INFORME CASOS DE USO DE IA APLICABLES A LA REALIDAD MUNICIPIO DE IPIALES

ELABORADO POR:

DIEGO FERNANDO AGUADO MURIEL

JHON JAIR AGREDO MESA

JOSE ALEJANDRO CHAMORRO LOPEZ

BRAYAM SEFF GRAVINI ANGULO

RODRIGO DIAZ CHANTRE

DARIO ACOSTA PARRA

Municipio de Ipiales – Nariño

DFA CONSULTORES SAS

TABLA DE CONTENIDO

1. RESUMEN EJECUTIVO 5
2. CONTEXTO MUNICIPAL 5
3. PROBLEMA ACTUAL 5
3.1. Rezago documental y falta de digitalización: 5
3.2. Base de datos tributaria incompleta 6
3.3. Procesos manuales en cobro coactivo. 6
3.4. Riesgos de corrupción. 6
3.5. Analítica inexistente de datos. 6
4. OBJETIVO DEL INFORME 7
5. METODOLOGÍA UTILIZADA 7
5.1. Fuentes de información 8
5.2. Criterios de evaluación de casos de uso 8
6. CONTEXTO MUNICIPAL 8
6.1. Brechas tecnológicas relevantes para Hacienda 9
6.2. Oportunidades para IA en Hacienda 9
7. IDENTIFICACIÓN DE CASOS DE USO DE IA 9
7.1. CASO DE USO 1: Digitalización Inteligente de Expedientes Tributarios (OCR + IA) .. 9
7.1.1. Descripción del Caso 9
7.1.2. Dependencia Responsable 10
7.1.3. Situación Actual 10
7.1.4. Propuesta Basada en IA 10
7.1.5. Requerimientos Técnicos 10
7.1.6. Beneficios Esperados 11
7.1.7. Implementación técnica propuesta 11
7.1.8. Matriz de Viabilidad del Caso 13
7.2. CASO DE USO 2: Motor Predictivo de Recaudo y Riesgo de Morosidad 14
7.2.1. Descripción del Caso 14
7.2.2. Dependencia Responsable 15
7.2.3. Situación Actual 15
7.2.4. Propuesta Basada en IA 15
7.2.5. Requerimientos Técnicos 15
7.2.6. Beneficios Esperados 15
7.2.7. Implementación técnica propuesta 16
7.2.8. Matriz de Viabilidad del Caso 17
7.3 CASO DE USO 3: Validación Automática de Datos Tributarios (Barrido de base actual) 18

7.3.1. Descripción del Caso 18
7.3.2. Dependencia Responsable 19
7.3.3. Situación Actual 19
7.3.4. Propuesta Basada en IA 19
7.3.5. Requerimientos Técnicos 19
7.3.6. Beneficios Esperados 19
7.3.7. Implementación técnica propuesta: 20
7.3.8. Matriz de Viabilidad del Caso 21

7.4 CASO DE USO 4: Chatbot Tributario con IA 22
7.4.1. Descripción del Caso 22
7.4.2. Dependencia Responsable 23
7.4.3. Situación Actual 23
7.4.4. Propuesta Basada en IA 23
7.4.5. Requerimientos Técnicos 23
7.4.6. Beneficios Esperados 23
7.4.7. Implementación técnica propuesta 23
7.4.8. Matriz de Viabilidad del Caso 25

7.5 CASO DE USO 5: Detección de Anomalías y Riesgos de Corrupción (barrido de base de datos) 26
7.5.1. Descripción del Caso 26
7.5.2. Dependencia Responsable 26
7.5.3. Situación Actual 27
7.5.4. Propuesta Basada en IA 27
7.5.5. Requerimientos Técnicos 27
7.5.6. Beneficios Esperados 28
7.5.7. Implementación técnica propuesta 28
7.5.8. Matriz de Viabilidad del Caso 30

7.6 CASO DE USO 6: Validación Automática de Procesos de Gestión Tributaria (IA + BPM) 31
7.6.1. Descripción del Caso 31
7.6.2. Dependencia Responsable 32
7.6.3. Situación Actual 32
7.6.4. Propuesta Basada en IA 32
7.6.5. Requerimientos Técnicos 32
7.6.6. Beneficios Esperados 33
7.6.7. Implementación técnica propuesta 33

8 PROPUESTA DE CASOS DE USO RECOMENDADOS (LISTA BASE PARA IPIALES) 33

8.1. Digitalización Inteligente de Expedientes Tributarios (OCR + IA) (Caso #1) 34

8.2. Validación Automática de Datos Tributarios (Barrido de base actual) (Caso #3) 34

8.3. Motor Predictivo de Recaudo y Riesgo de Morosidad (Caso #2) 34

8.4. Detección de Anomalías y Riesgos de Corrupción (Caso #5) 34

8.5. Chatbot Tributario con IA (Caso #4) 34

9 HOJA DE RUTA DE IMPLEMENTACIÓN (ROADMAP) 35
FASE 0: Preparación y Habilitación 35
FASE 1: Capa de Fundación (Digitalización y Calidad de Datos) 35
FASE 2: Motores de Valor (Analítica y Seguridad) 36
FASE 3: Despliegue, MLOps y Experiencia Ciudadana 37
10 RIESGOS Y CONSIDERACIONES ÉTICAS 37
10.1. Sesgo algorítmico 37
10.2. Gobernanza de IA 38
11 CONCLUSIONES 39
11.1. Solución Estratégica es la Interoperabilidad: 39
11.2. Cumplimiento de Meta PDM: 39
11.3. Prioridad de Implementación y Habilitadores (Hoja de Ruta) 39
11.4. Impacto Cuantificable y Riesgos Mitigados 40
11.4.1.Retorno de Inversión (ROI) Alto: 40
11.4.2. Reducción de Riesgo Institucional: 40
11.4.3. Riesgo Documental: 40
11.4.4.Riesgo Operacional: 40
11.4.5. Riesgo de Corrupción: 40
11.4.6. Beneficio Ciudadano Directo: 41
11.5. Recomendación Final: 41

1. RESUMEN EJECUTIVO

Este informe presenta un análisis detallado del estado actual de los procesos de recaudo y gestión tributaria del Municipio de Ipiales, exponiendo las principales falencias operativas y tecnológicas que afectan el cumplimiento de metas fiscales, la eficiencia administrativa y la transparencia institucional.

Asimismo, se plantean casos de uso basados en Inteligencia Artificial (IA) como estrategia para transformar la gestión hacendaria sin necesidad de reemplazar el sistema actual, priorizando soluciones de alta viabilidad técnica y financiera, sostenibles para el contexto local.

2. CONTEXTO MUNICIPAL

Según el Plan de Desarrollo Municipal 2024–2027 “Ipiales es de Todos”, la sostenibilidad financiera es uno de los pilares estratégicos para ejecutar proyectos en seguridad, infraestructura, movilidad, educación, salud y demás sectores prioritarios. Sin embargo, la capacidad institucional para gestionar el recaudo se encuentra limitada por:

- Baja digitalización histórica.
- Infraestructura tecnológica insuficiente.
- Procesos manuales dependientes de archivos físicos y hojas de cálculo.
- Falta de integración entre áreas clave (Tesorería, Jurídica, Archivo, Sistemas).
- Cartera morosa significativa y en crecimiento.

En este escenario, la tecnología —y en particular la IA— se constituye en una herramienta estratégica para superar brechas estructurales y optimizar el recaudo.

3. PROBLEMA ACTUAL

El diagnóstico identifica problemas críticos:

3.1. Rezago documental y falta de digitalización:

- Miles de expedientes físicos (predial, acuerdos de pago, procesos coactivos, industria y comercio) no han sido digitalizados ni indexados.
- No existen escáneres de producción, ni personal asignado para la tarea.
- Dependencia de herramientas vulnerables
- El uso de Excel como repositorio temporal: Genera alto riesgo de manipulación y corrupción, no garantiza trazabilidad ni control, fragmenta la información.

