Agentes de código: lo básico

Cómo trabaja un agente como Claude Code y cómo mantenerlo bajo control, explicado para profesionales sanitarios.

José Manuel Ramírez Sánchez

jmramirez.engineer · noviembre de 2026

EL CASO

Un informe trimestral de HbA1c

Una carpeta con analíticas anonimizadas y ficticias. Pedimos al agente el informe del trimestre y lo hacemos siempre en cuatro pasos.

1

Explorar

Lee la carpeta y describe los datos. No cambia nada.

→

2

Planificar

Propone los pasos. Tú los corriges y apruebas.

→

3

Ejecutar

Escribe y ejecuta los scripts sobre copias.

→

4

Revisar

Compruebas el resultado. La responsabilidad es tuya.

En Claude Code 101 este flujo se llama explore → plan → code → commit.

Qué es un agente

Un chat

Contesta una vez. Si algo falla, tienes que darte cuenta tú.

Un agente

Repite un bucle hasta cumplir el objetivo: piensa, actúa con una herramienta y observa el resultado.

Vuelta	Piensa	Actúa	Observa
1	¿Qué archivos hay?	ls datos/	12 CSV de analíticas
2	Los uno con un script	python unir.py	Error: no existe «Fecha»
3	Corrijo el nombre de la columna	python unir.py	tabla_unida.csv creada

La carpeta de trabajo

```
proyecto-hba1c/
├─ AGENTS.md reglas
├─ datos/
│ ├─ originales/ no se tocan
│ └─ trabajo/ copias
├─ scripts/
└─ resultados/ se entrega
```

Es el escritorio del agente

Solo ve lo que hay dentro. Para leer o escribir fuera, te pide permiso.

Lo primero que lee: AGENTS.md

Y la estructura le dice qué es cada cosa.

Originales aparte, copias para trabajar, resultados aparte.

Los modos de permisos

Shift+Tab cambia de modo en mitad de la sesión.

Modo	Qué hace	Cuándo
Pedir permiso	Pregunta antes de editar un archivo o ejecutar un comando	Por defecto, y siempre al empezar
Aceptar ediciones	Edita sin preguntar; los comandos aún piden permiso	Cuando ya has aprobado el plan
Modo plan	Solo lee y propone un plan; no edita nada	Antes de tareas con varios pasos

Con datos clínicos, empieza pidiendo permiso y lee cada acción antes de aprobarla.

AGENTS.md y la regla CORTO

Un archivo de texto con las reglas del proyecto.
Se carga **entero** al empezar **cada** sesión.

Es contexto, no una orden garantizada: cuanto más corto y concreto, mejor se sigue.

Claude Code lee CLAUDE.md y, si no existe, AGENTS.md. /init crea uno inicial.

- **Concreto:** instrucciones que se pueden comprobar
- **Operativo:** qué hacer, no la historia
- **Revisado:** cada mes, borra lo obsoleto
- **Tamaño:** menos de 200 líneas
- **Ordenado:** títulos claros; lo específico, en otro archivo

Un AGENTS.md para el caso

```
# Proyecto HbA1c
## Reglas
- Responde en español.
- Fechas en AAAA-MM-DD.
- No modifiques datos/originales/.
- Antes de borrar filas, pregunta.
## Datos
- Anonimizados. Columnas: DATOS.md
## Salidas
- Tablas y gráficas en resultados/.
## Compact Instructions
- Conserva columnas y decisiones.
```

