



TRABAJO DE GRADO
Opción Seminario-Diplomado

Automatización Para Contratos Electrónicos Del Secop

Corporación Universitaria Remington

Facultad De Ingeniería

Ingeniería De Sistemas

Presentado por:

Richard Josue Zambrano Diaz

Tutor : Luis Camargo Ortega

Opción de Trabajo de grado Seminario

2026

TABLA DE CONTENIDO

DEDICATORIA	3
AGRADECIMIENTO	4
RESUMEN	4
AUTOMATIZACIÓN	5
INTRODUCCIÓN	5
PLANTEAMIENTO DEL PROBLEMA	6
MI JUSTIFICACIÓN PARA HACER ESTE PROYECTO	6
FLUJO EN N8N	7
Extracción de datos	7
Transformación (ETL)	8
Análisis con Inteligencia Artificial	10
Almacenamiento	10
Registro y visualización	11
CÓDIGO FUENTE	12
CONCLUSIÓN	32
REFERENCIAS	33

DEDICATORIA

Dedico este proyecto principalmente a Dios, por darme la fuerza, la sabiduría y la perseverancia necesarias para culminar esta etapa tan importante de mi vida.

A mi familia, por su apoyo incondicional, por creer en mí incluso en los momentos más difíciles y por ser el motor que me impulsa a seguir adelante.

Y finalmente, a mí mismo, por no rendirme, por enfrentar cada desafío y por demostrar que con disciplina y pasión por el desarrollo y la tecnología es posible construir soluciones reales.

AGRADECIMIENTO

Quiero expresar mi sincero agradecimiento al profesor **Luis Camargo Ortega** por su acompañamiento, orientación y apoyo durante el desarrollo de este proyecto. Sus conocimientos y recomendaciones fueron fundamentales para fortalecer cada una de las etapas del trabajo.

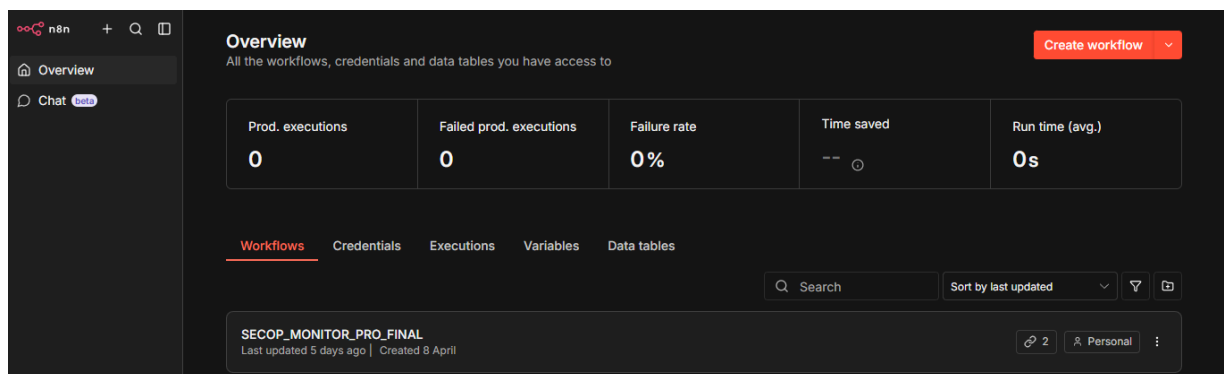
RESUMEN

Este trabajo presenta el diseño e implementación de un sistema automatizado orientado al análisis de contratos públicos, utilizando como fuente principal los datos abiertos del SECOP.

La solución integra un flujo de procesamiento que permite la recolección, depuración y transformación de datos (ETL), seguido de un análisis mediante modelos de inteligencia artificial para la identificación de posibles riesgos contractuales.

El sistema fue desarrollado empleando n8n como herramienta de automatización, articulando servicios como APIs REST, almacenamiento en bases de datos NoSQL (MongoDB) y mecanismos de visualización.

Como resultado, se obtiene una plataforma capaz de procesar información de manera autónoma, facilitando la detección de patrones relevantes y apoyando procesos de toma de decisiones basados en datos.



INTRODUCCIÓN

En los últimos años, la disponibilidad de datos abiertos ha impulsado avances significativos en transparencia y control en el ámbito de la contratación pública. Sin embargo, el volumen y la complejidad de estos datos dificultan su análisis manual, limitando su aprovechamiento.

Este proyecto aborda dicha problemática mediante el desarrollo de un sistema automatizado que permite capturar, procesar y analizar información proveniente del SECOP. A través de la integración de herramientas de automatización y modelos de inteligencia artificial, se busca transformar datos en información útil para la detección de riesgos.