3.2. Base de datos tributaria incompleta

El sistema tributario no cuenta con: información histórica consolidada, expedientes electrónicos asociados al contribuyente, estados reales de deuda o acuerdos. Por ello, no es posible generar reportes confiables sobre morosidad, cartera o proyección de recaudo.

3.3. Procesos manuales en cobro coactivo.

La fase jurídica depende de búsqueda manual de documentos, revisión física y digitación caso a caso, generando retrasos que impactan la recuperación de cartera.

3.4. Riesgos de corrupción.

La ausencia de controles digitales y trazabilidad aumenta riesgos de:

- Eliminación de expedientes.
- Manipulación de valores.
- Movimientos sin auditoría.

3.5. Analítica inexistente de datos.

A pesar de contar con la información, no se han detenido a analizarla punto a punto, no tienen dashboards o tablero de mando para revisar tendencias gráficamente.

A pesar de que la institución cuenta con información relevante proveniente de distintos procesos y registros, esta no ha sido analizada de manera sistemática ni estratégica. Actualmente, los datos se almacenan sin un proceso estructurado de interpretación que permita convertirlos en información útil para la toma de decisiones.

No se dispone de dashboards, tableros de control ni herramientas de visualización que faciliten el análisis gráfico de tendencias, comportamientos o patrones. Esto limita la posibilidad de realizar seguimiento oportuno a indicadores clave, evaluar el impacto de las acciones implementadas y anticipar riesgos o necesidades futuras.

La ausencia de una cultura de analítica de datos dificulta la planificación basada en evidencias, reduce la eficiencia en la gestión institucional y restringe el aprovechamiento del potencial que ofrece la información disponible para la mejora continua de los procesos.

No se ha realizado una revisión periódica de los datos que han sido modificados, ni se cuenta con un análisis sistemático de los registros o logs que permita identificar, documentar y evaluar las razones por las cuales se han efectuado dichas modificaciones. Esta situación

limita la capacidad de verificar la coherencia de la información, detectar errores recurrentes, prevenir modificaciones indebidas y fortalecer los mecanismos de control y transparencia en la gestión de los datos.

4. OBJETIVO DEL INFORME

Proponer un portafolio de casos de uso de Inteligencia Artificial que permita:

- Recuperar la información tributaria rezagada.
- Modernizar el proceso de recaudo.
- Aumentar la eficiencia y reducir costos operativos.
- Fortalecer la transparencia institucional.
- Optimizar el proceso de cobro coactivo.
- Mejorar la experiencia del contribuyente.

Beneficios esperados

- Digitalización total de expedientes en meses, no años.
- Eliminación del uso de Excel en procesos críticos.
- Información consolidada y trazable.
- Incremento del recaudo anual.
- Reducción de cartera morosa.
- Procesos coactivos más ágiles y automáticos.
- Atención al ciudadano más rápida y moderna.

Recomendación clave: NO cambiar las plataformas software actuales.

El problema no es el software del municipio, sino la falta de datos completos, digitales e interoperables.

Cambiar de sistema generaría:

- Alto costo financiero.
- Alto impacto operativo.
- Curva de adaptación riesgosa.

La solución estratégica es construir una capa de IA e interoperabilidad sobre el sistema actual mediante APIs.

5. METODOLOGÍA UTILIZADA

Breve descripción de cómo se construyó el análisis:

5.1. Fuentes de información

El análisis se elaboró mediante:

- Observación directa de procesos de Tesorería, Recaudo, Archivo y Cobro Coactivo.
- Entrevistas técnicas y funcionales con personal administrativo.
- Revisión del Plan de Desarrollo Municipal 2024–2027, especialmente sus metas de sostenibilidad financiera.
- Revisión del PETI 2024, que evidencia brechas en digitalización, infraestructura e integración.
- Análisis de formatos institucionales (resoluciones, mandamientos de pago, acuerdos de pago).
- Lineamientos de Gobierno Digital del MinTIC.

5.2. Criterios de evaluación de casos de uso

Cada posible caso se evalúa con:

Criterio	Descripción
Impacto ciudadano	Mejora directa para usuarios y comunidades
Viabilidad técnica	Infraestructura disponible y datos accesibles
Disponibilidad de datos	Calidad, volumen y actualidad
Viabilidad financiera	Costos de implementación
Riesgo normativo	Protección de datos, ética, sesgos
Alineación estratégica	Ajuste con políticas públicas y PGD

6. CONTEXTO MUNICIPAL

El Municipio de Ipiales, según el *Plan de Desarrollo Municipal*, enmarca su estrategia en seis líneas esenciales:

- Seguridad y convivencia
- Salud
- Gestión social e inclusión
- Productividad, ciencia y tecnología

- Infraestructura y movilidad
- Gestión ambiental

Sin embargo, todas dependen del fortalecimiento institucional y la sostenibilidad financiera, objetivo transversal que involucra directamente a la Secretaría de Hacienda.

6.1. Brechas tecnológicas relevantes para Hacienda

Del PETI y el diagnóstico se identifican:

- Procesos manuales que dependen de papel.
- Falta de integración entre sistemas de Hacienda, Archivo y Jurídica.
- Bajo nivel de analítica y ausencia de información histórica estructurada.
- Infraestructura tecnológica insuficiente.
- Trabajos en Excel sin trazabilidad.
- Ausencia de automatización en validaciones, cálculos tributarios y gestión coactiva.
- Falta de digitalización de expedientes clave.

Estas brechas impiden alcanzar metas como:

- Actualización catastral.
- Reducción de morosidad.
- Mejora del recaudo.

6.2. Oportunidades para IA en Hacienda

- Digitalización masiva inteligente.
- Analítica predictiva de recaudo.
- Automatización de procesos coactivos.
- Chatbots tributarios.
- Detección de anomalías y corrupción.
- Consolidación de expedientes electrónicos automatizados.

7. IDENTIFICACIÓN DE CASOS DE USO DE IA

7.1. CASO DE USO 1: Digitalización Inteligente de Expedientes Tributarios (OCR + IA)

7.1.1. Descripción del Caso

Actualmente, la Alcaldía tiene miles de carpetas con documentos físicos sobre predial, acuerdos de pago y cobro coactivo.
Estos papeles se pueden perder, dañar o manipular, y buscar información en ellos toma horas o días.

Este caso de uso permite:

- Convertir todos esos documentos en expedientes digitales organizados
- Extraer la información automáticamente, sin necesidad de digitar
- Tener todo guardado con seguridad y respaldo
- Encontrar un documento en segundos
- Eliminar el riesgo de pérdida, daño o manipulación

Es como pasar de un archivo físico desordenado a un repositorio digital inteligente que encuentra todo por sí solo.

7.1.2. Dependencia Responsable

- Secretaría Hacienda

7.1.3. Situación Actual

- Miles de expedientes sin digitalizar.
- Documentos dispersos y vulnerables.
- Información parcial e inconsistente.
- Uso de Excel como fuente temporal

7.1.4. Propuesta Basada en IA

- Digitalización masiva con un sistema automatizado.
- Clasificación inteligente de cada documento.
- Extracción automática de datos tributarios.
- Integración total con el sistema actual.

7.1.5. Requerimientos Técnicos

Recurso	Descripción
Datos	Miles de expedientes físicos rezagados (predial, acuerdos de pago, coactivo); se requiere definir el **Dataset** de entrenamiento (muestras de documentos) para los modelos de clasificación y extracción de datos.
Infraestructura	Adquisición de **Escáneres de Producción** de alto volumen y capacidad de almacenamiento masivo (*Data Lake* o Repositorio Documental en la Nube) para preservar los archivos PDF/A.

Recurso	Descripción
Software	Servicios Cognitivos de la Nube para OCR/ICR (ej. Google Cloud Document AI, Azure Document Intelligence) o desarrollo en Python con librerías para OCR y procesamiento de lenguaje (NLP).
Talento	Personal de archivo para el escaneo masivo, un Ingeniero de Datos o Desarrollador Python para la integración de OCR/NLP y un Arquitecto de Soluciones Cloud.
Integración	Desarrollo de **APIs RESTful** para consulta de expedientes digitales, permitiendo que el sistema tributario actual acceda a los documentos y la metadata extraída en tiempo real.

