

PROGRAMA DE SEIS SEMANAS PARA EQUIPOS DE DESARROLLO

Programa de productividad con IA

Que el equipo saque más trabajo de la IA, gaste menos tokens y lo pueda demostrar

Seis semanas con Claude Code y GitHub Copilot: mido el punto de partida, hacemos el workshop, reviso los pull requests del equipo y vuelvo a medir. Al final el equipo tiene un método común y dos mediciones que se pueden comparar.

6

semanas, con dos mediciones

5 a 8 h

de workshop presencial, según el equipo

2

herramientas: Claude Code y GitHub Copilot

5

semanas de seguimiento en remoto

EL PROBLEMA DE PARTIDA

El equipo usa IA a diario, cada uno a su manera, y nadie sabe si le sale a cuenta

Uno usa el modelo más caro para renombrar una variable. Otro lleva una sesión abierta desde el lunes que reenvía todo su historial en cada mensaje. Otro deja al agente en bucle toda la tarde. Y los pull requests llegan en verde porque el mismo agente que escribió el código decidió que estaba bien.

La factura de tokens crece y todo el equipo tiene la impresión de ir más rápido, pero nadie ha medido nada. La documentación de costes de Claude Code da una media de unos 13 dólares por desarrollador y día activo, y lo primero es saber dónde está el equipo.

Por eso aquí no prometo una cifra. Mido el antes y el después con indicadores que elegimos juntos en la llamada, y lo que cambie, o no cambie, queda por escrito.

Lo que el equipo se lleva en seis semanas

01 Más trabajo terminado

El modelo caro solo donde hay que pensar, y agentes en paralelo con encargos que no se pisan.

02 Menos tokens

Una caché que no se rompe, contexto bajo demanda y sesiones que terminan solas.

03 Confianza en producción

La IA prueba lo que genera y otro agente intenta romperlo antes del merge.

04 Un método común

La misma configuración, los mismos agentes y las mismas reglas en todos los equipos.

05 Dos mediciones comparables

Los mismos indicadores al principio y al final, para ver qué ha cambiado.

Todo lo trabajamos en Claude Code y en GitHub Copilot. El workshop de la semana 1 es el curso de productividad con IA, completo; lo demás es lo que lo rodea para que el cambio se vea.

RESULTADO DEL PROGRAMA

Qué tiene el equipo al cabo de seis semanas

Cada resultado se comprueba con algo que se puede enseñar a dirección, y no con una impresión.

01

Más trabajo terminado con el mismo equipo

Modelo por ronda, agentes que trabajan en paralelo con encargos disjuntos y sesiones con límite de turnos y de gasto que terminan solas.

Evidencia: la misma tarea con un modelo para todo y con modelo por ronda, con el coste de cada una.

02

Menos tokens por tarea terminada

Una caché que no se rompe, contexto que se carga bajo demanda y salidas filtradas antes de llegar al modelo.

Evidencia: tokens por tarea medidos al principio y al final, con el mismo método.

03

Confianza en lo que va a producción

La IA prueba su propio código: test que falla primero, tests protegidos y una suite que se ejecuta antes de dar nada por terminado. Un revisor en contexto limpio busca lo que el autor dio por bueno.

Evidencia: comentarios por escrito sobre dos o tres pull requests del equipo hechos con agentes.

04

Una forma de trabajar común entre equipos

Instrucciones, permisos, skills, hooks y agentes configurados en los repositorios del equipo y acordados entre todos.

Evidencia: configuraciones en una rama del repositorio de la empresa, revisadas por el equipo.

05

Cinco acciones priorizadas para dirección

Un informe de tres a cinco páginas con lo que vi, lo que trabajamos y qué hacer después, con una persona responsable por acción.

Evidencia: el informe, entregado en la semana 2.

Lo que pasa cada semana y quién lo hace

Yo hago casi todo el trabajo. Al equipo le pido tiempo para la encuesta, para el día de workshop y para pasarme los pull requests.