La propuesta no solo optimiza procesos operativos, sino que también introduce capacidades analíticas que pueden ser utilizadas en escenarios de auditoría, monitoreo y control.

PLANTEAMIENTO DEL PROBLEMA

El análisis de contratos públicos presenta múltiples desafíos asociados a:

- Alto volumen de información
- Estructuras de datos heterogéneas
- Procesos manuales poco eficientes
- Limitada capacidad de análisis en tiempo real

Estas condiciones generan dificultades para identificar posibles irregularidades o patrones de riesgo dentro de los procesos contractuales.

En consecuencia, existe la necesidad de implementar soluciones tecnológicas que permitan automatizar el tratamiento de estos datos, mejorando su calidad, accesibilidad y capacidad analítica.

MI JUSTIFICACIÓN PARA HACER ESTE PROYECTO

El desarrollo de este sistema se justifica en la necesidad de optimizar el análisis de información pública mediante herramientas tecnológicas modernas.

La automatización de procesos permite:

- Reducir la carga operativa manual
- Aumentar la eficiencia en el procesamiento de datos
- Mejorar la trazabilidad de la información
- Facilitar la identificación de riesgos mediante inteligencia artificial

Además, el uso de plataformas como n8n posibilita la creación de soluciones escalables sin requerir infraestructuras complejas, lo que representa una alternativa accesible y eficiente para el tratamiento de datos abiertos.

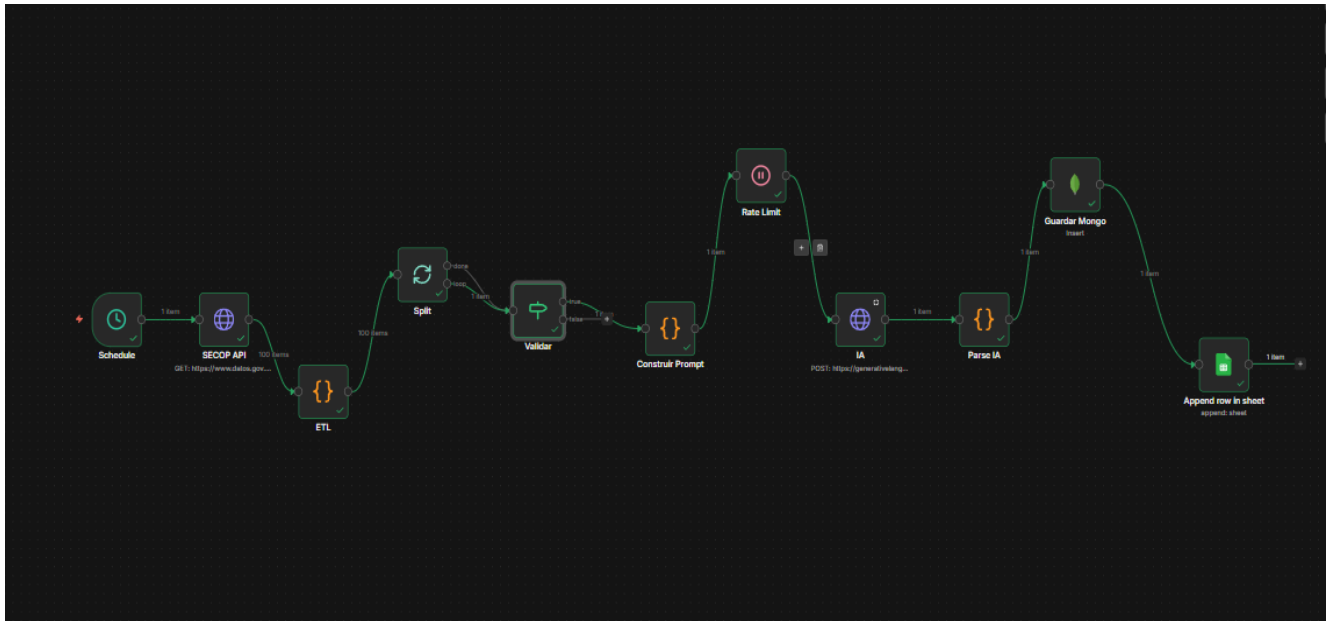
ARQUITECTURA DEL SISTEMA

El sistema propuesto se basa en una arquitectura compuesta por los siguientes elementos:

- **Fuente de datos:** API pública del SECOP
- **Motor de automatización:** n8n
- **Procesamiento:** Transformaciones ETL mediante JavaScript
- **Análisis:** Modelo de inteligencia artificial (Gemini)
- **Almacenamiento:** Base de datos NoSQL (MongoDB)
- **Visualización:** Herramientas tipo dashboard

El flujo sigue un enfoque secuencial que integra extracción, transformación, análisis y almacenamiento de la información, permitiendo un procesamiento continuo y automatizado.