7.1.6. Beneficios Esperados

- Datos estructurados listos para la analítica.
- Reducción del error humano.
- Mayor velocidad de procesamiento.
- Integridad documental.
- Expedientes disponibles en segundos.
- Menos tareas manuales y repetitivas.
- Eliminación del riesgo de pérdida de documentos.
- Mayor transparencia y control.

7.1.7. Implementación técnica propuesta

Etapa	Paso a Paso de Implementación	Tecnologías Clave	Consideraciones Adicionales
1. Preparación	**a.** Adquirir escáneres de producción de alto volumen con alimentador automático. **b.** Realizar un inventario y clasificación inicial de los tipos de documentos a	Escáneres de Producción de Alto Rendimiento (ej. Kodak Alaris, Fujitsu).	Definir la política de **preservación digital** y la validez legal de los documentos escaneados.

	digitalizar (ej. Predial, Coactivo).		
2. Captura y Proceso	**a.** Escaneo masivo de los expedientes físicos a formato PDF/A (preservación). **b.** Aplicar OCR/ICR para convertir imágenes a texto y estructurar datos.	**Servicios Cognitivos:** Google Cloud Document AI o Azure Document Intelligence (Form Recognizer). **Librerías de Desarrollo:** Python con Tesseract (para OCR básico) o SDKs de proveedores de nube.	La **calidad del escaneo** es crítica para la precisión del OCR; se requiere personal capacitado.
3. Inteligencia y Extracción	**a.** Desarrollar modelos de Machine Learning (clasificación) para identificar automáticamente el tipo de documento (factura, acuerdo, resolución). **b.** Usar Modelos de Lenguaje (NLP) para extraer campos clave: Cédula, Monto, Fecha, Dirección, Identificador Predial.	**Plataforma ML:** Azure Machine Learning o Google Vertex AI. **Lenguaje:** Python.	Establecer un flujo de **revisión humana** para los documentos con baja confianza de extracción (ej. 10%).

4. Integración y Almacenamiento	**a.** Almacenar los documentos digitalizados y su metadata extraída en el Repositorio Documental. **b.** Crear **APIs RESTful** para que el sistema tributario actual pueda consultar los expedientes digitales en tiempo real.	**Almacenamiento:** Repositorio Documental con APIs (ej. Alfresco, Nuxeo, o un Data Lake en Google Cloud Storage/Azure Blob Storage). **Integración:** APIs RESTful.	La integración con el sistema actual es la clave para la **interoperabilidad** y la eliminación del uso de Excel.

7.1.8. Matriz de Viabilidad del Caso

Criterio	Puntaje 1–5	Observaciones
Viabilidad técnica	5	Las tecnologías (OCR, Document AI, APIs) son maduras y probadas. La implementación se centra en la adquisición de hardware (escáneres) y la integración con el repositorio documental.
Viabilidad financiera	4	Requiere una inversión inicial significativa en escáneres de alto rendimiento y servicios cognitivos en la nube. El retorno es alto por la reducción de riesgo y la eficiencia operativa a largo plazo.
Impacto ciudadano	5	Impacto directo: acceso a expedientes en segundos, mayor transparencia y eliminación del riesgo de pérdida o manipulación de documentos personales.
Impacto institucional	5	**Fundamental.** Crea la base de datos estructurada y digital que es pre-requisito para todos los demás casos de IA.

		Elimina el riesgo de manipulación y la dependencia de Excel.
Riesgo ético / legal	4	Riesgo bajo si se define correctamente la política de **preservación digital** y la seguridad en el acceso a los datos. Cumplimiento de la Ley 1581 (Protección de Datos Personales).
Disponibilidad de datos	3	Los datos existen (miles de documentos físicos), pero están **inconsistentes y no estructurados**. Este proyecto es el que transforma el dato no disponible en dato estructurado.
Nivel de Prioridad	**ALTA**	Es la base habilitante para el resto de la modernización.

7.2. CASO DE USO 2: Motor Predictivo de Recaudo y Riesgo de Morosidad

7.2.1. Descripción del Caso

Se implementa un modelo de Machine Learning que:

- Analiza el historial de pagos de cada contribuyente
- Detecta patrones de comportamiento (puntual, ocasional, moroso crónico)
- Calcula probabilidades de pago
- Proyecta el recaudo mensual esperado
- Clasifica la cartera en niveles de riesgo

Los algoritmos pueden usar variables como:

- Tipo de predio
- Zona geográfica
- Histórico de pagos
- Acuerdos vigentes
- Morosidad comparable
- Comportamiento similar de otros contribuyentes

Esto permite generar:

- Alertas tempranas
- Prioridades de cobro

- Metas realistas de recaudo
- Estrategias diferenciadas por tipo de contribuyente

7.2.2. Dependencia Responsable

- Secretaría Hacienda

7.2.3. Situación Actual

- No existe capacidad de predicción.
- No se priorizan contribuyentes según riesgo.
- No hay alertas tempranas.

7.2.4. Propuesta Basada en IA

- Sistema que calcula riesgo de no pago.
- Priorización inteligente de cartera.
- Proyección mensual automatizada.

7.2.5. Requerimientos Técnicos

Recurso	Descripción
Datos	**Dataset histórico consolidado y limpio** de pagos y morosidad de contribuyentes (resultado del Caso #3). Se requiere volumen y antigüedad de datos para la **Ingeniería de Características**.
Infraestructura	Un *Data Warehouse* (ej. PostgreSQL robusto o BigQuery) para el pre-procesamiento de datos y Servidores en la Nube con Docker/Kubernetes para el despliegue de la API de Inferencia del modelo (entorno MLOps).
Software	Plataforma de Machine Learning (ej. Python con Scikit-learn, o Google Vertex AI/Azure ML) para entrenamiento y herramientas de **Visualización** (Power BI, Looker Studio) para crear *Dashboards* de riesgo.
Talento	Un **Científico de Datos** o Analista de Datos avanzado para el modelado y la Ingeniería de Características, y un Ingeniero MLOps para el despliegue del modelo en producción.
Integración	Una **API Gateway** o Microservicio para exponer el *score* de riesgo del contribuyente. Integración en el sistema de gestión de cartera para el accionamiento de alertas y prioridades.

7.2.6. Beneficios Esperados

- Modelos adaptativos.
- Optimización de cartera.
- Decisiones basadas en datos reales.

- Más recaudo.
- Menos morosidad.
- Gestión más rápida y efectiva.

7.2.7. Implementación técnica propuesta

Etapa	Paso a Paso de Implementación	Tecnologías Clave	Consideraciones Adicionales
1. Datos y Preparación	**a.** Consolidar la data histórica de pagos depurada (resultado del Caso #3). **b.** Realizar **Ingeniería de Características**: crear variables como "antigüedad de la mora", "frecuencia de acuerdos de pago", "zona catastral".	**Base de Datos:** Data Warehouse (ej. PostgreSQL, BigQuery, Snowflake) para análisis de gran volumen.	Asegurar la disponibilidad de la **información histórica consolidada** para el entrenamiento.
2. Modelado	**a.** Seleccionar y entrenar un algoritmo de clasificación (ej. Regresión Logística, Random Forest o Gradient Boosting) para predecir la probabilidad de mora. **b.** Clasificar a cada contribuyente en niveles de riesgo (Bajo, Medio, Alto).	**Plataforma ML:** Python (Scikit-learn) o entornos MLOps (Azure ML, Google Vertex AI).	El **balance de clases** (contribuyentes morosos vs. no morosos) debe ser manejado adecuadamente para evitar sesgos.
3. Despliegue y MLOps	**a.** Desplegar el modelo entrenado en un entorno de producción como un servicio de **API de Inferencia**. **b.** Automatizar la ejecución diaria/semanal para	**MLOps:** Servidores en la Nube con Docker/Kubernetes para el despliegue de modelos.	Implementar un sistema de **monitoreo de deriva del modelo** (model drift) para garantizar su

	generar puntajes de riesgo actualizados.	**Integración:** API Gateway.	precisión a lo largo del tiempo.
4. Accionamiento y Monitoreo	**a.** Integrar el *score* de riesgo en el sistema de gestión de cartera. **b.** Crear *Dashboards* para que los gestores visualicen las **prioridades de cobro** (ej. Top 100 de Alto Riesgo).	**Visualización:** Power BI o Google Looker Studio. **Alertas:** Sistemas de Notificación automatizada.	El resultado del modelo debe traducirse en **estrategias diferenciadas** de comunicación y cobro (ej. recordatorios vía Chatbot - Caso #4).

