo cumplen una función complementaria de seguimiento. Sus registros permiten consultar aspectos del uso de MiDocIA,

mientras que la evaluación detallada de la exactitud del análisis se realiza mediante la comparación con los documentos originales.

### 4.1.4. REQUERIMIENTOS NO FUNCIONALES

Además de las operaciones anteriores, se definieron condiciones relacionadas con la privacidad, el manejo de la información, el acceso y el mantenimiento de la aplicación.

**Privacidad — RNF-01**. Se estableció que el archivo original debía procesarse temporalmente, sin conservar una copia persistente en la infraestructura de MiDocIA. Esta condición orientó el manejo de los documentos en memoria durante las operaciones del backend.

**Separación del contenido y los registros de seguimiento — RNF-02**. Se definió que la base de datos debía recibir únicamente los datos autorizados para el seguimiento académico. El archivo, el texto extraído y el análisis generado quedaron excluidos de ese almacenamiento. Esta condición delimita la información registrada y forma parte de las medidas de protección del proyecto.

**Disponibilidad — RNF-03**. Se planteó que la aplicación estuviera publicada y pudiera utilizarse sin que el desarrollador tuviera que iniciar manualmente el servidor para cada sesión. Su disponibilidad efectiva depende del alojamiento, de la conexión a Internet y de los servicios externos utilizados.

**Rendimiento — RNF-04.** Se consideró necesario que el procesamiento ofreciera tiempos de respuesta adecuados para apoyar una actividad administrativa. Este aspecto se revisó mediante las mediciones de la evaluación, distinguiendo el tiempo automático del sistema del tiempo dedicado por la persona a comprobar el resultado.

**Compatibilidad — RNF-05**. Se definió el acceso mediante un navegador moderno, sin exigir la instalación de MiDocIA en el equipo del usuario. La interfaz debía permitir realizar las operaciones principales desde los dispositivos contemplados en el proyecto.

**Mantenibilidad — RNF-06.** Se estableció una organización modular del código, con separación entre las rutas, los controladores, los servicios y las funciones de validación. Esta distribución permite ubicar las responsabilidades de cada componente y facilita la revisión de errores o modificaciones.

**Consideración de la protección de datos — RNF-07.** Se incorporaron el consentimiento previo y la delimitación de la información conservada como medidas relacionadas con los referentes legales del proyecto. Su implementación se describe dentro del alcance de MiDocIA y no se presenta como una certificación integral de cumplimiento normativo.

### 4.1.5. ALCANCE DE LAS RESPUESTAS Y PARTICIPACIÓN DEL USUARIO

Durante la definición de requerimientos se distinguieron dos tipos de salida: la sugerencia de respuesta y la recomendación de acción. La primera ofrece un texto que el usuario puede revisar y adaptar para una comunicación. La segunda propone aspectos que conviene comprobar o atender de acuerdo con el documento.

Esta distinción permite que la salida se ajuste al propósito del archivo. Una carta o una PQRS puede requerir la preparación de una contestación, mientras que una factura o una cotización puede necesitar una revisión de valores, condiciones o datos antes de continuar con el trámite.

Se estableció que estas salidas debían permanecer sujetas a revisión humana. MiDocIA presenta la información y permite finalizar el análisis mediante una confirmación, pero no envía comunicaciones, realiza pagos ni aprueba contratos. También se delimitó el significado de las alertas. Estas señalan elementos que

merecen atención y deben contrastarse con el documento. Su generación no demuestra por sí sola la existencia de un incumplimiento, y la ausencia de alertas tampoco garantiza que el archivo esté libre de asuntos relevantes.

### 4.1.6. RESULTADO DEL ANÁLISIS DE REQUERIMIENTOS

Como resultado de esta etapa se definieron las categorías documentales, los datos que debía buscar la aplicación, las operaciones del flujo y las condiciones de procesamiento. Estos elementos sirvieron como base para diseñar la arquitectura y organizar la implementación.

Los requerimientos también permitieron establecer qué aspectos debían revisarse durante las pruebas. La clasificación se relacionó con la categoría asignada; la extracción, con los campos identificados; y la presentación de recomendaciones, con su utilidad y necesidad de ajustes. La confirmación humana y los registros de uso se revisaron como componentes adicionales del funcionamiento.

De esta manera, el análisis de requerimientos conectó la problemática inicial con las funciones desarrolladas y con los criterios utilizados para evaluar MiDocIA.

## 4.2. DISEÑO DE LA ARQUITECTURA DEL SISTEMA

La arquitectura de MiDocIA se diseñó bajo un modelo web cliente-servidor. Esta organización permitió distribuir las responsabilidades entre la interfaz que utiliza la persona, el servidor que coordina el procesamiento y los servicios externos que realizan operaciones específicas.

El diseño tomó como base los requerimientos definidos para la aplicación. Se consideró necesario recibir documentos, procesar temporalmente su contenido, presentar los resultados y mantener la revisión humana. También se estableció una

separación entre la información documental y los registros autorizados para el seguimiento académico.

A partir de estas condiciones se organizaron los componentes de presentación, aplicación e integración con servicios externos. Esta distribución permitió estructurar el recorrido de la información y ubicar las funciones de cada parte del sistema.

### *4.2.1.* ORGANIZACIÓN POR CAPAS

La capa de presentación corresponde a la interfaz desarrollada con React. Desde ella se gestiona la interacción con el usuario: consentimiento, selección del archivo, consulta del análisis, confirmación y valoración voluntaria.

La capa de aplicación corresponde al backend construido con Node.js y Express. Su responsabilidad es recibir las solicitudes de la interfaz, validar los archivos y coordinar las operaciones necesarias para obtener el análisis. También atiende las solicitudes de confirmación, encuesta y consulta administrativa.

Los servicios externos cumplen funciones diferenciadas. Google Cloud Document AI realiza la extracción de texto; Google Gemini clasifica y analiza el contenido; y Supabase conserva los registros autorizados de uso y evaluación.

La comunicación con estos servicios se concentra en el backend. Por tanto, la interfaz solicita el análisis a MiDocIA y recibe el resultado a través de su servidor. Esta organización permite mantener las integraciones y las credenciales de los proveedores dentro de la configuración del backend.

### 4.2.2. DISEÑO DE LA CAPA DE PRESENTACIÓN

La interfaz se organizó en componentes asociados con las actividades del usuario. El formulario permite seleccionar el documento, mientras que la vista de resultados presenta el contenido original junto con el análisis. Otros componentes atienden el consentimiento y la encuesta de utilidad.

Esta separación permite que cada parte de la interfaz tenga una responsabilidad concreta. Por ejemplo, el formulario reúne las acciones de selección y validación inicial del archivo, mientras que el componente de resultados organiza la categoría, el resumen, los datos relevantes, las alertas y la sugerencia generada. También se contemplaron los diferentes momentos de la interacción. La aplicación debe distinguir cuándo está esperando un archivo, cuándo se encuentra procesándolo y cuándo dispone de un resultado para mostrar. Si ocurre una dificultad, la interfaz presenta el mensaje correspondiente para que el usuario conozca el estado de la operación.

La presentación simultánea del documento y del análisis se incorporó para apoyar la revisión humana. Esta disposición permite contrastar nombres, fechas, valores y demás elementos sin depender únicamente del resumen o de la recomendación producida por el modelo.

### *4.2.3.* ORGANIZACIÓN INTERNA DEL BACKEND

El backend se estructuró mediante rutas, controladores, servicios y funciones intermedias de validación y manejo de errores. Las rutas identifican las operaciones que puede solicitar la interfaz. Entre ellas se encuentran analizar un documento, confirmar la revisión, registrar la encuesta y consultar las métricas administrativas. Esta organización permite dirigir cada solicitud hacia el componente encargado de atenderla. Los controladores coordinan las acciones necesarias para responder. En el análisis documental, el controlador recibe el archivo validado, solicita la extracción de texto, utiliza el contenido obtenido para el análisis con inteligencia artificial y devuelve la información a la interfaz. También coordina el registro de los metadatos correspondientes.

Los servicios reúnen la comunicación con Document AI, Gemini y Supabase. Esta separación permite ubicar la configuración y las operaciones de cada proveedor en componentes específicos. Por ejemplo, una modificación en la integración del OCR

puede revisarse en su servicio sin mezclarla con la presentación de resultados. Las funciones intermedias permiten comprobar las condiciones del archivo y gestionar errores durante la atención de las solicitudes. En conjunto, esta estructura distribuye las responsabilidades y facilita el seguimiento del proceso cuando se presenta una dificultad.

### *4.2.4.* COMUNICACIÓN ENTRE LOS COMPONENTES

El flujo de análisis inicia cuando la interfaz envía el documento al backend mediante una solicitud que contiene el archivo. El servidor verifica las condiciones de recepción y utiliza su contenido para solicitar la extracción de texto.
El texto obtenido mediante Document AI se emplea como entrada del análisis con Gemini. La integración contempla una primera llamada para identificar la categoría documental y una segunda para aplicar las instrucciones correspondientes a esa categoría.
Estas etapas tienen una dependencia entre sí. El análisis requiere el texto extraído, y las instrucciones especializadas dependen de la clasificación inicial. Por ello, el backend coordina el orden de las operaciones antes de devolver el resultado.
La respuesta dirigida a la interfaz contiene la información necesaria para presentar el análisis. El documento original continúa disponible en la vista de revisión, de manera que el usuario pueda compararlo con el resultado recibido.
La comunicación con Supabase cumple una función diferente. Este servicio recibe los datos autorizados de seguimiento, mientras que el contenido documental se utiliza temporalmente para completar el análisis. La separación evita que el registro de una actividad implique conservar también el archivo o la respuesta completa del modelo.

### *4.2.5.* PROCESAMIENTO TEMPORAL DEL CONTENIDO DOCUMENTAL

El manejo del archivo se diseñó de acuerdo con el enfoque Zero Storage definido para el proyecto. El backend recibe el documento como un búfer en memoria mediante multer.memoryStorage(), lo que permite utilizar su contenido durante la solicitud sin escribir una copia en el sistema de archivos del servidor. Cuando un PDF requiere división, los bloques se generan también en memoria. El texto obtenido de su procesamiento se reúne temporalmente para continuar con el análisis. Esta operación mantiene el mismo criterio de tratamiento del contenido durante las diferentes etapas.

Después de enviar la respuesta y dejar de utilizar las referencias temporales, la memoria puede ser recuperada por el entorno de ejecución. Este comportamiento no equivale a un procedimiento de borrado seguro certificado, sino al manejo temporal previsto en la arquitectura.

El alcance de Zero Storage corresponde a la infraestructura de MiDocIA. La aplicación envía información a Document AI y Gemini para realizar sus funciones, por lo que el procesamiento requiere circulación de datos hacia proveedores externos. La arquitectura tampoco conserva un historial de documentos o análisis completos que el usuario pueda recuperar posteriormente. Los registros de uso tienen un alcance distinto y se mantienen separados del contenido procesado.

### *4.2.6.* DISEÑO DE LOS REGISTROS DE SEGUIMIENTO

Los registros almacenados en Supabase permiten relacionar un análisis con las acciones posteriores del usuario. Cada registro cuenta con un identificador que se utiliza para asociar la confirmación y, cuando se diligencia, la encuesta de utilidad. Al obtener el análisis se registran datos como la fecha, la categoría documental y el tiempo de procesamiento. Cuando el usuario confirma la revisión, se actualiza el

estado correspondiente. Si responde la encuesta, se incorporan la calificación y los demás datos proporcionados en ese formulario.

Esta organización permite diferenciar tres eventos: obtener un resultado, confirmar su revisión y registrar una valoración voluntaria. La existencia del primer evento no implica que los otros dos se hayan completado.

El diseño de estos registros permite consultar métricas sin almacenar el documento original, el texto extraído ni el análisis completo. Sin embargo, algunos campos autorizados, como el correo electrónico o un comentario, pueden contener información personal; por ello, los registros no se describen como completamente anónimos. El panel administrativo consulta esta información a través del backend. Su acceso utiliza un token administrativo verificado por el servidor, separado de las operaciones ordinarias de análisis documental.

### 4.2.7. INCORPORACIÓN DE LA REVISIÓN HUMANA

El principio Human-in-the-Loop se incorporó mediante un paso de revisión y confirmación posterior a la presentación del análisis. La interfaz permite al usuario consultar el resultado y compararlo con el documento antes de finalizar el ciclo.

La acción de finalización se adapta al tipo de salida. Cuando se presenta una sugerencia de respuesta, el usuario puede copiarla al portapapeles y finalizar la revisión. Cuando se muestra una recomendación de acción, puede confirmar que la ha revisado mediante el control correspondiente.

La copia de una sugerencia permite utilizarla posteriormente en otro medio, pero no envía una comunicación desde MiDocIA. El usuario conserva la posibilidad de corregir, completar o descartar el texto antes de emplearlo.

En el backend, la confirmación se atiende mediante una solicitud diferente de la utilizada para analizar el documento. Esta separación permite registrar que la acción de confirmación ocurrió después de obtener el resultado.

El estado confirmado representa esa interacción dentro del flujo. No permite

comprobar por sí solo la profundidad de la revisión ni la exactitud de todos los campos, por lo que se interpreta como una métrica de uso y participación humana.

### *4.2.8.* ORGANIZACIÓN DEL DESPLIEGUE Y LA CONFIGURACIÓN

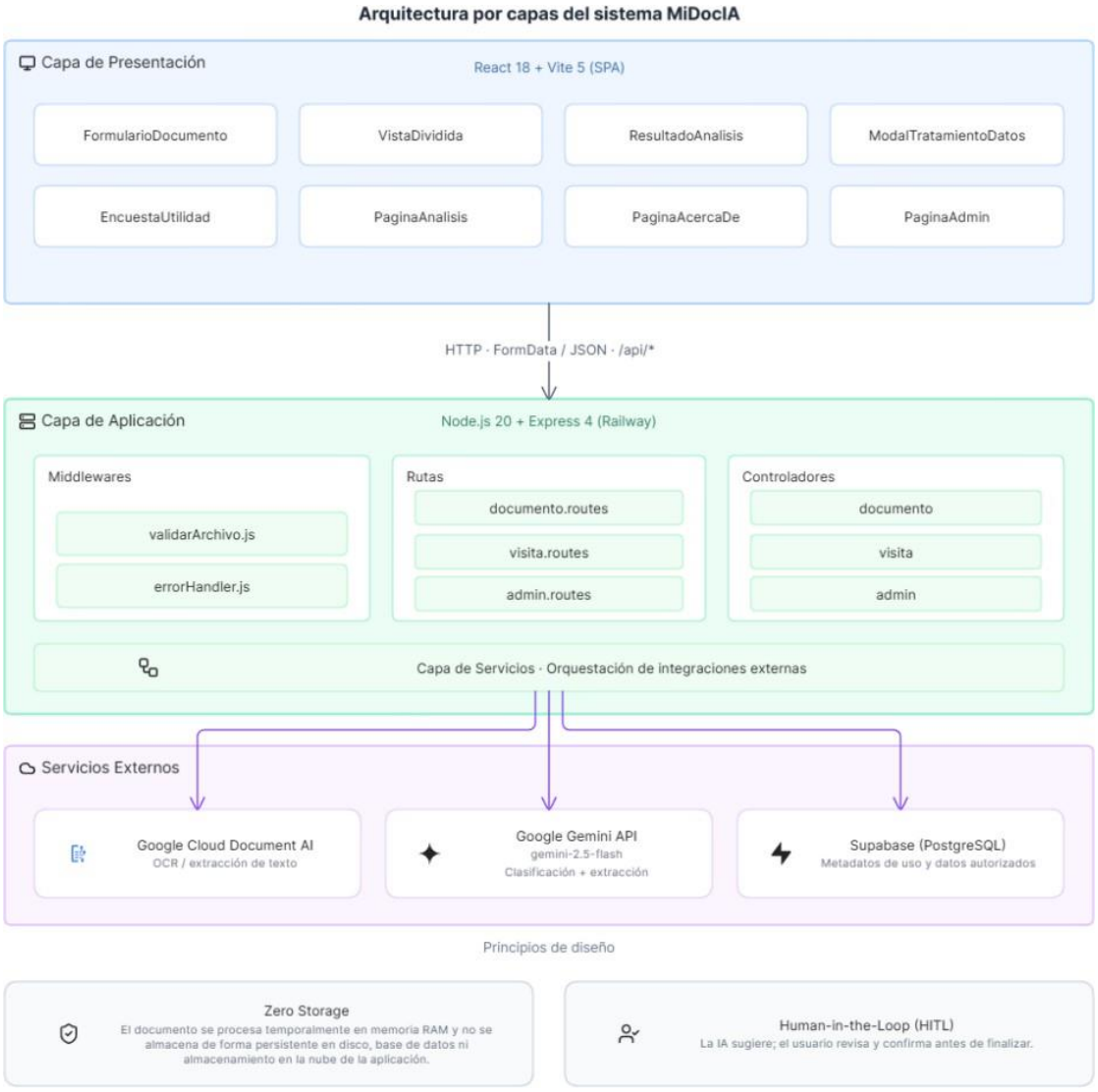
El proyecto reúne el frontend y el backend en un mismo repositorio, con directorios y dependencias diferenciados. Esta estructura permite mantener los componentes de la solución bajo una organización común.

En el entorno local, la interfaz y el servidor se ejecutan mediante procesos separados para facilitar el desarrollo. En producción, Express entrega los archivos compilados del frontend y atiende las solicitudes de la API desde un único servicio desplegado en Railway.

De esta forma, el navegador accede a la interfaz y a las operaciones del backend bajo el mismo dominio. Los servicios de OCR, inteligencia artificial y base de datos permanecen externos y son consultados por el servidor cuando se requieren. Las credenciales de estos proveedores se gestionan mediante la configuración del entorno. Esta separación permite configurar una instalación sin incorporar las claves de acceso en el código público ni en los archivos de la interfaz enviados al navegador.

La Ilustración 1 presenta la distribución general de los componentes y su relación con los servicios externos. También muestra los principios Zero Storage y Human-in-the-Loop, que orientan el manejo de la información y la participación del usuario dentro de la arquitectura.

**ILUSTRACIÓN 1 ARQUITECTURA GENERAL DE MIDOCIA**




**FUENTE: ELABORACIÓN PROPIA.**

## 4.3. IMPLEMENTACIÓN Y DESARROLLO

La implementación de MiDocIA comprendió la construcción de los componentes definidos en la arquitectura y su integración en un flujo funcional. Se desarrollaron las operaciones de recepción de documentos, extracción de texto, clasificación, análisis, presentación de resultados y confirmación humana. También se incorporaron el consentimiento, la encuesta voluntaria y el panel administrativo. El código se organizó en los directorios del frontend y del backend, manteniendo separados los componentes de la interfaz y los servicios encargados del procesamiento. Esta distribución permitió trabajar sobre cada función y posteriormente comprobar su interacción dentro de la aplicación.

### 4.3.1. CONFIGURACIÓN DEL ENTORNO DE TRABAJO

El entorno del proyecto se preparó con Node.js y npm para instalar las dependencias y ejecutar las tareas de desarrollo. La organización mediante npm workspaces permitió administrar desde la raíz del repositorio los paquetes correspondientes al frontend y al backend.

Para la ejecución local se configuraron las variables necesarias para acceder a Google Cloud Document AI, Gemini y Supabase. También se estableció el token utilizado para proteger las consultas administrativas.

Las credenciales se mantuvieron separadas del código fuente publicado. En el entorno local, la autenticación de Google Cloud utiliza un archivo de credenciales ubicado fuera de los archivos incluidos en el repositorio. En producción, esta información se proporciona mediante la configuración del servicio desplegado. Esta preparación permitió disponer de los accesos necesarios antes de ejecutar el flujo documental. Los procedimientos de instalación y configuración quedaron descritos en el Manual de Instalación y Mantenimiento para facilitar su reproducción.

### 4.3.2. CONSENTIMIENTO Y SELECCIÓN DEL DOCUMENTO

En la interfaz se incorporó un componente para presentar las condiciones del tratamiento de datos y solicitar la aceptación del usuario antes de utilizar el procesamiento documental. Este componente incluye el correo electrónico y el control de aceptación.

La aplicación conserva en el navegador un indicador del consentimiento, de acuerdo con el funcionamiento descrito en el manual de usuario. Este indicador permite reconocer la aceptación durante el uso posterior de la aplicación y se encuentra separado del contenido de los documentos.

Después de este paso, el usuario puede seleccionar un archivo mediante el formulario de carga. En el frontend se incorporaron comprobaciones del formato y del tamaño para informar si el archivo cumple las condiciones de entrada antes de enviarlo al servidor.

Cuando se solicita el análisis, la interfaz construye una solicitud FormData que contiene el archivo en el campo documento. Esta solicitud se dirige a la ruta del backend encargada de iniciar el procesamiento.

### *4.3.3.* RECEPCIÓN Y VALIDACIÓN DE ARCHIVOS EN EL BACKEND

La recepción se implementó mediante Multer configurado con memoryStorage(). Esta opción permite disponer del archivo como un búfer temporal en memoria, sin generar una copia persistente en el sistema de archivos del servidor.
El componente de validación comprueba la presencia del archivo, el tipo MIME admitido y el tamaño máximo de 10 MB. Los formatos contemplados son PDF, PNG, JPEG y TIFF.

Las comprobaciones del backend complementan las realizadas en el navegador. De esta manera, el servidor verifica las condiciones de entrada antes de entregar el archivo al controlador del análisis.

Cuando la solicitud no cumple estas condiciones, el proceso de validación devuelve un error y evita continuar hacia los servicios de OCR e inteligencia artificial. La interfaz presenta el mensaje correspondiente para que el usuario pueda revisar el archivo seleccionado.

Una vez superada la validación, el controlador utiliza el búfer recibido para iniciar la extracción de texto. El documento permanece dentro del procesamiento temporal previsto para la aplicación.

### 4.3.4. IMPLEMENTACIÓN DE LA EXTRACCIÓN DE TEXTO

La integración con Google Cloud Document AI se desarrolló en el servicio de OCR del backend. Este componente prepara la solicitud, obtiene la autenticación necesaria y envía el contenido al procesador configurado. Para el envío, el contenido binario del documento se codifica en base64 y se incorpora en la solicitud realizada mediante HTTPS. El servicio recibe la respuesta de Document AI y obtiene el texto que se utilizará posteriormente en el análisis. En los documentos PDF que superan catorce páginas se utiliza pdf-lib para generar bloques de hasta catorce páginas. Estos bloques se procesan de manera secuencial y el texto obtenido se concatena para continuar con la siguiente etapa. La división corresponde al manejo interno del archivo y se realiza en memoria. El usuario carga un documento y recibe un resultado integrado, sin tener que dividirlo manualmente antes de utilizar MiDocIA.

Este procedimiento permite trabajar con PDF extensos dentro de las condiciones de la aplicación. El tamaño máximo del archivo continúa siendo de 10 MB y el

procesamiento depende de los recursos disponibles y de la respuesta del servicio externo.

### 4.3.5. CLASIFICACIÓN DOCUMENTAL CON GEMINI

El componente de inteligencia artificial se implementó mediante el modelo gemini-2.5-flash. La primera llamada al modelo se utiliza para identificar la categoría del documento.
Para esta clasificación se prepara una muestra del texto extraído, tomando hasta 2.500 caracteres del inicio y 1.500 de su parte media. Esta selección permite incorporar información de dos zonas del documento, en lugar de depender únicamente de sus primeras líneas.

El resultado de esta etapa se utiliza para seleccionar las instrucciones del análisis especializado. Las categorías previstas son factura, contrato, PQRS, documento judicial, carta, cotización y otro.

La clasificación funciona, por tanto, como una decisión intermedia dentro del procesamiento. Su resultado orienta los campos que se solicitan en la siguiente llamada al modelo. Si la categoría identificada no corresponde al contenido principal, el análisis posterior puede enfocarse en información que no sea la más pertinente; por esta razón, la categoría también se muestra al usuario para su revisión.

