

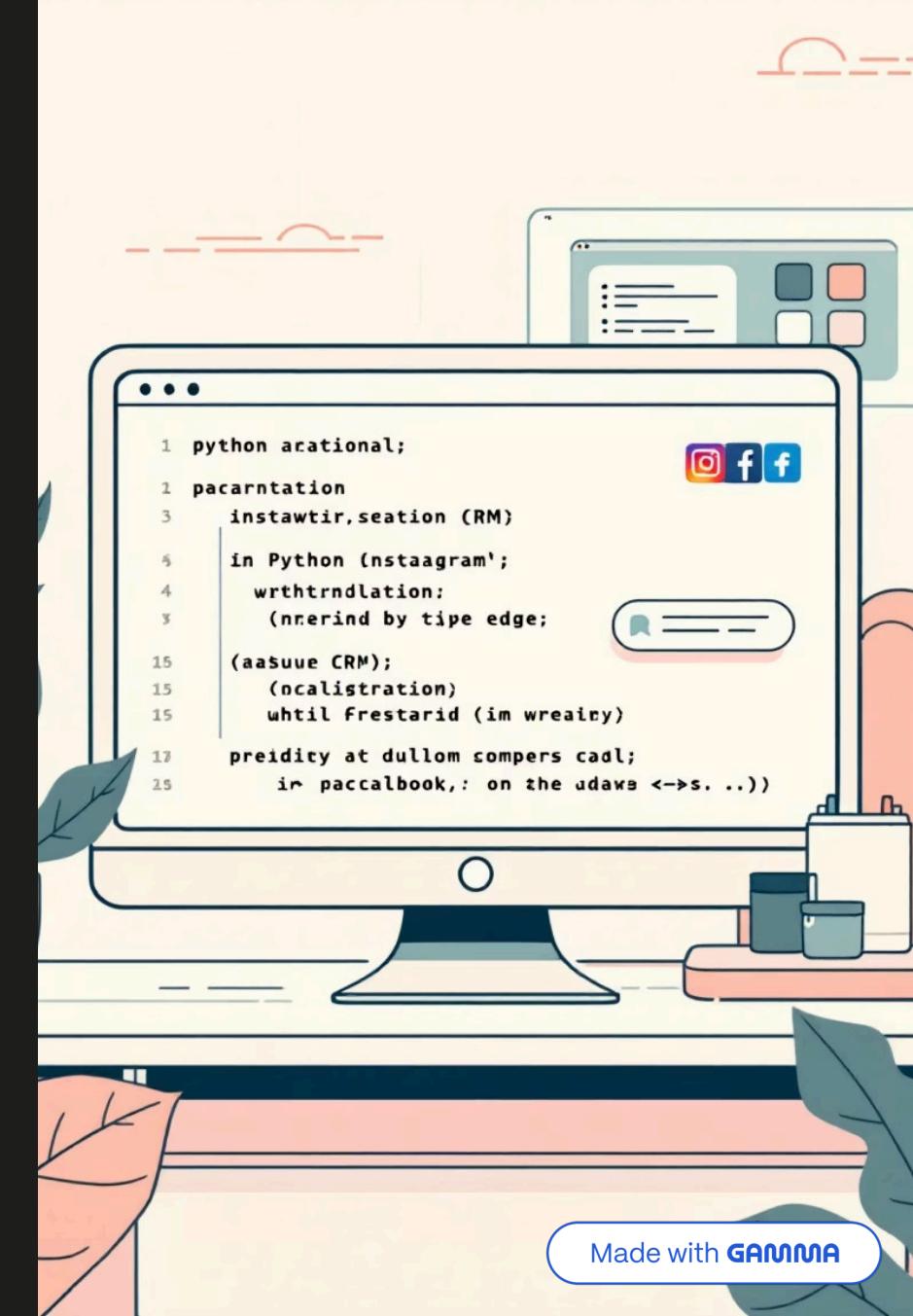
LangChain

Framework para construir aplicaciones con IA conectando modelos de lenguaje con datos y herramientas



¿Qué es LangChain?

Framework para construir aplicaciones con LLMs - como 'LEGO' para IA que permite conectar bloques pre-construidos en vez de programar desde cero.



5 Conceptos Clave

01

LLM (Large Language Model)

El "cerebro" - modelos como GPT-4, Claude o Llama que entienden y generan lenguaje natural.

02

Prompts

Las instrucciones que le das al LLM. La calidad del prompt determina la calidad de la respuesta.

03

Embeddings (Vectores)

Convertir texto en números que las computadoras pueden comparar. Palabras similares = vectores similares.

04

RAG (Retrieval Augmented Generation)

Patrón que combina búsqueda + generación para que el LLM responda basándose en TUS documentos.

05

Agents (Agentes)

Sistemas que razonan y deciden qué hacer automáticamente. Usan herramientas según sea necesario.

Flujo RAG

01

Usuario pregunta

03

Pasar documentos + pregunta al LLM

02

Buscar documentos relevantes

04

LLM genera respuesta basada en los documentos

Componentes Principales

Chains	Secuencias de operaciones	Traducir → Resumir → Formatear
Memory	Recordar conversaciones	Chatbot que recuerda tu nombre
Tools	Extender capacidades del LLM	Buscar Google, calcular, leer PDFs
Output Parsers	Estructurar respuestas	Convertir texto a JSON

Ejemplo Rápido

Imagina un asistente virtual que no solo entiende tus preguntas, sino que también puede buscar en internet para encontrar la respuesta más actual. Eso es LangChain en acción:

- Combina Modelos de Lenguaje Grandes (LLMs) con herramientas externas.
- Permite al LLM interactuar con el mundo exterior (ej. búsqueda web, bases de datos).
- Resultados: respuestas más precisas, actualizadas y contextuales.

Conclusión

LangChain se establece como un marco indispensable para el desarrollo de aplicaciones de inteligencia artificial de nueva generación. Su arquitectura permite a los Modelos de Lenguaje Grandes (LLMs) ir más allá de su entrenamiento inicial, conectándolos con fuentes de datos externas y herramientas del mundo real.

Esto resulta en:

- **Mayor precisión:** Respuestas basadas en información actualizada.
- **Interacción contextual:** La IA puede buscar y procesar datos específicos para cada consulta.
- **Potencial ilimitado:** La base para construir asistentes, agentes y aplicaciones de IA verdaderamente inteligentes y dinámicas.