7.2.8. Matriz de Viabilidad del Caso

Criterio	Puntaje 1–5	Observaciones
Viabilidad técnica	4	Las técnicas de Machine Learning (Clasificación/Regresión Logística, Random Forest) son estándar. El desafío técnico principal es la correcta **Ingeniería de Características**.
Viabilidad financiera	5	El retorno de inversión es el más alto del portafolio, ya que impacta directamente el **incremento del recaudo anual** y optimiza la labor de los gestores de cartera.
Impacto ciudadano	3	Impacto indirecto: permite estrategias de cobro más justas y diferenciadas, mejorando la comunicación a través de canales (e.g., *Chatbot*).

Impacto institucional	5	Permite tomar **decisiones basadas en datos** para la gestión de cartera, priorizando esfuerzos y estableciendo metas realistas de recaudo. Genera alertas tempranas.
Riesgo ético / legal	3	Riesgo medio por la posibilidad de **sesgos algorítmicos** que lleven a la priorización o discriminación involuntaria de ciertos grupos. Requiere monitoreo de deriva y balance de clases.
Disponibilidad de datos	3	**Dependencia clave:** No puede implementarse sin la **data histórica consolidada y depurada**, siendo un caso dependiente del éxito del Caso #3.
Nivel de Prioridad	**ALTA**	Impacto directo y cuantificable en la meta de **sostenibilidad financiera**.

7.3 CASO DE USO 3: Validación Automática de Datos Tributarios (Barrido de base actual)

7.3.1. Descripción del Caso

IA revisa la base de datos y detecta:

- Registros duplicados
- Expedientes faltantes
- Inconsistencias entre documentos
- Alteración de valores
- Fechas contradictorias
- Manipulación intencional o accidental
- Descuentos o exoneraciones atípicas

Utiliza técnicas como:

- Detección de anomalías
- Reglas de integridad
- Modelos estadísticos

El sistema revisa automáticamente toda la base de datos y detecta:

- Errores
- Datos repetidos

- Valores que no cuadran
- Movimientos sospechosos

7.3.2. Dependencia Responsable

- Secretaría Hacienda

7.3.3. Situación Actual

- Tenemos una base de datos que no ha sido alimentada correctamente
- Hay información que han manipulado previamente.

7.3.4. Propuesta Basada en IA

- Base de datos depurada
- Optimización de data para un sistema confiable

7.3.5. Requerimientos Técnicos

Recurso	Descripción
Datos	La **Base de Datos Tributaria actual**, sucia e inconsistente, es el insumo y el objeto a optimizar. Se requiere documentación formal de las **Reglas de Integridad** (unicidad, consistencia, validez).
Infraestructura	Servidores dedicados o recursos Cloud para la ejecución de scripts de limpieza y orquestación de *workflows* (ej. Apache Airflow), necesarios para el barrido nocturno y el monitoreo continuo.
Software	Herramientas **ETL** (*Extraction, Transformation, Load*) de código abierto (ej. Pentaho, scripts avanzados de Python con Pandas) y algoritmos de **Detección de Anomalías** (ej. Isolation Forest).
Talento	Un **Ingeniero de Datos** para el desarrollo de scripts de limpieza y un Analista Funcional/Jurídico para definir las reglas de negocio y aprobar las correcciones iniciales.
Integración	Acceso seguro a la Base de Datos tributaria (conexión SQL). Se requiere el desarrollo de **APIs** para la **validación en tiempo real** de nuevos registros de datos, evitando errores futuros.

7.3.6. Beneficios Esperados

- Auditoría automatizada.
- Optimización de datos

7.3.7. Implementación técnica propuesta:

Etapa	Paso a Paso de Implementación	Tecnologías Clave	Consideraciones Adicionales
1. Definición de Calidad	a. Documentar formalmente las Reglas de Integridad (Validez, Consistencia, Unicidad) para cada campo crítico. b. Identificar fuentes de datos primarias y secundarias.	**Herramientas de Documentación:** Confluence, Notion. SQL: Consultas para establecer el *baseline* de calidad.	La participación del equipo Jurídico y de Tesorería es esencial para definir las reglas de negocio.
2. Desarrollo de la Limpieza Inicial	a. Desarrollar scripts para la Deduplicación (identificación de registros repetidos de contribuyentes). b. Implementar algoritmos de Detección de Anomalías para identificar valores fuera de rango o cambios sospechosos.	**Data Quality Tools:** Herramientas ETL (ej. Talend, Pentaho) o scripts avanzados de Python (Pandas). Algoritmos: Isolation Forest o modelos estadísticos para detección de *outliers*.	El barrido inicial debe ejecutarse en un entorno de prueba para evitar afectar la operación, y requerirá intervención humana para la aprobación de correcciones.
3. Automatización	a. Configurar el proceso de validación para que se ejecute de forma programada (ej. cada noche). b. Implementar validaciones en tiempo real en las APIs de	**Orquestación:** Apache Airflow o servicios de *Workflow* en la nube (ej. Azure Data Factory). Integración: APIs para validación de entrada.	El objetivo es pasar de un "barrido" a un monitoreo continuo de la calidad de los datos.

	ingreso de datos para prevenir errores futuros.		
4. Monitoreo y Reporte	a. Generar *Dashboards* de Calidad de Datos que muestren la evolución de los errores, los duplicados y las anomalías detectadas. b. Enviar alertas automáticas a los responsables de área.	**Visualización:** Power BI, Looker Studio. Notificación: Email o sistemas de mensajería integrados.	El éxito se mide en la reducción del porcentaje de datos inconsistentes con cada ciclo de limpieza.

7.3.8. Matriz de Viabilidad del Caso

Criterio	Puntaje 1–5	Observaciones
Viabilidad técnica	4	Se utilizan herramientas ETL (Extracción, Transformación y Carga) y algoritmos de detección de anomalías estándar. El desafío radica en la definición de **reglas de negocio** con las áreas funcionales.
Viabilidad financiera	5	Inversión media en software de calidad de datos y desarrollo. El beneficio principal es la **prevención de pérdidas** y la reducción del riesgo de manipulación de valores, lo cual previene grandes costos futuros.
Impacto ciudadano	2	Impacto muy indirecto, pero crucial para garantizar que las obligaciones tributarias (facturas, acuerdos) sean calculadas de forma **precisa y justa**.

Impacto institucional	5	**Base fundamental.** Limpia y estandariza la base de datos principal, convirtiéndose en un **monitoreo continuo** de la calidad de la data. Habilita los Casos #2 y #5.
Riesgo ético / legal	5	Riesgo bajo (5 es bajo riesgo) ya que su propósito es reducir el riesgo de manipulación, fraude y errores en los cálculos, fortaleciendo la **transparencia y la auditoría**.
Disponibilidad de datos	4	Los datos existen en la base actual, pero el reto es su **calidad e inconsistencia**. La data es el insumo y el resultado a optimizar del mismo caso de uso.
Nivel de Prioridad	**ALTA**	Pre-requisito de la mayoría de los proyectos de IA.