### 4.3.6. ANÁLISIS ESPECIALIZADO Y ORGANIZACIÓN DE LA RESPUESTA

Después de identificar la categoría, el backend selecciona el prompt correspondiente y envía el texto extraído para su análisis. Las instrucciones especifican los campos que deben buscarse y solicitan el resumen, las alertas y la sugerencia de respuesta o recomendación de acción.
Esta separación permite adaptar la solicitud al tipo documental. En una factura se

orienta la extracción hacia sus datos comerciales y valores; en un contrato se consideran las partes, el objeto y las obligaciones; y en una comunicación se revisan el asunto y la acción solicitada.

También se incorpora la fecha del sistema como información de contexto para el análisis temporal. Este dato puede apoyar la interpretación de fechas, pero no aporta por sí mismo información que falte en el documento, como una fecha de recepción o notificación.

La temperatura del modelo se configuró en cero con el propósito de reducir la variabilidad de las respuestas. Esta configuración forma parte de los ajustes de generación y no garantiza que todos los resultados sean idénticos o estén libres de errores. Para integrar la respuesta con la interfaz se utiliza una estructura JSON. El servicio incluye una función de limpieza que retira delimitadores de bloques de código cuando aparecen en la respuesta, antes de convertirla en un objeto de JavaScript.
Este tratamiento permite organizar los datos para su presentación. La posibilidad de interpretar correctamente la estructura recibida se distingue de la exactitud de su contenido, que debe revisarse frente al documento original.

### *4.3.7.* CONSTRUCCIÓN DE LA VISTA DE RESULTADOS

En el frontend se desarrolló una vista dividida para presentar el documento original y el análisis. La visualización utiliza el archivo seleccionado por el usuario, mientras que el panel de resultados muestra la información devuelta por el backend. Para representar el documento se utiliza una dirección temporal generada en el navegador. Esta permite mostrar el archivo sin crear un repositorio permanente de documentos dentro de la aplicación. El componente contempla la liberación de esa dirección cuando deja de utilizarse.

El resultado se organiza en secciones de categoría documental, resumen, datos relevantes, alertas y sugerencia de respuesta o recomendación de acción. También se incorporó el tratamiento de valores que pueden llegar como listas u objetos, para presentarlos de forma legible.

La interfaz se adapta al contenido recibido y al tipo documental identificado. Esta organización permite que el usuario ubique los datos y los contraste con el archivo antes de utilizar la información generada.

### *4.3.8.* REGISTRO DEL ANÁLISIS, CONFIRMACIÓN Y ENCUESTA

Al completar el análisis, el backend registra en Supabase los datos autorizados de seguimiento y obtiene el identificador del registro creado. Este identificador se devuelve a la interfaz junto con el resultado para asociar las acciones posteriores. El registro inicia con el estado de confirmación pendiente. Después de revisar el documento y el análisis, el usuario utiliza el control de finalización y la interfaz envía una solicitud a la ruta de confirmación. El backend actualiza entonces el campo correspondiente.

Para las sugerencias de respuesta se implementó la opción de copiar el texto al portapapeles y finalizar. En las recomendaciones de acción se utiliza el control de confirmación previsto para ese tipo de salida. La copia permite trasladar el texto a otro medio, pero no realiza su envío desde MiDocIA.

Después de la confirmación se presenta la encuesta voluntaria. Esta permite registrar la calificación mediante estrellas, la respuesta sobre tiempo y un comentario de hasta 200 caracteres. Los datos suministrados se asocian con el mismo registro del análisis.

La aplicación diferencia así el procesamiento, la confirmación y la valoración. Un análisis registrado puede quedar sin confirmación o sin encuesta si el usuario no

completa esas acciones. Esta diferencia se conserva al interpretar las métricas de uso.

### 4.3.9. IMPLEMENTACIÓN DEL PANEL ADMINISTRATIVO

Se desarrolló un panel para consultar los registros de uso y las métricas calculadas por el backend. La interfaz administrativa realiza solicitudes a rutas protegidas mediante el token configurado para esta función.

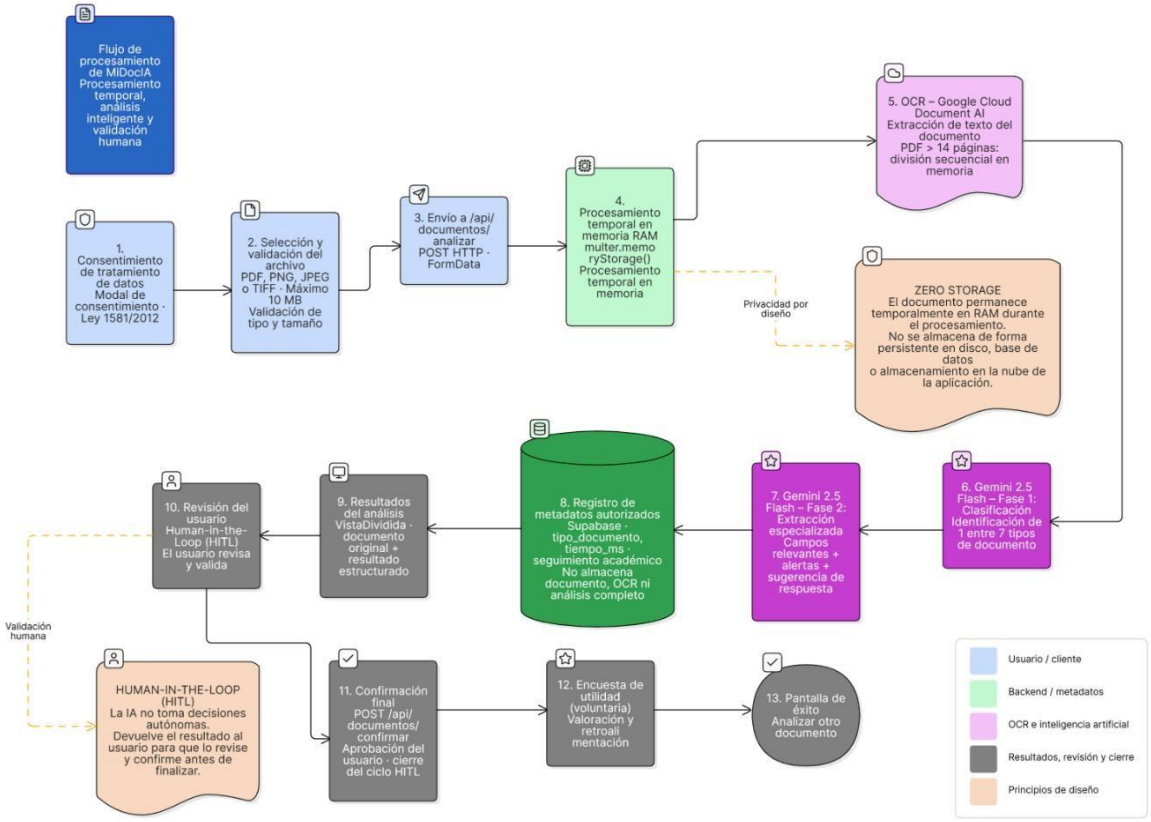
El panel permite consultar el número de análisis, las confirmaciones, las valoraciones y los tiempos de procesamiento. También presenta la distribución por categoría documental y la actividad reciente registrada.
Estas consultas se realizan sobre los datos de seguimiento almacenados en Supabase. El panel no permite recuperar el documento original ni el análisis completo, porque estos contenidos no forman parte de los registros persistentes de MiDocIA.
La información administrativa complementa la evaluación manual realizada con el instrumento del proyecto. Su función es mostrar el uso registrado por la aplicación, mientras que la revisión detallada de campos, alertas y recomendaciones se documenta en la matriz de evaluación.

### *4.3.10.* DESPLIEGUE Y DOCUMENTACIÓN DE LA SOLUCIÓN

La aplicación se desplegó en Railway a partir del repositorio del proyecto. Durante la preparación del servicio se configuraron las variables de entorno y las tareas necesarias para instalar dependencias, compilar la interfaz e iniciar el backend. En producción, Express entrega los archivos generados por Vite y atiende las rutas de la API. Esta configuración permite que el frontend y el backend funcionen desde el mismo dominio y que el usuario acceda a MiDocIA mediante un navegador.
La implementación se acompañó de dos documentos. El Manual de Usuario explica

el consentimiento, la carga, la interpretación de resultados y la confirmación humana. El Manual de Instalación y Mantenimiento describe la preparación del entorno, la configuración de servicios, el despliegue y las actividades de soporte. Como resultado se integraron las funciones del sistema en un recorrido que inicia con el consentimiento y termina con la revisión, la confirmación y la valoración voluntaria. La Ilustración 2 presenta este flujo y permite observar la relación entre las acciones del usuario, el procesamiento del backend y los servicios externos.

**ILUSTRACIÓN 2 FLUJO DE PROCESAMIENTO DE MIDOCIA.**




**FUENTE: ELABORACIÓN PROPIA.**

## 4.4. PRUEBAS Y OPTIMIZACIONES

Durante el desarrollo de MiDocIA se realizaron pruebas funcionales para revisar los componentes y su integración. Estas actividades permitieron identificar dificultades en la comunicación con los servicios externos, ajustar el procesamiento documental y comprobar el recorrido de la información hasta su presentación al usuario. Las pruebas de desarrollo se utilizaron para corregir el funcionamiento de la aplicación antes de la evaluación formal. En esta etapa se revisaron las operaciones del sistema y se realizaron los ajustes descritos a continuación. La evaluación con el conjunto documental y sus resultados se presentan en el capítulo 5.

### *4.4.1.* AJUSTE DE LA COMUNICACIÓN CON EL SERVICIO OCR

Durante la integración inicial con Google Cloud Document AI se presentó una dificultad relacionada con la serialización del parámetro fieldMask mediante el cliente gRPC utilizado. Este parámetro permite indicar los campos que se solicitan en la respuesta del servicio, pero en la configuración empleada no se transmitía de la forma esperada.

Para resolver esta dificultad se implementaron las solicitudes mediante el módulo HTTPS de Node.js y la API REST de Document AI. Este cambio permitió construir directamente el contenido de la solicitud y controlar los parámetros enviados al procesador. El ajuste se concentró en la comunicación entre el backend y el servicio OCR. Se mantuvo la función de extraer el texto del documento y entregarlo al componente de análisis, así como el procesamiento temporal del archivo en memoria. La dificultad descrita corresponde a la integración utilizada durante el proyecto. No se plantea como una limitación general de todas las versiones del cliente gRPC ni de todos los entornos que emplean Document AI.

### *4.4.2.* ADECUACIÓN DEL PROCESAMIENTO DE PDF EXTENSOS

El manejo de archivos PDF requirió considerar tanto el tamaño permitido por MiDocIA como las condiciones de procesamiento del servicio OCR. Para atender documentos extensos se incorporó la división mediante pdf-lib en bloques de hasta catorce páginas.

Cada bloque se envía de manera secuencial al servicio y el texto obtenido se reúne para continuar con el análisis. La división se realiza internamente, de modo que el usuario mantiene la operación de cargar un único archivo.

Este ajuste permitió conservar el enfoque Zero Storage durante el tratamiento de documentos extensos. Tanto los bloques como el texto extraído se manejan temporalmente en memoria, sin crear carpetas de archivos intermedios dentro del servidor. La cantidad de bloques influye en las operaciones necesarias para completar el OCR. Por ello, el procesamiento de un documento extenso puede requerir más tiempo que el de uno corto, aunque ambos cumplan el límite de 10 MB establecido para la carga.

### 4.4.3. SEPARACIÓN DE LA CLASIFICACIÓN Y EL ANÁLISIS ESPECIALIZADO

En las pruebas iniciales del componente de inteligencia artificial se utilizó una solicitud que reunía la clasificación y la extracción de información. Con esta organización se observaron resultados menos consistentes, porque las instrucciones generales debían atender simultáneamente documentos con propósitos y campos diferentes.

El ajuste consistió en dividir el procesamiento en dos llamadas. La primera identifica la categoría documental y la segunda utiliza instrucciones específicas para esa categoría. De esta forma, la solicitud de análisis puede concentrarse en los datos pertinentes para una factura, un contrato, una PQRS u otro tipo de documento.

También se utilizó una muestra de texto procedente del inicio y de la parte media para apoyar la clasificación. Esta decisión buscó incorporar información adicional a las primeras líneas, que pueden contener encabezados o elementos secundarios. La separación de las llamadas permitió organizar con mayor claridad las instrucciones enviadas al modelo. Sin embargo, el proyecto no presenta una comparación cuantitativa independiente entre ambas configuraciones; los indicadores del capítulo de resultados corresponden a la versión utilizada en la evaluación formal.

### *4.4.4.* AJUSTES EN LA GENERACIÓN Y PREPARACIÓN DE LAS RESPUESTAS

Durante el desarrollo se identificaron variaciones en la respuesta generada por Gemini. Entre las situaciones documentadas se encontraba la aparición de la sugerencia de respuesta en algunos análisis y su ausencia en otros. Para atender esta dificultad se reforzó la solicitud de ese campo dentro de las instrucciones y se configuró la temperatura del modelo en cero. El propósito fue reducir la variabilidad y favorecer una respuesta que incluyera los elementos requeridos por la aplicación.

También se incorporó una función para retirar delimitadores de bloques de código cuando aparecían alrededor del JSON generado. Esta preparación facilita que el backend interprete la respuesta y entregue sus datos a la interfaz. Los ajustes atienden aspectos diferentes: las instrucciones orientan el contenido solicitado, la temperatura modifica la configuración de generación y la limpieza prepara la estructura recibida para su procesamiento. Ninguno de estos mecanismos reemplaza la comparación del resultado con el documento original.

### *4.4.5.* COMPROBACIÓN DEL FLUJO COMPLETO

Después de integrar los componentes se realizaron pruebas del recorrido completo de la aplicación. Se revisaron la carga del documento, la extracción de texto, la clasificación, la identificación de información relevante y la presentación de alertas y recomendaciones.

La comprobación del flujo permitió revisar que el resultado de una etapa estuviera disponible para la siguiente. Por ejemplo, el texto obtenido mediante OCR debía llegar al servicio de análisis y la respuesta generada debía presentarse en las secciones correspondientes de la interfaz.

También se verificaron la confirmación del usuario y el registro de métricas. Estas operaciones forman parte del cierre del proceso y permiten distinguir entre un análisis obtenido y uno cuya revisión fue confirmada.

La revisión del funcionamiento completo complementó las pruebas sobre componentes individuales. Un servicio podía responder correctamente de manera aislada y aun requerir ajustes para que su resultado fuera utilizado por el resto de la aplicación.

### *4.4.6.* ALCANCE DE LAS OPTIMIZACIONES REALIZADAS

Las optimizaciones se concentraron en resolver las dificultades identificadas durante la construcción e integración de MiDocIA. Comprendieron la comunicación con el OCR, el tratamiento de PDF extensos, la organización de las llamadas al modelo y la preparación de las respuestas para su presentación.

En este contexto, la optimización se entiende como el ajuste del funcionamiento para completar el proceso previsto. No se atribuyen porcentajes de mejora a cada modificación, porque no se documentó una medición comparativa independiente para todos los cambios.

Asimismo, las pruebas funcionales permitieron revisar las operaciones

contempladas en el proyecto, pero sus resultados no constituyen una garantía de ausencia de errores en cualquier documento o condición de uso. La exactitud de la información y la utilidad de las recomendaciones se analizaron posteriormente mediante el instrumento de evaluación.

### *4.4.7.* VERSIÓN UTILIZADA PARA LA EVALUACIÓN

Después de los ajustes se obtuvo la versión funcional utilizada en la fase formal de evaluación. Esta versión integró la recepción del archivo, el OCR, el análisis especializado, la visualización del documento y del resultado, la confirmación humana y los registros autorizados de seguimiento.
La evaluación se realizó con documentos reales suministrados de manera confidencial por la empresa colaboradora y documentos sintéticos preparados para las pruebas académicas. Su propósito fue observar el desempeño de la aplicación frente al conjunto seleccionado y registrar los aciertos, las dificultades y los tiempos obtenidos. La separación entre las pruebas de desarrollo y la evaluación formal permite identificar la función de cada actividad. Las primeras apoyaron la construcción y corrección del sistema; la segunda aportó los datos utilizados para valorar su desempeño y utilidad. Las evidencias y los resultados de esta evaluación se presentan en el capítulo siguiente.

## 5. RESULTADOS

Como resultado del proyecto se obtuvo una versión funcional de MiDocIA, desplegada en un entorno web, que integra la recepción de documentos, la extracción de texto mediante OCR y el análisis con inteligencia artificial. La aplicación presenta la categoría documental, un resumen, datos relevantes, alertas y una sugerencia de respuesta o recomendación de acción.
El producto desarrollado incorpora la revisión humana como parte del proceso. El usuario puede consultar el documento original junto con el análisis, comparar la

información y confirmar la finalización de la revisión. También dispone de una encuesta voluntaria para valorar la utilidad del resultado.
La evaluación se realizó con 32 documentos, de los cuales 20 fueron reales de origen empresarial y 12 correspondieron a documentos sintéticos. Los resultados se obtuvieron mediante la comparación manual registrada por la persona evaluadora de la empresa colaboradora y mediante los registros de uso generados por MiDocIA.

En este capítulo se presenta la aplicación obtenida, la composición del conjunto documental y los resultados de la evaluación. Se distinguen las mediciones de clasificación, extracción y utilidad de los datos relacionados con los tiempos, la confirmación humana y la encuesta voluntaria.

## 5.1. APLICACIÓN WEB OBTENIDA

MiDocIA quedó implementada como una aplicación web que permite realizar el análisis documental desde un navegador. La versión desarrollada reúne las funciones necesarias para seleccionar un archivo, procesarlo y consultar la información generada dentro de una misma interfaz.

El acceso se realiza sin instalar la aplicación en el equipo del usuario. La interfaz contempla su utilización desde computadores, tabletas y dispositivos móviles, mientras que el procesamiento se coordina desde el servidor y los servicios externos integrados.

### 5.1.1. INTERFAZ DE CARGA Y ESTADO DEL PROCESAMIENTO

La pantalla principal incluye el formulario para seleccionar documentos en formato PDF, PNG, JPEG o TIFF, con un tamaño máximo de 10 MB. En esta sección se muestran los formatos admitidos, el control de selección y el nombre del archivo cargado.

La Ilustración 3 presenta la interfaz durante el procesamiento de un documento. En la captura se observa el archivo seleccionado, el botón con el estado “Analizando…” y un mensaje que informa que la operación se encuentra en curso. Estos elementos permiten identificar el documento enviado y reconocer que la aplicación está atendiendo la solicitud. La captura evidencia una etapa del funcionamiento de MiDocIA, previa a la presentación del resultado. La pantalla también reúne los accesos al análisis documental, la información del proyecto y el apartado administrativo. Esta distribución mantiene la operación principal de carga en el área central de la interfaz.

### 5.1.2. PRESENTACIÓN DEL DOCUMENTO Y DEL ANÁLISIS

Una vez obtenido el resultado, la aplicación presenta el documento original junto con el análisis generado. Esta vista permite consultar ambas fuentes dentro de la misma pantalla y comparar la información antes de utilizarla. La Ilustración 4 muestra un archivo clasificado por la aplicación como carta comercial. En el panel izquierdo aparece el documento y en el derecho se observan el tipo identificado, el resumen, una alerta y los datos relevantes extraídos. La disposición de estos elementos permite seguir la relación entre el archivo y su interpretación. El usuario puede revisar, por ejemplo, si el asunto resumido corresponde al contenido de la carta o si las fechas y personas identificadas aparecen en el documento original.

La captura muestra cómo se presenta el resultado de un caso concreto. La exactitud de la clasificación y de los campos se determina mediante su comparación con el documento, por lo que la visualización del análisis y la valoración de su contenido se consideran aspectos diferentes.

### 5.1.3. ORGANIZACIÓN DE LA INFORMACIÓN GENERADA

El resultado se distribuye en secciones que permiten consultar diferentes aspectos del documento. La categoría identifica el tipo asignado por el sistema, mientras que el resumen presenta una descripción de su contenido principal. Los datos relevantes reúnen la información extraída de acuerdo con la categoría. Esta organización permite ubicar elementos como remitentes, destinatarios, fechas, valores o solicitudes sin depender únicamente de una respuesta extensa en texto continuo.

Las alertas se presentan en un espacio diferenciado para señalar asuntos que pueden requerir atención. En el ejemplo de la Ilustración 4, la aplicación muestra una alerta relacionada con un plazo de respuesta identificado en la carta. Su presentación permite localizar ese asunto para revisarlo frente al documento. La salida también incluye una sugerencia de respuesta o una recomendación de acción según el tipo documental. Este contenido ofrece un punto de partida que el usuario puede ajustar antes de emplearlo en una actividad administrativa.

### 5.1.4. REVISIÓN Y FINALIZACIÓN DEL ANÁLISIS

La versión obtenida incorpora controles para que el usuario confirme la revisión después de consultar el resultado. Cuando se presenta una sugerencia de respuesta, dispone de la opción de copiar el texto y finalizar. En los casos de recomendación de acción, puede confirmar su lectura mediante el control correspondiente.

Esta interacción mantiene una intervención humana entre la generación del análisis y su utilización posterior. La aplicación no envía comunicaciones, realiza pagos ni aprueba documentos de manera automática.

Después de la confirmación se presenta la encuesta voluntaria. Sus respuestas permiten registrar la percepción de utilidad y complementar la información de

seguimiento. La confirmación y la encuesta constituyen acciones distintas, por lo que se analizan por separado en los resultados de uso.

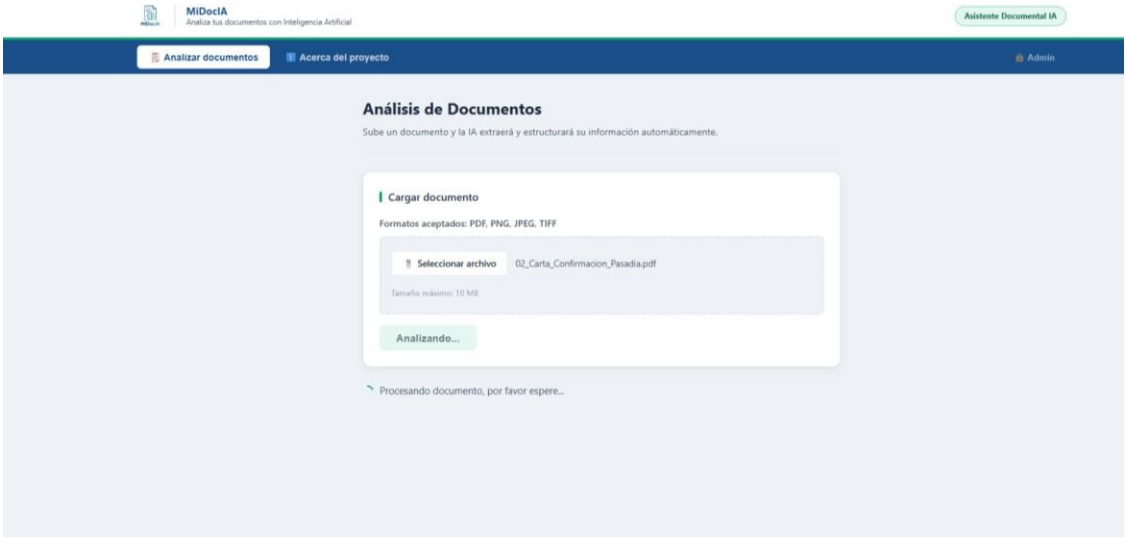
### 5.1.5. ALCANCE DEL PRODUCTO OBTENIDO

El resultado funcional del proyecto es una herramienta que reúne en un mismo flujo la lectura automatizada del documento, la organización de información y la presentación de orientaciones para su revisión.

El procesamiento temporal permite realizar estas operaciones sin conservar de forma persistente el archivo original, el texto OCR o el análisis completo en la infraestructura de MiDocIA. Los registros almacenados corresponden a los datos autorizados de seguimiento, que pueden consultarse desde el panel administrativo.