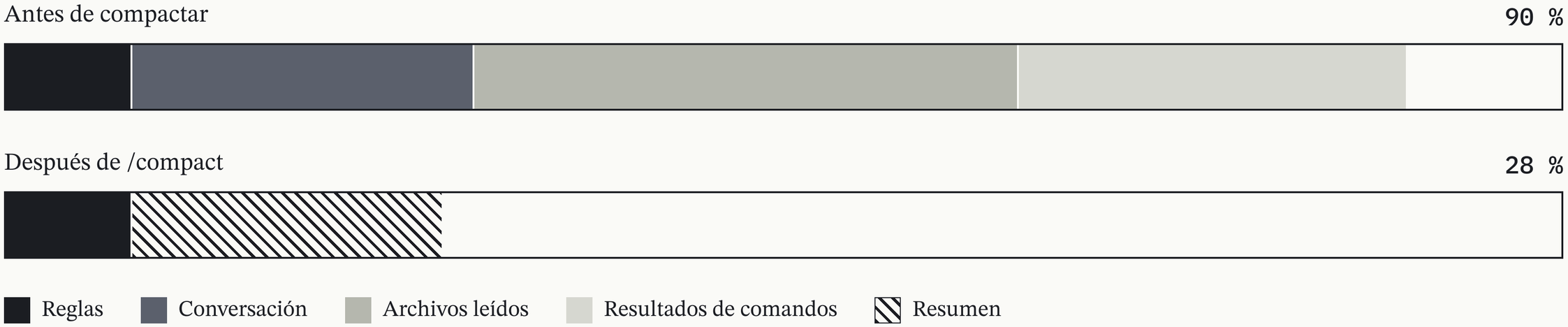
De vago a concreto

Vago	Concreto
«Ordena bien los datos»	Fechas en AAAA-MM-DD
«Ten cuidado»	Antes de borrar filas, pregunta
«Sé breve»	Respuestas de menos de 10 líneas

Nunca escribas aquí datos de pacientes ni contraseñas: el archivo se envía al modelo en cada sesión.

La ventana de contexto

Todo lo que el modelo ve a la vez. Tiene un tamaño fijo, medido en tokens.



Se conserva: las reglas enteras, tus peticiones y lo importante.

Puede perderse: instrucciones largas del principio y resultados antiguos.

Memoria a corto y a largo plazo

El modelo no recuerda nada entre sesiones. Lo que perdura está en archivos.

Momento	Ventana (corto plazo)	AGENTS.md (largo plazo)
Sesión 1	«Pon las fechas como AAAA-MM-DD»	Sin cambios
Sesión 1	«Apúntalo en AGENTS.md»	+ Fechas en AAAA-MM-DD
/clear	Se vacía por completo	Sigue igual
Sesión 2	Empieza vacía	Se carga: ya conoce el formato

Si quieres que lo recuerde mañana, escríbelo en un archivo.

Cinco comandos básicos

Comando	Qué hace	Cuándo
/init	Mira la carpeta y crea un archivo de reglas inicial	Al empezar un proyecto
/clear	Conversación nueva; las reglas siguen ahí	Al cambiar de tarea
/compact	Resume y libera ventana; puedes decir qué conservar	Cuando la ventana se llena
/rewind	Vuelve a un punto anterior: código, conversación o ambos	Si un paso sale mal
/context	Muestra qué ocupa la ventana	Si parece ignorar una regla

Una sesión por tarea: /clear al cambiar, /compact al llenarse, /rewind si algo falla.

Atajos que ahorran tiempo y ventana

Atajo	Qué hace
Esc	Detiene al agente y conserva lo hecho; lo rediriges sin empezar de cero
Shift+Tab	Cambia de modo de permisos, incluido el modo plan
@archivo	Señala un archivo concreto: lee solo ese, no toda la carpeta
/btw	Pregunta lateral que no entra en el historial ni ocupa ventana
claude --resume	Retoma una sesión anterior; --continue abre la última

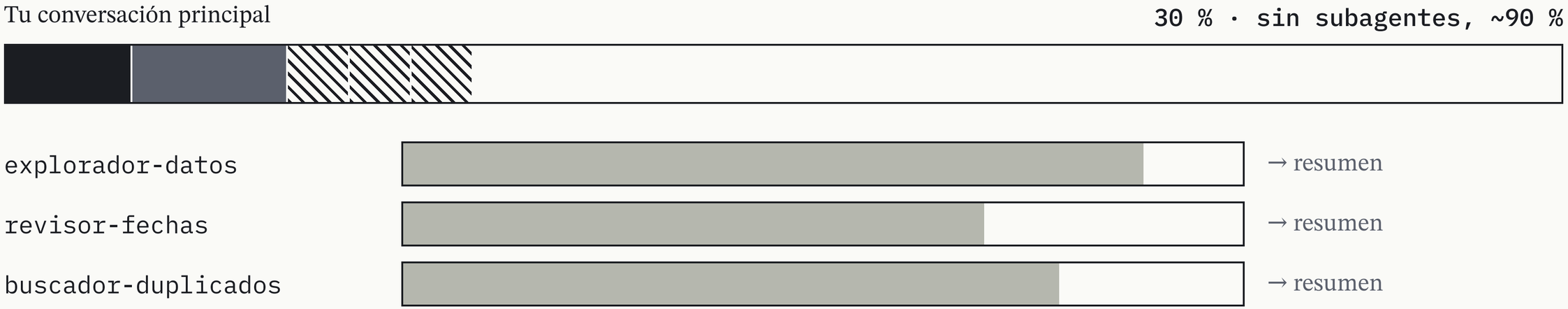
Si tras dos correcciones sigue fallando: /clear y escribe una petición mejor.

