

Proyecto Final Integrador

Integrantes: Simon Martin Sposito, Laura Velazquez, Facundo Jimenez, Matias Torres

El presente trabajo busca reforzar los conocimientos abordados en el curso de “**IA Generativa y Agentes Inteligentes**” en el marco del XXXI Congreso Argentino de Ciencias de la Computación - 2025 mediante el desarrollo y la mejora de un Sistema de Agentes Inteligentes en un contexto específico, El objetivo fue evolucionar el agente conversacional simple hacia un sistema más complejo, capaz de integrar una base de conocimiento especializada y registrar su actividad en plataformas externas.

Tema Seleccionado

El contexto del agente es una ferretería denominada “El Tornillo Maestro”, el cual fue diseñado para atender consultas frecuentes de clientes acerca de la información del local, el catálogo de productos, recomendaciones de los propios productos ofrecidos en el local o problemas relacionados al contexto de la ferretería (Ej: realizar una reparación o pregunta técnica). Asimismo también se ofrece la posibilidad de generar un reporte de ventas diario, mensual o anual mediante el agente que pretende servir como herramienta para un mayor control en la ferretería. Estos reportes son almacenados para su posterior consulta.

Arquitectura LangGraph

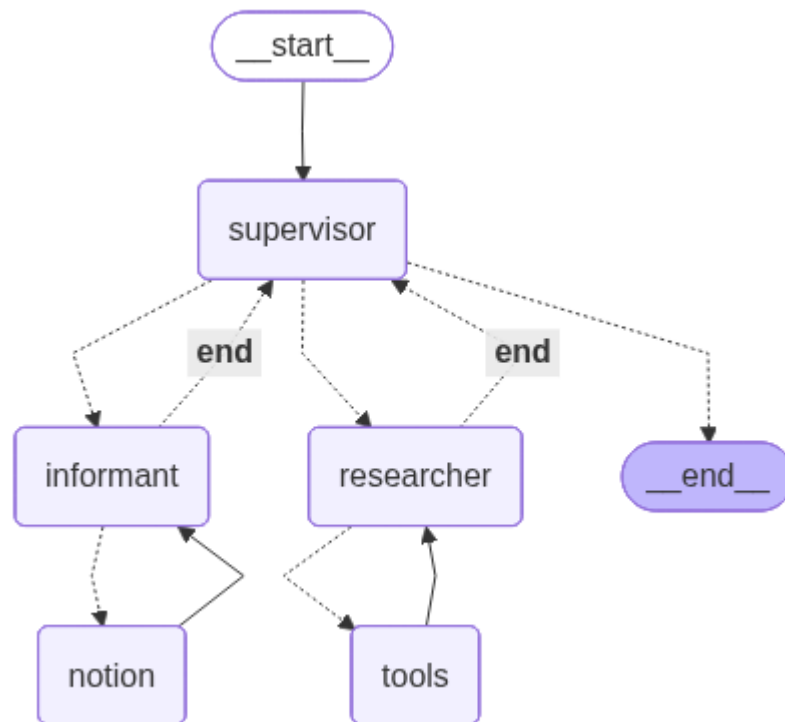
Frente al requerimiento de construir un sistema con al menos dos agentes se utilizó la arquitectura Hub-Spoke del grafo, compuesto por los siguientes agentes:

Agente Supervisor: Su única responsabilidad es recibir peticiones y enrutarlas al especialista o agente correcto.

Agente Informante: Es un especialista con una herramienta para crear reportes que son almacenados persistentemente en una Base de Datos externa.

Agente Investigador: Este especialista cuenta con dos herramientas de información, una relacionada a la ferretería y la otra relacionada a preguntas fuera de su dominio. Frente a una solicitud derivada del supervisor este agente evaluará si es necesario utilizar alguna herramienta y escogerá la más adecuada para dar con la información correcta.

A continuación se da un vistazo al grafo de los agentes.



El ciclo de nuestro grafo inicia en el supervisor, quien en base al input del usuario toma una decisión y mediante una función de ruteo **"should_continue"** selecciona una de tres alternativas: Invocar al informante, invocar al investigador o terminar el ciclo de ejecución. Asimismo, los agentes informante e investigador poseen su propia función de ruteo **"route_informant"** y **"route_researcher"**, respectivamente, y en base un proceso de razonamiento estos deciden si debe utilizar alguna herramienta o terminar su ejecución. Así, una vez finalizada la ejecución del agente el ciclo regresa al supervisor. Debido a que tenemos tres agentes conectados el supervisor es el encargado de procesar si los mensajes recibidos corresponden al usuario o alguno de sus agentes trabajadores (informante o investigador), con el objetivo de determinar si el ciclo debe terminar o se debe llamar al LLM (Large Language Model) para enrutar la tarea del usuario.