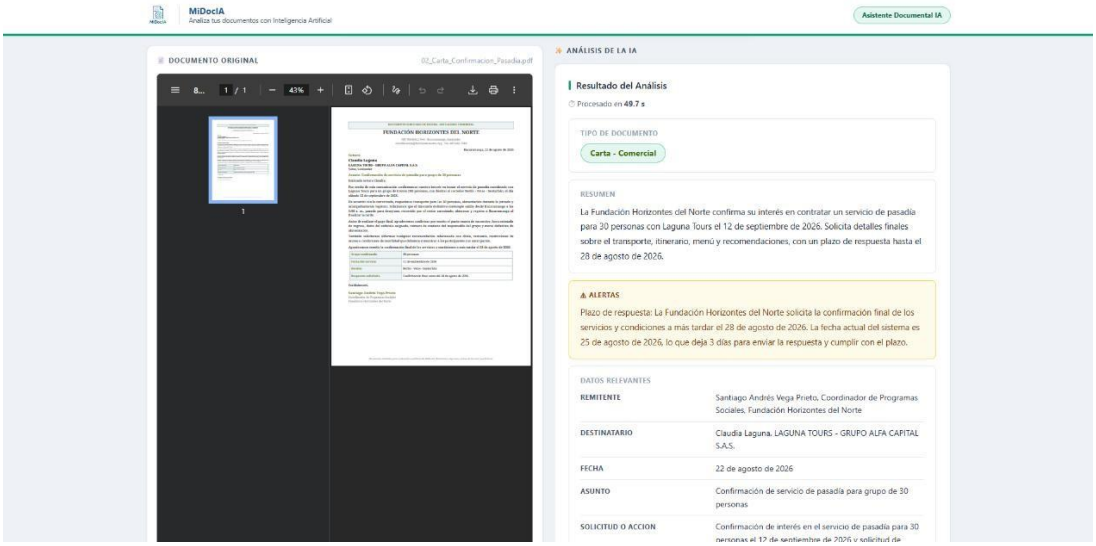
Las ilustraciones 3 y 4 evidencian la interfaz de carga y la presentación del análisis. La valoración del desempeño de la aplicación se complementa en los siguientes apartados con los datos del conjunto documental, las calificaciones del instrumento y las métricas registradas durante la evaluación.

**ILUSTRACIÓN 3 INTERFAZ PRINCIPAL DE MIDOCIA**




**FUENTE: ELABORACIÓN PROPIA.**

**ILUSTRACIÓN 4 VISUALIZACIÓN DEL DOCUMENTO Y RESULTADO DEL ANÁLISIS**




**FUENTE: ELABORACIÓN PROPIA.**

## 5.2. CONJUNTO DOCUMENTAL UTILIZADO PARA LA EVALUACIÓN

La evaluación de MiDocIA se realizó con 32 documentos, organizados en el Instrumento de Validación del proyecto. De este total, 20 correspondieron a documentos reales de origen empresarial y 12 a documentos sintéticos controlados. Esta combinación permitió revisar el funcionamiento de la aplicación con información utilizada en actividades administrativas y complementar las pruebas con casos preparados para las categorías contempladas en el desarrollo. La composición del conjunto es importante para interpretar los resultados, porque permite reconocer qué documentos tuvieron mayor participación y bajo qué condiciones se realizaron las pruebas. Los indicadores presentados en los siguientes apartados corresponden a estos 32 casos y deben entenderse dentro de ese alcance.

### 5.2.1. ORIGEN DE LOS DOCUMENTOS

Los documentos reales correspondieron a información empresarial de Laguna Tours. En la matriz se registraron 18 cotizaciones y dos documentos relacionados con propuestas turísticas, cuya categoría esperada fue otro. Estos archivos permitieron revisar el análisis de contenidos asociados con servicios, valores, condiciones comerciales e información sobre actividades turísticas. Los 12 documentos sintéticos controlados complementaron el conjunto mediante casos de cartas, PQRS, contratos, documentos judiciales, facturas y otros documentos. Su utilización permitió establecer previamente la información esperada y contrastarla con los resultados presentados por MiDocIA.

La Tabla 3 muestra la distribución de los documentos según su origen.

**TABLA 3 COMPOSICIÓN DEL CONJUNTO DE EVALUACIÓN**

| Origen | Cantidad | Participación |
|---|---|---|
| Documentos reales | 20 | 62,50 % |
| Documentos sintéticos controlados | 12 | 37,50 % |
| **Total** | **32** | **100 %** |

Fuente: elaboración propia a partir de la matriz de validación.

Los documentos reales representaron la mayor parte del conjunto. Por su parte, los documentos sintéticos permitieron ampliar la variedad de contenidos evaluados. Ambos grupos aportaron información para revisar el comportamiento de la aplicación, aunque los casos sintéticos no equivalen a una validación de todas las situaciones que pueden presentarse en una empresa.

### 5.2.2. DISTRIBUCIÓN POR CATEGORÍA DOCUMENTAL

La distribución se estableció a partir de la categoría esperada registrada en el banco documental. Esta corresponde al tipo de documento utilizado como referencia para la evaluación y se distingue de la categoría detectada por MiDocIA durante el procesamiento.

El conjunto incluyó 18 cotizaciones, cuatro documentos de la categoría otro, tres facturas, dos cartas, dos PQRS, dos contratos y un documento judicial. De esta manera, se incorporaron casos de las siete categorías definidas para el proyecto. Las cotizaciones representaron el 56,25 % de los documentos evaluados. Esta participación se relaciona con la composición de los archivos empresariales disponibles y debe considerarse al analizar los resultados generales. Las demás categorías tuvieron una cantidad menor de casos, especialmente la de documento judicial, que contó con un solo archivo.

Por esta razón, la evaluación permitió observar el funcionamiento de MiDocIA en

las siete categorías, pero no ofrece la misma cantidad de evidencia para cada una. El resultado global tiene una mayor participación de las cotizaciones y no debe interpretarse como una medida uniforme del desempeño en todos los tipos documentales.

### 5.2.3. FORMATO, EXTENSIÓN Y CALIDAD VISUAL

Los 32 documentos registrados en la matriz se encontraban en formato PDF. Aunque MiDocIA permite cargar otros formatos, los resultados de este conjunto corresponden específicamente al procesamiento de archivos PDF.
En cuanto a su extensión, 27 documentos tenían una página, tres tenían dos páginas, uno tenía catorce páginas y otro tenía veinte páginas. La mayoría de los casos correspondió, por tanto, a documentos breves, acompañados de dos archivos de mayor extensión.

Esta distribución permitió incluir contenidos con diferentes cantidades de información. Sin embargo, la presencia de un documento de veinte páginas no constituye una prueba de rendimiento para cualquier extensión posible, ya que el comportamiento también depende del tamaño del archivo, de su contenido y de las condiciones del procesamiento.

En el banco documental, la calidad visual de los 32 archivos se registró como alta. Esta característica debe tenerse en cuenta al interpretar la evaluación del reconocimiento de texto, debido a que el conjunto no permite establecer cómo se comportaría la aplicación ante documentos borrosos, deteriorados o de baja legibilidad.

### 5.2.4. ORGANIZACIÓN DE LA INFORMACIÓN PARA LA EVALUACIÓN

Cada documento se identificó mediante un código dentro del instrumento. Este código permitió relacionar su registro en el banco documental con la evaluación general y con la revisión individual de los campos esperados.
El banco documental reunió información sobre el origen, el tipo de documento, la categoría esperada, el formato, el número de páginas, la calidad visual y las fechas de recepción y prueba. La hoja de evaluación registró aspectos como la clasificación obtenida, la calidad del OCR, la extracción de información, la utilidad de las recomendaciones y los tiempos empleados.

La hoja de campos evaluados permitió detallar qué información se identificó correctamente, cuál se obtuvo de manera parcial y cuál no fue detectada. Esta organización facilitó la comparación entre lo esperado y lo presentado por MiDocIA, manteniendo la relación de cada resultado con el documento correspondiente.

### 5.2.5. CONFIDENCIALIDAD Y ALCANCE DEL CONJUNTO

Los documentos reales utilizados durante la evaluación no se reproducen en el informe debido al carácter confidencial de su contenido. La información consolidada de su evaluación se presenta en la matriz anexa con autorización de la empresa colaboradora.
El conjunto utilizado permitió realizar una evaluación práctica de MiDocIA con documentos empresariales y casos sintéticos controlados. Su composición delimita el alcance de los resultados: corresponde a 32 archivos PDF de calidad visual alta, con predominio de cotizaciones y participación de una empresa colaboradora. Estas condiciones permiten valorar el funcionamiento de la aplicación en los casos examinados. Para ampliar la evidencia sería necesario incorporar documentos de

otras empresas, aumentar la cantidad de casos por categoría y considerar archivos con diferentes formatos y condiciones de legibilidad.

## 5.3. RESULTADOS DE LA EVALUACIÓN MANUAL

Cada uno de los 32 análisis generados por MiDocIA fue comparado con el documento correspondiente por la persona designada por la empresa colaboradora. La valoración se registró en el instrumento de validación del proyecto y comprendió la clasificación documental, la calidad del texto obtenido mediante OCR, la extracción de campos, las alertas, la utilidad de las recomendaciones y los tiempos empleados.

Esta revisión permitió identificar tanto los aciertos de la aplicación como los aspectos que requirieron intervención humana. La Tabla 4 presenta los resultados generales; posteriormente se explica su significado y se describen algunas de las observaciones registradas durante las pruebas.

**TABLA 4 RESULTADOS GENERALES DE LA EVALUACIÓN MANUAL**

| Indicador | Resultado |
|---|---|
| Documentos evaluados | **32** |
| Clasificación exacta | **30 de 32** |
| Precisión de clasificación | **93,75 %** |
| Clasificaciones parciales | **2 de 32** |
| Calidad promedio del OCR | **4,97 / 5** |
| Campos individuales evaluados | **356** |
| Precisión promedio de extracción | **82,33 %** |
| Utilidad promedio de la recomendación | **3,56 / 5** |
| Tiempo manual promedio | **12,16 min** |
| Tiempo total promedio utilizando MiDocIA | **5,91 min** |
| Tiempo promedio ahorrado | **6,25 min** |
| Reducción promedio del tiempo | **50,65 %** |

Fuente: elaboración propia a partir de la matriz de evaluación.

Los indicadores de tiempo se incluyen como parte del resumen general de la evaluación. Su cálculo e interpretación se desarrollan en el siguiente apartado.

### 5.3.1. CLASIFICACIÓN DOCUMENTAL

En la valoración de la clasificación, 30 documentos recibieron una calificación de correcta y dos fueron considerados parciales. La precisión se obtuvo al dividir los 30 casos correctos entre los 32 documentos evaluados y multiplicar el resultado por cien, obteniendo un 93,75 %.

Los dos casos parciales permiten reconocer dificultades concretas. En el documento ITI-R-001, correspondiente a una propuesta turística, MiDocIA asignó la categoría cotización, mientras que la categoría esperada en el instrumento era otro. Aunque el análisis recuperó parte del propósito y de los servicios descritos, la clasificación orientó algunas alertas hacia elementos propios de una cotización. El segundo caso fue CAR-S-001, una carta comercial que contenía una solicitud de modificación de reserva. La aplicación la identificó como PQRS — Petición. La valoración fue parcial porque reconoció el carácter de solicitud, pero no coincidió con la categoría esperada para esa comunicación.

Estos casos muestran que documentos con contenidos cercanos pueden generar confusiones entre categorías. También permiten observar que la clasificación influye en las etapas posteriores: cuando se asigna un tipo documental diferente al esperado, las alertas y recomendaciones pueden responder a criterios que no corresponden completamente al archivo.

En la columna de valoración no se registraron clasificaciones calificadas como totalmente incorrectas. Esto no significa que todas las categorías asignadas coincidieran con las esperadas, pues las dos diferencias descritas quedaron registradas como resultados parciales.

### 5.3.2. CALIDAD DEL TEXTO OBTENIDO MEDIANTE OCR

La calidad del OCR alcanzó una valoración promedio de 4,97 sobre 5. De los 32 documentos, 31 recibieron una calificación de cinco y uno recibió una calificación de cuatro.

Este resultado indica que el texto obtenido fue valorado favorablemente durante las pruebas. Sin embargo, corresponde a una calificación realizada mediante el instrumento y no a una medición de exactitud carácter por carácter. Por ello, no debe interpretarse como un porcentaje de transcripción perfecta. También debe considerarse que todos los archivos del conjunto fueron registrados con calidad visual alta. La valoración obtenida describe el comportamiento observado bajo esa condición y no permite anticipar el mismo resultado en documentos borrosos o con dificultades de lectura.
La calidad del OCR y la extracción de campos representan aspectos diferentes. Contar con un texto legible facilita el análisis, pero la aplicación todavía debe reconocer cuáles datos son relevantes, relacionarlos con su significado y mostrarlos correctamente al usuario.

### 5.3.3. EXTRACCIÓN DE INFORMACIÓN

La evaluación comprendió 356 campos individuales. De estos, 271 fueron identificados correctamente, 41 de manera parcial y 44 no fueron detectados. Esta distribución muestra que la aplicación recuperó una parte importante de la información esperada, aunque permanecieron datos incompletos o ausentes que debían revisarse en el documento original.

Para calcular la precisión de extracción, el instrumento asignó un punto a cada campo correcto y medio punto a cada campo parcial. La suma se dividió entre el número de campos esperados del documento. Los campos no detectados no aportaron puntos al cálculo.

El indicador se obtuvo primero para cada documento y luego se promediaron los resultados de los 32 casos. Mediante este procedimiento se alcanzó una precisión promedio de 82,33 %. Por tanto, esta cifra incluye la valoración parcial y otorga el mismo peso a cada documento dentro del promedio.

Entre las observaciones se encontraron datos resumidos sin todo el detalle esperado, información no identificada y dificultades de presentación. En el caso ITI-R-001, por ejemplo, algunos elementos aparecieron en la interfaz como “[object Object]”, lo que impidió revisar su contenido de forma comprensible. Este hallazgo muestra que la evaluación también depende de cómo se presenta la información. Aunque un dato forme parte de la respuesta del procesamiento, su utilidad disminuye si el usuario no puede leerlo o relacionarlo con el documento.

### 5.3.4. ALERTAS Y OBSERVACIONES SOBRE EL ANÁLISIS

En el criterio de alertas detectadas, la matriz registró 18 respuestas afirmativas, seis parciales y ocho negativas. Estas cantidades describen los registros del instrumento, pero no constituyen por sí solas una medida de precisión de las alertas. La ausencia de una alerta debe interpretarse de acuerdo con el contenido del documento. En la factura FAC-S-003, por ejemplo, la observación de la evaluación consideró adecuada esa ausencia porque el documento registraba el pago completo y un saldo de cero.

En otros casos se encontraron alertas o indicaciones que requerían corrección. En CAR-S-001 se señaló la incorporación de un plazo y una fecha límite que no se derivaban de la comunicación original. Esta observación evidenció la necesidad de distinguir entre los datos presentes en el archivo y las interpretaciones añadidas durante el análisis.

Las alertas resultaron útiles para señalar asuntos que merecían atención, pero su contenido debía contrastarse con el documento antes de utilizarlo para orientar una actuación.

### 5.3.5. UTILIDAD DE LAS RECOMENDACIONES Y CORRECCIÓN HUMANA

La utilidad de las recomendaciones obtuvo un promedio de 3,56 sobre 5. Se registraron 17 valoraciones de tres, 12 de cuatro y tres de cinco. Estos resultados muestran que las recomendaciones ofrecieron apoyo para la tarea, aunque su calidad y nivel de ajuste variaron entre documentos.

En la pregunta sobre si se utilizaría la recomendación, 28 respuestas indicaron que podía emplearse con ajustes y cuatro señalaron que podía utilizarse sin esa condición. Esto permite reconocer su utilidad como punto de partida, manteniendo la necesidad de revisar el contenido antes de aplicarlo.

El instrumento también valoró por separado la corrección humana necesaria en el análisis. En este criterio se registraron 17 casos con corrección importante, 14 con corrección menor y uno sin corrección. Esta valoración comprende las necesidades de ajuste del análisis y no equivale únicamente a la disposición de utilizar la recomendación.

Entre las dificultades observadas estuvo la perspectiva desde la cual se redactaba la respuesta. En los documentos turísticos ITI-R-001 e ITI-R-002 se indicó que la recomendación estaba formulada como si un cliente respondiera a la empresa,

cuando se esperaba una orientación para la revisión interna del documento. En conjunto, los resultados muestran que MiDocIA pudo apoyar la identificación y organización de información, así como ofrecer recomendaciones aprovechables. La revisión humana permitió detectar omisiones, corregir interpretaciones y adaptar las salidas al propósito de cada documento.

## 5.4. RESULTADOS RELACIONADOS CON EL TIEMPO DE PROCESAMIENTO

Uno de los aspectos evaluados fue el tiempo requerido para revisar un documento y establecer una posible acción a partir de su contenido. Para realizar la comparación, el instrumento registró el tiempo del procedimiento manual, el tiempo de procesamiento de MiDocIA y el tiempo dedicado por la persona a revisar el resultado generado.

La comparación incluyó la participación humana dentro del proceso apoyado por la aplicación. Esto permitió valorar el tiempo requerido para obtener y revisar el análisis, considerando que las respuestas de MiDocIA necesitan ser contrastadas con el documento original.

### 5.4.1. COMPONENTES DEL TIEMPO EVALUADO

El tiempo manual registrado correspondió a la actividad de revisión realizada sin el apoyo de MiDocIA. En los 32 documentos evaluados, este valor presentó un promedio de 12,16 minutos por documento.

Para el proceso con MiDocIA se consideraron dos componentes. El primero fue el tiempo de procesamiento consignado en la matriz, cuyo promedio fue de 22,6

segundos. El segundo fue el tiempo de revisión humana, que alcanzó un promedio de 5,53 minutos.

El tiempo total con MiDocIA se calculó para cada documento convirtiendo los segundos de procesamiento a minutos y sumándolos al tiempo de revisión humana. Mediante este procedimiento se obtuvo un promedio de 5,91 minutos por documento.

Esta separación permite comprender qué representa el resultado. La generación automática del análisis ocupó una parte menor del tiempo total, mientras que la mayor parte correspondió a la revisión realizada por la persona. Por ello, los 22,6 segundos de procesamiento no representan la duración completa de la actividad evaluada.

### 5.4.2. AHORRO PROMEDIO OBTENIDO

El ahorro se calculó para cada documento restando el tiempo total con MiDocIA al tiempo manual registrado. Posteriormente se promediaron las diferencias de los 32 casos, obteniendo un ahorro de 6,25 minutos por documento.
La comparación muestra que el tiempo promedio pasó de 12,16 minutos en el procedimiento manual a 5,91 minutos con el apoyo de la aplicación. Este resultado se obtuvo incluyendo el tiempo dedicado a comprobar el análisis, por lo que la reducción observada mantuvo la revisión humana como parte de la actividad.
En todos los documentos evaluados se registró un tiempo total con MiDocIA inferior al tiempo manual. Sin embargo, la magnitud del ahorro fue diferente entre casos. Las diferencias registradas estuvieron aproximadamente entre 2,52 y 12,56 minutos por documento.

Esta variación muestra que el beneficio no fue idéntico para todos los archivos. Algunos análisis necesitaron más tiempo de revisión y ajuste, mientras que otros

pudieron revisarse en menos tiempo. Los valores permiten describir esa diferencia, aunque la evaluación no aisló el efecto de cada característica documental sobre el ahorro obtenido.

### 5.4.3. CÁLCULO DE LA REDUCCIÓN PORCENTUAL

La reducción porcentual se calculó individualmente para cada documento. Para ello, el tiempo ahorrado se dividió entre el tiempo manual del mismo caso y el resultado se multiplicó por cien. Luego se obtuvo el promedio de los 32 porcentajes, cuyo valor fue de 50,65 %.

Este procedimiento considera la reducción relativa de cada documento antes de calcular el resultado general. Por esa razón, el 50,65 % no se obtiene directamente al dividir el ahorro promedio de 6,25 minutos entre el tiempo manual promedio de 12,16 minutos. Esa operación corresponde a una relación entre promedios, diferente del cálculo utilizado en el instrumento.

La distinción permite mantener la coherencia entre la matriz y los resultados del informe. El ahorro promedio expresa cuántos minutos se redujeron por documento, mientras que la reducción porcentual promedio expresa la disminución relativa calculada para cada caso.

### 5.4.4. PARTICIPACIÓN DE LA REVISIÓN HUMANA

El promedio de 5,53 minutos destinado a la revisión humana muestra que el usuario mantuvo una participación importante en el proceso. La disponibilidad de un análisis automático no eliminó la necesidad de consultar el archivo, comprobar los datos y ajustar las recomendaciones.

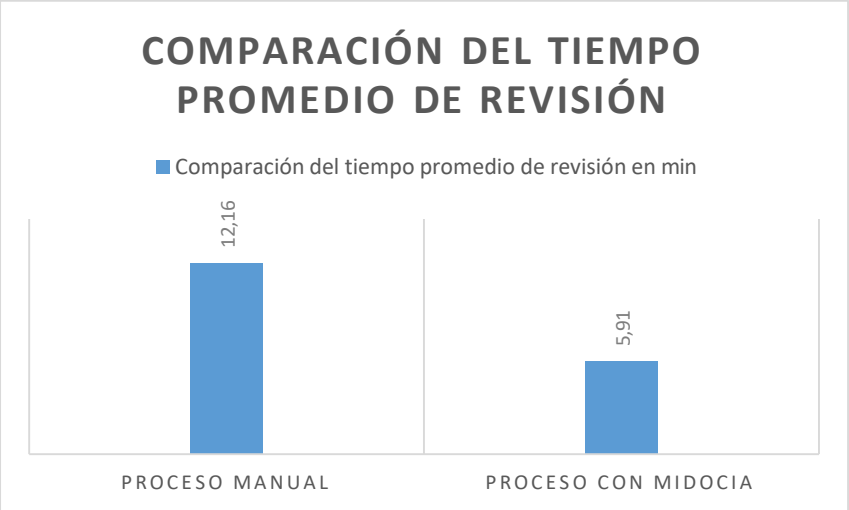
Este resultado se relaciona con las observaciones de la evaluación manual, en las que se identificaron campos incompletos, interpretaciones que requerían corrección y respuestas que debían adaptarse al propósito del documento.
El apoyo de MiDocIA permitió disponer de información organizada para iniciar esa revisión. Aun así, el tiempo necesario para verificarla formó parte del cálculo total.
De esta manera, la comparación contempla una condición central del funcionamiento de la aplicación: el análisis generado debe ser revisado antes de utilizarse.

### 5.4.5. ALCANCE DE LOS RESULTADOS DE TIEMPO

Los resultados muestran una reducción del tiempo registrado en los 32 casos evaluados. Su interpretación corresponde al conjunto documental utilizado y a las condiciones bajo las cuales se desarrollaron las pruebas.
Los valores de este apartado provienen del instrumento de validación. El tiempo de procesamiento allí consignado se distingue de la medición automática registrada por el aplicativo, que se presenta en el siguiente apartado. Estas mediciones se conservan por separado para identificar el origen de cada indicador.
El ahorro observado constituye evidencia del apoyo que MiDocIA ofreció durante la evaluación. Su magnitud puede variar al utilizar otros documentos, al cambiar las condiciones de procesamiento o según el tiempo que cada persona necesite para verificar el resultado.

La Ilustración 5 presenta la comparación entre el tiempo manual promedio y el tiempo total promedio con MiDocIA. En este último se incluyen tanto el procesamiento como la revisión humana.

**ILUSTRACIÓN 5 COMPARACIÓN DEL TIEMPO PROMEDIO DE REVISIÓN**




Fuente: elaboración propia a partir de la matriz de evaluación.

## 5.5. MÉTRICAS REGISTRADAS AUTOMÁTICAMENTE POR MIDOCIA

Además de la evaluación manual, MiDocIA registró información de uso que permitió consultar el número de análisis realizados, el tiempo de procesamiento y la confirmación del usuario. También conservó las respuestas voluntarias de la encuesta de utilidad, cuando fueron diligenciadas. Estos registros complementaron la matriz de validación y permitieron observar el uso de las funciones implementadas. Parte de la información se generó automáticamente durante el procesamiento, mientras que la confirmación y las respuestas de la encuesta dependieron de la participación del usuario.

La Tabla 5 presenta los valores obtenidos a partir de los 32 registros de uso.

