

🔊 LUISTER DIT VERSLAG

-10%

1,0x

+10%

+25%

1,0x

↓ Download naar telefoon

▶ 📖 Lees mee met de podcast

Soort: vergadering · **Duur:** onbekend · **Deelnemers:** [Spreker 1] [Spreker 2]

Let op — sprekers niet bevestigd. Plaud nummert sprekers (Speaker 1, 2, ...) maar koppelt geen namen. Ga niet uit van wie wat zei; "ik/wij" hoeft niet dezelfde persoon te zijn.

Notulen

Degradatie van AI-modellen

Sinds 5 augustus presteren bepaalde AI-modellen aantoonbaar slechter en maken ze meer domme fouten dan voorheen.

Spreker 1 wijst op brede online klachten over deze "downgrade".

Spreker 2 bevestigt vergelijkbare problemen: zijn model vertoont sinds 5 augustus prestatieachteruitgang.

Architectuur en 'source of truth'

Spreker 2 vergelijkt zijn systeemarchitectuur met een boot die door voortdurende toevoegingen aan één kant te zwaar wordt, wat instabiliteit veroorzaakt.

Conclusie: er zijn te veel "sources of truth", wat herhaling en fouten in de hand werkt.

Spreker 1 heeft zijn systeem herzien naar één centrale "source of truth"; elke klant krijgt een eigen playbook zodat de AI exact weet waar informatie te vinden is.

Gebruik en testen van verschillende bots en modellen

Spreker 1 draait een "Hermes bot" op zijn Mac via het Codex-model met toegang tot alle mappen.

Hij plant een terminalbot met "Kimmie" te bouwen voor klanttesten.

Communicatie met de Hermes bot verloopt via Telegram; volgens Spreker 2 verbruikt dit veel tokens en is het ineffectief voor contextteruglezing.

Hoog tokenverbruik (tot 400 miljoen per dag) ontstaat doordat bij simpele vragen steeds de volledige context wordt meegestuurd. Voorgestelde oplossing: sessies opsplitsen per taak.

Proactiviteit en automatisering van AI

Spreker 2 vindt zijn AI-model (Linky) onvoldoende proactief; het zou zelfstandig analyses moeten uitvoeren en taken initiëren.

Het model van Spreker 1 genereert wel taakopsommingen, waarna hij handmatig akkoord geeft.

Voorgestelde oplossing: een prompt ontwikkelen die de AI toestemming geeft om bepaalde acties uit verslagen automatisch uit te voeren.

Veiligheid en beheer van API-keys

Er is een terugkerend probleem met openbaar geworden API-keys, wat dagelijks tot beveiligingsmeldingen leidt.

Spreker 1 heeft een alert ingesteld die direct waarschuwt bij verdachte activiteiten; hij moet vervolgens bevestigen of hij de actie zelf heeft uitgevoerd.

Automatische key-rotatie wordt als belangrijk gezien, maar is nog niet geïmplementeerd.

Ontwikkeling van een AI-cursus

Spreker 2 ontwikkelt een AI-cursus voor beginners ("rookies").

De cursus dekt onder meer werken met de terminal en moet de leercurve die zij zelf doorliepen voor anderen vereenvoudigen.

Suggestie: prijsrange €500–€900.

Volgende afspraken

Spreker 1: bouw en test een terminalbot met Kimmie voor klantgebruik.

Spreker 2: laat het AI-systeem zichzelf analyseren alsof het een senior developer is, en verbeter op basis daarvan de architectuur/structuur.

Spreker 1: herontwerp een nieuw dashboard dat aansluit bij de huidige werkzaamheden.

Beiden: definieer en implementeer taakgebaseerde sessies om tokenverbruik te verlagen (opsplitsen per taak).

Beiden: ontwerp een prompt- en toestemmingsframework voor automatische uitvoering van geselecteerde acties.

Beiden: plan en implementeer automatische rotatie van API-keys en bijbehorende monitoring.

AI-suggesties

AI-suggesties

AI heeft de volgende problemen geïdentificeerd die niet zijn afgesloten in de vergadering of waarvoor duidelijke actiepunten ontbreken; gelieve hierop te letten:

Er is een significant hoog tokenverbruik (tot 400 miljoen per dag) vastgesteld, wat leidt tot hoge kosten. De voorgestelde oplossing om sessies op te splitsen per taak is nog niet in een concreet actiepunt omgezet.

De AI-modellen zijn niet proactief genoeg. Er is besproken dat de AI zelfstandig taken zou moeten uitvoeren, maar er is geen duidelijk plan om dit te implementeren, behalve het idee om een prompt te maken.

Er is een terugkerend beveiligingsrisico met openbare API-keys. Het automatisch roteren van keys werd als belangrijk genoemd, maar er is geen besluit genomen om dit te implementeren.