Guía práctica

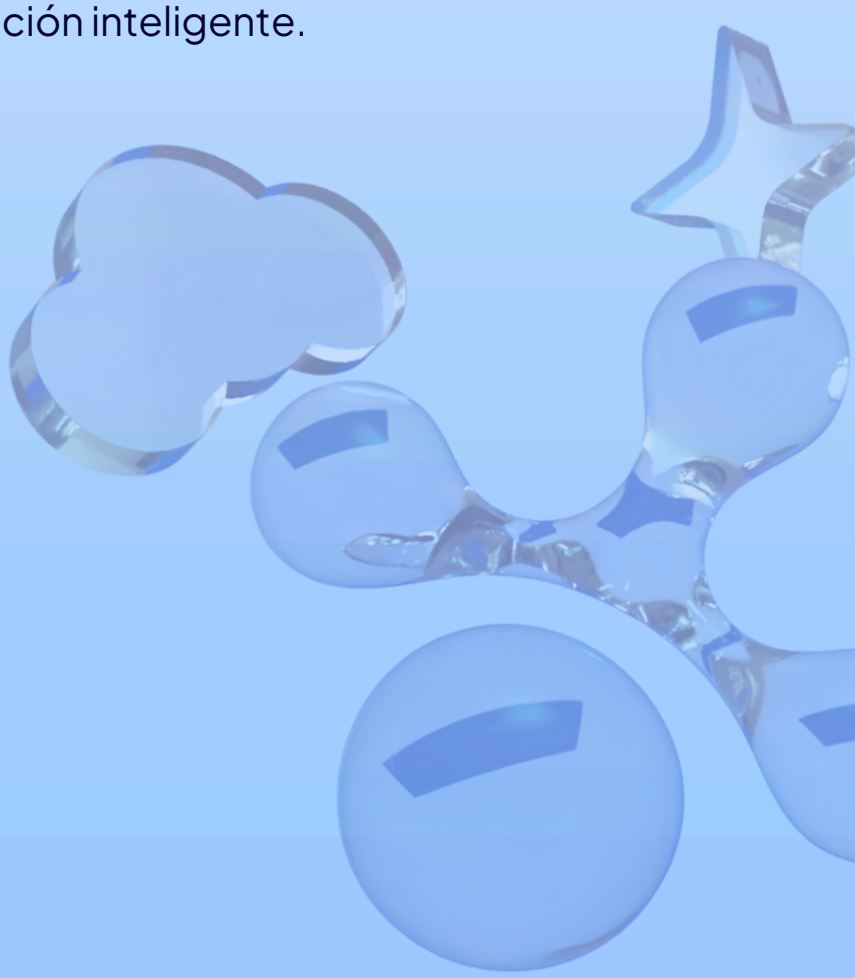
para construir tu primer agente de IA

merkur.ia

Guía práctica

para construir tu primer agente de IA

Una guía para dominar los fundamentos, las mejores prácticas y los secretos de la automatización inteligente.



Índice

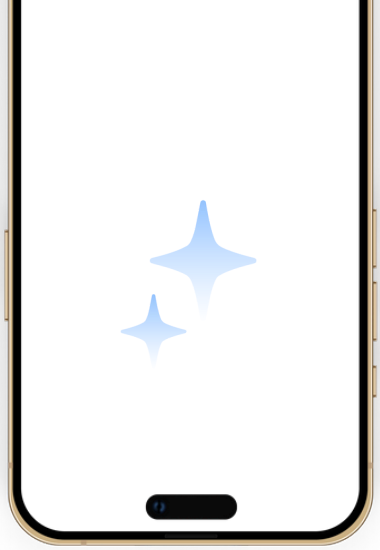
01 ¿QUÉ ES UN AGENTE DE IA? p5.

02 LOS 3 PILARES p8.

03 ORQUESTACIÓN: DE AGENTE SOLITARIO A EQUIPO DE ESPECIALISTAS p11.

04 CONSTRUYENDO AGENTES SEGUROS Y CONFIABLES p14.

05 TU ÉXITO p16.



Healthier, happier & prolific humans...**thanks to AI**

Nuestra obsesión

Programamos agentes de inteligencia artificial que potencian las capacidades humanas.

Somos expertos en **INTELIGENCIA ARTIFICIAL.**
Desarrollamos agentes IA con distintas funciones: identificar emociones, atender usuarios y resolver conflictos.
Usamos nuestros modelos decisionales para hacerlos interactuar entre sí, resolviendo retos reales.