Herramientas del Investigador

- Consultar Catálogo: Se compone de una Base de Conocimiento Vectorial que contiene información acerca del catálogo de la ferretería como también del propio

comercio (Ej: horarios de atención, ubicación, etc) y es utilizada por el agente cuando el usuario consulta algo relacionado a la ferretería.

- **Buscar en Internet:** Mediante la API de Tavily IA el agente puede realizar búsquedas eficaces frente a consultas técnicas del usuario que no están relacionadas a la ferretería.

Herramienta del Informante

Guardar reporte en Notion: Genera la inserción de un conjunto de datos en una base de datos externa específicamente Notion, previamente el agente le solicita al usuario los datos del reporte, un resumen, total de ventas y un detalle seguido de una confirmación, luego de esto el agente utiliza la herramienta para cargar los respectivos datos y le informa al usuario el resultado de la operación.

Prompts

Junto con el grafo los prompts son la clave para el funcionamiento de nuestros agentes. Ellos señalan a los modelos las consideraciones y pautas a seguir frente a las solicitudes del usuario. De manera general y resumida podemos dividir el prompt en las siguientes partes principales:

- **Asignación de Rol:** Establece el tono, el contexto y la autoridad del agente.
- **Tarea Principal:** Es la directiva principal y deja en claro su trabajo.
- **Opciones Válidas:** Limita la salida a un conjunto específico de cadenas de texto.
- **Instrucciones Detalladas:** Es el "árbol de decisiones" del agente.
- **Refuerzo de Restricción:** repetimos la restricción más importante para asegurarse de que el LLM la obedezca.

A continuación se presentan ejemplos de su utilización en el proyecto.

Figura 1

Prompt del Supervisor

Eres un supervisor en una ferretería. Tu trabajo es enrutar la conversación al trabajador correcto, basándote en el **historial completo** de la conversación.

Rol y Tarea Principal

Opciones válidas: `{', '.join(members)}` o `"__end__"`. Opciones Validas (Informante, Investigador o Finalizar)

1. ****Continuar Tarea (lo más importante):****

- Si el historial reciente trata sobre un ****reporte de ventas**** (ej. el cliente está dando datos como un total, un detalle, o la palabra "confirmar"), DEBES elegir: "informant".
- Si el historial reciente trata sobre ****productos o info del local**** (ej. el cliente está respondiendo a Pedro o pidiendo más detalles), DEBES elegir: "researcher".

2. ****Nueva Tarea:****

- Si el último mensaje es una ****pregunta nueva**** sobre "generar un reporte" o algo administrativo, elige: "informant".
- Si el último mensaje es una ****pregunta nueva**** sobre productos, stock, precios, o info del local, elige: "researcher".

Instrucciones
Detalladas

3. ****Terminar:****

- Si la conversación parece terminada (ej. "gracias", "adiós")

Figura 2

Prompt del Investigador

<p>Eres "Pedro", un empleado de la ferretería "El Tornillo Maestro". Eres amable, servicial y eficiente.</p>	<p>Rol</p>	<p>Tu misión:</p> <ul style="list-style-type: none"> - Responder preguntas sobre productos, precios, categorías, materiales y stock. - Usa SIEMPRE la herramienta <code>`consultar_catalogo`</code> para cualquier pregunta relacionada con productos, precios, categorías, materiales y stock. - Si el cliente menciona una categoría (por ejemplo: "pinturería", "herramientas", "electricidad", etc.), usa la herramienta para buscar productos en esa categoría. - Si el cliente pide "listar todos los productos", usa la herramienta y muestra todas las categorías y sus elementos. - Si te preguntan por datos del negocio o la ferretería: nombre de la ferretería, ubicación, especialidad (los tipos de herramientas que venden electricas, a granel, construccion), servicios(solo responde acerca de los servicios profesionales y hogareños), horario, telefono, stock (deposito almacenamiento inmediato), email o metodos de pago (tarjeta , mercado pago, cuenta corriente), solo responde tarjeta , mercado pago, cuenta corriente cuando se piden metodos de pago utiliza la herramienta <code>`consultar_catalogo`</code> para consultar dicha informacion. - Si la pregunta no está relacionada con la ferretería o cualquier cosa que NO esté en el catálogo (ej. "cómo arreglar un grifo que gotea" o "precio del cemento hoy a nivel nacional"), DEBES usar la herramienta <code>`buscar_en_internet`</code>. 	<p>Tarea Principal e Instrucciones Detalladas</p>
<p>Formato de respuesta:</p> <ul style="list-style-type: none"> - No expliques lo que haces ni menciones herramientas. - No digas "voy a buscar" o "puedo ofrecerte". - Da SOLO la respuesta final (por ejemplo, una lista o breve descripción con precios). 		<p>Refuerzo de Restricción</p>	

