

Token management: 7 gewoontes die je **AI-limiet sparen**

Loop je met Claude (of een andere AI-assistent) tegen je gebruikslimiet aan? Eén mentaal model verklaart bijna alles: **bij elk bericht wordt de hele voorgaande conversatie opnieuw meegestuurd en meegeteld**. Een lang gesprek wordt dus per bericht steeds "duurder". De goedkoopste upgrade is daarom niet een zwaarder abonnement — het is slimmer gebruik. Deze zeven gewoontes maken het verschil.

01 Nieuw onderwerp = nieuwe chat

De grootste winst. Een verse chat slaapt geen dure geschiedenis mee. Context nodig uit het oude gesprek? Vraag daar een **korte samenvatting** en plak die als eerste bericht in de nieuwe chat. Vuistregel: na ±15–20 berichten ben je toe aan een nieuwe.

02 Plan je vraag — en corrigeer door te bewerken

Bundel gerelateerde vragen in één bericht en geef context en data direct mee: dat voorkomt heen-en-weer-berichten. Ging er iets mis? **Bewerk je vorige bericht** in plaats van een correctie eronder te sturen — dan groeit het gesprek niet.

03 Zet terugkerende documenten in een project

Projectkennis wordt **gecached** en telt bij hergebruik niet steeds opnieuw mee; bovendien laadt een project per vraag alleen de relevante stukken. Dus niet elke chat dezelfde briefing plakken — één keer in het project. Houd projectinstructies kort en ruim ongebruikte bestanden op.

04 Zet features bewust aan of uit

Extended thinking uit voor routinewerk — het uitgebreide redeneren kost extra. En beperk tools en connectors die je in dat gesprek niet nodig hebt: die zijn token-intensief, ook als je ze niet gebruikt.

05 Kies het juiste model per taak

Misschien wel de grootste knop. Het zwaarste model voor een simpele vraag is zonde van je limiet: **topmodel voor complex denk- en planwerk, middenmodel voor dagelijks werk, licht model voor snelle klusjes**. Wisselen kan gewoon per chat.

06 Bereid je bestanden voor

Lever alleen de relevante pagina's of tabbladen aan, niet het hele rapport. Plak tekst als tekst — **een screenshot van tekst kost een veelvoud** van dezelfde tekst als tekst. Grote datasets: eerst laten samenvatten of filteren, daarna doorvragen.

07 Grote klus? Laat het dure model denken en een goedkoop model doen

Grote projecten hoef je niet in één model te draaien. Hak ze op: het topmodel doet het denk- en planwerk één keer, een licht model voert uit. Zo betaal je het dure denkwerk niet bij elke stap opnieuw.

Het recept in drie stappen

- 1 Plannen (topmodel):** laat het de klus analyseren en een plan maken.
- 2 Overdragen (topmodel):** laat het een overdrachtsdocument schrijven met doel & context, exacte stappen en structuur, stijlregels en valkuilen — en vooral: **één volledig uitgewerkt voorbeeld**.
- 3 Uitvoeren (licht model):** geef per deeltaak het overdrachtsdocument mee (tip: zet het in een project) en laat het plan stap voor stap afwerken.

Voorbeeldprompt — de planner

"Analyseer deze 3 voorbeeldpagina's en onze tone of voice. Maak een plan om alle 12 productpagina's te herschrijven. Schrijf het als overdrachtsdocument voor een collega die de uitvoering doet, inclusief 1 uitgewerkt voorbeeld."

Voorbeeldprompt — de uitvoerder

"Hierbij het overdrachtsdocument. Herschrijf productpagina [X] exact volgens dit document. Wijk niet af van de structuur en stijlregels."

Vuistregel voor het overdrachtsdocument: als een nieuwe collega er zelfstandig mee aan de slag zou kunnen, kan een licht AI-model dat ook.

De checklist

- ✓ Nieuw onderwerp = nieuwe chat
- ✓ Vragen bundelen, context direct mee
- ✓ Bericht bewerken i.p.v. corrigeren
- ✓ Terugkerende documenten → project
- ✓ Projectbestanden opgeruimd houden
- ✓ Extended thinking uit bij routinewerk
- ✓ Ongebruikte tools & connectors uit
- ✓ Juiste model per taak
- ✓ Alleen relevante bestanden, tekst als tekst
- ✓ Grote klus → planner/worker-recept

Dit spiekbriefje hoort bij het artikel "**AI implementeren in je bedrijf: waarom top-down niet werkt**" op roelandvandervliet.nl. Wil je je hele team leren werken met AI — inclusief dit soort gewoontes vanaf dag één? Kijk dan naar de **incompany AI-cursus**: roelandvandervliet.nl/ai-cursus.