



# Debrief — openai/gpt-5.4-mini

---

- Instanser: 9
  - Tokens inn: 20,394
  - Tokens ut: 8,589
  - Kostnad: \$0.1900
  - Kjørt: 2026-04-14 16:15:43.995423+00:00
- 

## Sverm-debrief

---

### Konsensus

1. **Azure OpenAI + egen lettvektssorkestrering på Azure er beste hovedspor** for en helse-streng, Microsoft-sentrisk, cloud-only organisasjon. Dette bevarer Microsoft-investeringen og gir nok kontroll for multi-agent.
  2. **Copilot Studio/Power Automate er nyttig, men ikke nok som swarm-motor.** Det egner seg til enkle workflows og human-in-the-loop, ikke 100-agents parallellanalyse eller dynamisk case-routing.
  3. **Ikke bruk ett identisk prompt 100 ganger.** Agentene må ha tydelige roller, smale kontekstpakker og strukturert output, ellers blir resultatet “100 meninger”.
  4. **Data må holdes innen tenant/grense** via Entra ID, Managed Identity, Private Endpoints, Purview/DLP og intern RAG mot SharePoint/Dataverse/Azure AI Search. Ingen rå helse-data til eksterne API-er.
  5. **Observerbarhet og koststyring må bygges inn fra dag 1:** tokens, latency, feilrate, retrieval-hit-rate, confidence og human override-rate.
- 

### Dissens

- **Hvor langt man bør gå utenfor Microsoft-økosystemet:** flere instanser avviste proprietære swarm-rammeverk som primærplattform, men noen åpnet for dem kun i sandbox/prototyping.
- **Tempoet i utrulling:** noen foreslo 4–6 uker til pilot, andre 8–12 uker eller 3–6 måneder til produksjonsklar løsning.



- **Hvordan agentene organiseres:** noen anbefalte 8–12 faste rollefamilier, andre 10–20 roller med underroller og hierarkisk clustering. Fellesnevneren var likevel strukturert differensiering.
- **Copilot Studio sin rolle:** fra “bare front-end” til “lett orkestrering for enkle workflows”.

---

## Blindsoner avdekket

- At sværm ikke bare er et skalerings spørsmål, men et **governance- og risikodesignproblem**.
- At **routing/triage** ofte gir mer ROI enn bedre prompts.
- At **kaoset** i 100-agent-oppsett må håndteres med en meta-arkitektur: clustering, arbiters, confidence scoring og schema-validering.
- At sværm kan være feil løsning hvis én god retrieval + én sterk analytiker løser oppgaven bedre og billigere.

---

## Anbefalinger

1. **Bygg en pilot i Azure:** Azure OpenAI + Service Bus + Functions/Durable Functions + App Insights + Purview.
2. **Start med to konkrete use cases:** 10-agent perspektivanalyse og queue-basert small-batch processing.
3. **Definer JSON-outputschema og eval-metrikker før implementering:** claims, evidence, confidence, risk, recommendation.
4. **Lag en arbiter/meta-agent** som dedupliserer, vekter og oppsummerer; ikke la alle agenter være like.
5. **Kjør en kontrollert benchmark mot dagens Copilot-prosess** for kost, kvalitet og etterlevelse før skalering.