ecoteers

15

TIPKAART

AI verstandig gebruiken: 15 concrete tips

Praktische keuzes voor ondernemers en marketingteams die AI dagelijks gebruiken. In tien minuten gelezen, vandaag toe te passen.

AI is nuttig. De vraag is niet of je het gebruikt, maar hoe: waarvoor, hoe vaak, met welk model en met welke afspraken. Deze tipkaart brengt vijftien keuzes samen die je verbruik, je wachttijd en je factuur tegelijk verlagen, zonder dat je kwaliteit inlevert. Ze staan op volgorde van effect. Bij een deel staat een gemeten cijfer met een bronnummer, zodat je kunt nakijken waar het vandaan komt. Begin bovenaan en pak er vandaag drie uit.

Modelkeuze en redeneerstand

1 Kies het lichtste model dat volstaat

Zet niet elke taak bij je zwaarste model neer. Samenvatten, herschrijven, vertalen en sorteren lukt prima met een lichter model. In een meting over 88 modellen zit tussen de lichtste en de zwaarste taak ruwweg een factor 1.450 [19]. Je merkt het verschil in wachttijd en op je factuur.

DOE VANDAAG
Zet in je teamchat welk model het uitgangspunt is per taaksoort.

2 Zet de redeneerstand standaard uit

De redeneerstand laat een model hardop doorrekenen voordat het antwoordt. Dat kost volgens de gemeten bronnen 4 tot 30 keer meer energie, omdat er 300 tot 800 keer meer tokens worden geproduceerd [18][20]. Voor het herschrijven van een alinea heb je dat niet nodig. Zet hem aan bij analyses met meerdere stappen.

DOE VANDAAG
Controleer nu in je AI-tool of de redeneerstand aanstaat en zet hem uit.

3 Gebruik een filter waar dat kan

Sorteren, labelen en opzoeken hoeft niet door een taalmodel. Een zoekopdracht, een filter in je CRM of een formule in een spreadsheet geeft een reproduceerbaar antwoord dat je niet hoeft na te lopen. Classificeren kost in de metingen ongeveer 0,002 Wh per keer, ver onder een generatieve aanroep [19].

DOE VANDAAG
Zoek één terugkerende AI-taak die ook een filter of formule kan zijn.

Beeld en video

4 Check eerst of bestaand beeld volstaat

Een gegenereerde afbeelding kost tussen 0,63 en 2,9 Wh, afhankelijk van model en meetgrens [19][22]. Belangrijker nog: een eigen foto, een screenshot of een simpele grafiek is vaak geloofwaardiger dan gegenereerd beeld. Loop dus eerst je eigen beeldbank langs voordat je een prompt intikt.

DOE VANDAAG
Open je beeldbank en zoek eerst daar naar een bruikbare foto.

5 Laat video altijd door iemand aftekenen

Dit is de enige categorie waar één klik echt telt. Een gegenereerde video van vijf seconden kost ongeveer 944 Wh, ruwweg gelijk aan vierduizend korte tekstvragen [22]. Spreek af dat een tweede persoon meekijkt voordat er video wordt gegenereerd, en dat er een concreet doel ligt.

DOE VANDAAG
Wijs één collega aan die meekijkt voordat jullie video genereren.

6 Maak beeld dat je vaker gebruikt

Genereer je toch beeld, maak er dan iets van dat je meer dan één keer inzet. Een beeld dat past bij een terugkerend thema kun je in een nieuwsbrief, een blog en een socialpost gebruiken. Sla het op in je beeldbank met een duidelijke naam, zodat een collega het terugvindt.

DOE VANDAAG
Geef het laatste gegenereerde beeld een vindbare naam en zet het in de beeldbank.

Prompts en context

7 Plak alleen de passages die ertoe doen

Lange invoer telt zwaar mee. Een gewone vraag komt volgens een schatting uit op ongeveer 0,3 Wh, met 10.000 tokens context op ongeveer 2,5 Wh en met 100.000 tokens op ongeveer 40 Wh [21]. Knip de twee relevante artikelen uit het contract in plaats van het hele document.

DOE VANDAAG
Knip bij je volgende vraag het document terug tot de relevante alinea's.

8 Vraag om een lengte en vorm

Aan de uitvoerkant is de lengte van het antwoord de directe kostendrijver [20]. Een model dat niet weet hoeveel woorden je wilt, levert standaard te veel. Zeg er dus bij: 120 woorden, drie alinea's, geen inleiding. Je scheelt jezelf bovendien het schrappen achteraf.

DOE VANDAAG
Voeg aan je standaardprompt een regel toe met gewenst aantal woorden en vorm.

9 Regenereer hooguit één keer

Opnieuw genereren is een volledige nieuwe aanroep, inclusief het opnieuw verwerken van alle context. Vijf keer regenereren en dan de derde kiezen kost vijf keer zoveel als één keer goed herformuleren. Schrijf na een teleurstellend antwoord op wat er mis was en pas daarmee je prompt aan.

