

Europæisk AI: billig til chat, dyr til nat-agenter

Til chat er forskellen småpenge. Til agenter, der kører hele natten, er regnestykket et andet.

GUIDE · SEPTEMBER 2026

Stefano Vincenti er AI-træner, konsulent og ekstern lektor ved ITU og DIS Copenhagen og co-founder af BotTellMe. Han hjælper nordiske ledelsesteams og fagfolk med at gå fra AI-nysgerrige til konkret, ansvarlig brug.

Følg på LinkedIn

Tilmeld nyhedsbrevet

linkedin.com/in/stefanovincenti · stefanovincenti.substack.com · aitrainer.dk

Jeg har lige bygget en lille app, der hjælper dig med at vælge europæisk AI, opgave for opgave. Da den var færdig, satte jeg mig ned og regnede på det spørgsmål, jeg får oftest lige nu: hvad koster det os at gå europæisk?

Det korte svar er, at det kommer an på, hvad I bruger AI'en til. Til chat og dokumenter er forskellen småpenge. Til agenter, der skal køre en hel nat i jeres systemer, er det en anden snak. Her er regnestykket, lag for lag.

■ 1. MODELLEN ER DEN BILLIGE DEL

En AI-regning har tre lag: modellen, harness'en rundt om den, og den compute, det hele trækker, når det kører. De fleste kigger kun på det første.

De tre lag i en AI-regning

De fleste kigger kun på det første

MODEL *betales pr. token*

- Pris pr. million tokens ind og ud
- Modellen er den billige del

HARNESS *alt det udenom modellen*

- Hukommelse på tværs af kontekstvinduer
- Selverifikation
- Delagenter
- Guardrails, der holder hele natten

COMPUTE *det, det hele trækker*

- Den compute, det hele trækker, når det kører

Ved I, hvilket af de tre lag der driver jeres regning?

Model 1. En AI-regning har tre lag: model, harness og compute.

Modellen betaler I for pr. token. De amerikanske frontier-modeller, Claude Fable 5.1 og GPT-6 Astra, koster begge omkring 10 dollar pr. million tokens ind og 50 ud. En europæisk open-model som Mistral Medium 3.5 ligger på 1,50 og 7,50, altså omkring en syvendedel af tokenprisen. Fable og Astra scorer også højere på Artificial Analysis' intelligens-index, men til daglig chat er det ikke der, forskellen betyder mest.

Model	Ind (USD pr. mio. tokens)	Ud (USD pr. mio. tokens)
Claude Fable 5.1	10	50
GPT-6 Astra	10	50
Mistral Medium 3.5	1,50	7,50

Tabel 1. Tokenpriser pr. september 2026, USD ekskl. moms. Kilder: Mistral AI, OpenAI, Artificial Analysis.

For daglig chat er det nemlig ikke prisen pr. million tokens, der afgør regningen. Det er, hvor mange millioner tokens en medarbejder faktisk bruger på en måned, og det er mindre, end de fleste tror.

■ 2. TIL CHAT OG DOKUMENTER ER EUROPÆISK BILLIGERE, OGSÅ PÅ FAKTURAEN

For en medarbejder, der bruger AI til at skrive, opsummere og slå op i løbet af dagen, er forbruget beskedent. Til den slags er en europæisk open-model som Mistral Medium 3.5 en nem beslutning. I betaler en brøkdelt pr. token, jeres data bliver i Europa, og I giver ikke køb på noget, der betyder noget i praksis.

■ 3. TIL NAT-AGENTER ER TOKENPRISEN IKKE DET, DER DRIVER REGNINGEN

Springet kommer et andet sted fra. En agent, der skal arbejde, mens I sover, sender ikke én prompt. Den kører i timer. Den læser, planlægger, verificerer sit eget arbejde, prøver igen, holder styr på en opgave hen over mange kontekstvinduer.

Det er alt det udenom modellen, der afgør, om den kan det: hukommelse på tværs af vinduer, delagenter, selvverifikation og guardrails, der holder hele natten. Og det er præcis der, de amerikanske frontier-modeller er foran lige nu. De kan køre den slags langt løbende arbejde ud af boksen.

Vil I have det samme oven på en europæisk model i dag, skal I selv bygge mere af stilladset, og I betaler for det i udviklertid og i, hvor pålideligt løbet er. Tokenprisen er lavere. Den samlede pris på et nat-løb, målt hele vejen, er ikke nødvendigvis.