**TABLA 5 MÉTRICAS REGISTRADAS POR EL APLICATIVO**

| Indicador | Resultado |
|---|---|
| Documentos analizados | **32** |
| Análisis confirmados mediante HITL | **31** |
| Tasa de confirmación HITL | **96,88 %** |
| Encuestas con calificación | **30** |
| **Indicador** | **Resultado** |

| | |
|---|---|
| Calificación promedio | **3,8 / 5** |
| Tiempo automático promedio de procesamiento | **19,5 s** |
| Respuestas con tiempo manual estimado entre 5 y 15 minutos | **24** |
| Respuestas con tiempo manual estimado entre 15 y 30 minutos | **5** |
| Total de respuestas sobre tiempo manual estimado | **29** |

Fuente: elaboración propia a partir de los registros de uso de MiDocIA.

### 5.5.1. INFORMACIÓN REGISTRADA DURANTE EL USO

Los registros incluyeron un identificador, la fecha y hora, el tipo documental detectado y el tiempo de procesamiento. También incorporaron el estado de confirmación, la calificación de utilidad, el rango de tiempo seleccionado y el comentario, cuando estos últimos fueron proporcionados. El correo electrónico formó parte de la información autorizada para el seguimiento académico.
El almacenamiento de estos datos se mantuvo separado del contenido documental. Los registros de uso no incluyeron el archivo original, el texto obtenido mediante OCR ni el análisis completo generado por la aplicación.
Esta separación permitió consultar indicadores del funcionamiento sin construir un repositorio de los documentos procesados. Sin embargo, la presencia del correo electrónico y de comentarios voluntarios significa que los registros no deben describirse como información completamente anónima.
Los 32 registros corresponden a análisis realizados. Su cantidad no representa necesariamente 32 personas diferentes, porque una misma persona puede analizar varios documentos.

### 5.5.2. CONFIRMACIÓN DE LA REVISIÓN HUMANA

De los 32 análisis registrados, 31 presentaron confirmación mediante el mecanismo Human-in-the-Loop. La tasa se calculó dividiendo los análisis confirmados entre el total de registros, obteniendo un 96,88 %.

Este indicador permite establecer que se realizó la acción prevista para finalizar la revisión en 31 casos. En el registro restante no quedó marcada esa confirmación. La información disponible no permite determinar por sí sola el motivo de su ausencia.

La confirmación tampoco demuestra que todos los datos del análisis fueran correctos ni permite conocer la profundidad con la que se revisó cada documento. Su significado se limita al registro de la acción realizada dentro de la aplicación. La comprobación de la clasificación, los campos extraídos y las recomendaciones se efectuó mediante el instrumento de validación. Por esta razón, la tasa de confirmación se presenta como una métrica de uso del flujo y se mantiene separada de los indicadores de precisión.

### 5.5.3. CALIFICACIÓN DE UTILIDAD

En 30 de los 32 registros se encontró una calificación mediante estrellas. La distribución correspondió a diez valoraciones de tres estrellas, dieciséis de cuatro y cuatro de cinco. El promedio de estas respuestas fue de 3,8 sobre 5.

Los dos registros sin calificación se excluyeron del cálculo. La ausencia de una respuesta no equivale a una valoración de cero, dado que la encuesta era voluntaria.

El resultado describe la percepción de utilidad expresada después de utilizar la aplicación. Aunque el panel presenta el contador bajo el rótulo “Usuarios encuestados”, la cifra corresponde a registros con calificación y no demuestra que

hayan participado 30 personas distintas. Este promedio también debe diferenciarse del 3,56 sobre 5 obtenido en la evaluación manual de las recomendaciones. La encuesta valoró la utilidad del análisis mediante una pregunta general, mientras que la matriz incluyó un criterio específico sobre la recomendación. Los valores provienen de preguntas y cantidades de respuestas diferentes, por lo que se presentan por separado.

### 5.5.4. ESTIMACIÓN DEL TIEMPO DE REVISIÓN MANUAL

La pregunta de tiempo de la encuesta solicitó indicar cuánto habría tardado la persona en realizar la actividad sin utilizar MiDocIA. Se obtuvieron 29 respuestas:
24 ubicaron ese tiempo entre 5 y 15 minutos y cinco entre 15 y 30 minutos. Estos datos corresponden a una estimación del usuario expresada mediante rangos. No representan una medición exacta del tiempo ni permiten calcular directamente cuántos minutos se ahorraron en cada análisis.
En el panel administrativo, la distribución aparece bajo el rótulo “Tiempo ahorrado declarado por usuarios”. Para interpretar los resultados de este informe se toma como referencia la pregunta de la encuesta, que se refiere al tiempo que habría requerido la actividad sin la aplicación.
El ahorro promedio de 6,25 minutos presentado en el apartado anterior procede de la comparación registrada en la matriz de validación. Las respuestas de la encuesta aportan información complementaria sobre la percepción del usuario, pero no sustituyen ese cálculo.

### 5.5.5. TIEMPO AUTOMÁTICO DE PROCESAMIENTO

El promedio del tiempo registrado automáticamente por el aplicativo fue de 19,5 segundos. Este valor se obtuvo a partir del campo de duración del procesamiento,

almacenado en milisegundos y convertido a segundos para su presentación. La medición corresponde al procesamiento registrado por el sistema y no incluye el tiempo posterior que la persona dedica a leer, comparar y ajustar el resultado. Por ello, no equivale al tiempo total de revisión con MiDocIA.

En la matriz de validación, el promedio de procesamiento consignado durante las pruebas fue de 22,6 segundos. Ambos valores provienen de registros distintos y se conservan identificados según su fuente. La diferencia observada no permite atribuir una causa específica sin comprobar los puntos de inicio y finalización utilizados en cada medición.

Para calcular el tiempo total con MiDocIA y el ahorro del apartado 5.4 se utilizaron los valores de la matriz. El promedio automático de 19,5 segundos se presenta como una métrica complementaria del funcionamiento del aplicativo.

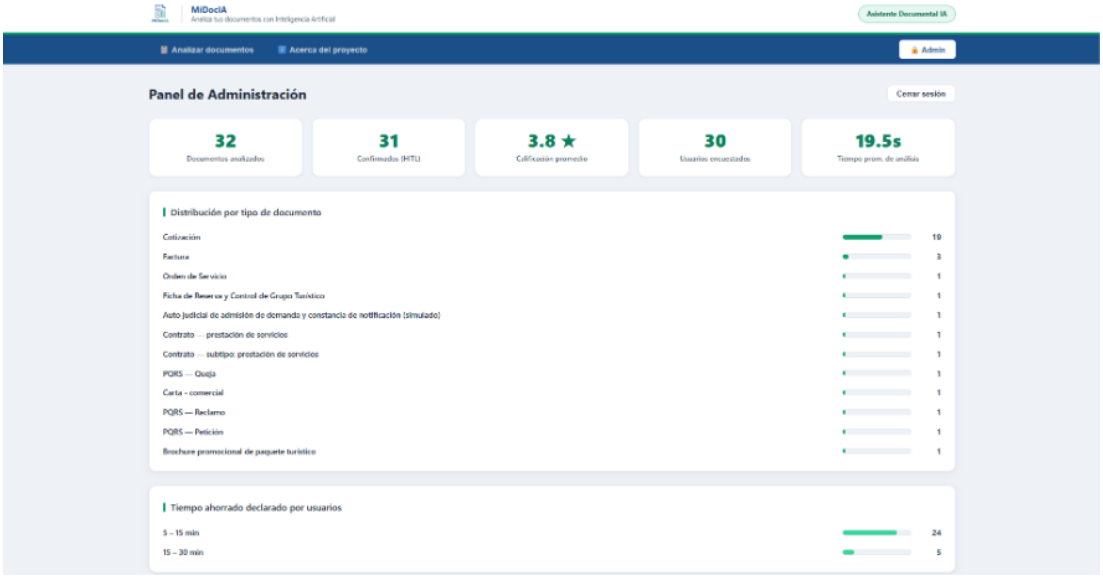
### 5.5.6. CONSULTA E INTERPRETACIÓN DEL PANEL ADMINISTRATIVO

El panel reunió los indicadores de uso y permitió consultar la distribución de los tipos documentales detectados. Esta distribución corresponde a las etiquetas producidas por MiDocIA y puede diferir de las categorías esperadas del banco documental.

Por ejemplo, el panel registró 19 análisis como cotización, mientras que el conjunto de evaluación contenía 18 documentos cuya categoría esperada era cotización. Esta diferencia es coherente con el caso de la propuesta turística clasificada como cotización, descrito en la evaluación manual.
La comparación muestra por qué los registros automáticos requieren interpretación. El panel permite conocer qué procesó y registró la aplicación; la matriz permite contrastar esos resultados con lo esperado para cada documento.

La Ilustración 6 presenta el panel administrativo utilizado para consultar estas métricas durante la evaluación.

**ILUSTRACIÓN 6 PANEL DE ADMINISTRACIÓN CON MÉTRICAS DE EVALUACIÓN DE MIDOCIA**




Fuente: elaboración propia.

## 5.6. ANÁLISIS GENERAL DE LOS RESULTADOS

La evaluación de MiDocIA permitió revisar de manera conjunta el funcionamiento de la aplicación, la calidad de sus resultados y el tiempo requerido para utilizarla. La matriz de validación aportó la comparación con los documentos originales, mientras que los registros del aplicativo permitieron consultar el procesamiento, la confirmación humana y las respuestas voluntarias de utilidad.
Al relacionar estas fuentes, se observa que la aplicación ofreció apoyo para organizar y revisar información documental. También se identificaron dificultades que deben considerarse al utilizar sus resultados, especialmente en la recuperación completa de los campos, la interpretación del contenido y la preparación de recomendaciones acordes con el propósito de cada documento.

### 5.6.1. RELACIÓN ENTRE CLASIFICACIÓN, EXTRACCIÓN Y PRESENTACIÓN

La clasificación obtuvo una valoración correcta en 30 de los 32 documentos evaluados, equivalente al 93,75 %. Este resultado muestra que MiDocIA reconoció adecuadamente el tipo documental en la mayoría de los casos. Sin embargo, los dos resultados parciales evidenciaron que una clasificación diferente de la esperada puede influir en las alertas y recomendaciones posteriores.

Por esta razón, la clasificación debe analizarse como parte de una secuencia. La categoría asignada orienta la búsqueda de campos y las instrucciones utilizadas para interpretar el contenido. Si esa categoría no corresponde al propósito del documento, el análisis puede incluir observaciones que resulten poco apropiadas, aunque recupere algunos datos correctamente.

La valoración del OCR, de 4,97 sobre 5, y la precisión promedio de extracción, de 82,33 %, también muestran que disponer de texto legible no garantiza recuperar toda la información esperada. Después del reconocimiento de texto, la aplicación debe identificar los datos relevantes, conservar su significado y presentarlos de forma comprensible.

Las observaciones sobre campos incompletos y problemas de visualización permiten reconocer que la utilidad depende de todo ese recorrido. Para el usuario, la información debe estar disponible y ser clara en la interfaz, además de corresponder con el archivo original.

### 5.6.2. RELACIÓN ENTRE AHORRO DE TIEMPO Y CALIDAD DEL ANÁLISIS

El tiempo total promedio con MiDocIA fue de 5,91 minutos, frente a los 12,16 minutos registrados para el procedimiento manual. La reducción porcentual promedio por documento fue de 50,65 %, incluyendo el tiempo de revisión humana dentro del proceso apoyado por la aplicación.

Este resultado permite observar una reducción del tiempo en las condiciones evaluadas, incluso cuando fue necesario comprobar y ajustar las respuestas. La aplicación proporcionó un análisis inicial que podía utilizarse como base para continuar la revisión.

Sin embargo, el ahorro de tiempo debe interpretarse junto con los resultados de calidad. La existencia de campos no detectados y recomendaciones que necesitaron ajustes demuestra que una respuesta rápida todavía puede requerir trabajo adicional antes de utilizarse.

La evaluación resulta más completa al considerar ambos aspectos: el tiempo indica cuánto duró la actividad registrada y la revisión manual permite identificar qué tan aprovechable fue el resultado. De esta manera, el beneficio observado se relaciona con la posibilidad de apoyar la tarea manteniendo la comprobación del contenido.

### 5.6.3. UTILIDAD DE LAS RECOMENDACIONES Y PARTICIPACIÓN HUMANA

Las recomendaciones obtuvieron una utilidad promedio de 3,56 sobre 5 en la matriz de validación. En 28 casos se consideraron utilizables con ajustes y en cuatro se indicó que podían utilizarse sin esa condición. Estos resultados muestran que las salidas ofrecieron una base aprovechable, aunque frecuentemente requirieron adaptación.

Las observaciones permitieron identificar necesidades concretas, como completar información, corregir la perspectiva desde la cual se redactaba una respuesta y revisar indicaciones que no se desprendían del documento. Estas tareas requieren comprender el contexto en el que se utilizará el análisis.

La participación humana tuvo, por tanto, una función práctica durante la evaluación. Permitió contrastar lo presentado por MiDocIA y decidir qué debía conservarse, modificarse o descartarse.

Los 31 registros con confirmación HITL muestran que la acción de finalización fue utilizada en la mayoría de los análisis. No obstante, la evidencia sobre las correcciones necesarias proviene de la matriz de validación. La confirmación registrada por el sistema y la calidad del resultado son aspectos relacionados con el proceso, pero se evalúan mediante criterios diferentes.

### 5.6.4. MANEJO DE LA INFORMACIÓN Y SEGUIMIENTO DEL USO

La implementación permitió mantener separados el contenido documental y los datos utilizados para el seguimiento académico. Los archivos, el texto extraído y los análisis completos no formaron parte de los registros persistentes de uso de MiDocIA.

Esta organización hizo posible consultar tiempos, confirmaciones y valoraciones sin conservar los documentos como parte de un archivo empresarial dentro de la aplicación. Su alcance corresponde al manejo de la información en la infraestructura propia de MiDocIA, considerando que el procesamiento utiliza servicios externos de reconocimiento de texto e inteligencia artificial.

El registro de metadatos tampoco equivale a la ausencia de datos personales, pues incluye información autorizada como el correo electrónico. Por ello, la separación del contenido documental constituye una medida concreta de manejo de información, cuyo alcance debe explicarse con precisión.

La evaluación presentada aporta evidencia sobre el funcionamiento y el uso de la aplicación. Sus indicadores de clasificación, extracción y tiempo no constituyen por sí mismos una verificación integral de seguridad o privacidad.

### 5.6.5. ALCANCE DE LA EVALUACIÓN Y APORTE DEL PROYECTO

Los resultados corresponden a 32 documentos, de los cuales 20 fueron reales y 12 sintéticos controlados. Aunque se incluyeron las siete categorías previstas, la

distribución tuvo un predominio de cotizaciones y todos los archivos evaluados fueron PDF de calidad visual alta.

Estas características delimitan la interpretación de los indicadores. La evaluación permitió observar el comportamiento de MiDocIA en un contexto empresarial concreto y en casos complementarios preparados para las pruebas. La cantidad de documentos por categoría no permite atribuir el mismo nivel de desempeño a todos los tipos documentales ni extender los resultados a todas las Mipymes de Bucaramanga.

Dentro de ese alcance, el proyecto mostró que es posible integrar la recepción del archivo, el reconocimiento de texto, el análisis mediante inteligencia artificial y la revisión humana en una aplicación web. La comparación realizada aportó evidencia de reducción del tiempo y permitió identificar aspectos que requieren mejora. El aporte observado se concentra en facilitar la revisión inicial y ofrecer información organizada para que la persona continúe la actividad. Las pruebas no midieron directamente la calidad de las decisiones empresariales posteriores, por lo que el apoyo a la toma de decisiones se entiende como la disponibilidad de datos, alertas y recomendaciones sujetos a comprobación.

## 6. CONCLUSIONES

El desarrollo de MiDocIA permitió integrar reconocimiento óptico de caracteres y modelos de lenguaje en una aplicación web orientada al análisis documental. El producto obtenido reúne la carga del archivo, la extracción de texto, la clasificación, la identificación de información relevante y la generación de alertas y recomendaciones. Estas funciones permitieron ofrecer al usuario un análisis inicial organizado para apoyar la revisión de documentos administrativos.

El análisis de la problemática permitió establecer los requerimientos funcionales y de información de la aplicación. La definición de siete categorías documentales

sirvió para organizar los campos y las instrucciones de análisis según el propósito de cada documento. Este trabajo delimitó las actividades que podía apoyar MiDocIA y mantuvo la comprobación de los resultados y la decisión sobre su uso bajo responsabilidad del usuario.

El diseño de la arquitectura permitió separar la interfaz, el procesamiento documental y los servicios externos utilizados. La aplicación incorporó el procesamiento temporal del contenido y excluyó el archivo original, el texto extraído y el análisis completo de su almacenamiento persistente. Esta medida se aplicó en la infraestructura propia de MiDocIA y se complementó con el consentimiento y la separación de los metadatos autorizados para seguimiento académico. Su implementación constituye una medida de protección de la información dentro del alcance del proyecto.

El módulo de procesamiento permitió recibir archivos y obtener su contenido mediante OCR. En la evaluación, realizada con documentos PDF de calidad visual alta, el reconocimiento de texto alcanzó una valoración promedio de 4,97 sobre 5. Este resultado indica que el contenido obtenido fue valorado favorablemente para continuar el análisis, aunque su alcance corresponde a las condiciones de los archivos evaluados.

La integración del componente de inteligencia artificial permitió clasificar documentos y presentar información relacionada con su contenido. De los 32 casos evaluados, 30 obtuvieron una clasificación valorada como correcta, equivalente al 93,75 %. La precisión promedio de extracción por documento fue de 82,33 %, considerando la ponderación de los campos parciales establecida en el instrumento. Los resultados muestran que la aplicación recuperó información útil, pero también presentó omisiones y datos incompletos que debían contrastarse con el archivo original.

La evaluación permitió identificar que la utilidad del análisis depende de la relación entre sus etapas. Una clasificación diferente de la esperada puede orientar las alertas hacia criterios poco adecuados para el documento. Asimismo, un texto correctamente reconocido puede perder utilidad cuando los campos se presentan de forma incompleta o poco comprensible. Por ello, el funcionamiento debe valorarse desde la recepción del archivo hasta la información que finalmente puede consultar el usuario.

El uso de MiDocIA presentó una reducción del tiempo registrado durante las pruebas. El promedio del procedimiento manual fue de 12,16 minutos por documento y el promedio con la aplicación, incluida la revisión humana, fue de 5,91 minutos. El ahorro promedio alcanzó 6,25 minutos y la reducción porcentual promedio calculada por documento fue de 50,65 %. Estos resultados aportan evidencia de que la aplicación pudo agilizar la actividad evaluada manteniendo la comprobación humana dentro del proceso.

Las recomendaciones generadas obtuvieron una utilidad promedio de 3,56 sobre 5. En 28 de los 32 casos se consideraron utilizables con ajustes, lo que demuestra que su aprovechamiento frecuentemente requirió completar información o adaptar la respuesta al contexto. El mecanismo Human-in-the-Loop permitió incorporar una acción explícita de confirmación, mientras que la evaluación manual evidenció las correcciones necesarias. La utilidad de la herramienta se encuentra, por tanto, en ofrecer una base de trabajo que la persona pueda revisar y adaptar. La combinación de 20 documentos reales y 12 sintéticos controlados permitió evaluar las siete categorías contempladas. Sin embargo, el predominio de cotizaciones, la participación de una empresa colaboradora y las características de los archivos delimitan el alcance de los resultados. La evidencia obtenida describe el comportamiento de MiDocIA en el conjunto examinado y no permite atribuir el mismo desempeño a todos los contextos empresariales.

Frente al objetivo general, se obtuvo una aplicación web funcional que organiza información documental y propone posibles acciones, con evidencia de reducción del tiempo en las pruebas realizadas. El apoyo a la toma de decisiones se concretó en la presentación de datos, alertas y recomendaciones para la revisión del usuario. La evaluación no midió el efecto de su uso sobre decisiones empresariales posteriores, por lo que el aporte demostrado corresponde al apoyo en la revisión y preparación de la información necesaria para continuar el trámite.

## 7. RECOMENDACIONES

Se recomienda priorizar los ajustes relacionados con los campos identificados de manera parcial o no detectados durante la evaluación. La matriz de validación permite ubicar estas dificultades por documento y puede utilizarse como base para revisar las instrucciones de extracción. Después de cada modificación, conviene comparar nuevamente los resultados con los campos esperados para comprobar si el cambio mejora la recuperación de información.

También se recomienda revisar la clasificación de documentos con propósitos cercanos, especialmente las propuestas turísticas, las cotizaciones y las cartas que contienen solicitudes. Los casos parciales mostraron que una categoría diferente de la esperada puede modificar las alertas y recomendaciones posteriores. Para evaluar los ajustes, sería conveniente incluir ejemplos adicionales que permitan comprobar cómo distingue la aplicación estos contenidos.

En la presentación de los resultados, se recomienda verificar que los campos compuestos, como listas de servicios, actividades o condiciones, se visualicen de forma comprensible. Las observaciones registradas sobre datos mostrados como “[object Object]” evidencian la necesidad de revisar la transformación de la respuesta antes de presentarla en la interfaz. Esta comprobación debe incluir tanto

la existencia del dato como la posibilidad de que el usuario pueda leerlo y compararlo con el archivo original.

Para mejorar las recomendaciones, se propone precisar en las instrucciones del modelo la perspectiva desde la cual debe analizarse el documento. Cuando el contenido no permita establecer si el usuario actúa como emisor, destinatario o responsable de una revisión interna, la aplicación debería evitar asumir una posición sin fundamento. Esta mejora ayudaría a reducir respuestas redactadas desde un rol diferente al requerido.

Se recomienda reforzar la distinción entre información extraída e interpretación generada por la aplicación. Las fechas, valores y condiciones presentes en el documento deben diferenciarse de las sugerencias o inferencias añadidas durante el análisis. En particular, las indicaciones sobre plazos requieren una revisión cuidadosa para evitar presentar como dato del archivo una fecha o un término que no se encuentre sustentado en su contenido.

Debe mantenerse la revisión humana como parte del funcionamiento de MiDocIA. La confirmación debería continuar acompañada de orientaciones claras para consultar el original, comprobar los datos relevantes y ajustar las recomendaciones antes de utilizarlas. La acción de finalizar el análisis debe seguir interpretándose como un registro del flujo de uso, sin presentarse como garantía de exactitud del resultado.

Para ampliar la evaluación, se recomienda incorporar documentos de otras empresas y aumentar la cantidad de casos en las categorías con menor participación. También conviene incluir los demás formatos admitidos por la aplicación y archivos con diferentes niveles de legibilidad, extensión y complejidad. Esta ampliación permitiría conocer mejor las condiciones en las que se mantiene el desempeño observado y aquellas que generan mayores dificultades.

En futuras pruebas de tiempo, se recomienda definir y documentar con precisión el inicio y la finalización de cada medición. El procesamiento automático, la revisión humana y la actividad manual deben conservarse como componentes diferenciados. Registrar estas condiciones facilitaría la comparación entre los tiempos consignados en el instrumento y los almacenados por el aplicativo. Asimismo, se recomienda ajustar los rótulos del panel administrativo para que correspondan con la información recopilada. El indicador denominado “Usuarios encuestados” debería identificar los registros con valoración, salvo que se implemente un conteo específico de personas distintas. El apartado de “Tiempo ahorrado declarado por usuarios” debería denominarse de acuerdo con la pregunta actual de la encuesta, que solicita estimar cuánto habría tardado la actividad sin MiDocIA. Estos cambios facilitarían la interpretación de las métricas.
En futuras versiones debe conservarse la separación entre el contenido documental y los datos autorizados para seguimiento. Los cambios en registros de diagnóstico, almacenamiento o servicios externos deberían revisarse para evitar que incorporen de manera involuntaria archivos, texto extraído o análisis completos. También conviene definir los periodos de conservación y los permisos de acceso aplicables a los correos y comentarios registrados.