7.4 CASO DE USO 4: Chatbot Tributario con IA

7.4.1. Descripción del Caso

Un chatbot integrado con APIs del sistema permite:

- Consultar predial
- Descargar facturas
- Acceder a certificados
- Revisar acuerdos
- Generar recordatorios personalizados
- Registrar PQRS

Usa NLP (procesamiento natural del lenguaje).

Es un asistente virtual que responde preguntas y ayuda a los ciudadanos las 24 horas, como:

- "¿Cuánto debo de predial?"
- "Descargue mi factura"
- "Estoy en acuerdo de pago, ¿cuándo vence mi cuota?"

Reduce filas, tiempos y llamadas.

7.4.2. Dependencia Responsable

- Secretaría Hacienda

7.4.3. Situación Actual

- Tenemos 2 fuentes de consulta, la página web de tributaria y la ventanilla de atención al usuario, las cuales cada una tiene limitaciones.
- Desconocimiento de canales de información

7.4.4. Propuesta Basada en IA

- Atención al público de manera automatizada, y con intervención humana cuando sea necesaria

7.4.5. Requerimientos Técnicos

Recurso	Descripción
Datos	Base de **Conocimiento (FAQ)** extensa sobre temas tributarios y acceso a la data transaccional (deuda, acuerdos) a través de APIs del sistema actual. Se requiere un **Lenguaje (Español)** entrenado en modismos locales.
Infraestructura	Plataforma de NLU/NLP en la nube (ej. Google Dialogflow CX, Azure Bot Service) para el motor conversacional y un **API Gateway** para gestionar la seguridad de las APIs transaccionales.
Software	Plataforma de desarrollo de Bots con capacidad de **NLU/NLP** y conectores a múltiples canales de atención (Web, WhatsApp, etc.). Se requiere un sistema de ***Hand-off*** para traspaso a agentes humanos.
Talento	Un **Diseñador Conversacional** para mapear flujos de diálogo y un Desarrollador de Bots o Integrador de APIs para conectar el motor NLU/NLP con el sistema tributario.
Integración	Desarrollo de **APIs de Microservicios** seguros (con autenticación OAuth/API Gateway) que permitan al chatbot realizar consultas y generar documentos (facturas) directamente del sistema tributario.

7.4.6. Beneficios Esperados

- Respuestas más oportunas ante el ciudadano
- Mejora el interés del ciudadano en poder interactuar con la alcaldía municipal

7.4.7. Implementación técnica propuesta

Etapa	Paso a Paso de Implementación	Tecnologías Clave	Consideraciones Adicionales

1. Diseño de Conversación	**a.** Mapear los casos de uso más frecuentes y transaccionales (ej. "Consultar Deuda", "Descargar Factura"). **b.** Diseñar la base de conocimiento (FAQ) y los flujos de diálogo (árboles de decisión).	**Plataforma NLU/NLP:** Google Dialogflow CX, Azure Bot Service, o RASA.	El diseño conversacional debe ser natural, claro y debe contemplar variaciones en el lenguaje (modismos locales).
2. Desarrollo del Motor NLP	a. Entrenar el modelo de NLU para reconocer las Intenciones (ej. consultar_deuda) y extraer las Entidades (ej. numero_cedula). b. Programar las respuestas basadas en la base de conocimiento.	**NLP/NLU:** El motor interno de la plataforma seleccionada (Dialogflow, Azure). **Lenguaje:** Español.	Enfocarse inicialmente en alto volumen y baja complejidad para una rápida puesta en marcha.
3. Integración Transaccional	a. Desarrollar o consumir APIs de Microservicios que permitan al chatbot interactuar con el sistema tributario (consultas de estado, generación de facturas). b. Asegurar la autenticación y seguridad de las APIs.	**Integración:** APIs RESTful. **Seguridad:** OAuth, API Gateway.	La seguridad de las APIs que exponen datos personales y transaccionales es una prioridad máxima (Cumplimiento Ley 1581).
4. Despliegue y *Hand-off*	a. Integrar el chatbot en los canales de atención (Web, WhatsApp, etc.). b. Implementar la funcionalidad de traspaso a agente humano cuando el chatbot no pueda resolver la	**Canales:** Integración vía *webhooks* o conectores específicos de la plataforma.	Monitorear las métricas de tasa de resolución por la IA y tiempo de espera en el *hand-off* humano.

	consulta o se trate de una PQRS compleja.		

7.4.8. Matriz de Viabilidad del Caso

Criterio	Puntaje 1–5	Observaciones
Viabilidad técnica	4	Las plataformas NLU/NLP (ej. Dialogflow) son fáciles de implementar, pero la mayor complejidad es la **integración transaccional segura** con el sistema tributario actual vía *APIs de Microservicios*.
Viabilidad financiera	4	Inversión en plataforma conversacional y desarrollo de integración. Genera un rápido ahorro de costos por la **reducción de la carga** en ventanilla y llamadas de atención al ciudadano.
Impacto ciudadano	5	**Alto y directo.** Ofrece atención 24/7, reduce tiempos de espera, facilita la descarga de documentos (facturas) y resuelve dudas básicas de manera inmediata.
Impacto institucional	4	Permite concentrar al personal de atención en casos complejos (*hand-off* a agente humano) y genera una **imagen de modernización** en la interacción con el ciudadano.
Riesgo ético / legal	3	Riesgo medio-alto. La **seguridad de las APIs** es una prioridad máxima (OAuth/API Gateway) para evitar fugas de información personal y transaccional sensible, cumpliendo con la Ley 1581.

Disponibilidad de datos	4	Se basa en una base de conocimiento (FAQ) bien diseñada y en el acceso a datos en tiempo real (consultas de estado) a través de APIs.
Nivel de Prioridad	**MEDIA-ALTA**	Gran impacto ciudadano y eficiencia operativa en atención.

7.5 CASO DE USO 5: Detección de Anomalías y Riesgos de Corrupción (barrido de base de datos)

7.5.1. Descripción del Caso

El sistema monitorea:

- Eliminación de registros
- Cambios drásticos en valores
- Alteraciones de fechas
- Movimientos atípicos

Utiliza modelos de:

- Detección de fraude
- Auditoría algorítmica
- Sensores de integridad documental

La IA detecta automáticamente:

- Si alguien borra un expediente
- Si se altera un valor
- Si se cambia una fecha
- Si se manipula una factura o un acuerdo

7.5.2. Dependencia Responsable

- Secretaría Hacienda

7.5.3. Situación Actual

- No existen mecanismos automáticos de auditoría.
- Eliminación de datos sin trazabilidad.
- Cambios manuales sin registro adecuado.
- Expedientes que desaparecen o se alteran.
- Riesgo latente de corrupción interna.

7.5.4. Propuesta Basada en IA

- Monitoreo continuo y automático.
- Identificación inmediata de acciones sospechosas.
- Alertas cuando se borran registros.
- Alertas por cambios inusuales de valores, fechas o estados.
- Reportes automáticos para control interno y disciplinario.

7.5.5. Requerimientos Técnicos

Recurso	Descripción
Datos	**Logs de Auditoría** inmutables (quién, qué, cuándo) de todas las transacciones críticas del sistema. Este *dataset* es nuevo y debe ser generado. Se requiere data histórica de movimientos normales para definir patrones.
Infraestructura	Implementación de un **Sistema de Logs** centralizado (ej. ELK Stack o Splunk) y Servidores dedicados para el monitoreo continuo y la ejecución de modelos de auditoría algorítmica.
Software	Herramientas de gestión de *logs* (ELK Stack) y desarrollo de modelos de **Detección de Anomalías** (ML, Python) y **Auditoría Algorítmica** para clasificar el riesgo de cada evento.
Talento	Un **Ingeniero de Seguridad/Auditoría** para la implementación de la trazabilidad y un Científico de Datos para el entrenamiento y despliegue de los modelos de detección de fraude.

Integración	Conexión directa y en tiempo real del sistema tributario a la plataforma de *Logs* (vía *stream* o *webhooks*). Integración con el sistema de control interno para el envío automatizado de **alertas**.

