

Mi stack exacto de AI. 11 herramientas. Un solo sistema.

CATEGORÍA · PANORAMA IA, INVESTIGACIÓN Y TRÁFICO

El stack completo que uso en 2026 para ejecutar solo y mover como equipo. Cada herramienta explicada: para qué sirve, cómo la configuro, cuándo la uso y cómo se conectan entre sí.

QUÉ VAS A ENCONTRAR

- 01 Las 11 herramientas con explicación detallada de cada una
- 02 Por qué cada herramienta está en el stack y cuál es su rol
- 03 Cómo se conectan entre sí para formar un sistema
- 04 La regla del 2do monitor: cómo duplicar el output sin sumar equipo
- 05 Por dónde empezar si no tienes ninguna todavía

FICHA DEL RECURSO

Qué es, para qué sirve y cómo se implementa

FUNCIÓN

El stack completo que uso en 2026 para ejecutar solo y mover como equipo. Cada herramienta explicada: para qué sirve, cómo la configuro, cuándo la uso y cómo se conectan entre sí.

PARA QUÉ SIRVE

Es el más efectivo. La mayoría de la gente usa una o dos herramientas de AI de forma aislada. Abren ChatGPT, hacen una pregunta, cierran. Eso no es un sistema. Eso es usar una calculadora para reemplazar a un equipo de finanzas. Un stack es diferente. Cada herramienta tiene un rol específico. Algunas son para velocidad, otras para privacidad, otras para planificación. Cuando las conectas correctamente, el output que produces solo es indistinguible del de un equipo de 5 personas. Este es el stack que uso hoy. No es el más barato ni el más caro. Es el que mejor funciona para ejecutar rápido, proteger datos sensibles y no depender de una sola herramienta. 11 tools

DÓNDE APLICARLO

Cubre: Las 11 herramientas con explicación detallada de cada una; Por qué cada herramienta está en el stack y cuál es su rol; Cómo se conectan entre sí para formar un sistema; La regla del 2do monitor: cómo duplicar el output sin sumar equipo; Por dónde empezar si no tienes ninguna todavía.

CÓMO IMPLEMENTARLO

- 1 Por Qué Importa: No es el stack más caro.
- 2 Antes De Empezar: No instales todo el mismo día.
- 3 Herramienta 01 · Codex App: Código generado antes de que termines de formularlo.
- 4 Herramienta 02 · Hermes Agent: El agente que trabaja mientras duermes.
- 5 Herramienta 03 · Openclaw: Múltiples agentes.
- 6 AI privada. Sin internet.
- 7 Herramienta 05 · Chatgpt 5.5 Pro: Planificación con el mejor razonamiento disponible.

POR QUÉ IMPORTA

No es el stack más caro.

Es el más efectivo. La mayoría de la gente usa una o dos herramientas de AI de forma aislada. Abren ChatGPT, hacen una pregunta, cierran. Eso no es un sistema. Eso es usar una calculadora para reemplazar a un equipo de finanzas. Un stack es diferente. Cada herramienta tiene un rol específico. Algunas son para velocidad, otras para privacidad, otras para planificación. Cuando las conectas correctamente, el output que produces solo es indistinguible del de un equipo de 5 personas. Este es el stack que uso hoy. No es el más barato ni el más caro. Es el que mejor funciona para ejecutar rápido, proteger datos sensibles y no depender de una sola herramienta. 11 tools 1 persona UN SOLO SISTEMA OUTPUT DE EQUIPO

ANTES DE EMPEZAR

No instales todo el mismo día.

La tentación es instalar las 11 herramientas en una tarde y quedarte paralizado sin saber por dónde empezar. No hagas eso. Cada herramienta requiere configuración y tiempo de práctica para que entre al flujo de trabajo. La recomendación es empezar por la capa de ejecución (Codex + Claude Code) en la primera semana, agregar la capa de inteligencia (OpenClaw + Gemma 4) en la segunda, y construir el sistema completo del 2do monitor en la tercera.

ESTRUCTURA DEL STACK

CAPA DE EJECUCIÓN Codex App + Hermes Agent
CAPA DE INTELIGENCIA OpenClaw + Gemma 4
CAPA DE ESTRATEGIA ChatGPT 5.5 Pro
CAPA DE BUILD Claude Design + Claude Code
SISTEMA OPERATIVO Image Gen 2 + Spotify + 2do monitor

HERRAMIENTA 01 · CODEX APP

Código generado antes de que termines de formularlo.