Se recomienda realizar el mantenimiento de la aplicación siguiendo los procedimientos documentados en el Manual de Instalación y Mantenimiento. Las actualizaciones de dependencias, modelos y servicios externos deben acompañarse de pruebas sobre la carga de archivos, el procesamiento, la presentación del análisis, la confirmación y la encuesta. Esto permitiría detectar cambios de comportamiento antes de utilizar una nueva versión en actividades de evaluación.
Finalmente, conviene mantener actualizados el manual de usuario, el manual técnico y el instrumento de validación cuando se modifiquen las funciones de MiDocIA. Registrar la versión utilizada en cada nueva evaluación facilitaría

relacionar los resultados con los cambios implementados y conservar una documentación coherente con el funcionamiento de la aplicación.

## 8. REFERENCIAS BIBLIOGRÁFICAS

Cavoukian, A. (2012). *Operationalizing Privacy by Design: A guide to implementing strong privacy practices*. Information and Privacy Commissioner of Ontario. https://www.ipc.on.ca/en/media/1769/download

Congreso de Colombia. (2012). *Ley 1581 de 2012. Por la cual se dictan disposiciones generales para la protección de datos personales.* https://www.funcionpublica.gov.co/eva/gestornormativo/norma.php?i=49981

Congreso de Colombia. (2015). *Ley 1755 de 2015. Por medio de la cual se regula el Derecho Fundamental de Petición y se sustituye un título del Código de Procedimiento Administrativo y de lo Contencioso Administrativo.* https://www.funcionpublica.gov.co/eva/gestornormativo/norma.php?i=65334

Congreso de Colombia. (1913, 20 de agosto). *Ley 4 de 1913. Sobre régimen político y municipal***.** *Diario Oficial No. 15.012.* https://www.funcionpublica.gov.co/eva/gestornormativo/norma.php?i=8426

Express. (s. f.-a). *Express: Node.js web application framework*. https://expressjs.com/

Express. (s. f.-b). *Multer*. https://expressjs.com/en/resources/middleware/multer/

Fernández Villalobos, W. A. (2025). *Evolución de la ciencia de datos y Big Data en la gestión documental* [Monografía, Universidad Nacional Abierta y a Distancia]. Repositorio Institucional UNAD. https://repository.unad.edu.co/handle/10596/67192

Google Cloud. (2026). *Document AI overview*. Google Cloud Documentation. https://docs.cloud.google.com/document-ai/docs/overview

Google for Developers. (2025). *LLMs: Fine-tuning, distillation, and prompt engineering*. Machine Learning Crash Course. https://developers.google.com/machine-learning/crash-course/llm/tuning

Google for Developers. (2026a). *Introduction to large language models*. Machine Learning Crash Course. https://developers.google.com/machine-learning/crash-course/llm

Google for Developers. (2026b). *LLMs: What's a large language model?* Machine Learning Crash Course. https://developers.google.com/machine-learning/crash-course/llm/transformers

Google. (2026). *Text generation*. Google AI for Developers. https://ai.google.dev/gemini-api/docs/text-generation

Guerrero Urango, J. E. (2025). *Sistema automatizado de gestión y respuesta de PQR en propiedades horizontales con inteligencia artificial y automatización de procesos* [Tesis de maestría, Universidad Autónoma de Bucaramanga]. Repositorio Institucional UNAB. https://repository.unab.edu.co/handle/20.500.12749/31591

International Energy Agency. (2025). *Energy and AI*. https://www.iea.org/reports/energy-and-ai

Node.js. (s. f.). *Introduction to Node.js*. https://nodejs.org/en/learn/getting-started/introduction-to-nodejs

Peña Del Aguila, E. (2024). *Automatización inteligente en la gestión documental de compras públicas: un enfoque basado en inteligencia artificial (IA)* [Tesis de grado, Universidad Tecnológica del Perú]. Repositorio Institucional UTP. https://alicia.concytec.gob.pe/vufind/Record/UTPD_38dcf85c8794fb656cd1121a1bf2a518/Details

Presidencia de la República de Colombia. (2015). *Decreto 1074 de 2015. Decreto Único Reglamentario del Sector Comercio, Industria y Turismo.* https://www.funcionpublica.gov.co/eva/gestornormativo/norma.php?i=76608

Railway. (s. f.). *Deploy an Express app*. https://docs.railway.com/guides/express

React. (s. f.). *Quick start*. https://react.dev/learn

Ruales Aldás, J. I. (2024). *Análisis, diseño y desarrollo de un sistema de automatización de la gestión documental en empresas a través de una interfaz gráfica intuitiva* [Proyecto técnico de titulación, Universidad Politécnica Salesiana]. Repositorio Institucional UPS. https://dspace.ups.edu.ec/handle/123456789/27200?locale=es

Supabase. (s. f.). *Database*. https://supabase.com/docs/guides/database/overview

Tabassi, E. (2023). *Artificial Intelligence Risk Management Framework (AI RMF 1.0)* (NIST AI 100-1). National Institute of Standards and Technology. doi:10.6028/NIST.AI.100-1. https://www.nist.gov/publications/artificial-intelligence-risk-management-framework-ai-rmf-10

Vite. (s. f.). *Getting started*. https://vite.dev/guide/

# 9. APÉNDICES

## 9.1. APÉNDICE A. INSTRUMENTO DE EVALUACIÓN DE MIDOCIA

Para la evaluación del funcionamiento de MiDocIA se utilizó una matriz diseñada para registrar y comparar los resultados obtenidos durante el procesamiento documental. El instrumento permitió identificar el origen y categoría de cada documento, verificar la clasificación realizada por el sistema, valorar la calidad de la extracción de texto, comparar los campos esperados con los obtenidos, evaluar las alertas y recomendaciones generadas y registrar los tiempos asociados al análisis. La matriz se estructuró en cuatro componentes: banco documental, evaluación general, campos evaluados y resumen de resultados. Para los documentos sintéticos se establecieron previamente valores esperados que permitieron realizar una comparación controlada con la información extraída por MiDocIA. El instrumento fue diseñado por el autor del proyecto y diligenciado durante la fase formal de evaluación por la persona designada por la empresa colaboradora. Los documentos reales utilizados durante la evaluación no se reproducen en el informe. La matriz completa utilizada como instrumento de evaluación se presenta en el Anexo C con autorización de la empresa colaboradora.

### 9.2. APÉNDICE B. ENCUESTA DE UTILIDAD

Después de la confirmación Human-in-the-Loop, MiDocIA presenta de manera voluntaria una encuesta orientada a conocer la percepción del usuario sobre el análisis realizado. El instrumento está compuesto por tres preguntas:

1. ¿Qué tan útil fue el análisis realizado por MiDocIA? **Escala:** 1 a 5 estrellas.
2. ¿Cuánto tiempo estima que habría tardado en realizar esta actividad sin utilizar MiDocIA? **Respuesta mediante rangos de tiempo.**
3. Comentario u observación sobre el análisis realizado. **Respuesta abierta, máximo 200 caracteres.**

La encuesta forma parte de las funcionalidades implementadas en la versión final y sus respuestas se utilizan exclusivamente para el seguimiento y la evaluación académica del proyecto.

## 10. ANEXOS

### 10.1. ANEXO A. MANUAL DE USUARIO DE MIDOCIA

Documento versión 1.0 que describe el acceso y uso funcional de MiDocIA, incluyendo consentimiento para el tratamiento de datos, selección y análisis de documentos, interpretación de resultados, confirmación Human-in-the-Loop, encuesta de utilidad, privacidad y recomendaciones de uso.

### 10.2. ANEXO B. MANUAL DE INSTALACIÓN Y MANTENIMIENTO DE MIDOCIA

Documento versión 1.0 dirigido al personal técnico, en el cual se describen los requisitos, configuración de servicios externos, instalación local, despliegue en

Railway, mantenimiento, actualización, rotación de credenciales, solución de problemas y verificaciones asociadas a Zero Storage y Human-in-the-Loop.

## 10.3. ANEXO C. MATRIZ COMPLETA DE VALIDACIÓN Y EVALUACIÓN DE MIDOCIA

Archivo de Excel utilizado para registrar y consolidar la evaluación de los 32 documentos procesados. Contiene el banco documental, la evaluación general, la revisión de campos y el resumen de resultados.

La matriz se incorpora de manera completa como evidencia del proceso de evaluación, con autorización de la empresa colaboradora. Los documentos reales utilizados durante las pruebas no forman parte de los anexos ni son reproducidos en el informe.

Tabla 6 C1 Ficha de validación

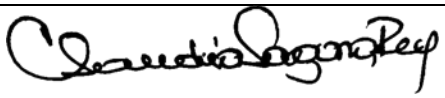
| Campo | Información |
|---|---|
| **Proyecto** | MiDocIA — Asistente de Análisis Documental con Inteligencia Artificial |
| **Código del proyecto** | 125-2026-002 |
| **Empresa participante** | Laguna Tours |
| **Responsable de la validación** | Claudia Patricia Laguna Rey |
| **Cargo** | Coordinación administrativa |
| **Rol en la evaluación** | Usuario evaluador |
| **Periodo de evaluación** | 24/06/26-22/08/26 |
| **Instrumento utilizado** | Instrumento de Validación de MiDocIA |
| **Documentos evaluados** | 32 |
| **Elaboración del instrumento** | Mario Andrés Jácome Mantilla |
| **Diligenciamiento de la evaluación** | Personal autorizado de Laguna Tours |
| **Constancia de diligenciamiento:** Se deja constancia de que los resultados registrados en las hojas “Evaluación” y “Campos evaluados” del presente instrumento corresponden a la valoración realizada durante las pruebas de MiDocIA por parte del personal autorizado de Laguna Tours. El autor del proyecto diseñó y estructuró el instrumento de evaluación y brindó el acompañamiento técnico necesario para la ejecución de las pruebas, sin sustituir el criterio del usuario evaluador en la calificación de los resultados. | |
| Nombre: Claudia Patricia Laguna Rey | |
| Cargo: Coordinación administrativa | |
| Firma:  | |
| CC: 32750715 | |
| Fecha: 27/08/26 | |

Tabla 7 C2 Banco Documental: identificación y origen

| Código | Nombre del documento | Grupo | Origen | Tipo documental |
|---|---|---|---|---|
| **ITI-R-001** | Pasadía Oro bajo Frailejones — Vetas | Empresarial | Real | Itinerario / propuesta turística |
| **ITI-R-002** | Plan Princesa Sora — Vetas, 2 días y 1 noche | Empresarial | Real | Brochure promocional / propuesta turística |
| **COT-S-001** | Servicio de transporte terrestre en buseta de 20 pasajeros | Empresarial | Real | Cotización comercial de servicio de transporte terrestre |
| **COT-S-002** | Transporte terrestre en buseta de 14 pasajeros | Empresarial | Real | Cotización comercial de servicio de transporte terrestre |
| **COT-S-003** | Transporte terrestre en buseta de 27 pasajeros | Empresarial | Real | Cotización comercial de servicio de transporte terrestre |
| **COT-S-004** | Transporte terrestre en busetón de 30 pasajeros | Empresarial | Real | Cotización comercial de servicio de transporte terrestre |
| **COT-S-005** | Transporte terrestre en busetón de 30 pasajeros | Empresarial | Real | Cotización comercial de servicio de transporte terrestre |
| **COT-S-006** | Transporte terrestre en buseta de 22 pasajeros | Empresarial | Real | Cotización comercial de servicio especial de transporte |
| **COT-S-007** | Alojamiento rural para 14 personas | Empresarial | Real | Cotización de alojamiento rural |
| **COT-S-008** | Hospedaje en cabañas para 20 personas | Empresarial | Real | Cotización de alojamiento rural |

| Código | Nombre del documento | Grupo | Origen | Tipo documental |
|---|---|---|---|---|
| **COT-S-009** | Alojamiento y desayuno para 22 personas | Empresarial | Real | Cotización de alojamiento y desayuno |
| **COT-S-010** | Alojamiento con desayuno para 27 personas | Empresarial | Real | Cotización de alojamiento con desayuno |
| **COT-S-011** | Alojamiento hotelero para 30 personas | Empresarial | Real | Cotización de alojamiento hotelero |
| **COT-S-012** | Alojamiento en villas eco-sostenibles para 30 personas | Empresarial | Real | Cotización de alojamiento rural |
| **COT-S-013** | Plan de alimentación para pasadía de 20 personas | Empresarial | Real | Cotización de servicio de alimentación |
| **COT-S-014** | Servicio de alimentación para pasadía de 14 personas | Empresarial | Real | Cotización de servicio de alimentación |
| **COT-S-015** | Plan completo de alimentación para 22 | Empresarial | Real | Cotización de servicio de alimentación |
| **COT-S-016** | Alimentación para grupo turístico de 27 personas | Empresarial | Real | Cotización de servicio de alimentación |
| **COT-S-017** | Plan de alimentación para pasadía de 30 personas | Empresarial | Real | Cotización de servicio de alimentación |

| Código | Nombre del documento | Grupo | Origen | Tipo documental |
|---|---|---|---|---|
| **COT-S-018** | Servicio integral de alimentación para 30 personas | Empresarial | Real | Cotización de servicio de alimentación |
| **CAR-S-001** | Solicitud de modificación de reserva grupal de 22 a 27 personas | Empresarial | Sintético controlado | Carta comercial de modificación de reserva |
| **PQR-S-001** | Reclamo por retraso en servicio de transporte | Empresarial | Sintético controlado | PQRS — Reclamo por afectación del servicio |
| **PQR-S-002** | Queja por novedades en alojamiento y solicitud de compensación | Empresarial | Sintético controlado | PQRS — Queja y solicitud de compensación |
| **CON-S-001** | Acuerdo de reserva y alojamiento para grupo turístico | Empresarial | Sintético controlado | Acuerdo contractual de reserva y alojamiento |
| **CON-S-002** | Contrato de prestación de servicio especial de transporte turístico | Empresarial | Sintético controlado | Contrato de prestación de servicio especial de transporte turístico |
| **JUD-S-001** | Auto de admisión de demanda y constancia secretarial de notificación | Empresarial | Sintético controlado | Auto judicial de admisión de demanda y constancia secretarial de notificación simulada |

| Código | Nombre del documento | Grupo | Origen | Tipo documental |
|---|---|---|---|---|
| **FAC-S-001** | Factura electrónica por servicio especial de transporte turístico | Empresarial | Sintético controlado | Factura electrónica de venta por servicio de transporte |
| **FAC-S-002** | Factura electrónica por alojamiento grupal y desayunos | Empresarial | Sintético controlado | Factura electrónica de venta por servicios de alojamiento |
| **FAC-S-003** | Factura electrónica por servicios de alimentación para 30 personas | Empresarial | Sintético controlado | Factura electrónica de venta por servicios de alimentación |
| **OTR-S-001** | Ficha de reserva y control de grupo turístico | Empresarial | Sintético controlado | Ficha interna de reserva y control operativo de grupo turístico |
| **OTR-S-002** | Orden de servicio e itinerario operativo para grupo de 30 pasajeros | Empresarial | Sintético controlado | Orden de servicio con itinerario turístico operativo |
| **CAR-S-002** | Confirmación de servicio de pasadía para grupo de 30 personas | Empresarial | Sintético controlado | Carta comercial de confirmación de servicio |

Tabla 8 C3 Banco documental: características y datos de evaluación

| Código | Cat. esp. | Formato | N° pág. | Cal. vis. | F. rec. | F. pru. |
|---|---|---|---|---|---|---|
| **ITI-R-001** | otro | PDF | 14 | Alta | 24/06/2026 | 24/06/2026 |
| **ITI-R-002** | otro | PDF | 20 | Alta | 25/06/2026 | 25/06/2026 |
| **COT-S-001** | cotizacion | PDF | 1 | Alta | 27/07/2026 | 28/07/2026 |
| **COT-S-002** | cotizacion | PDF | 1 | Alta | 27/07/2026 | 28/07/2026 |
| **COT-S-003** | cotizacion | PDF | 1 | Alta | 27/07/2026 | 29/07/2026 |
| **COT-S-004** | cotizacion | PDF | 1 | Alta | 27/07/2026 | 30/07/2026 |
| **COT-S-005** | cotizacion | PDF | 1 | Alta | 27/07/2026 | 30/07/2026 |
| **COT-S-006** | cotizacion | PDF | 1 | Alta | 27/07/2026 | 31/07/2026 |
| **COT-S-007** | cotizacion | PDF | 1 | Alta | 01/08/2026 | 04/08/2026 |
| **COT-S-008** | cotizacion | PDF | 1 | Alta | 01/08/2026 | 04/08/2026 |
| **COT-S-009** | cotizacion | PDF | 1 | Alta | 01/08/2026 | 05/08/2026 |
| **COT-S-010** | cotizacion | PDF | 1 | Alta | 01/08/2026 | 05/08/2026 |
| **COT-S-011** | cotizacion | PDF | 1 | Alta | 01/08/2026 | 11/08/2026 |
| **COT-S-012** | cotizacion | PDF | 1 | Alta | 01/08/2026 | 11/08/2026 |
| **COT-S-013** | cotizacion | PDF | 1 | Alta | 12/08/2026 | 13/08/2026 |
| **COT-S-014** | cotizacion | PDF | 1 | Alta | 12/08/2026 | 13/08/2026 |
| **COT-S-015** | cotizacion | PDF | 1 | Alta | 12/08/2026 | 13/08/2026 |
| **COT-S-016** | cotizacion | PDF | 1 | Alta | 12/08/2026 | 18/08/2026 |
| **COT-S-017** | Cotizacion | PDF | 1 | Alta | 12/08/2026 | 18/08/2026 |

| Código | Cat. esp. | Formato | N° pág. | Cal. vis. | F. rec. | F. pru. |
|---|---|---|---|---|---|---|
| **COT-S-018** | cotizacion | PDF | 1 | Alta | 12/08/2026 | 18/08/2026 |
| **CAR-S-001** | Carta | PDF | 1 | Alta | 20/08/2026 | 20/08/2026 |
| **PQR-S-001** | PQRS | PDF | 1 | Alta | 18/08/2026 | 20/08/2026 |
| **PQR-S-002** | PQRS | PDF | 1 | Alta | 19/08/2026 | 21/08/2026 |
| **CON-S-001** | Contrato | PDF | 2 | Alta | 10/08/2026 | 21/08/2026 |
| **CON-S-002** | Contrato | PDF | 2 | Alta | 10/08/2026 | 21/08/2026 |
| **JUD-S-001** | documento_judicial | PDF | 2 | Alta | 19/08/2026 | 21/08/2026 |
| **FAC-S-001** | Factura | PDF | 1 | Alta | Pendiente / No determinada | 21/08/2026 |
| **FAC-S-002** | Factura | PDF | 1 | Alta | Pendiente / No determinada | 21/08/2026 |
| **FAC-S-003** | Factura | PDF | 1 | Alta | Pendiente / No determinada | 21/08/2026 |
| **OTR-S-001** | otro | PDF | 1 | Alta | Pendiente / No determinada | 21/08/2026 |
| **OTR-S-002** | otro | PDF | 1 | Alta | Pendiente / No determinada | 21/08/2026 |
| **CAR-S-002** | Carta | PDF | 1 | Alta | 22/08/2026 | 22/08/2026 |

Tabla 9 C4 Banco documental: observaciones

| Código | Observaciones |
|---|---|
| **ITI-R-001** | Documento interno real de LagunaTours. MiDocIA lo clasificó incorrectamente como cotización. |
| **ITI-R-002** | Documento interno real de LagunaTours. MiDocIA lo identificó correctamente como brochure promocional de paquete turístico. |
| **COT-S-001** | Documento interno real de LagunaTours. MiDocIA lo identificó correctamente como una cotizacion |
| **COT-S-002** | Documento interno real de LagunaTours. MiDocIA lo identificó correctamente como una cotizacion |
| **COT-S-003** | Documento interno real de LagunaTours. MiDocIA lo identificó correctamente como una cotizacion |
| **COT-S-004** | Documento interno real de LagunaTours. MiDocIA lo identificó correctamente como una cotizacion |
| **COT-S-005** | Documento interno real de LagunaTours. MiDocIA lo identificó correctamente como una cotizacion |
| **COT-S-006** | Documento interno real de LagunaTours. MiDocIA lo identificó correctamente como una cotizacion |
| **COT-S-007** | Documento interno real de LagunaTours. MiDocIA lo identificó correctamente como una cotizacion sobre posada |
| **COT-S-008** | Documento interno real de LagunaTours. MiDocIA lo identificó correctamente como una cotizacion sobre posada |
| **COT-S-009** | Documento interno real de LagunaTours. MiDocIA lo identificó correctamente como una cotizacion sobre posada |
| **COT-S-010** | Documento interno real de LagunaTours. MiDocIA lo identificó correctamente como una cotizacion sobre posada |
| **COT-S-011** | Documento interno real de LagunaTours. MiDocIA lo identificó correctamente como una cotizacion sobre posada |
| **COT-S-012** | Documento interno real de LagunaTours. MiDocIA lo identificó correctamente como una cotizacion sobre posada |
| **COT-S-013** | Documento interno real de LagunaTours. MiDocIA clasificó correctamente el documento como cotización y extrajo adecuadamente proveedor, destinatario, número, fecha, ítems, cantidades, precios unitarios, total, vigencia y condiciones de pago. |
| **COT-S-014** | Documento interno real de LagunaTours. MiDocIA clasificó correctamente el documento como cotización y extrajo adecuadamente proveedor, destinatario, número, fecha, ítems, cantidades, precios unitarios, total, vigencia y condiciones de pago. |

| Código | Observaciones |
|---|---|
| **COT-S-015** | Documento interno real de LagunaTours. MiDocIA clasificó correctamente el documento como cotización y extrajo proveedor, cliente, número, fecha, ítems, cantidades, precios unitarios, total, vigencia y condiciones de pago. |
| **COT-S-016** | Documento interno real de LagunaTours. MiDocIA clasificó correctamente el documento como cotización y extrajo proveedor, cliente, número, fecha, ítems, cantidades, precios unitarios, total, vigencia y condiciones de pago. La recomendación es útil |
| **COT-S-017** | Documento interno real de LagunaTours. MiDocIA clasificó correctamente el documento como cotización y extrajo proveedor, cliente, número, fecha, ítems, cantidades, precios unitarios, total, vigencia y condiciones de pago. supone que se procederá con el anticipo del 50 % una vez definida la fecha, decisión que no está confirmada en el documento. |
| **COT-S-018** | MiDocIA clasificó correctamente el documento como cotización y extrajo proveedor, destinatario, número, fecha, ítems, cantidades, precios unitarios, total, vigencia y condiciones de pago. |
| **CAR-S-001** | MiDocIA clasificó el documento como “PQRS — Petición”. La clasificación se considera parcialmente correcta, ya que formalmente corresponde a una carta comercial. El sistema recuperó adecuadamente los principales datos de la comunicación y la acción requerida. |
| **PQR-S-001** | MiDocIA clasificó correctamente el documento como “PQRS — Reclamo” y recuperó adecuadamente el remitente, la fecha, el motivo del reclamo, la afectación presentada y las solicitudes de explicación, compensación y confirmación escrita. |
| **PQR-S-002** | MiDocIA clasificó correctamente el documento como “PQRS — Queja” e identificó adecuadamente a la remitente, la fecha de radicación, el motivo y las cuatro solicitudes formuladas. |
| **CON-S-001** | MiDocIA clasificó correctamente el documento como contrato y extrajo con precisión las partes, Las cláusulas de riesgo fueron recuperadas parcialmente. Las alertas sobre el vencimiento del saldo el 22 de agosto y la condición de no reembolso del anticipo son pertinentes. |
| **CON-S-002** | MiDocIA clasificó correctamente el documento como contrato de prestación de servicios. Las obligaciones de las partes fueron recuperadas parcialmente, ya que se omitieron algunas obligaciones expresamente contenidas en el contrato. |

| Código | Observaciones |
|---|---|
| **JUD-S-001** | MiDocIA clasificó correctamente el documento como auto judicial de admisión de demanda. |
| **FAC-S-001** | MiDocIA clasificó correctamente el documento como factura. |
| **FAC-S-002** | MiDocIA clasificó correctamente el documento como factura. La alerta sobre el vencimiento del 22 de agosto de 2026 es correcta y oportuna. |
| **FAC-S-003** | MiDocIA clasificó correctamente el documento como factura. a ausencia de alertas es coherente con el documento. La recomendación resulta útil como base |
| **OTR-S-001** | MiDocIA clasificó correctamente el documento como ficha de reserva y control de grupo turístico. Las alertas fueron pertinentes al identificar el saldo pendiente, la confirmación operativa requerida |
| **OTR-S-002** | MiDocIA clasificó correctamente el documento como Orden de Servicio. La alerta es parcialmente pertinente |
| **CAR-S-002** | MiDocIA clasificó correctamente el documento como Carta - Comercial e identificó adecuadamente remitente. El contenido fue recuperado |

