



GUIA DOCUMENTO DE REQUISITOS

INFORME PROYECTO:

LINA LIZETH LONDOÑO MARÍN: 2226650
LEANDRO RIVERA RÍOS: 2226651
BALMER VALENCIA BANGUERO: 2227097

CAPÍTULO 1: VISIÓN Y ALCANCE DEL PROYECTO

DOCENTE:
SANDRA LUCIA GUAÑARITA FERNANDEZ

25/01/2025

TABLA DE CONTENIDO

TABLA DE CONTENIDO 2

1.1 CONTEXTO ANALISIS 3

 1.1.1 Organización 3

 1.1.2 Proceso actual..... 3

1.2 PROBLEMA 4

 1.2.1 Planteamiento del problema 4

 1.2.2 Formulación del problema 5

1.3 DEFINICIÓN DE LA SOLUCIÓN 5

 1.3.1 Descripción proyecto 5

 1.3.2 Concepción producto 7

 1.3.3 Proceso futuro 8

REFERENCIAS: 9

1.1 CONTEXTO ANALISIS

1.1.1 Organización

- **Nombre de la organización:** Bancolombia S.A.
- **Sector:** Banca
- **Ubicación:** Medellín
- **Misión:** Proveemos acceso a productos y servicios financieros y no financieros para personas, empresas, pymes y gobiernos que abren oportunidades al crecimiento y a la prosperidad.
- **Visión:** Lograr la más alta recomendación siendo una Banca más Humana.

1.1.2 Proceso actual

Nombre del Proceso: Generación de Consultas SQL para Responder a Requerimientos de Datos

¿Qué se hace?

El proceso consiste en la generación de consultas SQL personalizadas para obtener datos específicos almacenados en diferentes bases de datos. Los analistas del área de integración son responsables de interpretar los requerimientos de las diferentes áreas de la organización, convertirlos en consultas SQL y ejecutar estas consultas en el sistema para extraer los datos necesarios.

¿Quiénes intervienen?

Analistas del área de integración de datos: Encargados de redactar, depurar y ejecutar consultas SQL basadas en los requerimientos recibidos.

Áreas solicitantes (usuarios de negocio): Departamentos o equipos internos que necesitan acceder a datos para sus operaciones, análisis o reportes.

Sistemas de almacenamiento de datos: Bases de datos, data lakes y otros repositorios gestionados por la organización.

¿Dónde se está realizando?

El proceso se lleva a cabo en el área de integración de datos, que pertenece al departamento de tecnología y operaciones de la organización. Las bases de datos están alojadas en infraestructuras locales y en la nube, gestionadas por diferentes equipos de soporte técnico.

¿Cómo se está realizando?

- **Recepción de Requerimientos:** Los usuarios de negocio comunican sus necesidades de datos a los analistas a través de correos electrónicos, reuniones o HUs internos.

- Interpretación del Requerimiento: Los analistas analizan los objetivos y las preguntas planteadas para determinar qué datos son relevantes y cómo estructurar la consulta SQL.
- Diseño y Redacción de Consultas: Los analistas redactan manualmente las consultas SQL, tomando en cuenta la estructura de las bases de datos y posibles restricciones de acceso.
- Ejecución y Validación: Las consultas se ejecutan en las bases de datos correspondientes, y los analistas validan los resultados para asegurarse de que cumplen con el requerimiento.
- Entrega de Resultados: Los datos se formatean en hojas de Excel y se envían a los usuarios de negocio.

1.2 PROBLEMA

1.2.1 Planteamiento del problema

El proceso manual de creación de consultas SQL presenta dificultades significativas que impactan negativamente en el flujo de trabajo del área de integración de datos y en la operación general de la organización. Estas dificultades se evidencian por los siguientes síntomas:

- Retrasos en la entrega de información clave: Las solicitudes de datos provenientes de diversas áreas enfrentan tiempos prolongados de espera, ya que los analistas deben interpretar manualmente los requerimientos y generar las consultas correspondientes.
- Sobrecarga de los analistas: La dependencia exclusiva de los analistas para la creación de consultas incrementa su carga de trabajo, limitando su capacidad para realizar tareas de análisis avanzado y estratégico.
- Baja autonomía de los usuarios de negocio: La falta de herramientas intuitivas para interactuar directamente con los datos deja a los usuarios de negocio dependientes del equipo de analistas, ralentizando sus procesos de toma de decisiones.

