

Avances del Proyecto: Agente de IA Autónomo

Arquitectura, Memoria y Capacidades Multimodales

Prof. César Rodríguez Gemini Code Assist

Equipo de Investigación 'Tecnología Venezolana'

17 de febrero de 2026

Contenido

- 1 Visión y Objetivo
- 2 Arquitectura Técnica
- 3 Capacidades Implementadas
- 4 Interacción y Demo
- 5 Próximos Pasos

Objetivo del Proyecto

Misión

Desarrollar un Agente de IA autónomo, residente en infraestructura local (Linux), capaz de apoyar en tareas de investigación académica y gestión de conocimiento para el equipo.

Principios de Diseño:

- **Privacidad:** Ejecución y memoria locales.
- **Fiabilidad:** Separación entre razonamiento (LLM) y ejecución (Código).
- **Velocidad:** Uso de motores de inferencia de alta velocidad (Groq).

Arquitectura de 3 Capas

Hemos implementado un diseño robusto que evita alucinaciones y errores comunes:

- ❶ **Capa 1: Directivas (YAML)**

'El Manual de Operaciones'. Define qué se debe hacer (SOPs).

- ❷ **Capa 2: Orquestación (LLM)**

'El Cerebro'. Toma decisiones, enruta tareas y gestiona errores.
Actualmente potenciado por **Llama 3 (vía Groq)**.

- ❸ **Capa 3: Ejecución (Python)**

'Las Manos'. Scripts deterministas para tareas específicas (búsqueda web, manejo de archivos, cálculos).

Capacidades Clave

- **Memoria a Largo Plazo (RAG):** Uso de *ChromaDB* para recordar datos, preferencias y soluciones a errores entre sesiones.
- **Interfaz Telegram Completa:**
 - Comandos de voz (transcripción automática).
 - Análisis de imágenes (Visión por Computadora).
 - Gestión remota del servidor.
- **Velocidad Extrema:** Integración de **Groq** logrando respuestas en **0.6 segundos** (vs 2.6s de modelos estándar).
- **Sistema de Respaldo (Fallback):** Si Groq falla, el sistema cambia automáticamente a Google Gemini sin interrupción.

Comandos Disponibles

El agente responde a comandos naturales y estructurados:

- `/investigar` Realiza búsquedas en la web y genera resúmenes.
- `/recordar` Guarda información en la base de datos vectorial.
- `/status` Reporte de salud del servidor (CPU, RAM, Disco).
- `/modo` Cambio de personalidad (Serio, Sarcástico, Pirata...).
- `/reiniciar` Limpia el contexto y restablece la identidad.

Hoja de Ruta

- Ampliación de herramientas de ejecución (Sandbox de código).
- Integración con calendarios y herramientas de productividad.
- Refinamiento del sistema de auto-curación de errores.

¡Gracias!