



id: SI\_005

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

seeds:

04\_risiko: mulighet

03\_relasjon: makt

05\_handling: implementer

01\_perspektiv: strategisk

02\_tidhorisont: kortsiktig

## Instans SI\_005

**Seeds:** strategisk | kortsiktig | makt | mulighet | implementer

### Arkitektonisk Analyse: Fra Copilot til Swarm-dominans

**1. Arkitektur-valg: Hybrid kontroll** For en helseorganisasjon er Copilot Studio/Power Automate for rigid og begrenset for 100 agenter. **Strategisk valg:** Bygg orkestrering i **Azure OpenAI (AOAI) med Python (LangGraph eller Semantic Kernel)**.

- **Makt-aspekt:** Ved å eie orkestreringen via Azure, beholder du kontroll over data-residency og Entra ID-autentisering, mens du unngår “black box”-begrensningene i Copilot Studio.
- **Implementer:** Bruk Azure Functions/Container Apps for å kjøre agent-logikken. Dette gir skalering som Power Automate mangler.

**2. Spesialisering og Syntese (Case 1)** For å unngå 100 identiske svar, må du implementere en **Hierarkisk Swarm**.

- **Metode:** Ikke gi 100 agenter samme prompt. Bruk en Persona Factory som injiserer unike “Expert Profiles” (f.eks. `Role: Compliance_Officer, Constraint: GDPR/Helsekrav`).
- **Syntese:** Du trenger en Aggregator Agent (Master-instans) som bruker en “Map-Reduce”-tilnærming: Agenter genererer (Map), og en dedikert konsoliderings-agent destillerer (Reduce) til en strukturert rapport.

**3. Distribuert Processing (Case 2)** Bruk en **Queue-based architecture** (Azure Service Bus).

- **Implementer:** Case lander i kø → Dispatcher-agent evaluerer kompleksitet → Allokere til spesifikk agent-type → Resultat skrives til Dataverse.



- **Consistency:** Bruk Schema Enforcement (Pydantic) for alle agent-outputs for å sikre at variable case-typer alltid returnerer validerte JSON-objekter.

**4. Data-innsprøytning og Sikkerhet Makt/Mulighet:** Bruk **Azure AI Search** som en RAG-motor.

- **Sikkerhet:** Ingen data forlater Azure-tenanten. Bruk Managed Identities slik at agenter aksesserer SharePoint/Dataverse uten hardkodede credentials. Dataene “trekkes” inn i kontekst-vinduet kun ved behov (Just-in-time injection).

## 5. Observerbarhet og Kostnad

- **Implementer:** Integrer **Azure Monitor/Application Insights**. Du må tracke `token_count` per `agent_id` i sanntid.
- **Strategisk ROI:** Bruk GPT-4o for “Master/Aggregator” og GPT-4o-mini for de 100 spesialiserte arbeider-agentene. Dette kutter kostnader med >80%.

### Neste steg:

1. **Proof of Concept (PoC):** Sett opp en Azure Container App som kjører 5 spesialiserte agenter via LangGraph mot en lukket Azure OpenAI-instans.
2. **Data Mapping:** Identifiser de 3 viktigste SharePoint-kildene som krever RAG-tilgang.
3. **Governance-audit:** Bekreft med IT-sikkerhet at bruk av Managed Identities for agent-aksess er innenfor deres compliance-rammeverk.