Codex corre 5.5 medium en modo rápido. Es la herramienta para sprints de código corto donde la velocidad importa más que la profundidad. Le das el contexto del problema en una línea y genera la solución. No es para proyectos complejos de arquitectura. Es para las 50 veces al día que necesitas una función, un script, un bloque de código específico sin abrir un editor completo y perder el contexto de lo que estás haciendo.

CUÁNDO USARLO

- Scripts rápidos de automatización
- Funciones específicas en cualquier lenguaje
- Debugging de bloques de código cortos
- Conversión de formato (JSON, CSV, SQL, etc.)
- Generación de regex, queries, comandos de terminal

HERRAMIENTA 02 · HERMES AGENT

El agente que trabaja mientras duermes.

Hermes es un agente autónomo powered by modelo local. Cero costos de API, cero datos saliendo de tu máquina. Ejecuta flujos complejos con razonamiento secuencial: puede tomar decisiones, iterar sobre resultados y completar tareas en múltiples pasos sin que estés presente. La diferencia entre Hermes y un chat normal es que Hermes no espera tu siguiente mensaje. Lo configuras con un objetivo, le das las herramientas que necesita, y trabaja solo hasta completarlo. Lo configuras una vez y trabaja solo.

CASOS DE USO REALES

- Procesar y clasificar documentos en batch
- Investigación autónoma con múltiples fuentes
- Análisis de datos que no pueden salir de tu máquina
- Flujos de QA automatizados sobre archivos locales
- Cualquier tarea repetitiva que requiera razonamiento

HERRAMIENTA 03 · OPENCLAW

Múltiples agentes.

Un solo operador. OpenClaw usa ChatGPT 5.5 con OAuth integrado. Su función principal es coordinar varios agentes especializados en paralelo desde una sola interfaz. En vez de abrir 4 chats separados con diferentes contextos, tienes un orquestador que delega tareas y consolida los resultados. Piénsalo como el gerente de operaciones del stack. No ejecuta directamente: asigna, monitorea y consolida. Es la herramienta que convierte 4 agentes individuales en un equipo coordinado.

DIFERENCIA CON UN CHAT NORMAL

Chat normal: una tarea, un modelo, un resultado

OpenClaw: una tarea compleja dividida entre agentes especializados

Ejemplo: investigar un mercado, analizar competencia, redactar el reporte y formatearlo — todo en paralelo, consolidado en un solo output.

HERRAMIENTA 04 · GEMMA 4 EN MAC MINI

AI privada. Sin internet.

Sin suscripción. Gemma 4 de Google corriendo local en Mac Mini vía Ollama. 9.6 GB en disco, cero costos de API, cero datos saliendo del hardware. Para trabajo con información sensible de clientes, contratos, datos financieros o cualquier cosa que no puede ir a servidores externos. El modelo más fuerte disponible localmente hoy. En benchmarks independientes compite con versiones anteriores de GPT-4. La diferencia con los modelos en la nube: lo que le dices a Gemma 4 se queda en tu máquina. Punto.

INSTALACIÓN EN 2 PASOS

```
# 1. Instalar Ollama
curl -fsSL https://ollama.ai/install.sh | sh
# 2. Descargar y correr Gemma 4
ollama run gemma4
Cierra Chrome y apps pesadas antes de correr el modelo.
```

HERRAMIENTA 05 · CHATGPT 5.5 PRO

Planificación con el mejor razonamiento disponible.

ChatGPT 5.5 Pro tiene una sola función en el stack: planificación estratégica. La semana se diseña aquí. Los movimientos se mapean aquí. Las prioridades se bloquean aquí antes de que empiece la ejecución. La regla es simple: para planificar usas el modelo más inteligente disponible. Para ejecutar usas los más rápidos. Usar tu herramienta de planificación para tareas de ejecución es desperdiciar su capacidad de razonamiento en trabajo que no lo necesita.

LO QUE VA AQUÍ

- Diseño de la semana de trabajo
- Estrategia de contenido y ángulos de video
- Análisis de problemas complejos sin solución obvia
- Mapeo de movimientos comerciales (propuestas, outreach)
- Decisiones que van a impactar las próximas 4 semanas

HERRAMIENTAS 06 · 07 · CLAUDE DESIGN + CLAUDE CODE

Del mockup al código sin cambiar de herramienta.