Semana 0

Diagnóstico y primera medición

El equipo contesta una encuesta de 10 minutos sobre herramientas, hábitos y frustraciones. Yo miro los ficheros de instrucciones y permisos del equipo, y calculamos juntos la línea base de los indicadores que elijamos.

Entregable: Línea base escrita y laboratorios adaptados al stack del equipo.

Semana 1

El workshop

Entre cinco y ocho horas en persona sobre modelo por ronda, caché, contexto bajo demanda, ejecución finita, orquestación y verificación, con el último bloque en el repositorio de cada participante. Dejo las configuraciones en una rama del repositorio de la empresa.

Entregable: Configuraciones en rama y repositorio de laboratorios.

Semana 2

Informe a dirección

Preparo el informe de tres a cinco páginas con lo que vi en el diagnóstico, lo que trabajamos, los riesgos en las configuraciones (sin auditar nada) y cinco acciones priorizadas.

Entregable: Informe a dirección.

Semanas 1 a 5

Canal de dudas y revisión de pull requests

Un canal con el equipo, abierto hasta el final de la semana 5, para dudas de configuración y de uso. Reviso por escrito dos o tres pull requests que me señale el equipo.

Entregable: Respuestas por escrito y comentarios sobre los pull requests.

Semana 6

Revisión y segunda medición

Repetimos las mediciones de la semana 0 con el mismo método y dedicamos 90 minutos en remoto a ver qué ha cambiado, qué se ha atascado y qué falta.

Entregable: Comparativa antes y después, y siguiente paso acordado.

La semana 1 es el workshop completo. Su temario, prácticas y agenda están en danielalfocea.com/formacion/workshop-productividad-con-ia

Qué se mide y por qué no prometo nada

Elegimos entre cuatro y seis indicadores en la llamada, según lo que se pueda extraer de los sistemas de la empresa sin montar nada nuevo. Estos son candidatos y la lista no está cerrada. Con seis semanas no sale una conclusión estadística, y el informe lo dice así.

Tokens por tarea terminada

Lo que cuesta cada cambio aceptado, sacado de /usage o del panel de la herramienta. Es el indicador más barato de medir y el que más se mueve al ordenar el contexto.

Tiempo entre abrir y fusionar un pull request

Sale del historial del repositorio. No distingue quién ha escrito qué, pero sí si el flujo se atasca en la revisión.

Cambios revertidos o corregidos después

Cuántos cambios fusionados se revierten o necesitan un arreglo en los días siguientes. Dice si lo que sube a producción aguanta.

Esfuerzo de revisión

Rondas de comentarios y tamaño de los diffs en pull requests hechos con agente. Un diff enorme que alguien aprueba sin leer es una señal.

Percepción del equipo

La misma encuesta corta al principio y al final. La pregunta clave es cuánto tiempo creen ahorrar, y sirve para contrastarla con lo medido.

Workshop o programa

NECESIDAD

WORKSHOP

PROGRAMA

Que el equipo comparta una forma de trabajar con agentes

Con el workshop vale: un día, configurado en los repositorios del equipo.

También vale, y suma cinco semanas de seguimiento.

Saber si ha cambiado algo medible

El workshop enseña a medir la sesión, pero no hay segunda medición.

Para eso está el programa: dos mediciones comparables.

Rendir cuentas ante dirección

Cada participante sale con sus números de antes y después.

Añade el informe de tres a cinco páginas con cinco acciones priorizadas.

Corregir el rumbo con el código real del equipo

Solo en el último bloque del día.

Reviso dos o tres pull requests del equipo en las semanas siguientes.

Los temas, una vez vistos los beneficios

El contenido completo es el del workshop de la semana 1. Cada tema se explica en Claude Code y en GitHub Copilot, y en las semanas siguientes lo aplicamos sobre los pull requests del equipo.

01 Ahorro: modelo, caché y contexto

Menos tokens sin perder calidad.

