

Antes de pegarle algo a Claude: el *checklist*

Para tenerlo a mano. Vale para Claude y para cualquier otro chat de IA.

Las 7 cosas que no se pegan

NO PEGUES

HAZ ESTO EN SU LUGAR

1. Contraseñas

Escribe [CONTRASEÑA]. Claude no la necesita para ayudarte.

2. Códigos de verificación

Los que te llegan por mensaje no se comparten con nadie. Ningún banco ni servicio te los pide por chat.

3. Claves de API y tokens

Guárdalas en un archivo `.env` y dile a Claude el nombre de la variable, nunca el valor.

4. Tarjetas y cuentas completas

Con los últimos 4 dígitos basta. Tápalos en las fotos.

5. Documentos de identidad

Tapa el número y la foto antes de mandar una imagen. Si es por un trámite, descríbelo.

6. Datos de otras personas

Cambia los nombres por "Persona A" y quita teléfonos, correos y direcciones.

7. Lo confidencial de tu trabajo

Mira primero qué permite tu empresa. Usa un ejemplo inventado con la misma forma.

Antes de pegar, 3 preguntas

- ☐ ¿Hay alguna contraseña, código o clave?
- ☐ ¿Hay datos de alguien que no me dio permiso?
- ☐ ¿Me metería en un problema si esto saliera de aquí?

Si alguna es sí, tapa o cambia ese dato antes de pegar.

Si ya lo pegaste

- **Contraseña:** cámbiala ya, y en todos los sitios donde uses la misma.
- **Clave de API:** revócala en la consola del servicio y crea una nueva.
- **Tarjeta o código de verificación:** llama a tu banco al número oficial (el de la tarjeta o el de su web) y bloquéala.
- **Borra el chat.** Sale de tu historial al momento y Anthropic dice que lo borra de sus servidores en 30 días, salvo excepciones de su política (por ejemplo, si el chat se marcó por incumplir sus normas).

Lo que comprobé (19 de octubre de 2026)

Pegué datos falsos en tres chats nuevos de Claude:

- **Una clave y una contraseña:** me avisó, pero al final de la respuesta, cuando ya estaban en el chat.
- **Una tarjeta y un código de verificación:** me avisó primero y me dijo que bloqueara la tarjeta.
- **Un chat de grupo con teléfonos y una dirección de otras personas:** hizo el resumen sin decir nada.

El aviso llega tarde o no llega. El filtro eres tú.

Dos ajustes que vale la pena mirar

- **Entrenamiento:** en Configuración > Privacidad está el interruptor de ayudar a mejorar los modelos de Claude. Si lo apagas, tus chats nuevos no se usan para entrenar. Revisa qué elegiste. El nombre exacto puede cambiar con las versiones de la app.
- **Chat incógnito:** en un chat nuevo, el icono del fantasma arriba a la derecha. No queda en tu historial ni en la memoria de Claude y no se usa para entrenar, pero Anthropic lo guarda 30 días.

Si usas Claude Code

- Claude Code manda a los servidores de Anthropic lo que tú escribes y los archivos que lee.
- Las claves van en `.env`, y puedes impedir que Claude Code lo lea con una regla en `~/claude/settings.json`:

COPIA Y PEGA

```
{
  "permissions": {
    "deny": ["Read(./env)", "Read(./env.*)"]
  }
}
```

- Ojo: la regla frena las herramientas de Claude, no cualquier programa que abra ese archivo por su cuenta. Para un bloqueo total, la documentación recomienda activar el aislamiento (sandbox).

Fuentes oficiales (consultadas el 19-oct-2026)

- Ajuste de entrenamiento:
<https://privacy.claude.com/en/articles/12109829-how-do-i-change-my-model-improvement-privacy-settings>
- Chats incógnito: <https://support.claude.com/es/articles/12260368-usar-chats-incognito>
- Cuánto se guardan los datos:
<https://privacy.claude.com/es/articles/10023548-cuanto-tiempo-almacenas-mis-datos>
- Reglas de permisos de Claude Code:
<https://code.claude.com/docs/en/settings> y
<https://code.claude.com/docs/en/permissions>
- Qué envía Claude Code: <https://code.claude.com/docs/en/sandbox-environments>

Es una lista de buenas costumbres, no asesoría legal ni de seguridad.