

Sicher arbeiten mit einem Agenten

Die Prüfung, die zwischen „ausprobieren“ und „Projekt zerlegt“ steht.

Ein Agent, der Dateien ändern und Befehle ausführen darf, ist ein mächtiges Werkzeug. Diese Liste ist das Netz darunter. Druck sie aus oder häng sie neben den Bildschirm — die ersten Wochen lohnt sich der Blick darauf.

Hinweis: Stand 2026. Die Regeln gelten unabhängig von Versionen; einzelne Befehle und Bezeichnungen können sich ändern.

Einmal einrichten

- ☐ Projekt steht unter **Git**
 - ☐ Es gibt eine **Kopie außer Haus** (GitHub o. ä., auch privat)
 - ☐ `/doctor` läuft ohne Beanstandung
 - ☐ **CLAUDE.md** existiert und ist unter 200 Zeilen
 - ☐ Abschnitt „Fallen in diesem Projekt“ angelegt
 - ☐ **Keine Zugangsdaten** in CLAUDE.md oder in versionierten Dateien
 - ☐ Tests vorhanden für das, was nicht kaputtgehen darf
 - ☐ Test-Hook eingerichtet (läuft nach Änderungen automatisch)
-

Vor jedem größeren Auftrag

- ☐ **Arbeitsstand sauber?** Alles Offene ist committet
 - ☐ Ich kann in einem Schritt zurück
 - ☐ Der Auftrag nennt: **was herauskommen soll, woran man Erfolg erkennt, was nicht kaputtgehen darf**
 - ☐ Bei mehr als einer Datei: **Plan Mode**
 - ☐ Plan gelesen — Umfang, Dateien, fehlende Tests, Reihenfolge geprüft
 - ☐ Mindestens eine Sache am Plan korrigiert oder bewusst bestätigt
-

Während der Arbeit

- ☐ Rückfragen **gelesen**, nicht durchgeklickt
 - ☐ Bei Unklarheit gefragt: „Was macht dieser Befehl genau?“
 - ☐ Bei Abweichung vom Auftrag: **gestoppt**, nicht bestätigt
-

Nach dem Auftrag

- ☐ Tests laufen — **grün?**
- ☐ Das Ergebnis selbst angesehen (nicht nur die Zusammenfassung gelesen)
- ☐ Funktioniert, was vorher funktioniert hat?
- ☐ **Committed**, weil es funktioniert
- ☐ Ist etwas schiefgegangen? → eine Zeile unter „Fallen in diesem Projekt“

Die Ampel für Autonomie

Auto-Accept oder autonome Schleifen **nur**, wenn alle drei zutreffen:

	zutreffend?
Git sauber, Rückweg in einem Schritt	☐
Tests laufen automatisch	☐
Nichts wirkt nach außen (kein Deployment, kein Versand, keine Produktivdaten)	☐

Fehlt eines: manueller Modus. Der Zeitgewinn liegt bei Minuten, der mögliche Schaden bei Tagen.

Die roten Linien

Nie	Warum
Agent in einem Projekt ohne Git	kein Rückweg
Echte Kundendaten zum Üben	Datenschutz, und es ist unnötig
Zugangsdaten in versionierten Dateien	bekommt man aus der Historie kaum wieder heraus
Automatisches Ausrollen ohne dich	kaputter Stand vor Kunden
Autonome Schleife bei „schöner machen“	kein Abbruchkriterium

Der Satz vor jedem Start

Wenn das jetzt völlig schiefgeht — komme ich zurück?

Ist die Antwort nein, ist das die Aufgabe vor allen anderen.