- Modelo y esfuerzo por ronda: planificar, explorar, ejecutar y revisar no piden lo mismo.
 - Qué rompe la caché, rebobinar frente a compactar y salidas filtradas con hooks.
 - Contexto bajo demanda, sin RAG: un fichero índice, reglas por ruta y documentación que el agente recorre solo.
-

02 Ejecución finita y orquestación

Un agente sin límite de turnos ni de gasto sale caro.

- Límite de turnos y de gasto, condiciones de parada y el bucle Ralph con tope de iteraciones.
 - Worktrees por agente, panel de expertos con roles disjuntos, juez y agente adversario.
 - Cuándo no montar un sistema multiagente: Anthropic midió un 90 % más de rendimiento a cambio de unas quince veces más tokens.
-

03 Verificación

Un cambio está terminado cuando lo confirma una comprobación.

- Test que falla primero, tests protegidos contra ediciones del agente y una puerta de parada con la suite.
 - Revisor en contexto limpio y validación cruzada entre modelos distintos.
 - Un hook detiene lo que una instrucción no garantiza.
-

04 Configuración y herramientas

Cada mecanismo resuelve un problema distinto.

- Modo plan, skills y comandos, hooks, MCP y permisos: cuándo usar cada uno.
- Configuración gestionada para la organización y un catálogo propio de skills y agentes.
- Una sola carpeta de configuración que leen Claude Code y Copilot.

MATERIAL QUE SE QUEDA EN LA EMPRESA

Cuando acaba el programa, queda esto

Todo está hecho para vivir en el repositorio y en los documentos de la empresa, y no en una carpeta de formación.

Línea base y segunda medición

Los dos conjuntos de indicadores y el método para repetirlos.

Configuraciones en rama

Instrucciones, permisos, hooks y subagentes para Claude Code y Copilot, listos para revisar y fusionar.

Repositorio de laboratorios con soluciones

Licencia de uso interno durante un año.

Informe a dirección

De tres a cinco páginas, con cinco acciones priorizadas y responsable de cada una.

Recetas de ejecución finita

Bucle Ralph con su fichero de progreso, límites de turnos y de gasto y condiciones de parada.

Comentarios sobre los pull requests del equipo

Por escrito, con qué se acepta y por qué.

QUÉ INCLUYE ADEMÁS

Tres piezas que van con el workshop

Son las mismas que lleva el workshop suelto. En el programa se entregan dentro del calendario de seis semanas.

+1

Encuesta previa al equipo

Diez minutos por persona para adaptar los laboratorios al stack y a las herramientas del equipo.

+2

Configuraciones en una rama del repositorio de la empresa

Ficheros de instrucciones, permisos, hooks y subagentes con su modelo y su límite de turnos, listos para revisar y fusionar.

+3

Repositorio de laboratorios con soluciones

Enunciados, ficheros de partida y soluciones, con licencia de uso interno durante un año.

Para equipos que quieren saber si la IA les sale a cuenta

Encaja especialmente con

- Equipos que ya usan Claude Code o GitHub Copilot a diario y no tienen datos de qué efecto tiene.
- Tech leads y engineering managers que tienen que explicar a dirección qué se ha ganado con las licencias y los tokens.
- Equipos de plataforma que quieren una configuración y unos agentes comunes para varios equipos.
- Equipos que prefieren cambiar una cosa cada vez y comprobar el efecto antes de seguir.

No es la formación adecuada si

- La empresa busca un solo día de formación sin seguimiento. Para eso está el workshop.
- Lo que se espera es una cifra de ahorro garantizada. Yo no la doy y desconfiaría de quien la dé.
- Lo que hace falta es una auditoría del código o de las configuraciones.
- El equipo todavía no usa un asistente de código sobre un repositorio real.