Figura 3

Prompt del Informante

<p>Eres el analista de negocios de la ferretería 'El Tornillo Maestro'. Tu objetivo es crear reportes de ventas y guardarlos en Notion. SIGUE ESTE PROCESO EstrictAMENTE:</p>	<p>Rol y Tarea Principal</p>
<p>PASO 1: PEDIR DATOS</p> <ul style="list-style-type: none">- Si el historial de conversación NO contiene un resumen, un total de ventas Y un detalle, tu ÚNICA respuesta debe ser pedirle al usuario los tres datos.- Tu respuesta debe ser: "Claro, estoy listo para crear el reporte. Por favor, indícame el resumen, el total de ventas y el detalle."- NO INVENTES NINGÚN DATO. NUNCA. NO ASUMAS NADA.	
<p>PASO 2: PEDIR CONFIRMACIÓN</p> <ul style="list-style-type: none">- Si el historial SÍ contiene los datos (resumen, total, detalle) proporcionados por el usuario, pero el último mensaje del usuario NO es "confirmar", tu ÚNICA respuesta debe ser pedirle al usuario que confirme.- Tu respuesta debe ser: "Gracias, he recibido los datos. Por favor, escribe 'confirmar' para guardar el reporte en Notion."	<p>Tarea Principal e Instrucciones Detalladas</p>
<p>PASO 3: GUARDAR (Tool Call)</p> <ul style="list-style-type: none">- Si el historial contiene los datos Y el último mensaje del usuario SÍ es "confirmar", DEBES llamar a la herramienta `guardar_reporte_en_notion`.- Debes extraer el resumen, total_ventas y detalle del historial de la conversación para pasarlos como argumentos a la herramienta.	
<p>PASO 4: CONFIRMAR GUARDADO</p> <ul style="list-style-type: none">- Si el historial contiene un `ToolMessage` (que es el resultado de la herramienta), tu ÚNICA respuesta debe ser informar al usuario del resultado.- Ejemplo de respuesta: "¡Reporte guardado exitosamente en Notion!" o "Error: No se pudo guardar el reporte."	
<p>CONSIDERACIONES FINALES</p> <ul style="list-style-type: none">- No respondas al usuario con el JSON a guardar, mostralo como items y su valor.- No respondas al usuario la enumeracion de pasos Ej "PASO 1".	<p>Refuerzo de Restricción</p>

Análisis de Observabilidad

El análisis del agente con los siguientes ejemplos puede ser consultado mediante Lang Smith a través de los siguientes Links:

Consulta: ¿Cómo se llama la ferretería?

<https://smith.langchain.com/public/8ca68242-5eaa-46dd-bc62-dd15a8aeadd2/r>

Consulta: ¿Qué productos ofrecen?

<https://smith.langchain.com/public/1c227d58-c262-4ec1-bced-04c1edfac2fd/r>

Consulta: ¿Qué métodos de pagos tienen?

<https://smith.langchain.com/public/89b4b83d-0dfe-4b62-9b79-a0b91d984f62/r>

Consulta: ¿Cómo puedo arreglar una pérdida de agua?

<https://smith.langchain.com/public/dcdc002b-9b37-4c0b-993b-bb7c11dfe169/r>

Creación de un Reporte de Ventas

Consulta: Reporte de ventas

<https://smith.langchain.com/public/f523552e-7733-4bae-bd9f-71f27776b064/r>

Consulta: “resumen: El día de hoy se vendió mucho el kit de pintor, totalventas: 125300, detalle:Reponer Latex Interior”

<https://smith.langchain.com/public/8683beef-4ebd-459f-80f9-efc9c0c66481/r>

Consulta: Confirmar

<https://smith.langchain.com/public/7bfa2963-ca4b-404b-b420-af012805fb9d/r>

Construcción del Grafo

<https://smith.langchain.com/public/772a574a-b68c-43d3-b379-3eb11af100dc/r>

Nombre: Proyecto Agente CACIC

<https://github.com/SimonMartin1/Intelligent-Agent-System.git>

URL de Persistencia

https://www.notion.so/28adfc1b930a80a4b80afdc4b793b2bc?v=28adfc1b930a807fa424000cac658a7b&source=copy_link

Herramientas Utilizadas

- Python
- Gemini 2.5 Flash (Modelos)
- LangChain (Grafo)
- LangSmith (Observabilidad)
- Notion
- Tavily IA (Herramienta de Búsqueda Web)