FLUJO EN N8N - AUTOMATIZACIÓN



Extracción de datos

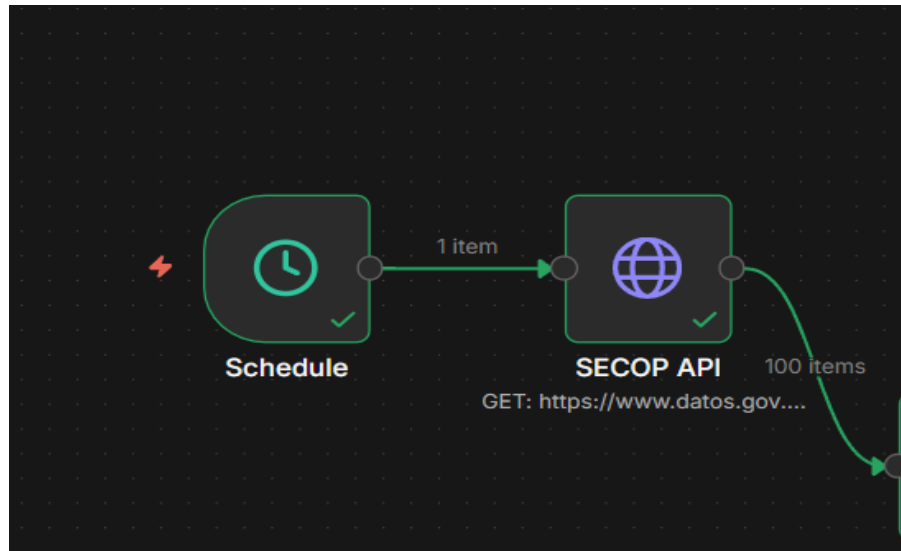
La obtención de información se realiza de forma automatizada mediante solicitudes a la API del SECOP. Este proceso se ejecuta de manera periódica, garantizando la actualización constante de los datos.

Este proceso se implementó mediante los siguientes nodos:

- **Nodo Schedule:**
Se encarga de ejecutar el flujo de manera automática en intervalos definidos (por ejemplo, cada cierto tiempo). Esto permite mantener el sistema actualizado sin intervención manual.
- **Nodo HTTP Request:**
Realiza la consulta a la API de SECOP, desde donde se obtienen los datos de contratos públicos en formato JSON.

¿Por qué hice esto así?

Porque permite trabajar con datos en tiempo real, garantizando que la información analizada sea actualizada y relevante.



Transformación (ETL)

En esta etapa se lleva a cabo el procesamiento de los datos obtenidos, incluyendo:

- Limpieza de información inconsistente
- Eliminación de registros duplicados
- Normalización de formatos

Estas transformaciones permiten estructurar los datos de manera adecuada para su posterior análisis.



¿Por qué importa esto?

Porque los datos provenientes de APIs suelen ser inconsistentes o desordenados, y sin esta etapa el análisis posterior sería incorrecto.

Validación de datos

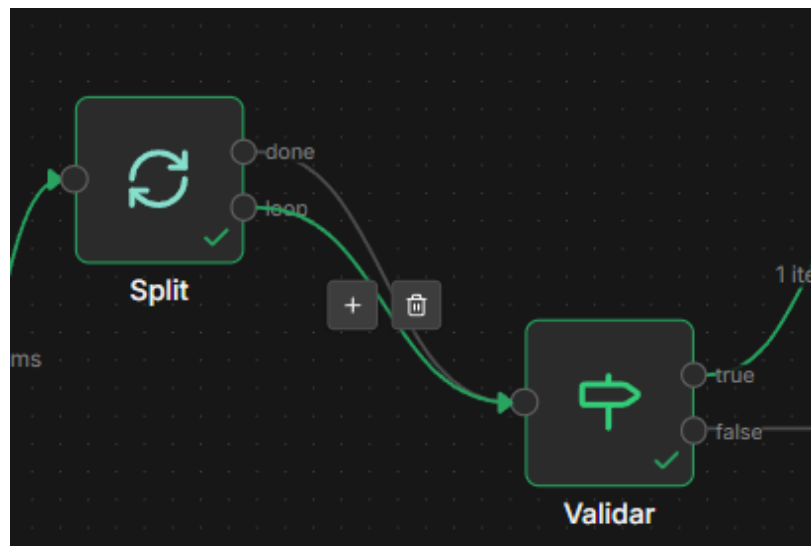
Se implementan mecanismos de control para asegurar la calidad de la información, tales como:

- Verificación de identificadores únicos
- Validación de campos obligatorios

Esto garantiza que únicamente los datos válidos continúen en el flujo.

