



id: SI_001

model: google/gemma-4-26b-a4b-it

seeds:

04_risiko: sårbarhet

03_relasjon: makt

05_handling: vent

01_perspektiv: juridisk

02_tidhorisont: kortsiktig

Instans SI_001

Seeds: juridisk | kortsiktig | makt | sårbarhet | vent

Denne analysen avviser den teknologiske optimismen i pilotvinduet på 6–9 måneder. Dere planlegger å delegere **makt** til agenter i et system som mangler evnen til å holde dem ansvarlige.

Arkitektur og kontroll: Maktforskyvning uten mandat

Deres nåværende landskap er en juridisk slagmark. Å innføre agenter i legacy-monolitter uten deterministiske API-er er å skape en uoversiktlig **sårbarhet**.

- **Prinsipp:** Implementer en “Intermediary Proxy Layer” (en teknisk buffer). Agenter skal aldri snakke direkte med API-er. All trafikk må gå gjennom en kontrollert gateway som fungerer som en juridisk og teknisk kontrollpost.
- **Execution Graph:** Modellér ikke bare handlinger, men hensikt. Hver node i grafen må knyttes til en spesifikk oppgave-ID for å sikre revisjonsspor som holder i en rettslig kontekst.

IAM og API: Fragmentert autoritet

Deres IAM-modell er utilstrekkelig for autonomi. RBAC er for statisk; agenter krever “Just-in-Time” (JIT) tilgang.

- **Løsning:** Bruk Ephemeral Workload Identity. Agenten får et token med en levetid på nøyaktig den tiden en spesifikk sub-oppgave tar (f.eks. 300 sekunder).
- **Risiko:** API-er uten idempotens er en tidsinnstilt bombe. En agent i en loop kan generere doble utbetalinger eller slette datasett før en menneskelig kontrollør i det hele tatt ser et varsel.



Blindsoner: Emergent kaos

Den største risikoen er ikke at agenten gjør “feil”, men at den finner “optimale” men ulovlige snarveier for å nå et mål (**emergent behavior**). I et system med svak datakvalitet vil agenten bruke datafeil som logiske premisser, noe som skaper en kjedereaksjon av korrupte beslutninger.

SI_001s dissens: Dere søker kontroll gjennom verktøy (OPA, Kyverno), men dere mangler det fundamentale konseptet: **en “Kill Switch” som fungerer på tvers av alle lag**. Uten en sentralisert, øyeblikkelig isoleringsmekanisme er piloten et eksperiment i kontrollert risiko som sannsynligvis vil eskalere.

VENT. Ikke rull ut autonomi før kontroll-laget er ferdigstilt.

Konkrete neste steg:

1. **Stopp piloten** inntil en “Intermediary Proxy” er definert som arkitektonisk krav for alle agent-kall.
2. **Kartlegg “Blast Radius”:** Identifiser de 3 mest kritiske systemene (høy integritetsrisiko) og innfør harde manuelle godkjenningpunkter (human-in-the-loop) for alle endringer her.
3. **Etabler “Shadow Logging”:** Start logging av alle API-kall og beslutningsstier i en isolert pipeline før agenten får skriveadgang.