7.5.6. Beneficios Esperados

- Respuestas más oportunas ante el ciudadano
- Mejora el interés del ciudadano en poder interactuar con la alcaldía municipal

7.5.7. Implementación técnica propuesta

Etapa	Paso a Paso de Implementación	Tecnologías Clave	Consideraciones Adicionales
1. Implementación de Trazabilidad	**a.** Habilitar un sistema de **Logging y Auditoría** en el sistema tributario para registrar todas las acciones de modificación, borrado o acceso a expedientes. **b.** Almacenar los *logs* en una base de datos inmutable.	**Sistemas de Logs:** ELK Stack (Elasticsearch, Logstash, Kibana) o Splunk. **Base de Datos:** Data Lake/Data Warehouse.	El *logging* debe incluir **quién, qué, cuándo y desde dónde** se realizó la acción para una trazabilidad completa.
2. Definición de Reglas y Patrones	**a.** Codificar reglas de negocio críticas (ej. "Nadie fuera del área Jurídica puede borrar un expediente coactivo"). **b.** Analizar el historial para definir los **patrones de comportamiento**	**SQL** para reglas de negocio. **Estadísticas Descriptivas** para el análisis de patrones normales.	La definición de "anomalía" debe ser un acuerdo entre el área de Sistemas, Jurídica y Control Interno.

	normal del usuario y del sistema.		
3. Desarrollo de Modelos de Detección	**a.** Implementar modelos de **Detección de Anomalías** para identificar acciones que se desvíen del patrón normal (ej. un usuario realizando 100 cambios en valores en una hora). **b.** Desarrollar modelos de *scoring* para asignar un nivel de riesgo a cada acción.	**Algoritmos de ML:** Isolation Forest, Detección de *Outliers* en series de tiempo (Python).	Priorizar la detección de anomalías en **expedientes de alto valor** o en los que implican el **borrado o descuento** de deuda.
4. Alerta y Respuesta	**a.** Configurar el sistema para generar **alertas inmediatas** (email, SMS) a Control Interno cuando se detecte un evento de Alto Riesgo. **b.** Crear un *Dashboard* que muestre las alertas en tiempo real y la trazabilidad completa del evento.	**Visualización:** Power BI/Looker Studio para el Dashboard de Auditoría. **Notificación:** Sistemas de Alerta (ej. PagerDuty, SendGrid).	Se debe establecer un **protocolo de acción** para cada nivel de alerta (ej. bloqueo temporal del usuario, inicio de investigación disciplinaria).

7.5.8. Matriz de Viabilidad del Caso

Criterio	Puntaje 1–5	Observaciones
Viabilidad técnica	3	La implementación es compleja porque depende de la creación y habilitación de un sistema de **Logging y Auditoría inmutable** en el sistema tributario (Etapa 1), que actualmente no existe (AS-IS).
Viabilidad financiera	5	Inversión en infraestructura de *logs* (ELK/Splunk) y desarrollo de modelos. Previene **pérdidas significativas** por fraude interno, siendo una inversión de alto valor estratégico.
Impacto ciudadano	2	Impacto muy indirecto, aunque es fundamental para la **confianza institucional** y la transparencia de la gestión de la Secretaría de Hacienda.
Impacto institucional	5	**Crítico.** Aborda directamente el problema de los **Riesgos de Corrupción** y la ausencia de trazabilidad, generando un monitoreo constante y alertas inmediatas a Control Interno.
Riesgo ético / legal	5	Riesgo muy bajo (5 es bajo riesgo). Es una herramienta de auditoría y control que **mitiga** los riesgos legales y disciplinarios al identificar comportamientos anómalos.
Disponibilidad de datos	2	La principal limitación. Los **logs de auditoría** necesarios (quién, qué, cuándo, desde dónde) **no existen** en el sistema actual de forma estructurada, y deben crearse desde cero.
Nivel de Prioridad	**ALTA**	Aborda un riesgo institucional crítico (corrupción).

7.6 CASO DE USO 6: Validación Automática de Procesos de Gestión Tributaria (IA + BPM)

7.6.1. Descripción del Caso

Los procesos de gestión tributaria del Municipio —determinación, fiscalización, liquidación y cobro— se ejecutan actualmente de forma manual, fragmentada y con bajo nivel de integración tecnológica. Las actuaciones (emplazamientos, aforos, sanciones, mandamientos de pago, medidas cautelares, cierres de expedientes) se realizan sin flujos uniformes, sin trazabilidad y con alta dependencia del trabajo humano, lo que genera retrasos, errores y vulnerabilidad frente a manipulación documental.

El caso propone implementar un **Sistema Inteligente de Validación Automática** que:

- Integra analítica de datos y modelos de IA para identificar contribuyentes omisos o inconsistentes.
- Ejecuta automáticamente las actuaciones administrativas con sus plantillas, firmas y notificaciones.
- Controla los tiempos procesales y genera alertas de vencimiento.
- Crea y administra expedientes digitales con trazabilidad completa.
- Unifica el flujo operativo del proceso tributario con BPM + IA.

Esta solución transforma por completo la operación tributaria, reduciendo tiempos, aumentando la eficiencia y elevando la trazabilidad y transparencia institucional.

Ejemplo de Flujo Automatizado – Programa de Omisos

1. **Analítica Automática de Datos**
   - Cruce masivo de bases: predial, ICA, RUT, declaraciones previas, DIAN, Cámara de Comercio.
   - Identificación de omisos, inconsistencias o valores atípicos.
   - Detección de duplicados y errores.
2. **Generación del Expediente Digital**
   - Creación automática del expediente virtual.
   - Carga de soportes, cronología del proceso y trazabilidad completa.
3. **Actuaciones Automáticas**
   El sistema genera y notifica secuencialmente:
   - Invitación a declarar
   - Emplazamiento
   - Resolución sanción por no declarar
   - Liquidación oficial de aforo
   - Traslado a cobro persuasivo
   - Mandamiento de pago
   - Medidas cautelares
   - Liquidación del crédito
   - Cierre del expediente

Cada actuación queda registrada con:

- Fecha – responsable – fundamento normativo – estado – notificación – firma (si aplica)

Funciones de Validación Automática

- Verificación de **consistencia de datos**
- Detección de **contribuyentes duplicados**
- Identificación de **expedientes faltantes**
- Alertas por **cambios sospechosos**
- Identificación de **movimientos atípicos** mediante ML
- Control automático de **términos procesales**
- Monitoreo continuo de **riesgos operativos y normativos**

7.6.2. Dependencia Responsable

Secretaría de Hacienda Municipal, con apoyo de:

- Oficina Jurídica
- Oficina de Sistemas
- Oficina de Control Interno

7.6.3. Situación Actual

- Etapas del proceso tributario sin integración.
- Documentos elaborados manualmente y sin estandarización.
- Baja trazabilidad (no existe línea de tiempo por expediente).
- Demoras en notificaciones y actuaciones.
- Dependencia excesiva de archivos físicos y Excel.
- Vulnerabilidad ante errores y manipulación.

7.6.4. Propuesta Basada en IA

- Motor central de reglas y flujos automáticos (BPM).
- Modelos de IA para segmentación, detección de omisos y análisis de riesgo.
- Sistema de generación masiva automatizada de actuaciones.
- Expediente digital con **trazabilidad completa**.
- Control de términos y alertas inteligentes.
- Notificación electrónica integrada.
- Integración por APIs con el sistema tributario.

7.6.5. Requerimientos Técnicos

Recurso	Descripción
Datos	Información tributaria consolidada, histórica y depurada. Requiere integración con DIAN, Cámara de Comercio y catastro.

Infraestructura	Servidores cloud para flujos BPM + motores IA; repositorio documental; servicios de notificación electrónica.
Software	BPM (Bizagi, Camunda, PowerAutomate), IA para clasificación y detección de anomalías, APIs REST.
Talento	Arquitecto de soluciones, ingeniero de datos, científico de datos, analista tributario, equipo jurídico.
Integración	API Gateway, conectores al sistema tributario y repositorio documental digital.