Tabla 10 C5 Evaluación: Resultados De Clasificación Documental Y Calidad Del Ocr

| Código | Categoría detectada | ¿Clasificación correcta? | Calidad del OCR, de 1 a 5 |
|---|---|---|---|
| **ITI-R-001** | Cotización | Parcial | 4 |
| **ITI-R-002** | Brochure promocional de paquete turístico | Si | 5 |
| **COT-S-001** | Cotización | Si | 5 |
| **COT-S-002** | Cotización | Si | 5 |
| **COT-S-003** | Cotización | Si | 5 |
| **COT-S-004** | Cotización | Si | 5 |
| **COT-S-005** | Cotización | Si | 5 |
| **COT-S-006** | Cotización | Si | 5 |
| **COT-S-007** | Cotización | Si | 5 |
| **COT-S-008** | Cotización | Si | 5 |
| **COT-S-009** | Cotización | Si | 5 |
| **COT-S-010** | Cotización | Si | 5 |
| **COT-S-011** | Cotización | Si | 5 |
| **COT-S-012** | Cotización | Si | 5 |
| **COT-S-013** | Cotización | Si | 5 |
| **COT-S-014** | Cotización | Si | 5 |
| **COT-S-015** | Cotización | Si | 5 |

| Código | Categoría detectada | ¿Clasificación correcta? | Calidad del OCR, de 1 a 5 |
|---|---|---|---|
| **COT-S-016** | Cotización | Si | 5 |
| **COT-S-017** | Cotización | Si | 5 |
| **COT-S-018** | Cotización | Si | 5 |
| **CAR-S-001** | PQRS — Petición | Parcial | 5 |
| **PQR-S-001** | PQRS — Reclamo | Si | 5 |
| **PQR-S-002** | PQRS — Queja | Si | 5 |
| **CON-S-001** | Contrato — Subtipo: Prestación de Servicios | Si | 5 |
| **CON-S-002** | Contrato — Prestación de Servicios | Si | 5 |
| **JUD-S-001** | Auto Judicial de Admisión de Demanda y Constancia de Notificación (Simulado) | Si | 5 |
| **FAC-S-001** | Factura | Si | 5 |
| **FAC-S-002** | Factura | Si | 5 |
| **FAC-S-003** | Factura | Si | 5 |
| **OTR-S-001** | Ficha de Reserva y Control de Grupo Turístico | Si | 5 |
| **OTR-S-002** | Orden de Servicio | Si | 5 |
| **CAR-S-002** | Carta - Comercial | Si | 5 |

Tabla 11 C6 Evaluación: Resultados De Extracción De Campos Y Detección De Alertas

| Código | Cam. esperados | Cam. correctos | Cam. parciales | Cam. no detectados | Precisión de extracción | Alertas detectadas |
|---|---|---|---|---|---|---|
| **ITI-R-001** | 12 | 3 | 9 | 0 | 62,50% | Parcial |
| **ITI-R-002** | 14 | 11 | 2 | 1 | 85,71% | Parcial |
| **COT-S-001** | 12 | 7 | 3 | 2 | 70,83% | Si |
| **COT-S-002** | 12 | 7 | 3 | 2 | 70,83% | Si |
| **COT-S-003** | 12 | 7 | 3 | 2 | 70,83% | Parcial |
| **COT-S-004** | 12 | 9 | 2 | 1 | 83,33% | Si |
| **COT-S-005** | 12 | 8 | 2 | 2 | 75,00% | Si |
| **COT-S-006** | 12 | 8 | 2 | 2 | 75,00% | Si |
| **COT-S-007** | 12 | 9 | 1 | 2 | 79,17% | No |
| **COT-S-008** | 12 | 8 | 1 | 3 | 70,83% | Si |

| Código | Cam. esperados | Cam. correctos | Cam. parciales | Cam. no detectados | Precisión de extracción | Alertas detectadas |
|---|---|---|---|---|---|---|
| **COT-S-009** | 12 | 10 | 0 | 2 | 83,33% | No |
| **COT-S-010** | 12 | 10 | 0 | 2 | 83,33% | Si |
| **COT-S-011** | 12 | 10 | 0 | 2 | 83,33% | Si |
| **COT-S-012** | 12 | 10 | 0 | 2 | 83,33% | Si |
| **COT-S-013** | 12 | 10 | 0 | 2 | 83,33% | No |
| **COT-S-014** | 12 | 10 | 0 | 2 | 83,33% | No |
| **COT-S-015** | 12 | 10 | 0 | 2 | 83,33% | No |
| **COT-S-016** | 12 | 10 | 0 | 2 | 83,33% | Si |
| **COT-S-017** | 12 | 10 | 0 | 2 | 83,33% | Si |
| **COT-S-018** | 12 | 10 | 0 | 2 | 83,33% | No |
| **CAR-S-001** | 7 | 5 | 2 | 0 | 85,71% | Parcial |
| **PQR-S-001** | 7 | 6 | 1 | 0 | 92,86% | Parcial |
| **PQR-S-002** | 7 | 5 | 2 | 0 | 85,71% | No |
| **CON-S-001** | 9 | 7 | 1 | 1 | 83,33% | Si |
| **CON-S-002** | 9 | 7 | 1 | 1 | 83,33% | Si |
| **JUD-S-001** | 8 | 7 | 0 | 1 | 87,50% | Si |
| **FAC-S-001** | 12 | 11 | 0 | 1 | 91,67% | Si |
| **FAC-S-002** | 12 | 11 | 0 | 1 | 91,67% | Si |
| **FAC-S-003** | 12 | 11 | 0 | 1 | 91,67% | No |
| **OTR-S-001** | 12 | 7 | 4 | 1 | 75,00% | Si |
| **OTR-S-002** | 12 | 11 | 1 | 0 | 95,83% | Parcial |
| **CAR-S-002** | 7 | 6 | 1 | 0 | 92,86% | Si |

Tabla 12 C7 Evaluación: Utilidad de las recomendaciones y corrección humana requerida

| Código | Utilidad de la recomendación, de 1 a 5 | Corrección humana necesaria | ¿Usaría la recomendación? |
|---|---|---|---|
| **ITI-R-001** | 3 | Importante | Si con ajustes |
| **ITI-R-002** | 3 | Importante | Si con ajustes |
| **COT-S-001** | 4 | Menor | Si con ajustes |
| **COT-S-002** | 4 | Menor | Si con ajustes |
| **COT-S-003** | 5 | Niguna | Si |
| **COT-S-004** | 4 | Menor | Si con ajustes |
| **COT-S-005** | 4 | Menor | Si con ajustes |
| **COT-S-006** | 4 | Menor | Si |
| **COT-S-007** | 3 | Importante | Si con ajustes |

| | | | |
|---|---|---|---|
| **COT-S-008** | 3 | Importante | Si con ajustes |
| **Código** | **Utilidad de la recomendación, de 1 a 5** | **Corrección humana necesaria** | **¿Usaría la recomendación?** |
| **COT-S-009** | 4 | Menor | Si con ajustes |
| **COT-S-010** | 3 | Importante | Si con ajustes |
| **COT-S-011** | 4 | Menor | Si con ajustes |
| **COT-S-012** | 4 | Menor | Si con ajustes |
| **COT-S-013** | 4 | Menor | Si con ajustes |
| **COT-S-014** | 4 | Menor | Si con ajustes |
| **COT-S-015** | 3 | Importante | Si con ajustes |
| **COT-S-016** | 3 | Importante | Si con ajustes |
| **COT-S-017** | 3 | Importante | Si con ajustes |
| **COT-S-018** | 3 | Importante | Si con ajustes |
| **CAR-S-001** | 3 | Importante | Si con ajustes |
| **PQR-S-001** | 3 | Importante | Si con ajustes |
| **PQR-S-002** | 3 | Importante | Si con ajustes |
| **CON-S-001** | 4 | Menor | Si con ajustes |
| **CON-S-002** | 4 | Menor | Si con ajustes |
| **JUD-S-001** | 5 | Menor | Si |
| **FAC-S-001** | 3 | Importante | Si con ajustes |
| **FAC-S-002** | 3 | Importante | Si con ajustes |
| **FAC-S-003** | 3 | Importante | Si con ajustes |
| **OTR-S-001** | 5 | Menor | Si |
| **OTR-S-002** | 3 | Importante | Si con ajustes |
| **CAR-S-002** | 3 | Importante | Si con ajustes |

Tabla 13 C8 Evaluación: Tiempos de procesamiento, revisión y ahorro obtenido

| **Código** | **Tiempo manual en minutos** | **Tiempo de procesamiento de MiDocIA en segundos** | **Tiempo de revisión humana en minutos** | **Tiempo total con MiDocIA min** | **Tiempo ahorrado min** | **Reducción porcentual** |
|---|---|---|---|---|---|---|
| **ITI-R-001** | 10 | 44,4 | 3 | 3,74 | 6,26 | 62,60% |
| **ITI-R-002** | 15 | 38 | 6 | 6,63 | 8,37 | 55,78% |
| **COT-S-001** | 10 | 20,7 | 4 | 4,345 | 5,655 | 56,55% |
| **COT-S-002** | 10 | 24,5 | 4 | 4,408333333 | 5,591666667 | 55,92% |
| **COT-S-003** | 10 | 17,7 | 5 | 5,295 | 4,705 | 47,05% |

| Código | Tiempo manual en minutos | Tiempo de procesamiento de MiDocIA en segundos | Tiempo de revisión humana en minutos | Tiempo total con MiDocIA min | Tiempo ahorrado min | Reducción porcentual |
|---|---|---|---|---|---|---|
| **COT-S-004** | 10 | 16,3 | 5 | 5,271666667 | 4,728333333 | 47,28% |
| **COT-S-005** | 10 | 16,2 | 4 | 4,27 | 5,73 | 57,30% |
| **COT-S-006** | 8 | 13,1 | 3 | 3,218333333 | 4,781666667 | 59,77% |
| **COT-S-007** | 12 | 41,2 | 7 | 7,686666667 | 4,313333333 | 35,94% |
| **COT-S-008** | 12 | 20,2 | 7 | 7,336666667 | 4,663333333 | 38,86% |
| **COT-S-009** | 14 | 22,5 | 6 | 6,375 | 7,625 | 54,46% |
| **COT-S-010** | 10 | 18,8 | 5 | 5,313333333 | 4,686666667 | 46,87% |
| **COT-S-011** | 16 | 26,4 | 5 | 5,44 | 10,56 | 66,00% |
| **COT-S-012** | 15 | 18,9 | 4 | 4,315 | 10,685 | 71,23% |
| **COT-S-013** | 18 | 26,5 | 5 | 5,441666667 | 12,55833333 | 69,77% |
| **COT-S-014** | 16 | 16 | 5 | 5,266666667 | 10,73333333 | 67,08% |
| **COT-S-015** | 17 | 15 | 5 | 5,25 | 11,75 | 69,12% |
| **COT-S-016** | 15 | 29,2 | 6 | 6,486666667 | 8,513333333 | 56,76% |
| **COT-S-017** | 16 | 16,8 | 7 | 7,28 | 8,72 | 54,50% |
| **COT-S-018** | 14 | 28,5 | 5 | 5,475 | 8,525 | 60,89% |
| **CAR-S-001** | 6 | 25,4 | 3 | 3,423333333 | 2,576666667 | 42,94% |
| **PQR-S-001** | 8 | 16,8 | 5 | 5,28 | 2,72 | 34,00% |

| Código | Tiempo manual en minutos | Tiempo de procesamiento de MiDocIA en segundos | Tiempo de revisión humana en minutos | Tiempo total con MiDocIA min | Tiempo ahorrado min | Reducción porcentual |
|---|---|---|---|---|---|---|
| **PQR-S-002** | 10 | 21,3 | 5 | 5,355 | 4,645 | 46,45% |
| **CON-S-001** | 10 | 28,9 | 7 | 7,481666667 | 2,518333333 | 25,18% |
| **CON-S-002** | 15 | 21 | 10 | 10,35 | 4,65 | 31,00% |
| **JUD-S-001** | 20 | 26,1 | 15 | 15,435 | 4,565 | 22,83% |
| **FAC-S-001** | 10 | 14,5 | 5 | 5,241666667 | 4,758333333 | 47,58% |
| **FAC-S-002** | 10 | 18,3 | 5 | 5,305 | 4,695 | 46,95% |
| **FAC-S-003** | 10 | 12,7 | 5 | 5,211666667 | 4,788333333 | 47,88% |
| **OTR-S-001** | 10 | 14,4 | 5 | 5,24 | 4,76 | 47,60% |
| **OTR-S-002** | 15 | 26,8 | 8 | 8,446666667 | 6,553333333 | 43,69% |
| **CAR-S-002** | 7 | 26,1 | 3 | 3,435 | 3,565 | 50,93% |

Tabla 14 C9 Evaluación: observaciones cualitativas de la evaluación

| Código | Comentario |
|---|---|
| **ITI-R-001** | El sistema extrajo correctamente el propósito general del pasadía, el destino y la existencia de transporte, alimentación, actividades y seguro. Sin embargo, clasificó el documento como cotización cuando corresponde a un itinerario o propuesta turística, categoría que actualmente debe ubicarse en “otro”. La alerta sobre ausencia de precio, vigencia y condiciones de pago es comercialmente útil, pero surge de la clasificación incorrecta. Los ítems fueron devueltos como objetos, pero la interfaz los mostró como “[object Object]”, impidiendo al usuario revisar su contenido. La recomendación requiere una corrección importante porque está redactada desde la perspectiva de un cliente que responde a LagunaTours, en lugar de presentar una recomendación interna para completar el producto turístico. |

| Código | Comentario |
|---|---|
| **ITI-R-002** | El sistema clasificó correctamente el documento como brochure promocional de un paquete turístico y extrajo de forma precisa la empresa oferente, nombre del plan, duración, destinos, punto y hora de salida, alojamiento, actividades principales, seguro, primeros auxilios, contacto y registro turístico. La extracción de la alimentación y del itinerario fue parcial, ya que se resumieron de manera general sin detallar completamente las actividades de cada día. No se detectó la entrega de dos souvenirs. La recomendación identifica adecuadamente la necesidad de verificar disponibilidad, condiciones de reserva y cancelación, pero está redactada desde la perspectiva de una empresa interesada en contratar a LagunaTours y no como una recomendación interna para mejorar el brochure. |
| **COT-S-001** | El sistema clasificó correctamente el documento como cotización y extrajo adecuadamente el proveedor, destinatario, número, fecha, valor total, vigencia y condiciones de pago. El campo de ítems cotizados presentó el error , mientras que el precio unitario y el subtotal solo pudieron inferirse parcialmente; la cantidad y el IVA no fueron identificados. |
| **COT-S-002** | El sistema clasificó correctamente el documento como cotizacióny reconoció los principales datos de identificación, valores y condiciones comerciales. El campo de ítems cotizados presentó un error, msí recuperó el transporte base, la guía opcional y la alimentación pendiente de definición. |
| **COT-S-003** | El sistema correctamente el documento como cotización y extrajo adecuadamente el proveedor, destinatario, número, fecha, valor total, vigencia y condiciones de pago. |
| **COT-S-004** | El sistema clasifico correctamente el documento como cotización y extrajo adecuadamente los datos de identificación, el servicio cotizado, la cantidad, el precio unitario, el total, la vigencia y las condiciones de pago.Las alertas fueron pertinentes, aunque utilizaron expresiones valorativas, y la recomendación resultó útil, pero requiere completar la fecha del pasadía y agregar el cierre antes de enviarse. |
| **COT-S-005** | El sistema clasificó correctamente el documento como cotización y extrajo adecuadamente el proveedor, destinatario, número, fecha, servicio, total, vigencia y condiciones de pago. Las alertas fueron pertinentes y la recomendación resultó útil. |

| Código | Comentario |
|---|---|
| **COT-S-006** | El sistema clasificó correctamente el documento como cotización y extrajo adecuadamente el proveedor, destinatario, número, fecha, servicio, total, vigencia y condiciones de pago. Las alertas fueron pertinentes y la recomendación resultó útil. La recomendación requiere corregir el saludo y agregar el cierre antes de utilizarse. |
| | |
| **COT-S-007** | El sistema clasificó correctamente el documento como cotización sobre una posada y extrajo adecuadamente los datos de identificación, los dos ítems, cantidades, precios unitarios, vigencia y condiciones de pago. La recomendación es útil como borrador |
| **COT-S-008** | El sistema clasificó correctamente el documento como cotización sobre una posada y extrajo adecuadamente los datos de identificación, los dos ítems, cantidades, precios unitarios, vigencia y condiciones de pago. La recomendación es útil como borrador |
| **COT-S-009** | El sistema clasificó correctamente el documento ccomo cotización y extrajo adecuadamente los datos de identificación, ítems, cantidades, precios unitarios, total, vigencia y condiciones de pago. La recomendación es útil y está bien redactada. |
| **COT-S-010** | El sistema clasificó correctamente el documento ccomo cotización y extrajo adecuadamente los datos de identificación, ítems, cantidades, precios unitarios, total, vigencia y condiciones de pago. |
| **COT-S-011** | El sistema clasificó correctamente el documento ccomo cotización y extrajo adecuadamente los datos de identificación, ítems, cantidades, precios unitarios, total, vigencia y condiciones de pago.La alerta de vencimiento fue correcta y oportuna, y la recomendación resultó útil |
| **COT-S-012** | El sistema clasificó correctamente el documento como cotización y extrajo adecuadamente los datos de identificación, ítems, cantidades, precios unitarios, total, vigencia y condiciones de pago.La alerta de vencimiento fue correcta y oportuna, y la recomendación resultó útil |
| **COT-S-013** | El sistema clasificó correctamente el documento como cotización y extrajo adecuadamente proveedor, destinatario, número, fecha, ítems, cantidades, precios unitarios, total, vigencia y condiciones de pago. |

| Código | Comentario |
|---|---|
| **COT-S-014** | El sistema clasificó correctamente el documento como cotización y extrajo adecuadamente proveedor, destinatario, número, fecha, ítems, cantidades, precios unitarios, total, vigencia y condiciones de pago. La ausencia de alertas fue coherente con la vigencia de la oferta y la recomendación fue útil, aunque requiere ajustes menores de saludo y cierre. |
| **COT-S-015** | El sistema clasificó correctamente el documento como cotización y extrajo adecuadamente proveedor, cliente, número, fecha, ítems, cantidades, precios unitarios, total, vigencia y condiciones de pago. |
| **COT-S-016** | MiDocIA clasificó correctamente el documento como cotización y extrajo adecuadamente proveedor, destinatario, número, fecha, ítems, cantidades, precios unitarios, total, vigencia y condiciones de pago. La recomendación es útil como borrador |
| **COT-S-017** | MiDocIA clasificó correctamente el documento como cotización y extrajo adecuadamente proveedor, destinatario, número, fecha, ítems, cantidades, precios unitarios, total, vigencia y condiciones de pago. |
| **COT-S-018** | MiDocIA clasificó correctamente el documento como cotización y extrajo adecuadamente proveedor, destinatario, número, fecha, ítems, cantidades, precios unitarios, total, vigencia y condiciones de pago. La recomendación es útil |
| **CAR-S-001** | MiDocIA clasificó el documento como “PQRS — Petición”; la clasificación se considera parcial porque, aunque corresponde a una carta comercial, su contenido formula una solicitud de modificación de reserva. Se requiere una corrección importante debido a la incorporación de un plazo legal de 15 días hábiles y una fecha límite del 10 de septiembre de 2026 que no se derivan de la comunicación original. |
| **PQR-S-001** | MiDocIA clasificó correctamente el documento como PQRS — Reclamo. Extrajo de manera precisa los principales datos del caso. El único campo parcial corresponde al plazo legal de respuesta, debido a que indicó 10 días hábiles en lugar de los 15 establecidos para PQRS en el protocolo de evaluación. |
| **PQR-S-002** | MiDocIA clasificó correctamente el documento como PQRS — Queja y extrajo adecuadamente la identificación de la remitente, el motivo, las solicitudes formuladas y la fecha de radicación. La descripción del caso fue parcial al condensar varios detalles de la afectación, y el plazo legal fue parcial. |

| Código | Comentario |
|---|---|
| **CON-S-001** | MiDocIA clasificó correctamente el documento como contrato. Las cláusulas de riesgo fueron parciales porque simplificó las disposiciones sobre daños y consumos adicionales. |
| **CON-S-002** | MiDocIA clasificó correctamente el documento como contrato de prestación de servicios. as obligaciones de las partes fueron parciales debido a que se omitieron algunas disposiciones expresas. as alertas fueron pertinentes y la recomendación resulta útil |
| **JUD-S-001** | MiDocIA clasificó correctamente el documento como una actuación judicial. o detectó expresamente la consecuencia prevista por falta de contestación.Consistente en que el proceso continuará sin que el vencimiento del término suspenda el trámite. |
| **FAC-S-001** | MiDocIA clasificó correctamente el documento como factura. La alerta de vencimiento fue pertinente y coincide con la fecha del documento. |
| **FAC-S-002** | MiDocIA clasificó correctamente el documento como factura. La alerta de vencimiento fue pertinente. |
| **FAC-S-003** | MiDocIA clasificó correctamente el documento como factura. La ausencia de alertas es adecuada debido a que el documento registra el pago completo y saldo cero. |
| **OTR-S-001** | MiDocIA clasificó correctamente el documento como una ficha de reserva y control de grupo turístico. La extracción fue parcial en información de contacto, horarios, desglose de servicios y control interno, y no presentó el punto de encuentro |
| **OTR-S-002** | MiDocIA clasificó correctamente el documento como Orden de Servicio. El itinerario fue parcial porque el análisis no presentó el detalle de horas, ubicaciones y actividades de cada etapa. La alerta y la recomendación son parcialmente útiles |
| **CAR-S-002** | MiDocIA clasificó correctamente el documento como carta comercial y recuperó con precisión los principales datos de la comunicación y la acción requerida. La alerta del 28 de agosto es pertinente. |