EL CASO, EN UNA SESIÓN

Explorar, planificar, ejecutar, revisar

Paso	Qué pides	Qué pasa	Qué haces
Explorar	«Lee datos/trabajo/ y describe las columnas. No cambies nada.»	Resume 12 CSV	Compruebas que entiende los datos
Planificar	Shift+Tab, modo plan: «Propón cómo hacer el informe»	Plan de cuatro pasos	Lo corriges y lo apruebas
Ejecutar	«Adelante con el plan»	Borra filas que no eran duplicados	/rewind y «solo duplicados exactos»
Ejecutar	Sigue con la gráfica	La ventana llega al 85 %	/compact conserva las columnas
Revisar	«Compara los totales con los originales»	Cuadran	Lo revisas tú y va a resultados/

Subagentes: ayudantes con ventana propia



Cómo funciona: cada uno lee mucho en su ventana y a la tuya solo llega un resumen.

Cuándo: tareas que leen mucho. Para una tarea pequeña no compensa.

Un subagente es un archivo .md

```
.claude/agents/revisor-fechas.md
---
name: revisor-fechas
description: Revisa formatos
de fecha en datos/trabajo/
tools: Read, Grep
model: haiku
---
Comprueba que las fechas están
en AAAA-MM-DD. Devuelve un
resumen de 10 líneas como máximo.
No modifiques ningún archivo.
```

Campo	Para qué
description	Cuándo delegar en él
tools	Qué puede hacer: aquí, solo leer y buscar
model	Un modelo pequeño basta para leer
instrucciones	Qué hacer y cuánto devolver

Permisos mínimos: un revisor solo necesita leer, no escribir. Y pídele siempre un resumen corto.

09 · REUTILIZAR

Skills: recetas que se cargan a la carta

Una carpeta con un SKILL.md y, si hace falta, scripts y plantillas. Siempre a la vista solo está su descripción.

Al empezar la sesión **12 %**



Al pedir «haz el informe de HbA1c del trimestre» **37 %**



Reglas · Descripciones de las skills · La skill en uso, entera

`informe-hba1c · en uso` `limpiar-fechas` `cartel-congreso`

Anatomía de una skill

```
.claude/skills/informe-hba1c/
├─ SKILL.md
├─ plantilla.xlsx
└─ scripts/grafica.py
---
name: informe-hba1c
description: Crea el informe
trimestral de HbA1c. Úsala
si piden ese informe.
---
1. Lee la tabla de resultados/.
2. Calcula media y mediana.
3. Ejecuta scripts/grafica.py.
```

Dónde vive	Quién la usa
~/.claude/skills/	Tú, en todos tus proyectos
.claude/skills/	Quien trabaje en ese proyecto

Se activa sola si tu petición encaja con la descripción, o a mano con /informe-hba1c.

SKILL.md corto: menos de 500 líneas; el detalle, en archivos aparte.

La descripción lo decide todo: di qué hace y cuándo usarla.

08 + 09 · COMPARATIVA

¿AGENTS.md, skill o subagente?