PRINCIP

Tokenprisen er lavere. Den samlede pris på et nat-løb, målt hele vejen, er ikke nødvendigvis.

■ 4. ET REGNEEKSEMPEL

Sig, at et natligt agent-løb bruger 5 millioner tokens ind og 2 millioner ud. På Claude Fable 5.1 eller GPT-6 Astra koster det omkring 150 dollar. På Mistral Medium 3.5 er det omkring 22 dollar. Isoleret set er den europæiske model klart billigst, også selv om løbet må køres nogle gange ekstra.

Ét natligt agent-løb

Tokenprisen alene, før udviklertid og fejlede løb

5 + 2 mio.

tokens ind og ud i ét natligt agent-løb

Regneeksempel

ca. 150 USD

på Claude Fable 5.1 eller GPT-6 Astra

10 / 50 USD pr. mio. tokens

ca. 22 USD

på Mistral Medium 3.5

1,50 / 7,50 USD pr. mio. tokens

Men et nat-løb koster ikke kun tokens.

Det koster den udvikler, der bygger og passer stilladset, og det koster, når løbet ikke er kørt rent.

Model 2. Regneeksempel: 5 millioner tokens ind og 2 millioner ud. Priser pr. september 2026.

Men et natligt løb, der skal være færdigt og rigtigt, når I vågner, koster ikke kun tokens. Det koster den udvikler, der bygger og passer stilladset, og det koster, når løbet ikke er kørt rent. Det er der, den mere modne frontier-harness tjener sin højere tokenpris hjem. Ikke på fakturaen fra model-udbyderen, men på jeres egen tid og på, om arbejdet er lavet, når I møder ind.

■ 5. HVAD DET BETYDER FOR JERES BESLUTNING

Beslutningen er ikke europæisk eller amerikansk på tværs af hele huset. Den er opgave for opgave.

Chat, dokumenter og afgrænsede opgaver: gå europæisk, tag besparelsen og suveræniteten med. Langtløbende, autonome agenter i jeres kernesystemer: i dag er frontier-harness'en mere moden, og kræver jeres compliance en europæisk model, så budgettér for det stillads, I selv skal lægge oveni.

Tre spørgsmål er værd at stille på jeres næste ledelsesmøde.

1. Hvilke af jeres AI-arbejdsgange er chat, og hvilke er nat-agenter, der kører uden opsyn?
2. Hvad koster ét natligt agent-løb jer i dag, målt hele vejen, ikke kun på tokenprisen?
3. Og ved I, hvilket af de tre lag, model, harness eller compute, der driver jeres regning?

De fleste kender ikke svaret på det sidste. Det er den bedste anledning til at få set efter, før tallet vokser af sig selv.

■ KILDER

Priser verificeret september 2026 mod leverandørernes og Artificial Analysis' egne sider.

- Mistral AI, API-priser (mistral.ai/pricing/api): Mistral Medium 3.5, 1,50 / 7,50 dollar pr. mio. tokens ind/ud.
- OpenAI, GPT-6 Astra (openai.com): 10 / 50 dollar pr. mio. tokens ind/ud.
- Artificial Analysis, Intelligence Index og modelsider (artificialanalysis.ai): Claude Fable 5.1, 10 / 50 dollar ind/ud.
- Egen EU AI picker-app og modeloversigt, september 2026

© 2026 Stefano Vincenti · aitrainer.dk · Publiceret 12. september 2026

Stefano Vincenti

AI-træner, konsulent og ekstern lektor · 20+ år i software og digital transformation · Co-founder af BotTellMe · Ekstern lektor ved ITU og DIS Copenhagen · Partner hos TryZone.

LinkedIn

linkedin.com/in/stefanovincenti

Nyhedsbrev

stefanovincenti.substack.com

Mere

aitrainer.dk · tryzone.dk ·
bottellme.com

Dette er generel praktisk vejledning, ikke juridisk rådgivning eller compliance-rådgivning, og jeg er ikke advokat eller compliance-specialist. Brug det ikke som beslutningsgrundlag på egen hånd. Tag altid jeres egen compliance- eller juridiske afdeling og jeres DPO med på råd, og bekræft datoer og status mod primære kilder (EUR-Lex, Europa-Kommissionen, Datatilsynet). Alle priser er oplyst i USD ekskl. moms og kan være ændret; tjek altid leverandørernes egne prissider. Status pr. september 2026.

Stefano Vincenti · AI, Built Human