01

¿Qué es un Agente de IA?

¿Qué es un Agente de IA?

Los agentes son sistemas que realizan tareas y flujos de trabajo de forma independiente en lugar de un humano.

Flujos de trabajo

Es una secuencia de pasos que deben ejecutarse para cumplir el objetivo del usuario, ya sea resolver un problema de atención al cliente, hacer una reserva en un restaurante, realizar un cambio de código o generar un informe.

Capacidades principales de un Agente de IA

Aprovecha un **Modelo Largo de Lenguaje "LLM"** para gestionar la ejecución del flujo de trabajo y tomar decisiones. Reconoce cuando se completa un flujo de trabajo y puede corregir proactivamente sus acciones, si es necesario. En caso de fallo, puede detener la ejecución y devolver el control al usuario.

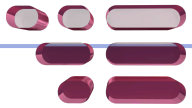
Tiene acceso a varias herramientas para interactuar con sistemas externos **–tanto para recopilar el contexto como para realizar acciones–** y selecciona dinámicamente las herramientas adecuadas en función del estado actual del flujo de trabajo, operando siempre dentro de unos límites claramente definidos.

¿Cuándo necesitas un agente de IA?

Un agente es la solución ideal cuando los métodos tradicionales fallan. Búscalo si tus flujos de trabajo son:

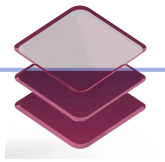
Complejos y con matices

Procesos que requieren juicio humano, como el análisis de un fraude o la aprobación de un reembolso complejo



Con base en reglas frágiles

Sistemas con miles de condiciones “si...entonces...” que son costosos y difíciles de mantener actualizados



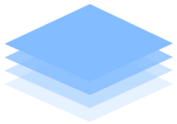
Dependientes de datos no estructurados

Cuando necesitas interpretar correos, documentos PDF o mantener conversaciones para extraer información clave

02

Los 3 pilares

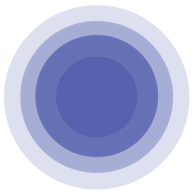
Los 3 Pilares Fundamentales



1. El Modelo

El cerebro

El LLM que dota de razonamiento al agente. La clave es elegir el modelo correcto para cada tarea, balanceando capacidad con costo y velocidad



2. Las Herramientas

Las manos

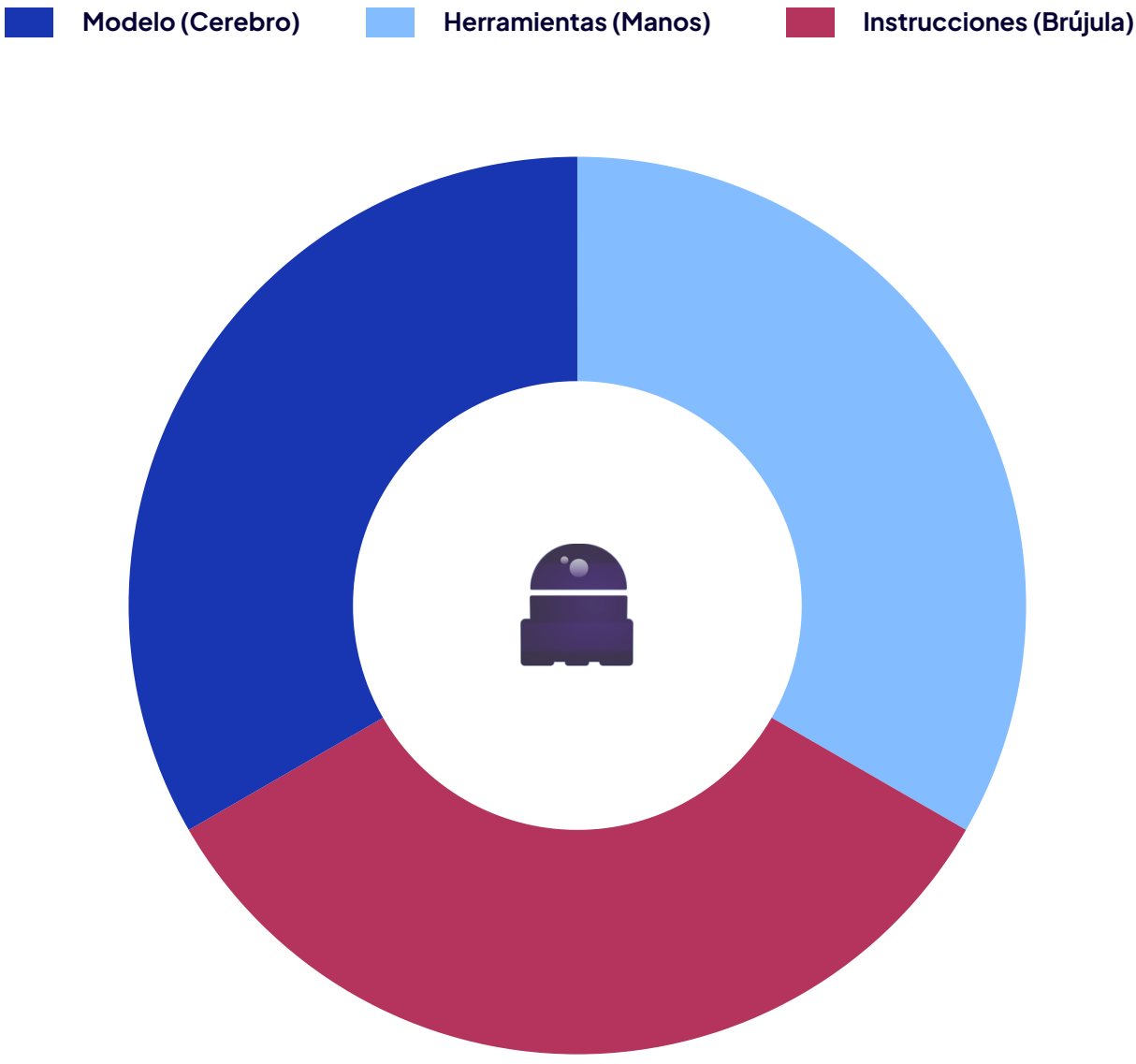
Las APIs y funciones que el agente usa para interactuar con el mundo exterior: obtener datos, ejecutar acciones y hasta llamar a otros agentes.