7.6.6. Beneficios Esperados

- Reducción de tiempos de trámite hasta 70%.
- Disminución de errores manuales y reprocesos.
- Trazabilidad total del expediente tributario.
- Mayor seguridad y menor riesgo de manipulación.
- Mejora significativa del recaudo.
- Unificación del proceso tributario para los próximos años.

7.6.7. Implementación técnica propuesta

Etapa	Actividades	Tecnologías
1. Preparación	Levantamiento de procesos, reglas, plantillas normativas, estructura de expedientes	BPM + Oficina Jurídica
2. Clasificación e Inteligencia	Segmentación de contribuyentes, detección de omisos, ML de anomalías	Python, ML, SQL
3. Automatización	Diseño del flujo BPM completo, generación automática de actos, integración	Camunda / Bizagi / PowerAutomate + APIs
4. Despliegue	Notificación electrónica, expediente digital, auditoría algorítmica	API Gateway + Firma Digital
5. Monitoreo	Alertas, trazabilidad, control de términos	Dashboard + Power BI

8 PROPUESTA DE CASOS DE USO RECOMENDADOS (LISTA BASE PARA IPIALES)

A continuación, se presentan los casos de uso más relevantes y de **Alta Prioridad**, enfocados en los pilares de **Recaudo, Cartera y Transparencia**, que constituyen la base

de la modernización institucional. Estos casos se detallan en la sección 4 y son habilitantes entre sí.

8.1. Digitalización Inteligente de Expedientes Tributarios (OCR + IA) (Caso #1)

Propósito: Base habilitante fundamental para transformar miles de documentos físicos rezagados en expedientes digitales organizados y con datos estructurados, eliminando el uso de archivos manuales.

Impacto: Permite la creación del *Dataset* limpio requerido para los casos de Analítica y Predicción.

8.2. Validación Automática de Datos Tributarios (Barrido de base actual) (Caso #3)

Propósito: Pre-requisito esencial para la confiabilidad de la data. Limpia y estandariza la base de datos principal, detectando duplicados, inconsistencias y anomalías, y estableciendo un monitoreo continuo de la calidad de los datos.

Impacto: Garantiza que la información utilizada en todos los demás proyectos y en los cálculos tributarios sea precisa y auditable.

8.3. Motor Predictivo de Recaudo y Riesgo de Morosidad (Caso #2)

Propósito: Estrategia clave para optimizar la gestión de cartera. Utiliza *Machine Learning* para priorizar los esfuerzos de cobro, calcular la probabilidad de morosidad y proyectar el recaudo mensual.

Impacto: Impacto directo y cuantificable en el incremento del recaudo anual y la eficiencia de los gestores de cartera.

8.4. Detección de Anomalías y Riesgos de Corrupción (Caso #5)

Propósito: Implementa un sistema de auditoría continua y algorítmica para monitorear transacciones y cambios de datos, identificando cualquier desviación del comportamiento normal (fraude, manipulación, errores).

Impacto: Fortalece la transparencia institucional, genera trazabilidad completa de los movimientos y previene grandes pérdidas financieras por fraude interno.

8.5. Chatbot Tributario con IA (Caso #4)

Propósito: Mejora directa y masiva en la experiencia del contribuyente. Ofrece atención 24/7 para consultas de deuda, descarga de facturas y revisión de acuerdos, liberando al

personal de atención en ventanilla para casos complejos.

Impacto: Reduce filas, minimiza el desconocimiento de canales de información y mejora la interacción del ciudadano con la Alcaldía

9 HOJA DE RUTA DE IMPLEMENTACIÓN (ROADMAP)

La estrategia se basa en la ejecución secuencial de proyectos habilitadores (*Foundation Layer*) antes de abordar los proyectos de alto impacto (*Value Layer*).

FASE 0: Preparación y Habilitación
(Duración Estimada: 2 Meses)

Esta fase se enfoca en establecer los pre-requisitos técnicos y de negocio para asegurar el éxito de los proyectos de datos y digitalización.

Actividades Clave	Actividades de Negocio	Entregables Principales
Adquisición de Infraestructura: Compra de escáneres de producción (Caso #1) y servicios *cloud* para almacenamiento (*Data Lake*).	**Documentación de Reglas:** Definición formal y exhaustiva de las Reglas de Integridad y negocio para la limpieza de la base de datos (Caso #3).	Hardware adquirido y configurado .
Arquitectura Base: Configuración inicial del Repositorio Documental con APIs (Caso #1) y el *Data Warehouse* central (Caso #2).	**Inventario Documental:** Clasificación inicial de los tipos de expedientes físicos para el escaneo masivo (Caso #1).	Reglas de Negocio aprobadas.

FASE 1: Capa de Fundación (Digitalización y Calidad de Datos)
(Duración Estimada: 4 Meses)

Fase intensiva de desarrollo donde se construye la base de datos limpia y digital sin la cual el resto de casos de IA no son posibles.

Caso de Uso	Actividades	Tareas Críticas	Objetivo Primario

#1 Digitalización	Desarrollo del *pipeline* de OCR/ICR y clasificación ML para la extracción de metadatos de los documentos.	Escaneo Masivo: Ejecución del escaneo físico de expedientes (la tarea que consume más tiempo del personal de archivo/expertos).	Expedientes 100% digitales y con metadatos extraídos.
#3 Validación Automática	Desarrollo de scripts ETL y de Deduplicación/Limpieza basados en las Reglas de Integridad definidas en Fase 0.	Aprobación de Correcciones: Revisión y aprobación manual del barrido inicial de datos y anomalías para validar los *scripts*.	Base de datos tributaria depurada, consistente y lista para ser analizada.

FASE 2: Motores de Valor (Analítica y Seguridad)
(Duración Estimada: 3 Meses)

Una vez que los datos están listos (Fase 1), los desarrolladores se enfocan en los modelos de alto valor institucional y riesgo.

Caso de Uso	Actividades	Tareas Críticas	Objetivo Primario
#2 Motor Predictivo	Ingeniería de Características: Creación de variables a partir de la data histórica limpia. Entrenamiento, validación y *tuning* del Modelo de Clasificación de riesgo.	Input de Dominio: Asesoría a los desarrolladores en la selección de variables y la interpretación inicial de los resultados del modelo.	Modelo de Riesgo entrenado y validado con un API de Inferencia.
#5 Detección de Anomalías	Implementación de la Infraestructura de *Logging* y Trazabilidad Inmutable (ELK/Splunk). Desarrollo de modelos ML para detectar *outliers* y anomalías.	Protocolos de Alerta: Definición de los umbrales de riesgo y los protocolos de acción para cada nivel de alerta (Control Interno/Jurídica).	Sistema de Auditoría Algorítmica en funcionamiento para prevención de fraude.

FASE 3: Despliegue, MLOps y Experiencia Ciudadana (Duración Estimada: 3 Meses)

Fase final centrada en la puesta en producción de los modelos de ML y la implementación del canal de atención ciudadana (Chatbot).

Caso de Uso	Tareas Críticas (2 Devs)	Tareas Críticas (5 Expertos Tributarios)	Objetivo Primario
#4 Chatbot Tributario	Integración de APIs: Desarrollo de Microservicios Seguros para consulta de deuda y descarga de facturas. Despliegue del motor NLU/NLP (ej. Dialogflow).	Diseño Conversacional: Elaboración exhaustiva de la Base de Conocimiento (FAQ) y los flujos de diálogo. Pruebas de usuario y validación de las respuestas transaccionales.	Canal de Atención 24/7 completamente operativo y con capacidad de *Hand-off* a agentes humanos.
MLOps y *Go-Live*	Despliegue de los modelos #2 y #5 en un entorno MLOps. Creación del API Gateway para asegurar la exposición de todos los servicios. Integración de *Dashboards* (Visualización).	Pruebas Integrales (*End-to-End*): Validación de que los sistemas se comunican correctamente y que las acciones automáticas se ejecutan según las reglas (cobro, alerta).	Sistemas de IA en producción, monitoreados y generando valor y transparencia.