Objetivo de esta etapa:

Que solo los datos válidos continúen en el flujo, evitando errores en las siguientes fases.

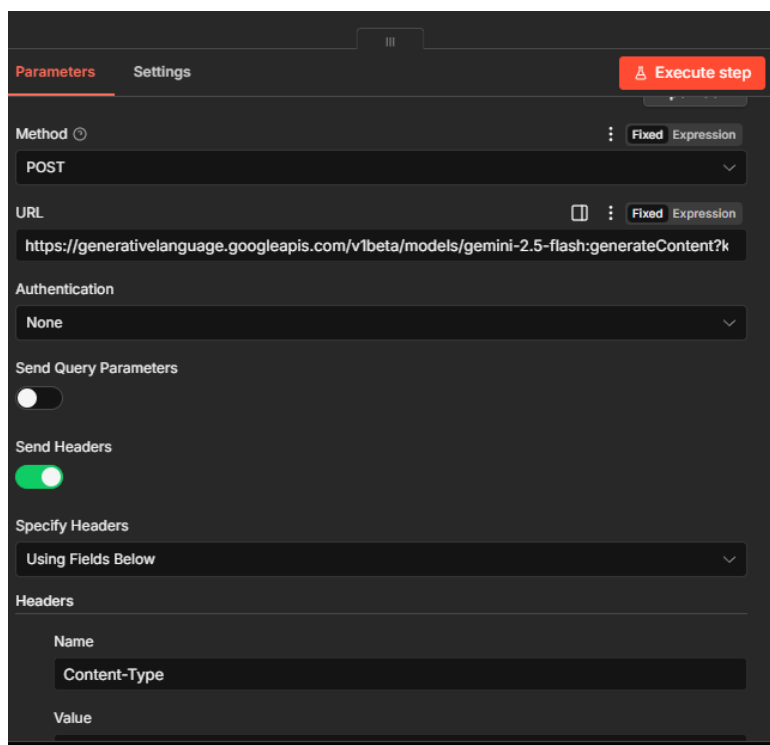
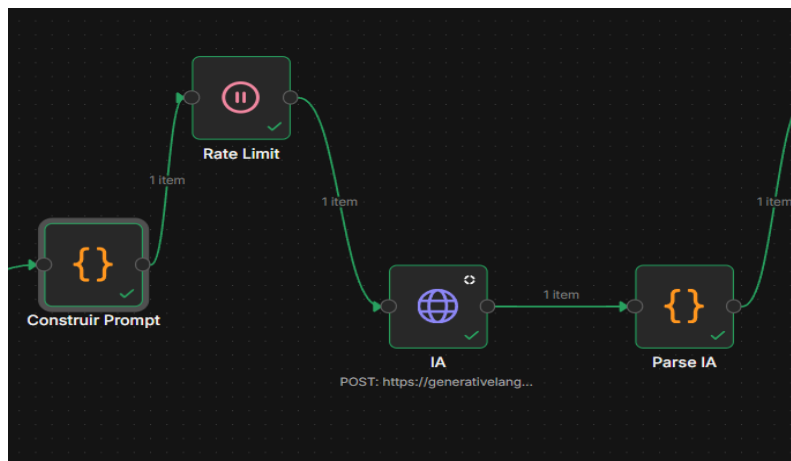


Análisis con inteligencia artificial

Se utiliza un modelo de lenguaje para evaluar la información de cada contrato. El sistema genera una interpretación que incluye:

- Nivel de riesgo (alto, medio, bajo)
- Puntuación asociada
- Breve descripción del análisis

Este enfoque permite identificar posibles irregularidades de manera automatizada



Almacenamiento

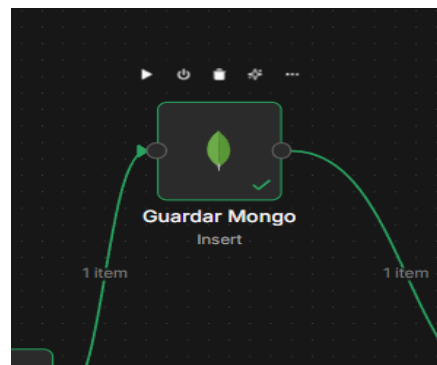
Los datos procesados son almacenados en una base de datos NoSQL, lo que permite:

- Manejo flexible de estructuras
- Escalabilidad
- Compatibilidad con formatos JSON

Además, se implementan mecanismos para evitar duplicidad de registros.