3. Las Instrucciones

La brújula

Las directrices explícitas que definen su comportamiento, límites y personalidad. La claridad aquí es crucial para evitar errores.

Un agente balancea estos tres pilares
de forma armónica



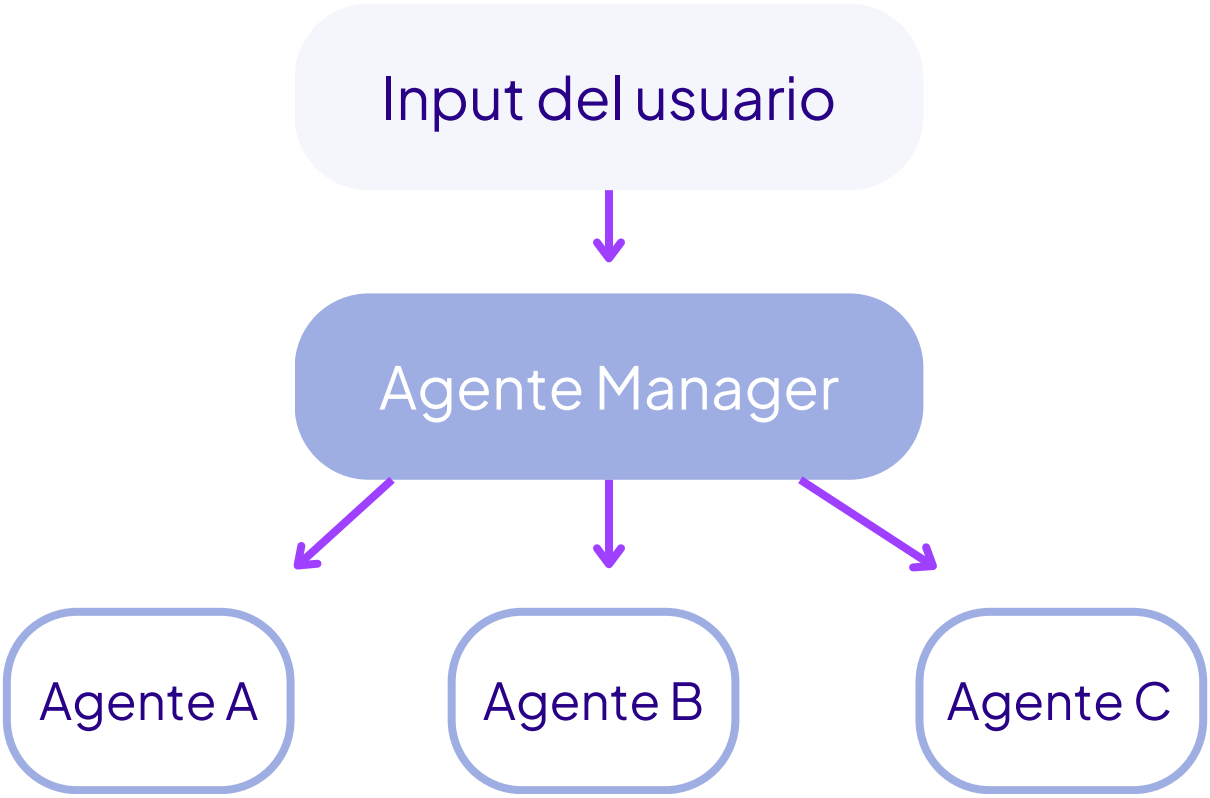
03

Orquestación: de agente solitario a equipo de especialistas

La orquestación define cómo se gestiona el flujo de trabajo. Empieza simple y escala solo cuando la complejidad lo requiera.

Patrón Manager – Especialista

Un agente “director” centraliza la tarea y la delega a agentes especializados.



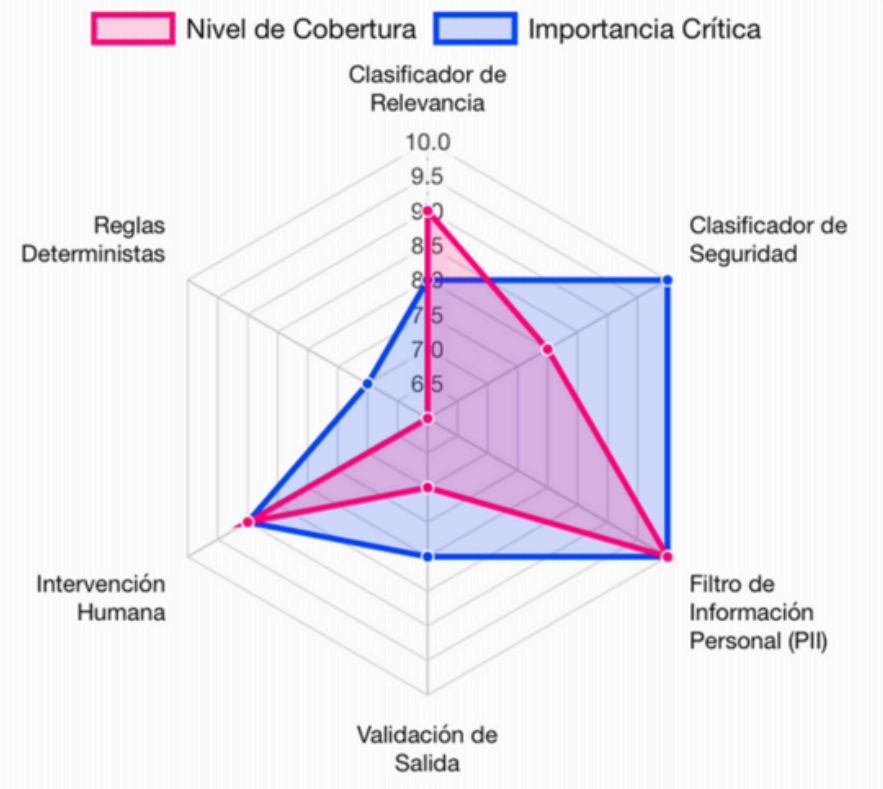
Patrón Decentralizado

Un agente de “clasificación” pasa el control total de la tarea a otro agente

04

Construyendo agentes seguros y confiables

Un agente debe operar dentro de límites claros.
Implementa una defensa en capas para mitigar riesgos.



Las barreras de protección (“vallas”) son tu red de seguridad. Combina múltiples tipos para una protección robusta.

- **Clasificador de relevancia:** Mantiene el agente enfocado en su tarea.
- **Clasificador de seguridad:** Detecta intentos de manipulación (jailbreaks).
- **Filtro de PII:** Protege la información personal sensible.
- **Reglas Deterministas:** Bloquea términos prohibidos o patrones maliciosos.
- **Validación de Salida:** Asegura que las respuestas se alineen con la marca.
- **Intervención Humana:** La barrera definitiva para acciones de alto riesgo.

05

Tu éxito

Tu hoja de ruta hacia el éxito

somos

merkur.ia