QUIÉN LO IMPARTE

De la ciberseguridad de BBVA a la IA de Santander



Soy Daniel Alfocea y llevo veinte años en desarrollo y seguridad. En Banco Santander fui Principal Engineer del departamento de IA cuando se puso en marcha el área. En BBVA fui Tech Lead de ciberseguridad. En 42Crunch, en Reino Unido, fui responsable de uno de sus productos de seguridad de APIs.

He presentado investigación en RSA Conference y en OWASP AppSec EU, y he sido ponente habitual de RootedCON. El Instituto Nacional de Ciberseguridad (INCIBE) me ha llamado varias veces como juez de sus concursos. Fundé Navaja Negra y el capítulo de OWASP Madrid. Soy autor de herramientas de seguridad muy usadas, como GoLismero, Plecost y Dockerscan, y de LAPIS, una investigación publicada en arXiv sobre cómo describirle una API a un modelo con menos tokens. Con Alfonso Muñoz escribí *MCP Seguro. Guía práctica para desarrolladores, arquitectos software y tech leads*.

Este programa sale de cómo trabajo yo: mido mis sesiones, reparto el trabajo entre agentes con su modelo, su límite de turnos y su informe, y no acepto un cambio sin un test que haya visto fallar.

Lo que arreglo sin cobrar más

Las configuraciones funcionan

Si las configuraciones que dejo en el repositorio de la empresa no funcionan, hago una sesión remota gratuita para arreglarlas.

Una hora si el equipo se atasca

Si a los 30 días el equipo se ha atascado aplicando algo del curso, le dedico una hora en remoto, gratis.

Esto no es una auditoría y no prometo porcentajes de ahorro. Me comprometo a medir igual las dos veces y a contar lo que salga tal cual.

Preguntas que mejor resolver ahora

¿Cuánto cuesta?

Depende de la duración y del tamaño del equipo. Lo vemos en la primera conversación y después mando una propuesta cerrada.

¿Cuánto mejorará el equipo?

No lo sé y no me voy a inventar una cifra. Lo que entrego es una comparación del propio equipo, hecha igual al principio y al final.

¿Seis semanas bastan para ver algo?

Para algunos indicadores sí, como los tokens por tarea o el tamaño de los diffs. Para otros, como los cambios revertidos, puede que no haya datos suficientes y el informe lo dirá. Prefiero un «no se puede concluir» a una conclusión inflada.

¿Qué datos necesitas que te demos?

Las respuestas a la encuesta, acceso de lectura al historial del repositorio para los indicadores y los ficheros de instrucciones. No necesito el código fuente completo y todo lo hacemos en los portátiles y con las cuentas de la empresa.

¿Qué pasa si el equipo no tiene tiempo para el seguimiento?

El seguimiento pide poco: contestar una encuesta, señalarme dos o tres pull requests y reservar 90 minutos al final. Si algo no llega a tiempo, sigo con lo que haya y lo dejo escrito.

¿Se puede bonificar por FUNDAE?

Al pedir fecha paso la información que necesita la gestoría o el departamento de formación.

FORMATO

Seis semanas, un equipo y dos mediciones

El programa incluye el diagnóstico, el workshop completo, el informe, el seguimiento y la revisión final. La duración del workshop y el tamaño del grupo los cerramos en la primera llamada. Si solo hace falta el día, con el workshop vale.

- Participantes del mismo equipo o de equipos que trabajan juntos.
- Workshop presencial en español o en inglés y seguimiento en remoto.
- Sala, pantalla, wifi y portátiles a cargo de la empresa.
- Si hay que viajar fuera de Madrid, lo acordamos antes de cerrar la fecha.

Escríbeme a

yo@danielalfocea.com

En la llamada vemos qué se puede medir, cuántas personas participan y qué herramientas usa el equipo. Con eso mando una propuesta cerrada.

Bonificable a través de FUNDAE.

CORREO

yo@danielalfocea.com

PROGRAMA EN LA WEB

danielalfocea.com/formacion/programa-productividad-con-ia