Importante:

En esta etapa también se evita guardar duplicados mediante la validación previa del `id`.



```
Type a query: { field: 'value' } or Generate query ✨ Explain Reset Find Options ▶
```

ADD DATA UPDATE DELETE EXPORT CODE 25 1-1 of 1 ↺ ↻ ↷ ☰ { } 🏠

```
1  _id: ObjectId('69e04c5eb0e3d2d135383c7e')           ObjectId
2  id: "CO1.PCCNTR.6780953,"                          String
3  entidad: "CAJA DE RETIRO DE LAS FUERZAS MILITARES," String
4  departamento: "Distrito Capital de Bogotá,"        String
5  ciudad: "Bogotá,"                                  String
6  contratista: "Andres Mauricio Cepeda Triana,"      String
7  objeto: "Prestación de servicios profesionales especializados para apoyar a la " String
   └──────────────────────────────────────────────────────────────────────────────────┘
8  valor: 23300213                                     Int32
9  tipo: "Prestación de servicios,"                   String
10 modalidad: "Contratación directa,"                String
11 fecha_inicio: "2024-09-23T00:00:00.000,"           String
12 source: "SECOP_II,"                                String
13 ingested_at: "2026-04-16T02:21:20.663Z,"           String
14 analisis: Object                                  Object
```

Registro y visualización

Finalmente, los datos son registrados en herramientas que permiten su consulta y análisis.


- **Google Sheets:**
Se utiliza como una capa adicional de registro, permitiendo visualizar los datos de forma sencilla.

Además, se cree un **dashboard** que permite:


- Visualizar contratos analizados
- Identificar niveles de riesgo
- Analizar tendencias

¿Por qué esta capa aunque se sale del flujo n8n para mi es importante?

Porque transforma los datos procesados en información útil para la toma de decisiones.



 contratos_secop

<https://docs.google.com/spreadsheets/d/14qVTx1ly-GctudaVrIOOyoJoBP7metQTj-dXS-RyshA/edit?usp=sharing>

 **Contratos Públicos** Sistema de visualización de datos

Total Contratos	Valor Total	Score Promedio	Alto Riesgo
2	44.731.922 COP	0.00	0

Filtrar por Riesgo: | Buscar por Entidad: | Ordenar por:

Entidad	Objeto	Valor	Riesgo	Score	Fecha
CAJA DE RETIRO DE LAS FUERZAS MILITARES	Prestación de servicios profesionales espe...	23.300.213 COP	 Medio	0.0	23/9/2024
CORPORACIÓN AUTÓNOMA REGIONAL DE LA ORINOQUIA	Sin Descripción	21.431.709 COP	 Medio	0.0	-

CONCLUSIÓN

El desarrollo de este sistema me permitió ver de forma directa el impacto que pueden tener el uso de herramientas basadas en inteligencia artificial para examinar y comprender datos públicos. A lo largo del proceso, logré integrar diferentes tecnologías para construir una solución funcional, capaz de transformar grandes cantidades de datos en conocimiento útil y aplicable.

Este proyecto no solo representa el cumplimiento de los objetivos planteados, sino también un crecimiento personal en el área del desarrollo web, la automatización y el análisis de datos. Me permitió comprender cómo diseñar soluciones reales a problemas concretos, aplicando herramientas modernas y enfoques actuales.

Por lo tanto, considero que este trabajo deja una base sólida para futuras mejoras, como la implementación de modelos de análisis más avanzados, sistemas de alertas o nuevas integraciones. De esta manera, el proyecto no se limita a un ejercicio académico, sino que puede evolucionar hacia una herramienta con aplicación real en distintos contextos.

REFERENCIAS

Gobierno de Colombia. (s.f.). *Datos abiertos SECOP*. Recuperado de:

<https://www.datos.gov.co>

n8n GmbH. (2024). *n8n Documentation*. Recuperado de: <https://docs.n8n.io>

MongoDB Inc. (2024). *MongoDB Documentation*. Recuperado de:

<https://www.mongodb.com/docs>

Google. (2024). *Generative AI – Gemini API Documentation*. Recuperado de:

<https://ai.google.dev>

Kimball, R., & Ross, M. (2013). *The Data Warehouse Toolkit: The Definitive Guide to Dimensional Modeling*. Wiley.

Han, J., Kamber, M., & Pei, J. (2011). *Data Mining: Concepts and Techniques*. Elsevier.