Causas del problema

- Falta de herramientas automatizadas: El proceso depende de conocimientos técnicos específicos y de la intervención manual de los analistas, lo que dificulta la escalabilidad.
- Dependencia de habilidades técnicas avanzadas: Los usuarios de negocio carecen de las competencias necesarias para acceder directamente a los datos, ya que no tienen formación en SQL ni acceso simplificado a los esquemas de las bases de datos.
- Ausencia de integración eficiente: No se cuenta con una solución que facilite la generación dinámica de consultas basadas en lenguaje natural o que permita a los usuarios acceder a los datos de forma directa.

Actores afectados y cómo son afectados

1. Analistas del área de integración de datos:

- **Afectación:** Sobrecarga de trabajo, pérdida de enfoque en actividades de mayor valor agregado como análisis avanzado, y presión para entregar resultados en tiempos limitados.

2. Usuarios de negocio (equipos de otras áreas):

- **Afectación:** Retrasos en la obtención de información para la toma de decisiones, baja autonomía para acceder a los datos necesarios, y limitaciones en la agilidad operativa.

3. Organización en general:

- **Afectación:** Reducción de la eficiencia global, disminución de la capacidad para responder de manera oportuna a las necesidades del mercado, y posible pérdida de oportunidades estratégicas.

Anexo Storyboard: Planteamiento del problema

1.2.2 Formulación del problema

¿Cómo se puede reducir el tiempo requerido para generar consultas SQL personalizadas y permitir que los usuarios no técnicos interactúen directamente con los datos, mejorando la eficiencia operativa y la accesibilidad de la información en la organización?

1.3 DEFINICIÓN DE LA SOLUCIÓN

1.3.1 Descripción proyecto

Título del Proyecto: Chatbot para Interacción con Bases de Datos.

Nombre del Proyecto: DataGenie

Sigla: AI-DG

Finalidad del Proyecto:

Este proyecto tiene como propósito agregar valor a la organización mediante el desarrollo de un prototipo funcional que optimiza el acceso y la interacción con bases de datos y sistemas externos. A través de un agente inteligente que actúa como intermediario avanzado, se busca reducir la complejidad técnica para los usuarios, mejorar la precisión de las respuestas y aumentar la eficiencia en la obtención de información clave.

Objetivo General:

Desarrollar un agente inteligente basado en Recuperación Aumentada por Generación (RAG) que permita a los usuarios interactuar de manera natural y eficiente con bases de datos relacionales (MySQL), no relacionales (MongoDB) y sistemas externos, integrando tecnologías avanzadas de procesamiento de lenguaje natural y aprendizaje continuo.

Requisitos del Proyecto

Interacción de Usuario:

- Implementar una interfaz gráfica desarrollada en Next.js que permita a los usuarios ingresar preguntas o instrucciones en lenguaje natural.
- Garantizar una experiencia intuitiva y fluida para los usuarios.

Procesamiento Backend:

- Diseñar un backend en FastAPI que gestione solicitudes de usuarios, conecte con bases de datos MySQL y MongoDB, y recupere documentos relevantes mediante embeddings.
- Integrar una API de modelo de lenguaje grande (LLM) para generar respuestas precisas y contextuales.

Gestión del Agente Inteligente:

- Implementar lógica para que el agente planifique y ejecute múltiples pasos cuando las consultas lo requieran.
- Registrar interacciones en MySQL para su análisis y mejora.