DOE VANDAAG
Vervang je volgende druk op opnieuw genereren door één zin met wat beter moet.

Werkstroom en hergebruik

10 Cache je vaste context één keer

Systeeminstructies, merkrichtlijnen en productdata stuur je vaak bij elke vraag opnieuw mee. Met prompt-caching hergebruikt het model die vaste voorkant in plaats van hem telkens opnieuw te verwerken [26]. Je slaat dus rekenwerk over zonder iets in te leveren op kwaliteit.

DOE VANDAAG
Vraag je ontwikkelaar of je AI-tool prompt-caching ondersteunt voor jullie vaste instructies.

11 Zet wat kan wachten in een batch

Vierhonderd productteksten hoeven zelden vandaag klaar te zijn. Bij batchverwerking lever je het werk in en krijg je het terug binnen een venster van 24 uur, buiten de spits [27]. Dat haalt de piek uit je werkdag en uit het net, en je hoeft er niet bij te zitten.

DOE VANDAAG
Zet één bulkklus van deze week op de lijst voor batchverwerking.

12 Bewaar goede prompts op één plek

De goedkoopste aanroep is de aanroep die je niet doet omdat een collega het antwoord al had. Leg werkende prompts en bruikbare output vast op een vindbare plek, met een korte notitie waarvoor ze bedoeld zijn. Dat scheelt dubbel werk en houdt de toon van je merk consistent.

DOE VANDAAG
Maak één gedeeld document en plak er je drie beste prompts in.

Afspraken en inkoop

13 Leg vast wat nooit in een model gaat

De Autoriteit Persoonsgegevens kreeg meldingen van datalekken doordat medewerkers persoonsgegevens in een AI-chatbot plakten, vaak op eigen initiatief [17]. De toezichthouder adviseert duidelijke afspraken over wat medewerkers wel en niet mogen invoeren [17]. Eén lijstje met categorieën voorkomt hier meer schade dan elke besparing in dit overzicht.

DOE VANDAAG
Schrijf drie categorieën gegevens op die nooit in een extern model gaan.

14 Vraag je leverancier naar regio en PUE

Waar het rekenwerk draait, weegt zwaarder dan je prompt. Het Nederlandse net zat in 2025 op ongeveer 262 gram CO2-equivalent per kilowattuur, het Franse op ongeveer 32 gram [24][25]. En tussen een gemiddelde serverruimte en een moderne locatie zit flink verschil in overhead, uitgedrukt in PUE [3][23].

DOE VANDAAG
Mail je leverancier twee vragen: in welke regio draait dit en welke PUE rapporteert de locatie.

15 Wijs één eigenaar aan

Afspraken die van iedereen zijn, verwateren binnen een maand. Zet één naam bij deze lijst en plan een kwartaalmoment. Die persoon kijkt de factuur per model na, het aandeel aanroepen met de redeneerstand aan en het aantal gegenereerde video's. Dat zijn indicatoren, geen voetafdruk.

DOE VANDAAG
Zet één naam en een datum in het kwartaal onder deze lijst.

Meer weten?

Deze kaart geeft de keuzes, niet de onderbouwing. Wil je weten waar de cijfers vandaan komen, waarom bronnen elkaar tegenspreken en wat dat betekent voor je eigen beleid, lees dan de whitepaper Verstandig omgaan met AI. Daarin staan de gemeten waarden per taaksoort, een beslisschema van zes stappen en zevenentwintig bronnen.

Plan een kennismaking: ecoteers.nl/contact · info@ecoteers.nl · 030 - 227 0475

ONDER MEER VOOR

Bronnen (nummering gelijk aan de whitepaper)

3 Google, 2026 Environmental Report (2026)
17 Autoriteit Persoonsgegevens, Pas op: gebruik van AI-chatbot kan leiden tot datalekken (2024)
18 Hugging Face, AI Energy Score v2 (Luccioni en Gamazaychikov) (2025)
19 Luccioni, Jernite en Strubell (ACM FAccT), Power Hungry Processing: watts driving the cost of AI deployment? (2024)
20 Dauner en Socher (Frontiers in Communication), Energy costs of communicating with AI (2025)
21 Epoch AI, How much energy does ChatGPT use? (2025)
22 MIT Technology Review, We did the math on AI's energy footprint (2025)
23 Uptime Institute, Global Data Center Survey 2026 (2026)
24 Electricity Maps, Grid in Review 2025: the Netherlands (2026)
25 Electricity Maps, Grid in Review 2025: France (2026)
26 Anthropic, Prompt caching: prijsverhoudingen voor cache-lees- en schrijftokens (documentatie) (2026)
27 OpenAI, Batch API (documentatie) (2026)