

Judgment Loop'et

Beslutningsret i enterprise AI — og de fire spørgsmål der gør "human-in-the-loop" til mere end en floskel.

Det hårdeste problem i enterprise AI lige nu er et spørgsmål om beslutningsret. Det ser bare ikke ud som ét.

For et par uger siden sad jeg i et rum med en ledelse, der gjorde alt rigtigt. Strategi. Pilotprojekter i flere afdelinger. Copilot-udrulning. Governance-udvalg. Budgetlinje. To nye data scientists. Underskrevet kontrakt med en enterprise-leverandør.

Jeg stillede dem ét spørgsmål. Når modellen er i loopet, hvem bestemmer så hvad?

Stilhed.

Det er ikke usædvanligt. Jeg ser det alle steder, hvor jeg arbejder — retail, finans, det offentlige, life sciences. Det der dræber AI-projekter er næsten altid det samme. Et spørgsmål ingen tænkte på at stille højt.

Denne artikel handler om det spørgsmål. Og om den lille ramme, jeg nu tager med ind i hvert rum for at tvinge det op på bordet.

"Human-in-the-loop" betyder ikke længere noget

Sætningen står i hver eneste AI-politik, jeg har læst de seneste to år. Bestyrelsesnotater. Leverandørpitches. EU AI Act-compliancedokumenter.

Den fortæller dig, at der er et menneske et eller andet sted i systemet. Det er det.

Den fortæller dig ikke, hvad mennesket laver. Hvornår. Hvilken autoritet det har. Hvilken information det ser. Eller hvad der sker, når mennesket er uenig med maskinen.

Jeg har set organisationer fejre "menneskeligt tilsyn", mens mennesket var en juniormedarbejder, der trykkede godkend på 200 model-outputs i timen, fordi alternativet var at misse sin SLA.

Det er stempelarbejde med produktivitetsmål. At kalde det tilsyn ændrer ikke på det.

Løsningen er at være konkret om, hvilken dømmekraft vi rent faktisk beder mennesker om at udøve.

Judgment Loop'et

Fire spørgsmål. Ét billede. Designet til at blive besvaret af et ledelsesteam på under halvanden time, og gennemgået igen hver gang systemet ændres.

For enhver AI-understøttet beslutning i jeres organisation:

1. Hvad beslutter modellen?

En forudsigelse er ikke en beslutning. En credit risk-score er en forudsigelse. Beslutningen er, om lånet skal godkendes, hvilken rente der skal tilbydes, om sagen skal eskaleres. Bland de to sammen, og ansvaret forsvinder stille og roligt.

2. Hvad beslutter mennesket?

Næsten ingen besvarer det her eksplicit.

Beslutter mennesket, om modellens output skal accepteres? Om sagen skal undersøges nærmere? Om der skal overstyres baseret på kontekst, modellen ikke kan se?

Hver af dem er et forskelligt job. Forskellig information. Forskellig træning. Forskellig tid på opgaven.

"Mennesket reviewer det" er ikke et svar.

3. Hvad har mennesket brug for, for at træffe en god beslutning?

Det er her, de fleste deployments knækker.

Hvis jobbet er at fange model-fejl, skal mennesket se modellens confidence, hvilke features den vægtede, og helst en sammenligningscase.

Hvis jobbet er at tilføje kontekst, modellen mangler, skal den kontekst være synlig. Ikke begravet i et separat system.

Hvis jobbet er at tage ansvar, skal mennesket have autoriteten til faktisk at overstyre.

De fleste "human-in-the-loop"-implementeringer giver mennesket en knap og et stopur. Kald det, hvad det er. Teater.

4. Hvad sker der, når mennesket og modellen er uenige?

Det er her, det viser sig, om loopet er ægte.

Hvis uenigheden udløser en læringshændelse — logget, gennemgået, ført tilbage i modellen og processen — har I et Judgment Loop.

Hvis uenigheden udløser en forsinkelse, en KPI-rød, eller en stille overstyring der ikke ender nogen steder, har I et compliance-ritual. En klassisk svaghed i compliance-funktioner.

Det er det. Fire spørgsmål. Tegn dem på en whiteboard. Pilen går begge veje mellem menneske

og model. Det er hele pointen.

Hvad loopet afslører

Første gang et ledelsesteam arbejder sig ærligt gennem Judgment Loop'et, sker der typisk tre ting.

Først: stilhed. De ved faktisk ikke svaret på spørgsmål 2. De troede de gjorde. De har skrevet det ned et eller andet sted. Men når de skal artikulere det for den konkrete beslutning foran dem, kommer ordene ikke.

Dernæst: en kamp. Som regel mellem den funktion der ejer processen, og den funktion der ejer teknologien. Begge troede de havde beslutningsretten. Ingen af dem havde skrevet det ned.

Til sidst: lettelse. Fordi Loop'et oversætter "er vi ansvarlige nok med AI?" — en abstrakt bekymring — til et konkret stykke design-arbejde, der kan laves, gennemgås og revideres.

Det er det, god dømmekraft i drift ser ud som. Mere konkret menneskelig involvering, ikke mindre.

Derfor står det her ikke i jeres AI-strategi-deck

Der er en grund til, at Judgment Loop'et ikke står i jeres AI-strategi.

AI-strategi, som den sælges til de fleste organisationer, optimerer efter to ting. Hvad vi skal med AI. Og hvordan vi undgår at blive sagsøgt for det.

Det første lever i roadmap-slides. Det andet lever i policy-dokumenter.

Ingen af dem indeholder den operationelle virkelighed. Hvem beslutter hvad, med hvilken information, under hvilken autoritet, i det øjeblik modellen leverer.

Den virkelighed lever i procesdesign. Og procesdesign laves i de fleste virksomheder af mennesker, der ikke var med til AI-strategi-mødet.

Det er gabet. Judgment Loop'et er det værktøj, jeg bruger til at lukke det.

Værktøjer ændrer sig. Dømmekraft akkumulerer.

Den frontier model I udruller i dag er forældet om atten måneder. Det agent-framework I standardiserer på, bliver erstattet. Leverandøren I valgte, bliver opkøbt eller pivoterer.

Det I bygger omkring jeres AI — hvem beslutter hvad, med hvilken information, med hvilken autoritet — det akkumulerer.

Det bliver skarpere for hver uenighed I logger. Klarere for hver edge case jeres team arbejder sig

igennem. Over tid bliver det den egentlige konkurrencefordel.

Ikke modellen. Dømmekraften omkring modellen.

Det er det jeg mener, når jeg siger AI, Built Human. Loop'et er systemet. Tag menneskene ud, og I har ikke en autonom AI. I har en uansvarlig en.

Én ting at prøve i denne uge

Inden jeres næste AI-deployment review, tag én produktionsklar use case. Bare én.

Gå jeres team igennem de fire Judgment Loop-spørgsmål, højt, i rummet.

Hvis I kan svare på alle fire i klare sætninger, er I blandt det mindretal, der når i produktion med noget holdbart.

Hvis I ikke kan, har I lige fundet det arbejde, der faktisk skal laves.

Det er ikke en fiasko. Det er starten.

/Stefano

AI, Built Human er et bi-weekly nyhedsbrev om dømmekraften bag maskinen. Hvis denne ramme var brugbar, så send den videre til én leder, der står midt i et pilotprojekt lige nu.

Stefano Vincenti · AI Advisor & Trainer · aitrainer.dk