Tabla 15 C10 Campos evaluados: Validación detallada de los campos esperados por documento

| Código | Campo esperado | Resultado | Observación |
|---|---|---|---|
| **ITI-R-001** | Tipo de plan | Correcto | El resumen lo identifica como pasadía experiencial. |
| **ITI-R-001** | Destino | Correcto | Identificó Vetas, Santander. |
| **ITI-R-001** | Punto de salida | Parcial | El documento indica Glorieta del Estadio, pero el dato no fue visible debido a [object Object]. |
| **ITI-R-001** | Hora de salida | Parcial | El documento indica 5:00 a. m., pero el dato no fue visible en la interfaz. |
| **ITI-R-001** | Transporte | Correcto | El resumen reconoce que incluye transporte. |
| **ITI-R-001** | Alimentación | Parcial | Detectó comidas de forma general, pero no mostró desayuno, refrigerio y almuerzo por separado. |
| **ITI-R-001** | Actividades principales | Parcial | Detectó actividades, pero no permitió leer los ítems individuales. |

| Código | Campo esperado | Resultado | Observación |
|---|---|---|---|
| **ITI-R-001** | Degustación | Parcial | El documento incluye manjares vetanos, pero no fue visible en el resultado. |
| **ITI-R-001** | Guianza | Parcial | El documento incluye intérpretes locales, pero no apareció de forma legible. |
| **ITI-R-001** | Seguro | Parcial | Detectó seguro de manera general, pero no mostró claramente Colasistencia. |
| **ITI-R-001** | Primeros auxilios | Parcial | Está en el documento, pero no apareció legiblemente en el análisis. |
| **ITI-R-001** | Souvenir | Parcial | Está incluido, pero no fue visible en el resultado. |
| | | | |
| **ITI-R-002** | Tipo de documento | Correcto | Lo identificó como brochure promocional de un paquete turístico. |
| **ITI-R-002** | Empresa oferente | Correcto | Identificó Laguna Tours Turismo y Aventura. |
| **ITI-R-002** | Nombre del plan | Correcto | Identificó Plan Princesa Sora. |

| Código | Campo esperado | Resultado | Observación |
|---|---|---|---|
| **ITI-R-002** | Duración | Correcto | Identificó 2 días y 1 noche. |
| **ITI-R-002** | Destinos | Correcto | Identificó Vetas y California, Santander. |
| **ITI-R-002** | Punto y hora de salida | Correcto | Identificó Bucaramanga, 5:00 a. m., a convenir. |
| **ITI-R-002** | Alojamiento | Correcto | Identificó hotel en California. |
| **ITI-R-002** | Alimentación | Parcial | Menciona comidas y gastronomía, pero no detalla desayunos, almuerzos y refrigerios de cada día. |
| **ITI-R-002** | Actividades principales | Correcto | Incluyó senderismo, complejo lagunar, Pozo del Rey, mina, joyería y muestras culturales. |
| **ITI-R-002** | Itinerario por días | Parcial | Reconoce que existe un itinerario, pero no reproduce claramente la secuencia completa de los dos días. |
| **ITI-R-002** | Seguro y primeros auxilios | Correcto | Identificó Seguros Colasistencia y personal de primeros auxilios. |
| **ITI-R-002** | Contacto | Correcto | Identificó correctamente el número 322-3698404. |

| Código | Campo esperado | Resultado | Observación |
|---|---|---|---|
| **ITI-R-002** | Registro turístico | Correcto | Extrajo RTN-261074 tal como aparece escrito en el PDF. |
| **ITI-R-002** | Souvenirs | No detectado | El documento indica entrega de dos souvenirs, pero el análisis no lo menciona. |
| | | | |
| **COT-S-001** | Proveedor | Correcto | Identificó a Andrés Herrera como proveedor independiente de transporte. |
| **COT-S-001** | Destinatario | Correcto | Identificó a Laguna Tours, Grupo Alfa Capital S.A.S. y su NIT. |
| **COT-S-001** | Número de cotización | Correcto | Reconoció COT-001-2026, aunque solo aparece utilizado en la recomendación |
| **COT-S-001** | Fecha | Correcto | Extrajo correctamente el 27 de julio de 2026. |
| **COT-S-001** | Ítems cotizados | Parcial | El servicio fue descrito en el resumen, pero el campo estructurado mostró |

| Código | Campo esperado | Resultado | Observación |
|---|---|---|---|
| **COT-S-001** | Precio Unitario | Parcial | Recuperó el valor de $1.000.000, pero no lo presentó específicamente como precio unitario del servicio. |
| **COT-S-001** | Cantidad | No detectado | No indicó una cantidad. El documento tampoco contiene una columna específica de cantidad. |
| **COT-S-001** | Subtotal | Parcial | Recuperó el monto total, pero no diferenció un subtotal. |
| **COT-S-001** | IVA | No detectado | No identificó IVA. El documento original tampoco lo discrimina. |
| **COT-S-001** | Total | Correcto | Extrajo correctamente el valor total de 1000000 |
| **COT-S-001** | Vigencia | Correcto | Identificó que está sujeta a confirmación del proveedor. |
| **COT-S-001** | Condiciones de pago | Correcto | Identificó que están por acordar con el proveedor. |
| | | | |

| Código | Campo esperado | Resultado | Observación |
|---|---|---|---|
| **COT-S-002** | Proveedor | Correcto | Identificó a Irvin Arias como proveedor independiente de transporte. |
| **COT-S-002** | Destinatario | Correcto | Identificó a Laguna Tours, Grupo Alfa Capital S.A.S. y el NIT. |
| **COT-S-002** | Número de cotización | Correcto | Recuperó correctamente COT-002-2026 y lo utilizó en la recomendación. |
| **COT-S-002** | Fecha | Correcto | Extrajo el 27 de julio de 2026. |
| **COT-S-002** | Ítems cotizados | Parcial | El resumen identificó los tres conceptos, pero el campo estructurado mostró [object Object]. |
| **COT-S-002** | Precio Unitario | Parcial | Recuperó $600.000 y $120.000 en el resumen, pero no presentó un desglose estructurado por ítem. |
| **COT-S-002** | Cantidad | No detectado | No indicó una cantidad de unidades; los 14 pasajeros corresponden a capacidad máxima. |

| Código | Campo esperado | Resultado | Observación |
|---|---|---|---|
| **COT-S-002** | Subtotal | Parcial | Reconoció el valor base y el valor con guía, pero no los clasificó expresamente como subtotal. |
| **COT-S-002** | IVA | No detectado | No fue identificado; el documento original tampoco discrimina IVA. |
| **COT-S-002** | Total | Correcto | Indicó el valor base de $600.000 y explicó que con la guía el total sería de $720.000. |
| **COT-S-002** | Vigencia | Correcto | Identificó que está sujeta a confirmación del proveedor. |
| **COT-S-002** | Condiciones de pago | Correcto | Identificó que deben acordarse con el proveedor. |
| **COT-S-003** | Proveedor | Correcto | Identificó a Transportes Piedecuesta S.A. |
| **COT-S-003** | Destinatario | Correcto | Identificó a Laguna Tours, Grupo Alfa Capital S.A.S. y el NIT. |
| **COT-S-003** | Numero de cotización | Correcto | Recuperó correctamente COT-003-2026 en la recomendación. |

| Código | Campo esperado | Resultado | Observación |
|---|---|---|---|
| **COT-S-003** | Fecha | Correcto | Extrajo correctamente el 27 de julio de 2026. |
| **COT-S-003** | Ítems cotizados | Parcial | El resumen describió el transporte en buseta para 27 pasajeros, pero el campo estructurado mostró [object Object]. |
| **COT-S-003** | Precio Unitario | Parcial | Recuperó el valor de $1.200.000, pero no lo presentó expresamente como precio unitario del ítem. |
| **COT-S-003** | Cantidad | No detectado | No presentó una cantidad de unidades del servicio. Los 27 pasajeros corresponden a la capacidad del vehículo. |
| **COT-S-003** | Subtotal | Parcial | Reconoció el valor total, pero no lo identificó separadamente como subtotal. |
| **COT-S-003** | IVA | No detectado | No identificó IVA; el documento tampoco lo discrimina. |
| **COT-S-003** | Total | Correcto | Extrajo correctamente $1.200.000 COP. |

| Código | Campo esperado | Resultado | Observación |
|---|---|---|---|
| **COT-S-003** | VIgencia | Correcto | Identificó que está sujeta a confirmación del proveedor. |
| **COT-S-003** | Condiciones de pago | Correcto | Identificó que están por acordar con el proveedor. |
| | | | |
| **COT-S-004** | Proveedor | Correcto | Identificó a Transportes Piedecuesta S.A. |
| **COT-S-004** | Destinatario | Correcto | Identificó a Laguna Tours, Grupo Alfa Capital S.A.S. y el NIT. |
| **COT-S-004** | Numero de cotización | Correcto | Recuperó correctamente COT-004-2026 en la recomendación |
| **COT-S-004** | Fecha | Correcto | Extrajo el 27 de julio de 2026. |
| **COT-S-004** | Ítems cotizados | Correcto | Identificó el servicio de transporte en busetón para un máximo de 30 pasajeros. |
| **COT-S-004** | Precio Unitario | Correcto | Presentó correctamente el valor de $1.500.000. |

| Código | Campo esperado | Resultado | Observación |
|---|---|---|---|
| **COT-S-004** | Cantidad | Parcial | Indicó cantidad 1, pero este valor no aparece expresamente en el documento; fue inferido a partir de un único servicio cotizado. |
| **COT-S-004** | Subtotal | Parcial | El valor puede coincidir con el total por existir un solo ítem, pero no fue presentado expresamente como subtotal |
| **COT-S-004** | IVA | No detectado | No presentó información sobre IVA; el documento original tampoco lo discrimina. |
| **COT-S-004** | Total | Correcto | Extrajo correctamente $1.500.000. |
| **COT-S-004** | VIgencia | Correcto | Identificó que está sujeta a confirmación del proveedor. |
| **COT-S-004** | Condiciones de pago | Correcto | Identificó que están por acordar con el proveedor. |
| | | | |

| Código | Campo esperado | Resultado | Observación |
|---|---|---|---|
| **COT-S-005** | Proveedor | Correcto | Identificó a Jesús Hernando Arias. La ausencia de NIT fue indicada correctamente porque el documento no lo contiene. |
| **COT-S-005** | Destinatario | Correcto | Identificó a Laguna Tours, Grupo Alfa Capital S.A.S. y su NIT. |
| **COT-S-005** | Numero de cotización | Correcto | Recuperó correctamente COT-005-2026 en la recomendación. |
| **COT-S-005** | Fecha | Correcto | Extrajo correctamente el 27 de julio de 2026. |
| **COT-S-005** | Ítems cotizados | Correcto | Identificó el servicio de transporte en busetón para hasta 30 pasajeros. |
| **COT-S-005** | Precio Unitario | Parcial | Presentó $1.500.000 como precio unitario, pero el documento solo lo muestra como valor del servicio, sin una columna de cantidad. |

| Código | Campo esperado | Resultado | Observación |
|---|---|---|---|
| **COT-S-005** | Cantidad | Parcial | Indicó cantidad 1, aunque este dato no aparece expresamente en el documento y fue inferido. |
| **COT-S-005** | Subtotal | No detectado | No presentó un subtotal separado. |
| **COT-S-005** | IVA | No detectado | No identificó IVA; el documento tampoco lo discrimina. |
| **COT-S-005** | Total | Correcto | Extrajo correctamente $1.500.000. |
| **COT-S-005** | VIgencia | Correcto | Identificó que está sujeta a confirmación del proveedor. |
| **COT-S-005** | Condiciones de pago | Correcto | Identificó que están por acordar con el proveedor. |
| | | | |
| **COT-S-006** | Proveedor | Correcto | Identificó a Empresa Flota Cáchira Ltda. |
| **COT-S-006** | Destinatario | Correcto | Identificó a Laguna Tours, Grupo Alfa Capital S.A.S. y su NIT. |
| **COT-S-006** | Numero de cotización | Correcto | Recuperó COT-006-2026 en la recomendación. |

| Código | Campo esperado | Resultado | Observación |
|---|---|---|---|
| **COT-S-006** | Fecha | Correcto | Extrajo correctamente el 27 de julio de 2026. |
| **COT-S-006** | Ítems cotizados | Correcto | Identificó el transporte en buseta para hasta 22 pasajeros. |
| **COT-S-006** | Precio Unitario | Parcial | Presentó $1.500.000 como precio unitario, aunque el documento lo muestra como valor del servicio y no usa expresamente esa denominación. |
| **COT-S-006** | Cantidad | Parcial | Indicó cantidad 1, pero este dato fue inferido y no aparece expresamente en el documento. |
| **COT-S-006** | Subtotal | No detectado | No presentó un subtotal independiente. |
| **COT-S-006** | IVA | No detectado | No presentó información sobre IVA; el documento tampoco lo discrimina. |
| **COT-S-006** | Total | Correcto | Extrajo correctamente $1.500.000. |
| **COT-S-006** | VIgencia | Correcto | Identificó que está sujeta a confirmación del proveedor. |

| Código | Campo esperado | Resultado | Observación |
|---|---|---|---|
| **COT-S-006** | Condiciones de pago | Correcto | Identificó que están por acordar con el proveedor. |
| | | | |
| **COT-S-007** | Proveedor | Correcto | Identificó a Posada Rural Casa de los Nonos. |
| **COT-S-007** | Destinatario | Correcto | Identificó a Laguna Tours, Grupo Alfa Capital S.A.S. y su NIT. |
| **COT-S-007** | Numero de cotización | Correcto | Recuperó PRCN-074-2026 en la recomendación. |
| **COT-S-007** | Fecha | Correcto | Extrajo correctamente el 1 de agosto de 2026. |
| **COT-S-007** | Ítems cotizados | Correcto | Identificó la casa completa y el desayuno campesino opcional. |
| **COT-S-007** | Precio Unitario | Correcto | Extrajo $450.000 para la casa y $18.000 por desayuno. |
| **COT-S-007** | Cantidad | Correcto | Identificó una noche y 14 desayunos. |
| **COT-S-007** | Subtotal | No detectado | No presentó los subtotales de $450.000 y $252.000. |

| Código | Campo esperado | Resultado | Observación |
|---|---|---|---|
| **COT-S-007** | IVA | No detectado | No mostró información sobre IVA; el documento tampoco lo discrimina. |
| **COT-S-007** | Total | Parcial | Presentó $702.000 como total, pero no mostró separadamente el total base de $450.000 y el total con opcionales. |
| **COT-S-007** | VIgencia | Correcto | Calculó correctamente la fecha de vigencia hasta el 16 de agosto de 2026. |
| **COT-S-007** | Condiciones de pago | Correcto | Extrajo el anticipo del 50 % y el saldo antes del ingreso. |
| | | | |
| **COT-S-008** | Proveedor | Correcto | Identificó a Villas de la Montaña. |
| **COT-S-008** | Destinatario | Correcto | Identificó a Laguna Tours, Grupo Alfa Capital S.A.S. y su NIT. |
| **COT-S-008** | Numero de cotización | No detectado | No presentó el número VDM-118-2026. |
| **COT-S-008** | Fecha | Correcto | Extrajo el 1 de agosto de 2026. |

| Código | Campo esperado | Resultado | Observación |
|---|---|---|---|
| **COT-S-008** | Ítems cotizados | Correcto | Identificó cabañas, desayuno y fogata grupal. |
| **COT-S-008** | Precio Unitario | Correcto | Extrajo $420.000, $20.000 y $120.000. |
| **COT-S-008** | Cantidad | Correcto | Identificó 4 cabañas, 20 desayunos y 1 fogata. |
| **COT-S-008** | Subtotal | No detectado | No mostró $1.680.000, $400.000 y $120.000. |
| **COT-S-008** | IVA | No detectado | No presentó información sobre IVA; el documento tampoco lo discrimina. |
| **COT-S-008** | Total | Parcial | Mostró $2.200.000, pero no diferenció el total base de $2.080.000 del total con la fogata opcional. |
| **COT-S-008** | VIgencia | Correcto | Calculó correctamente el vencimiento para el 11 de agosto de 2026. |
| **COT-S-008** | Condiciones de pago | Correcto | Extrajo anticipo del 40 % y saldo tres días antes de la llegada. |
| | | | |
| **COT-S-009** | Proveedor | Correcto | Identificó a Posada Pirita Charta. |

| Código | Campo esperado | Resultado | Observación |
|---|---|---|---|
| **COT-S-009** | Destinatario | Correcto | Identificó a Laguna Tours, Grupo Alfa Capital S.A.S. y su NIT. |
| **COT-S-009** | Numero de cotización | Correcto | Recuperó PP-092-2026 en la recomendación. |
| **COT-S-009** | Fecha | Correcto | Extrajo correctamente el 31 de julio de 2026. |
| **COT-S-009** | Ítems cotizados | Correcto | Identificó alojamiento, desayuno típico y café de bienvenida |
| **COT-S-009** | Precio Unitario | Correcto | Extrajo $220.000, $18.000 y “Incluido”. |
| **COT-S-009** | Cantidad | Correcto | Identificó 8 habitaciones, 22 desayunos y 22 cafés. |
| **COT-S-009** | Subtotal | No detectado | No mostró los subtotales de $1.760.000 y $396.000. |
| **COT-S-009** | IVA | No detectado | No presentó información sobre IVA; el documento tampoco lo discrimina. |
| **COT-S-009** | Total | Correcto | Extrajo correctamente $2.156.000. |
| **COT-S-009** | VIgencia | Correcto | Calculó correctamente el vencimiento para el 15 de agosto de 2026. |

| Código | Campo esperado | Resultado | Observación |
|---|---|---|---|
| **COT-S-009** | Condiciones de pago | Correcto | Extrajo el anticipo del 50 % y el saldo al momento del ingreso. |
| | | | |
| **COT-S-010** | Proveedor | Correcto | Identificó a Hospedaje Real Santurbán y señaló correctamente que no se especifica NIT. |
| **COT-S-010** | Destinatario | Correcto | Identificó a Laguna Tours, Grupo Alfa Capital S.A.S. y su NIT. |
| **COT-S-010** | Numero de cotización | Correcto | Recuperó HRS-166-2026 en la recomendación. |
| **COT-S-010** | Fecha | Correcto | Extrajo el 31 de julio de 2026. |
| **COT-S-010** | Ítems cotizados | Correcto | Identificó habitaciones triples, desayuno y uso del salón común. |
| **COT-S-010** | Precio Unitario | Correcto | Extrajo $180.000, $20.000 e “Incluido”. |
| **COT-S-010** | Cantidad | Correcto | Identificó 9 habitaciones, 27 desayunos y 1 salón. |
| **COT-S-010** | Subtotal | No detectado | No presentó los subtotales de $1.620.000 y $540.000. |

| Código | Campo esperado | Resultado | Observación |
|---|---|---|---|
| **COT-S-010** | IVA | No detectado | No mostró información sobre IVA; el documento tampoco lo discrimina. |
| **COT-S-010** | Total | Correcto | Extrajo correctamente $2.160.000. |
| **COT-S-010** | VIgencia | Correcto | Calculó correctamente el vencimiento para el 12 de agosto de 2026. |
| **COT-S-010** | Condiciones de pago | Correcto | Extrajo el abono del 50 % y el saldo máximo 24 horas antes del ingreso. |
| | | | |
| **COT-S-011** | Proveedor | Correcto | Identificó a Hotel El Santuario de California S.A.S. |
| **COT-S-011** | Destinatario | Correcto | Identificó a Laguna Tours, Grupo Alfa Capital S.A.S. y su NIT. |
| **COT-S-011** | Numero de cotización | Correcto | Recuperó HES-2026-231 en la recomendación. |
| **COT-S-011** | Fecha | Correcto | Extrajo correctamente el 30 de julio de 2026. |
| **COT-S-011** | Ítems cotizados | Correcto | Identificó habitaciones triples, desayuno americano y salón compartido. |

| Código | Campo esperado | Resultado | Observación |
|---|---|---|---|
| **COT-S-011** | Precio Unitario | Correcto | Extrajo $220.000, $22.000 e “Incluido” |
| **COT-S-011** | Cantidad | Correcto | Extrajo $220.000, $22.000 e “Incluido”Identificó 10 habitaciones, 30 desayunos y 1 salón. |
| **COT-S-011** | Subtotal | No detectado | No presentó los subtotales de $2.200.000 y $660.000. |
| **COT-S-011** | IVA | No detectado | No presentó información sobre IVA; el documento tampoco lo discrimina. |
| **COT-S-011** | Total | Correcto | Extrajo correctamente $2.860.000. |
| **COT-S-011** | VIgencia | Correcto | Calculó correctamente el vencimiento para el 14 de agosto de 2026 |
| **COT-S-011** | Condiciones de pago | Correcto | Extrajo 50 % para confirmar y 50 % al ingreso del grupo |
| | | | |
| **COT-S-012** | Proveedor | Correcto | Identificó a Refugio Piedra Parada y señaló que no se especifica NIT. |

| Código | Campo esperado | Resultado | Observación |
|---|---|---|---|
| **COT-S-012** | Destinatario | Correcto | Identificó a Laguna Tours, Grupo Alfa Capital S.A.S. y su NIT. |
| **COT-S-012** | Numero de cotización | Correcto | Recuperó RPP-GR-308-2026 en la recomendación. |
| **COT-S-012** | Fecha | Correcto | Extrajo el 30 de julio de 2026. |
| **COT-S-012** | Ítems cotizados | Correcto | Identificó villas grupales, desayuno y acceso a zonas comunes. |
| **COT-S-012** | Precio Unitario | Correcto | Extrajo $950.000, $25.000 e “Incluido”. |
| **COT-S-012** | Cantidad | Correcto | Identificó 3 villas, 30 desayunos y 1 acceso a zonas comunes. |
| **COT-S-012** | Subtotal | No detectado | No presentó $2.850.000 y $750.000 como subtotales individuales. |
| **COT-S-012** | IVA | No detectado | No presentó información sobre IVA; el documento tampoco la discrimina. |
| **COT-S-012** | Total | Correcto | Extrajo correctamente $3.600.000. |
| **COT-S-012** | VIgencia | Correcto | Identificó correctamente la vigencia hasta el 9 de agosto de 2026. |

| Código | Campo esperado | Resultado | Observación |
|---|---|---|---|
| **COT-S-012** | Condiciones de pago | Correcto | Extrajo anticipo del 50 % y saldo cinco días antes del ingreso. |
| | | | |
| **COT-S-013** | Proveedor | Correcto | Identificó a Al Rancho Restaurante - Bar. |
| **COT-S-013** | Destinatario | Correcto | Identificó a Laguna Tours Grupo Alfa Capital S.A.S. |
| **COT-S-013** | Numero de cotización | Correcto | Recuperó ARR-0826-014 en la recomendación. |
| **COT-S-013** | Fecha | Correcto | Extrajo correctamente el 12 de agosto de 2026. |
| **COT-S-013** | Ítems cotizados | Correcto | Identificó los cuatro componentes del plan de alimentación. |
| **COT-S-013** | Precio Unitario | Correcto | Extrajo $18.000, $8.000, $32.000 y $8.000. |
| **COT-S-013** | Cantidad | Correcto | Identificó correctamente 20 unidades para cada servicio. |
| **COT-S-013** | Subtotal | No detectado | No mostró los subtotales de $360.000, $160.000, $640.000 y $160.000. |

| Código | Campo esperado | Resultado | Observación |
|---|---|---|---|
| **COT-S-013** | IVA | No detectado | No presentó IVA; el documento original tampoco lo discrimina. |
| **COT-S-013** | Total | Correcto | Extrajo correctamente $1.320.000. |
| **COT-S-013** | VIgencia | Correcto | Calculó la vigencia hasta el 27 de agosto de 2026. |
| **COT-S-013** | Condiciones de pago | Correcto | Extrajo anticipo del 50 % y saldo antes de iniciar el servicio. |
| | | | |
| **COT-S-014** | Proveedor | Correcto | Identificó a Restaurante Casa Vieja Charta. |
| **COT-S-014** | Destinatario | Correcto | Identificó a Laguna Tours Grupo Alfa Capital S.A.S. y su NIT. |
| **COT-S-014** | Numero de cotización | Correcto | Recuperó CVC-126-2026 en la recomendación. |
| **COT-S-014** | Fecha | Correcto | Extrajo el 12 de agosto de 2026. |
| **COT-S-014** | Ítems cotizados | Correcto | Identificó desayuno, refrigerio de mañana, almuerzo y refrigerio de tarde. |

| Código | Campo esperado | Resultado | Observación |
|---|---|---|---|
| **COT-S-014** | Precio Unitario | Correcto | Extrajo $15.000, $7.000, $26.000 y $7.000. |
| **COT-S-014** | Cantidad | Correcto | Identificó 14 para cada ítem. |
| **COT-S-014** | Subtotal | No detectado | No mostró $210.000, $98.000, $364.000 y $98.000. |
| **COT-S-014** | IVA | No detectado | No presentó IVA; el documento original tampoco lo discrimina. |
| **COT-S-014** | Total | Correcto | Extrajo $770.000. |
| **COT-S-014** | VIgencia | Correcto | Calculó correctamente hasta el 27 de agosto de 2026. |
| **COT-S-014** | Condiciones de pago | Correcto | Extrajo anticipo del 50 % y saldo antes de iniciar el servicio. |
| | | | |
| **COT-S-015** | Proveedor | Correcto | Identificó a Restaurante Media Bota. |
| **COT-S-015** | Destinatario | Correcto | Identificó a Laguna Tours Grupo Alfa Capital S.A.S. y su NIT. |
| **COT-S-015** | Numero de cotización | Correcto | Recuperó MB-2026-0812 en la recomendación. |