Claude Design genera el diseño de front end desde una descripción. Claude Code convierte ese diseño en código funcional. Son el mismo ecosistema, el contexto pasa automáticamente entre los dos. El flujo es: describes lo que quieres en Claude Design, obtienes el mockup completo, lo pasas a Claude Code y obtienes el código listo para producción. Sin re-explicar el contexto. Sin cambiar de herramienta. Sin perder el hilo. Lo que antes era trabajo de 2 personas (diseñador + desarrollador) en 2 días, ahora es una persona en 2 horas.

EL FLUJO EXACTO

1. Describir el componente o página en Claude Design
2. Revisar y aprobar el mockup generado
3. Pasar a Claude Code con el contexto del diseño
4. Claude Code genera el HTML/CSS/JS listo
5. Ajustes menores y a producción

HERRAMIENTA 08 · CHATGPT IMAGE GEN 2

Creatividad visual sin límite operativo.

ChatGPT Image Gen 2 va para más de lo que imaginas al principio. No solo imágenes decorativas: mockups de producto, assets para presentaciones de ventas, material de contenido, visualizaciones de datos, portadas de documentos, thumbnails. La velocidad de producción visual que antes requería un diseñador de tiempo completo o dos días en Canva, ahora es un prompt. El estándar de calidad de los assets visuales que produces sube sin que suba el tiempo invertido.

USOS REALES EN EL DÍA A DÍA

- Portadas de PDFs y lead magnets
- Assets para carruseles de Instagram
- Mockups de producto para propuestas
- Imágenes de apoyo para presentaciones de ventas
- Thumbnails para videos y contenido
- Visualizaciones de conceptos abstractos

HERRAMIENTAS 09 · 10 · 11 · EL SISTEMA OPERATIVO

Spotify + 2do monitor.

No es accesorio. Es infraestructura. Spotify en lofi bangers no es opcional. El estado de flujo tiene un soundtrack. La concentración profunda necesita un soundtrack que bloquee el ruido sin interrumpir el pensamiento. Sin música en el setup estás trabajando al 70%. El 2do monitor es donde viven todos los agentes, 24/7. Nunca los cierras. Trabajas en el primer monitor. Prompteas constantemente el segundo. No es multitasking — el 2do monitor es tu equipo. Tú ejecutas en el primero, ellos trabajan en el segundo. La regla del sistema: cualquier tarea que puedas delegarle a un agente, se la delegas. Tu tiempo va únicamente a lo que solo tú puedes hacer: tomar decisiones, crear relaciones, generar ideas.

LA CONFIGURACIÓN DEL 2DO MONITOR

- Izquierda: OpenClaw con contexto del proyecto activo
- Centro: Claude Code o Hermes en tarea en curso
- Derecha: ChatGPT 5.5 Pro abierto para consultas rápidas
- Regla: nunca cierras el 2do monitor durante la jornada.
- Siempre hay algo corriendo ahí.

CHECKLIST DE IMPLEMENTACIÓN

Por dónde empezar.

SEMANA 1 · CAPA DE EJECUCIÓN

Instalar Codex App y configurar con tu contexto de trabajo
Instalar Ollama y descargar Gemma 4 (ollama run gemma4)
Configurar Hermes Agent con modelo local
Primer flujo autónomo con Hermes en una tarea real

SEMANA 2 · CAPA DE INTELIGENCIA Y BUILD

Configurar OpenClaw con ChatGPT 5.5 OAuth
Primera sesión de planificación semanal en ChatGPT 5.5 Pro
Primer proyecto completo con Claude Design + Claude Code
Activar ChatGPT Image Gen 2 para assets de contenido

SEMANA 3 · EL SISTEMA COMPLETO

Conectar el 2do monitor con los agentes siempre activos
Playlist de lofi configurada y lista antes de arrancar
Establecer la regla: todo lo delegable va al 2do monitor
Revisar output de la semana vs semanas anteriores

SIGUIENTE PASO

El stack sin sistema son solo apps instaladas.

Las herramientas no cambian nada por sí solas. Lo que cambia es el sistema: saber qué le pides a cada una, cuándo, y cómo conectar los outputs entre sí. Cada semana publico casos de uso reales, flujos

MÁS GUÍAS GRATIS EN [377CREATIVE.COM/RECURSOS](https://377creative.com/recursos)