10 RIESGOS Y CONSIDERACIONES ÉTICAS

La implementación de sistemas de Inteligencia Artificial en la gestión tributaria exige un marco riguroso de gestión de riesgos éticos, legales y de seguridad, dado que los proyectos operan con datos sensibles y procesos de decisión críticos.Sesgos de Datos

10.1. Sesgo algorítmico

El riesgo más significativo en la aplicación de modelos predictivos reside en el Sesgo Algorítmico, particularmente en el Caso #2 (Motor Predictivo de Recaudo). Si el *dataset* histórico utilizado para el entrenamiento contiene un desequilibrio o refleja prácticas discriminatorias pasadas, el modelo podría priorizar o penalizar involuntariamente a contribuyentes de ciertos grupos demográficos o zonas geográficas. La mitigación de este riesgo es fundamental y requiere un monitoreo constante de la deriva del modelo (*Model*

Drift) y un esfuerzo consciente por asegurar el balance de clases (contribuyentes morosos vs. no morosos) y la equidad en las variables de entrada durante la etapa de Ingeniería de Características. La calidad de los datos generada por los Casos #1 y #3 es la primera línea de defensa contra este sesgo.Protección de Datos Personales

La información tributaria gestionada incluye datos personales sensibles (cédula, dirección, estado de cuenta), por lo que la Protección de Datos Personales es una prioridad absoluta y transversal a todo el portafolio. Casos como el Chatbot (Caso #4) y la Detección de Anomalías (Caso #5) requieren el intercambio de esta data mediante APIs. Por lo tanto, se deben implementar medidas de seguridad extremas, incluyendo la ofuscación de datos sensibles, el uso de protocolos de autenticación robustos como OAuth/API Gateway, y el cifrado de la información tanto en reposo como en tránsito. Este enfoque de seguridad mitiga el riesgo de Fuga de Información, que podría ocurrir a través de los nuevos vectores de acceso creados por la interoperabilidad.Transparencia Algorítmica

Para garantizar la confianza en los sistemas de IA, tanto el Motor Predictivo (Caso #2) como el de Detección de Anomalías (Caso #5) deben ser explicables (*XAI - Explainable AI*). No se permite la operación de "cajas negras". Los gestores de cartera y los auditores necesitan entender con claridad los factores y las variables que llevan al modelo a asignar un alto *score* de riesgo de morosidad o a marcar una transacción como sospechosa. Esta transparencia es vital para la toma de decisiones justas y para la defensa de las decisiones administrativas ante el contribuyente.Cumplimiento Ley 1581

El marco legal colombiano, especialmente la Ley 1581 (Protección de Datos Personales), rige la implementación de estos proyectos. Además de la protección de datos, el cumplimiento legal se extiende a dos áreas críticas:

- Validez de la Documentación Digital: En el Caso #1 (Digitalización), se debe garantizar que el proceso de escaneo y la política de preservación digital (PDF/A) mantengan la validez legal de los expedientes para su uso en procesos coactivos.
- Auditoría Legal: El Caso #5 (Detección de Anomalías) debe mitigar los Riesgos de Corrupción y fraude, fortaleciendo la auditoría. Para esto, la Trazabilidad inmutable de cada acción crítica en el sistema es un requisito legal y disciplinario.

10.2. Gobernanza de IA

La gestión exitosa y ética de la IA exige una estructura de Gobernanza de IA dentro de la Alcaldía. Esto implica:

- Comité de Supervisión: Establecer un comité ético o de gobernanza que supervise la definición de las Reglas de Integridad (Caso #3), apruebe los parámetros y el diseño de los modelos predictivos y supervise los protocolos de respuesta ante las alertas

de anomalías.

- Calidad de Datos: Reconocer que la Calidad y Veracidad de Datos es un problema de gobernanza. La falta de datos limpios y estructurados, si no se resuelve con los Casos #1 y #3, resulta en que los modelos de IA sean inútiles o generen resultados erróneos (*Garbage In, Garbage Out*), lo que constituye un riesgo ético y operativo mayor.
- Sostenibilidad: Asegurar la permanencia de los proyectos a través del tiempo mediante la transferencia de conocimiento al talento local y la adopción de prácticas MLOps (Operaciones de Machine Learning) para la automatización del reentrenamiento y el monitoreo continuo de los modelos en entornos de producción.

11 CONCLUSIONES

El análisis de la gestión tributaria del Municipio de Ipiales y la propuesta de casos de uso con Inteligencia Artificial arrojan las siguientes conclusiones clave:

11.1. Solución Estratégica es la Interoperabilidad:

La principal conclusión es que el problema de gestión y recaudo no reside en el *software* actual, sino en la falta de datos completos, digitales e interoperables. La solución estratégica más viable y de menor riesgo operativo es NO cambiar el sistema actual, sino construir una capa de Inteligencia Artificial e interoperabilidad mediante APIs, conectando sistemas y estructurando el dato rezagado.

11.2. Cumplimiento de Meta PDM:

La implementación de este portafolio de IA es el mecanismo directo para superar las brechas tecnológicas (Baja digitalización, Falta de integración) e impulsar el objetivo transversal de fortalecimiento institucional y sostenibilidad financiera definido en el Plan de Desarrollo Municipal 2024–2027.

11.3. Prioridad de Implementación y Habilitadores (Hoja de Ruta)

Los casos de uso de Alta Prioridad son interdependientes y deben ejecutarse en un orden lógico:

Prioridad	Caso de Uso	Rol en el Proyecto

Base	**Digitalización Inteligente (Caso #1)** y **Validación Automática (Caso #3)**	Habilitadores primarios. Crean la Base de Datos digital, estructurada y limpia, que es el pre-requisito para toda analítica.
Impacto	**Motor Predictivo de Recaudo (Caso #2)** y **Detección de Anomalías (Caso #5)**	Motores de Valor. El Caso #2 garantiza el retorno de inversión más alto (Recaudo), y el Caso #5 garantiza la Transparencia (Control y Auditoría).
Experiencia	**Chatbot Tributario (Caso #4)**	Mejora Ciudadana. Masifica la atención, reduce costos operativos de ventanilla y facilita el acceso a la información.

11.4. Impacto Cuantificable y Riesgos Mitigados

11.4.1.Retorno de Inversión (ROI) Alto:

El Caso #2 (Motor Predictivo) y el Caso #5 (Detección de Anomalías) tienen la mayor Viabilidad Financiera (5/5). El primero impacta el incremento de ingresos y el segundo previene pérdidas por fraude o manipulación, justificando ampliamente la inversión en infraestructura de datos y talento especializado.

11.4.2. Reducción de Riesgo Institucional:

El portafolio, en su conjunto, ataca los tres problemas críticos de riesgo:

11.4.3. Riesgo Documental:

Eliminación de expedientes físicos vulnerables y pérdida de información (Caso #1).

11.4.4.Riesgo Operacional:

Eliminación de la dependencia de Excel en procesos críticos (Casos #1 y #3).

11.4.5. Riesgo de Corrupción:

Monitoreo constante, trazabilidad y alertas automáticas sobre manipulación de valores (Caso #5).

11.4.6. Beneficio Ciudadano Directo:

La digitalización total de expedientes (Caso #1) y el Chatbot (Caso #4) aseguran que el contribuyente pueda acceder a sus documentos y consultar su estado de cuenta en segundos, mejorando la transparencia y la calidad de la atención 24/7.

11.5. Recomendación Final:

Se recomienda priorizar la inversión en la adquisición de Escáneres de Producción y la contratación/capacitación de talento en Ingeniería de Datos para ejecutar la Fase 1 (Casos #1 y #3), ya que la calidad de los datos es el pilar que sostiene el éxito de todos los proyectos de Machine Learning y Analítica avanzada propuestos.