| Código | Campo esperado | Resultado | Observación |
|---|---|---|---|
| **COT-S-015** | Fecha | Correcto | Extrajo 12 de agosto de 2026. |
| **COT-S-015** | Ítems cotizados | Correcto | Identificó desayuno, dos refrigerios y almuerzo. |
| **COT-S-015** | Precio Unitario | Correcto | Extrajo $14.000, $7.500, $25.000 y $7.500. |
| **COT-S-015** | Cantidad | Correcto | Identificó 22 para cada servicio. |
| **COT-S-015** | Subtotal | No detectado | No mostró $308.000, $165.000, $550.000 y $165.000. |
| **COT-S-015** | IVA | No detectado | No presentó IVA; el documento tampoco lo discrimina. |
| **COT-S-015** | Total | Correcto | Extrajo $1.188.000. |
| **COT-S-015** | VIgencia | Correcto | Calculó hasta el 27 de agosto de 2026. |
| **COT-S-015** | Condiciones de pago | Correcto | Identificó anticipo del 50 % y saldo antes del servicio. |
| | | | |
| **COT-S-016** | Proveedor | Correcto | Identificó a CASICAL. |
| **COT-S-016** | Destinatario | Correcto | Identificó a Laguna Tours Grupo Alfa Capital S.A.S. y su NIT. |

| Código | Campo esperado | Resultado | Observación |
|---|---|---|---|
| **COT-S-016** | Numero de cotización | Correcto | Recuperó CAS-245-26 en la recomendación. |
| **COT-S-016** | Fecha | Correcto | Extrajo 12 de agosto de 2026. |
| **COT-S-016** | Ítems cotizados | Correcto | Identificó desayuno, refrigerio de mañana, almuerzo especial y refrigerio de tarde. |
| **COT-S-016** | Precio Unitario | Correcto | Extrajo $14.000, $8.000, $28.000 y $8.000. |
| **COT-S-016** | Cantidad | Correcto | Identificó 27 unidades para cada servicio. |
| **COT-S-016** | Subtotal | No detectado | El documento contiene $378.000, $216.000, $756.000 y $216.000, pero MiDocIA no los mostró. |
| **COT-S-016** | IVA | No detectado | No presentó información sobre IVA; el documento tampoco la discrimina. |
| **COT-S-016** | Total | Correcto | Extrajo $1.566.000. |
| **COT-S-016** | VIgencia | Correcto | Calculó correctamente el 27 de agosto de 2026. |
| **COT-S-016** | Condiciones de pago | Correcto | Identificó anticipo del 50 % y saldo antes de iniciar el servicio. |
| | | | |

| Código | Campo esperado | Resultado | Observación |
|---|---|---|---|
| **COT-S-017** | Proveedor | Correcto | Identificó a Restaurante y Asadero Donde Javis - Zairotales y señaló correctamente que no se especifica NIT. |
| **COT-S-017** | Destinatario | Correcto | Identificó a Laguna Tours Grupo Alfa Capital S.A.S. y su NIT. |
| **COT-S-017** | Numero de cotización | Correcto | Recuperó DJ-0812-2026 en la recomendación. |
| **COT-S-017** | Fecha | Correcto | Extrajo el 12 de agosto de 2026. |
| **COT-S-017** | Ítems cotizados | Correcto | Identificó desayuno berlinero, refrigerio de mañana, almuerzo asadero y refrigerio de tarde. |
| **COT-S-017** | Precio Unitario | Correcto | Extrajo $16.000, $8.000, $30.000 y $8.000. |
| **COT-S-017** | Cantidad | Correcto | Identificó 30 unidades para cada servicio. |
| **COT-S-017** | Subtotal | No detectado | El documento contiene $480.000, $240.000, $900.000 y $240.000, pero MiDocIA no los mostró. |

| Código | Campo esperado | Resultado | Observación |
|---|---|---|---|
| **COT-S-017** | IVA | No detectado | No presentó IVA; el documento tampoco lo discrimina. |
| **COT-S-017** | Total | Correcto | Extrajo $1.860.000. |
| **COT-S-017** | Vigencia | Correcto | Calculó correctamente hasta el 27 de agosto de 2026. |
| **COT-S-017** | Condiciones de pago | Correcto | Identificó anticipo del 50 % y saldo antes de iniciar el servicio. |
| | | | |
| **COT-S-018** | Proveedor | Correcto | Identificó a Hotel y Restaurante La Berlinera. |
| **COT-S-018** | Destinatario | Correcto | Identificó a Laguna Tours Grupo Alfa Capital S.A.S. |
| **COT-S-018** | Numero de cotización | Correcto | Recuperó LB-318-2026 en la recomendación. |
| **COT-S-018** | Fecha | Correcto | Extrajo el 12 de agosto de 2026. |
| **COT-S-018** | Ítems cotizados | Correcto | Identificó desayuno, refrigerio de mañana, almuerzo y refrigerio de tarde. |
| **COT-S-018** | Precio Unitario | Correcto | Extrajo $17.000, $8.500, $29.000 y $8.500. |

| Código | Campo esperado | Resultado | Observación |
|---|---|---|---|
| **COT-S-018** | Cantidad | Correcto | Identificó 30 para cada servicio. |
| **COT-S-018** | Subtotal | No detectado | No mostró $510.000, $255.000, $870.000 y $255.000. |
| **COT-S-018** | IVA | No detectado | No presentó IVA; el documento tampoco lo discrimina. |
| **COT-S-018** | Total | Correcto | Extrajo correctamente $1.890.000. |
| **COT-S-018** | Vigencia | Correcto | Calculó correctamente hasta el 27 de agosto de 2026. |
| **COT-S-018** | Condiciones de pago | Correcto | Identificó anticipo del 50 % y saldo antes de iniciar el servicio. |
| | | | |
| **CAR-S-001** | Remitente | Correcto | Identificó a Laura Marcela Rincón Torres, su cargo, empresa, NIT y datos de contacto |

| Código | Campo esperado | Resultado | Observación |
|---|---|---|---|
| **CAR-S-001** | Destinatario | Parcial | La recomendación evidencia que identificó a Laguna Tours como la parte receptora, pero no presentó expresamente el destinatario completo ni a Claudia Laguna en los datos relevantes. |
| **CAR-S-001** | Fecha | Correcto | Extrajo correctamente el 20 de agosto de 2026, aunque lo denominó “fecha de radicación”. |
| **CAR-S-001** | Asunto | Correcto | El “motivo principal” reproduce adecuadamente el asunto: modificación de reserva grupal por aumento de participantes. |
| **CAR-S-001** | Cuerpo del mensaje | Parcial | Capturó el propósito, pero omitió datos como salida el 5 de septiembre, retorno el 6 de septiembre y punto de encuentro en Bucaramanga a las 5:30 a. m. |

| Código | Campo esperado | Resultado | Observación |
|---|---|---|---|
| **CAR-S-001** | Accion requerida | Correcto | Identificó el ajuste a 27 personas, transporte, alojamiento, alimentación, disponibilidad, valor adicional, condiciones de pago y respuesta antes del 24 de agosto. |
| **CAR-S-001** | Firma | Correcto | Recuperó a Laura Marcela Rincón Torres como Coordinadora de Gestión Humana de Innova Santander S.A.S. |
| | | | |
| **PQR-S-001** | Remitente | Correcto | Identificó correctamente a Laura Fernanda Moreno Díaz y sus datos de contacto. |
| **PQR-S-001** | Tipo de Solicitud | Correcto | Lo identificó expresamente como reclamo. |
| **PQR-S-001** | Motivo | Correcto | Identificó el retraso de 90 minutos y la afectación de la actividad programada. |

| Código | Campo esperado | Resultado | Observación |
|---|---|---|---|
| **PQR-S-001** | Descripción del caso | Correcto | Resumió adecuadamente la situación y la experiencia insatisfactoria ocasionada por el retraso. |
| **PQR-S-001** | Solicitud concreta | Correcto | Recuperó las tres solicitudes: explicación, compensación y confirmación escrita. |
| **PQR-S-001** | Fecha de radicación | Correcto | Extrajo el 18 de agosto de 2026. |
| **PQR-S-001** | Plazo legal de respuesta | Parcial | Detectó un plazo, pero indicó 10 días hábiles cuando el protocolo establece 15 días hábiles para PQRS. |
| | | | |
| **PQR-S-002** | Remitente | Correcto | Identificó correctamente a Diana Carolina Rueda Silva, identificación, correo y teléfono. |
| **PQR-S-002** | Tipo de solicitud | Correcto | La clasificó como queja, coherente con el contenido del documento. |

| Código | Campo esperado | Resultado | Observación |
|---|---|---|---|
| **PQR-S-002** | Motivo | Correcto | Identificó las novedades del alojamiento y el retraso del desayuno. |
| **PQR-S-002** | Descripción del caso | Parcial | Reconoció los problemas principales, pero omitió detalles como las dos habitaciones no listas, la habitación triple no disponible, los tres huéspedes reubicados temporalmente y los horarios específicos. |
| **PQR-S-002** | Solicitud concreta | Correcto | Recuperó las cuatro solicitudes, incluida la compensación de $300.000 o descuento equivalente. |
| **PQR-S-002** | Fecha de radicación | Correcto | Identificó correctamente el 19 de agosto de 2026. |
| **PQR-S-002** | Plazo legal de respuesta | Parcial | Identificó que existe un plazo, pero indicó 10 días hábiles cuando el protocolo establece 15. |
| | | | 0 |

| Código | Campo esperado | Resultado | Observación |
|---|---|---|---|
| **CON-S-001** | Partes (contratante y contratista) | Correcto | Identificó a Grupo Alfa Capital S.A.S. / Laguna Tours y Posada Senderos de Niebla S.A.S., incluyendo representante. |
| **CON-S-001** | Objeto del contrato | Correcto | Recuperó el hospedaje para 22 personas, una noche y los servicios incluidos. |
| **CON-S-001** | Valor | Correcto | Identificó correctamente $2.156.000 antes de impuestos, IVA de $409.640 y total de $2.565.640. |
| **CON-S-001** | Forma de pago | Correcto | Identificó anticipo del 50 % por $1.282.820 y saldo dentro de los cinco días calendario posteriores a la salida, previa factura |
| **CON-S-001** | Fecha de inicio | Correcto | Identificó vigencia desde el 10 de agosto de 2026. |
| **CON-S-001** | Fecha de vencimiento | Correcto | Identificó como límite el 22 de agosto de 2026. |

| Código | Campo esperado | Resultado | Observación |
|---|---|---|---|
| **CON-S-001** | Obligaciones de cada parte | Correcto | Recuperó obligaciones tanto de la Agencia como del Alojamiento. |
| **CON-S-001** | Cláusulas de riesgo | Parcial | Detectó la pérdida del anticipo y los daños, pero agrupó los consumos adicionales como responsabilidad del grupo; el contrato establece que corresponden a los huéspedes salvo autorización escrita de facturación centralizada. |
| **CON-S-001** | Firma/legalización | No detectado | El documento incluye espacios de firma para Claudia Laguna y Ángela Marcela Pardo Nieto, pero MiDocIA no los presentó entre los datos relevantes. |
| | | | |

| Código | Campo esperado | Resultado | Observación |
|---|---|---|---|
| **CON-S-002** | Partes (contratante y contratista) | Correcto | Identificó a Grupo Alfa Capital S.A.S. / Laguna Tours y Transportes Cordillera Oriental S.A.S., incluyendo representantes e identificaciones. |
| **CON-S-002** | Objeto del contrato | Correcto | Recuperó correctamente el servicio especial de transporte para 30 viajeros, vehículo y ruta principal. |
| **CON-S-002** | Valor | Correcto | Identificó el valor total de $1.500.000. |
| **CON-S-002** | Forma de pago | Correcto | Identificó pago del 100 % dentro de los siete días calendario siguientes a la prestación satisfactoria, previa factura. |
| **CON-S-002** | Fecha de inicio | Correcto | Identificó la suscripción del 1 de agosto y la ejecución del servicio el 2 de agosto de 2026. |

| Código | Campo esperado | Resultado | Observación |
|---|---|---|---|
| **CON-S-002** | Fecha de vencimiento | Correcto | Identificó correctamente la vigencia máxima hasta el 20 de agosto de 2026. |
| **CON-S-002** | Obligaciones de cada parte | Parcial | Recuperó varias obligaciones, pero omitió algunas: no sustituir el vehículo por uno de menor capacidad, informar cambios de horario o recorrido y velar por las instrucciones de seguridad de los pasajeros. |
| **CON-S-002** | Cláusulas de riesgo | Correcto | Identificó la penalidad del 20 %, la compensación del 10 % por cancelación imputable al contratista, responsabilidad y causales de terminación. |

| Código | Campo esperado | Resultado | Observación |
|---|---|---|---|
| **CON-S-002** | Firma/legalización | No detectado | En la página 2 aparecen los espacios de firma de Claudia Laguna y Felipe Andrés Mendoza Ríos, pero el análisis no los reportó. |
| | | | |
| **JUD-S-001** | Autoridad judicial | Correcto | Identificó el Juzgado Civil Municipal de Prueba No. 01 de Bucaramanga. |
| **JUD-S-001** | Número de expediente | Correcto | Recuperó el radicado 68001-40-03-099-2026-00417-00. |
| **JUD-S-001** | Partes (demandante/demandado) | Correcto | Identificó a Paula Andrea Mendoza Rueda como demandante y Turismo Cumbres de Santander S.A.S. como demandado. |
| **JUD-S-001** | Tipo de providencia | Correcto | Identificó el auto de admisión de demanda y además reconoció la constancia de notificación simulada. |

| Código | Campo esperado | Resultado | Observación |
|---|---|---|---|
| **JUD-S-001** | Orden principal | Correcto | Reconoció que la demanda fue admitida y que se corrió traslado a la parte demandada para contestarla. |
| **JUD-S-001** | Plazos | Correcto | Identificó correctamente el término de 10 días hábiles, inicio el 25 de agosto y vencimiento simulado el 7 de septiembre de 2026. |
| **JUD-S-001** | Consecuencias por incumplimiento | No detectado | El auto advierte expresamente que el vencimiento del término sin contestación no suspende el trámite y que el proceso continuará; MiDocIA no presentó esta consecuencia. |
| **JUD-S-001** | Fecha | Correcto | Identificó el auto del 19 de agosto, la notificación simulada del 24 de agosto y las demás fechas relacionadas con el término. |
| | | | |

| Código | Campo esperado | Resultado | Observación |
|---|---|---|---|
| **FAC-S-001** | Emisor | Correcto | Identificó a Transportes Cordillera Oriental S.A.S. |
| **FAC-S-001** | Receptor | Correcto | Identificó a Grupo Alfa Capital S.A.S. |
| **FAC-S-001** | NIT/CC | Correcto | Recuperó correctamente los NIT del emisor y receptor. |
| **FAC-S-001** | Número de factura | Correcto | Identificó FE-TV-1047. |
| **FAC-S-001** | Fecha de emisión | Correcto | Extrajo 05/08/2026. |
| **FAC-S-001** | Fecha de vencimiento | Correcto | Extrajo 12/08/2026. |
| **FAC-S-001** | Ítems/conceptos | Correcto | Describió correctamente el servicio de transporte en busetón para 30 pasajeros y la ruta. |
| **FAC-S-001** | Subtotal | Correcto | Extrajo $1.500.000. |
| **FAC-S-001** | IVA | Correcto | Identificó $0; el documento indica que el servicio está excluido de IVA. |
| **FAC-S-001** | Total | Correcto | Extrajo $1.500.000. |

| Código | Campo esperado | Resultado | Observación |
|---|---|---|---|
| **FAC-S-001** | Estado de pago | Correcto | Lo reportó como “vencida” y en la recomendación aclaró que está pendiente de pago; el original dice “PENDIENTE - VENCIDA”. |
| **FAC-S-001** | Condiciones de pago | No detectado | El documento indica “Forma de pago: Crédito 7 días”, pero MiDocIA no presentó este dato. |
| | | | |
| **FAC-S-002** | Emisor | Correcto | Identificó Posada Senderos de Niebla S.A.S. |
| **FAC-S-002** | Receptor | Correcto | Identificó Grupo Alfa Capital S.A.S. / Laguna Tours. |
| **FAC-S-002** | NIT/CC | Correcto | Recuperó correctamente los NIT de las partes. |
| **FAC-S-002** | Número de factura | Correcto | FE-SN-0286 |
| **FAC-S-002** | Fecha de emisión | Correcto | 17/08/2026. |
| **FAC-S-002** | Fecha de vencimiento | Correcto | 22/08/2026. |
| **FAC-S-002** | Ítems/conceptos | Correcto | Identificó alojamiento para 22 huéspedes, desayuno y bebida de bienvenida. |

| Código | Campo esperado | Resultado | Observación |
|---|---|---|---|
| **FAC-S-002** | Subtotal | Correcto | $2.156.000. |
| **FAC-S-002** | IVA | Correcto | 19 % equivalente a $409.640. |
| **FAC-S-002** | Total | Correcto | $2.565.640. |
| **FAC-S-002** | Estado de pago | Correcto | Aunque el dato relevante dice “pendiente”, la recomendación reconoce correctamente que existe un pago parcial de $1.282.820 y saldo por igual valor; coincide con el original. |
| **FAC-S-002** | Condiciones de pago | No detectado | El original indica “50% anticipo / saldo a 5 días”, pero MiDocIA no presentó esta condición como campo extraído. |
| | | | |
| **FAC-S-003** | Emisor | Correcto | Sabores del Frailejón S.A.S. |
| **FAC-S-003** | Receptor | Correcto | Grupo Alfa Capital S.A.S. / Laguna Tours. |
| **FAC-S-003** | NIT/CC | Correcto | Recuperó correctamente los NIT de ambas partes. |
| **FAC-S-003** | Número de factura | Correcto | FE-SF-7612. |

| Código | Campo esperado | Resultado | Observación |
|---|---|---|---|
| **FAC-S-003** | Fecha de emisión | Correcto | 18/08/2026. |
| **FAC-S-003** | Fecha de vencimiento | Correcto | 18/08/2026. |
| **FAC-S-003** | Ítems/conceptos | Correcto | Recuperó desayuno, refrigerio de la mañana, almuerzo y refrigerio de la tarde para 30 personas. |
| **FAC-S-003** | Subtotal | Correcto | $1.860.000. |
| **FAC-S-003** | IVA/impuesto | Correcto | Identificó correctamente el Impuesto Nacional al Consumo del 8 % = $148.800. |
| **FAC-S-003** | Total | Correcto | $2.008.800. |
| **FAC-S-003** | Estado de pago | Correcto | Identificó “pagada”; el original registra pago recibido por $2.008.800 y saldo $0. |
| **FAC-S-003** | Condiciones de pago | No detectado | El documento indica “Forma de pago: Contado”, pero MiDocIA no la presentó entre los datos relevantes. |
| | | | |
| **OTR-S-001** | Número de reserva | Correcto | Identificó RSV-2026-087. |
| **OTR-S-001** | Estado de la reserva | Correcto | Identificó que está CONFIRMADA. |

| Código | Campo esperado | Resultado | Observación |
|---|---|---|---|
| **OTR-S-001** | Cliente, responsable y contacto | Parcial | Identificó Soluciones Andinas Integradas S.A.S., NIT y María José León Castro, pero omitió teléfono y correo. |
| **OTR-S-001** | Número de personas | Correcto | Identificó correctamente 27 personas. |
| **OTR-S-001** | Destino | Correcto | Recuperó Berlín – Vetas – Santurbán. |
| **OTR-S-001** | Fechas y horarios del viaje | Parcial | Identificó salida el 5 y regreso el 6 de septiembre, pero omitió las horas de 5:30 a. m. y 5:00 p. m. |
| **OTR-S-001** | Punto de encuentro | No detectado | No presentó “Parqueadero empresarial, sector Cabecera, Bucaramanga”. |
| **OTR-S-001** | Servicios contratados y control de pagos | Parcial | Reconoció transporte, alojamiento, alimentación y guía, pero no presentó cantidades, valores y pagos de cada servicio. |

| Código | Campo esperado | Resultado | Observación |
|---|---|---|---|
| **OTR-S-001** | Valor total, abonado y saldo | Correcto | Recuperó $6.004.000, $2.912.000 y $3.092.000. |
| **OTR-S-001** | Fecha límite de pago | Correcto | Identificó el 31 de agosto de 2026. |
| **OTR-S-001** | Observaciones y pendientes operativos | Parcial | Identificó confirmación de pasajeros, restricciones alimentarias, distribución de habitaciones, menús y guía. |
| **OTR-S-001** | Control interno | Correcto | Identificó menús y guía pendientes, pero omitió que ya figuran lista de pasajeros recibida, anticipo verificado, vehículo preasignado y hospedaje prebloqueado. |
| | | | |
| **OTR-S-002** | Empresa emisora | Correcto | Identificó a Operadora Turística Sendas del Oriente S.A.S. y su NIT. |
| **OTR-S-002** | Número de orden de servicio | Correcto | Identificó OS-2026-0912-030. |

| Código | Campo esperado | Resultado | Observación |
|---|---|---|---|
| **OTR-S-002** | Cliente | Correcto | Identificó Fundación Caminos del Futuro y su NIT. |
| **OTR-S-002** | Grupo / pasajeros | Correcto | Identificó correctamente 30 pasajeros. |
| **OTR-S-002** | Fecha del servicio | Correcto | Recuperó el 12 de septiembre de 2026. |
| **OTR-S-002** | Ruta / destinos | Correcto | Recuperó Bucaramanga – Berlín – Vetas – sector Cunta – Bucaramanga. |
| **OTR-S-002** | Coordinador de grupo | Correcto | Identificó a Natalia Fernanda Pardo Ríos y su teléfono. |
| **OTR-S-002** | Responsable operativo | Correcto | Identificó a Diego Alejandro Suárez Mora y su teléfono. |
| **OTR-S-002** | Itinerario operativo | Parcial | Reconoció que existe un itinerario detallado, pero no presentó las nueve etapas con sus horas, ubicaciones, actividades y responsables. |
| **OTR-S-002** | Transporte asignado | Correcto | Identificó busetón para 30 pasajeros y al proveedor Rutas del Páramo S.A.S. |

| Código | Campo esperado | Resultado | Observación |
|---|---|---|---|
| **OTR-S-002** | Alimentación asignada | Correcto | Recuperó desayuno, refrigerio de mañana, almuerzo, refrigerio de tarde y proveedor Fogón Santurbán. |
| **OTR-S-002** | Guianza y observaciones operativas | Correcto | Identificó guía local Vetas–Cunta y requisitos de documento, chaqueta, calzado, hidratación y condiciones climáticas/de movilidad. |
| | | | |
| **CAR-S-002** | Remitente | Correcto | Identificó a Santiago Andrés Vega Prieto, su cargo y la Fundación Horizontes del Norte. |
| **CAR-S-002** | Destinatario | Correcto | Identificó a Claudia Laguna y a Laguna Tours - Grupo Alfa Capital S.A.S. |
| **CAR-S-002** | Fecha | Correcto | Extrajo correctamente el 22 de agosto de 2026. |
| **CAR-S-002** | Asunto | Correcto | Recuperó la confirmación del servicio de pasadía para 30 personas. |

| Código | Campo esperado | Resultado | Observación |
|---|---|---|---|
| **CAR-S-002** | Cuerpo del mensaje | Parcial | Captó la finalidad y los elementos principales, pero condensó detalles del servicio, itinerario y condiciones descritas en el documento. |
| **CAR-S-002** | Acción requerida | Correcto | Identificó punto de encuentro, regreso, vehículo, contacto, menú, recomendaciones y fecha máxima de respuesta. |
| **CAR-S-002** | Firma | Correcto | Identificó al firmante Santiago Andrés Vega Prieto y su cargo. |

Tabla 16 C11 Resumen general de los resultados de la evaluación de MiDocIA

| Meta de pruebas | Documentos registrados | Documentos evaluados | Documentos empresariales | Documentos sintéticos | Precisión de clasificación | Precisión promedio de extracción | Utilidad promedio | Tiempo promedio ahorrado | Reducción promedio |
|---|---|---|---|---|---|---|---|---|---|
| 32 | 32 | 32 | 20 | 12 | 93,75 % | 82,33% | 3,5625 | 6,25 | 50,65 % |