Aprendizaje Continuo:

- Desarrollar mecanismos que permitan optimizar las respuestas del agente en función del feedback recibido.
- Integrar modelos de aprendizaje continuo para mejorar el comportamiento del sistema.
- Restricciones del Proyecto

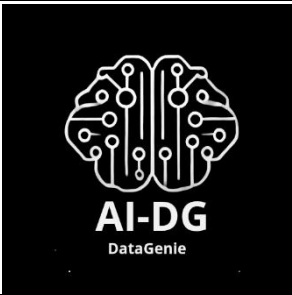
1.3.2 Concepción producto

El producto es una herramienta interactiva diseñada para optimizar el acceso y análisis de datos.

Elevator Pitch

Elemento	Descripción
For (Destinatarios)	Profesionales y empresas que necesitan optimizar el acceso y análisis de datos en sus operaciones.
Who (Necesidad u Oportunidad)	Quieren realizar consultas avanzadas y obtener insights sin necesidad de dominar SQL o lenguajes técnicos.
The (Nombre del Producto)	DataGenie
Is a (Categoría del Producto)	Herramienta de análisis y consulta de datos interactiva.
That (Beneficio Clave)	Permite acceder y analizar datos de forma intuitiva, rápida y precisa a través de lenguaje natural.
Unlike (Alternativa Competitiva)	Herramientas tradicionales que requieren conocimientos técnicos complejos.
Our Product (Diferenciación)	Proporciona un enfoque accesible con inteligencia artificial avanzada, mejorando precisión y experiencia del usuario.

Product Box

Elemento	Detalle
Nombre del Producto	DataGenie
Logo	
Beneficio Principal	Simplifica el análisis de datos mediante interacción en lenguaje natural.
Beneficio 2	Integra bases de datos como MySQL y MongoDB en tiempo real.

Beneficio 3	Proporciona una experiencia gráfica amigable y atractiva para usuarios no técnicos.
--------------------	---

NOT LIST

IN SCOPE	IN SCOPE	IN SCOPE
- Conexión a bases de datos SQL y NoSQL.	- Creación de bases de datos desde cero.	- Integración con herramientas de BI externas.
- Consultas en lenguaje natural (NLQ).	- Mantenimiento de bases de datos existentes.	- Soporte para bases de datos muy específicas.
- Generación de reportes básicos en tiempo real.	- Consultas extremadamente avanzadas de Big Data.	- Compatibilidad multilingüe.
- Visualización simple de datos procesados.	- Funcionalidad sin conexión (offline).	- Escalabilidad para uso masivo en la nube.

Anexo Boceto: Prototipo de baja Fidelidad

1.3.3 Proceso futuro

Con la implementación de esta solución, el proceso futuro transformará la manera en que los usuarios acceden y analizan los datos:

- Mayor Autonomía para los Usuarios: Los usuarios de negocio podrán interactuar directamente con el sistema para obtener información, reduciendo la dependencia de los analistas para tareas rutinarias de generación de consultas SQL.
- Reducción de Tiempos de Espera: Al automatizar la recuperación de datos y la generación de consultas, los usuarios obtendrán respuestas más rápidas, mejorando su eficiencia operativa.
- Enfoque en Análisis Avanzado: Los analistas podrán dedicar más tiempo a actividades estratégicas y de análisis avanzado, mientras que el chatbot maneja consultas básicas y operativas.
- Evolución Continuada: La solución se ajustará dinámicamente a las necesidades cambiantes de la organización, incorporando nuevas funcionalidades basadas en el feedback de los usuarios.

Anexo Storyboard: Proceso Futuro con la Nueva Solución

REFERENCIAS:

1. Tiangolo. FastAPI Documentación. Disponible en: <https://fastapi.tiangolo.com/>.
2. MySQL. MySQL Documentación. Oracle Corporation. Disponible en: <https://dev.mysql.com/doc/>.
3. MongoDB Inc. MongoDB Documentación. Disponible en: <https://www.mongodb.com/docs/>.
4. Vercel. Next.js: The React Framework for Production. Disponible en: <https://nextjs.org/>.
5. Hugging Face. A Guide to Retrieval-Augmented Generation (RAG). Disponible en: <https://huggingface.co/blog/rag>.