Pieza	Cuándo se carga	Para qué sirve	Ejemplo
AGENTS.md	Siempre, entero	Reglas del proyecto	«Fechas en AAAA-MM-DD»
Skill	A la carta, si encaja	Un procedimiento repetible	El informe trimestral de HbA1c
Subagente	Cuando se le delega	Trabajo pesado aparte; devuelve un resumen	Revisar 40 CSV sin llenar tu ventana

Lo que siempre vale, a AGENTS.md. Lo que haces a veces, a una skill. Lo que ensucia la ventana, a un subagente.

Local o cloud

Aspecto	Local	Cloud
Dónde corre	Tu ordenador	Una máquina aislada de Anthropic
Archivos	Los tuyos, reales	Una copia de tu repositorio de GitHub
Portátil cerrado	Se detiene	Sigue trabajando
Revisión	Ves cada cambio al momento	Revisas al terminar, también desde el móvil
Ideal para	Datos y tareas interactivas	Código y tareas largas

Disponibilidad según versión y plan. Con configuraciones HIPAA o de retención cero, las sesiones cloud pueden no estar disponibles.

EN LOCAL Y EN CLOUD

Lo que el agente lee se envía al modelo.

Anonimiza antes. Datos de pacientes: en local y anonimizados. En cloud, solo código o datos ficticios.

Qué modelo usar: tres reglas

1

El más pequeño que resuelva la tarea

Empieza con Haiku o Sonnet. Si falla dos veces, sube a Opus o Fable con /model.

2

Escribir cuesta más que leer

En Claude, cada token de salida cuesta 5 veces uno de entrada. Pide respuestas cortas.

3

No uses modelos antiguos

Cada generación da más por menos, y las antiguas se retiran: un script con un modelo fijo deja de funcionar.

A 11 de noviembre de 2026: Haiku, Sonnet, Opus y Fable, de más rápido a más capaz. Precios actuales en anthropic.com/pricing.

11 · MODELOS · TERMÓMETRO

El caso HbA1c, con el modelo que toca

1 · EXPLORAR

Leer la carpeta y listar columnas (subagente)

■□□□

Claude: **Haiku**

Abierto: 1–8B

2 · EXTRAER

Sacar la HbA1c de 200 informes

■□□□

Claude: **Haiku**

Abierto: 1–8B

3 · REDACTAR

Resumen del trimestre a partir de la tabla

■■□□

Claude: **Sonnet**

Abierto: 8–32B

4 · EJECUTAR

Informe completo con scripts y gráficos

■■■□

Claude: **Opus o Sonnet**

Abierto: 70B+

5 · REVISAR

¿Se sostienen las conclusiones?

■■■■

Claude: **Fable u Opus**

Abierto: 70B+, con dudas

Y DESPUÉS La decisión clínica es tuya: ningún nivel del termómetro llega ahí.

Orientativo, a 11 de noviembre de 2026 · abierto = en local, tamaño en parámetros · prueba con tu tarea

PRÁCTICA

Cuatro retos con la carpeta de ejemplo

```
agentes-practica/
├─ AGENTS.md
├─ DATOS.md
├─ datos/originales/
├─ datos/trabajo/
├─ resultados/
└─ .claude/
├─ agents/revisor-fechas.md
└─ skills/informe-hba1c/
```

1. Abre la carpeta con Claude Code y escribe /context. ¿Se cargó AGENTS.md?
2. En modo plan, pide un plan para el informe. Corrígelo antes de aprobar.
3. Delega la revisión de fechas en el subagente y compara tu ventana antes y después.
4. Lanza /informe-hba1c y comprueba los totales frente a los originales.

RESUMEN

Ocho ideas para llevarse

1 **Un agente actúa en bucle**
Tú pones el objetivo y revisas.

2 **Carpeta ordenada y permisos**
Originales aparte; aprueba cada acción al principio.

3 **Un AGENTS.md CORTO**
Reglas comprobables, menos de 200 líneas.

4 **La ventana es finita**
Lo que debe perdurar, a un archivo.

5 **Explorar, planificar, ejecutar, revisar**
Con /clear, /compact, /rewind y Esc a mano.

6 **Subagentes y skills**
Lo pesado se delega; lo repetido se reutiliza.

7 **Anonimiza siempre**
En local y en cloud, lo que lee va al modelo.

8 **El modelo más pequeño que sirva**
Y nunca uno antiguo.