

Nevrosymbolsk AI og Agentisk AI

ELMED219: Momentliste S01–S11

ELMED219

Vår 2026

1 Grunnleggende paradigmer

- S01: Kontrastere symbolsk og konneksjonistisk AI
- S02: Forklare konseptet nevrosymbolsk integrasjon

2 Kunnskapsgrafer og ontologier

- S03: Beskrive hva en kunnskapsgraf er
- S04: Kjenne til medisinske ontologier
- S05: Diskutere fordeler med nevrosymbolsk AI i medisin

3 Case: Nevrosymbolsk gliomdiagnose

- S06: Nevrosymbolsk AI for gliomdiagnostikk
- S07: Forklare hvordan kunnskapsgrafer kan validere prediksjoner

4 Agentisk AI

- S08: Definere agentisk AI og dens kjerneegenskaper
- S09: Beskrive hvordan en AI-agent kan orkestrere klinisk arbeidsflyt
- S10–S11: Human-in-the-loop og etiske utfordringer

S01: Kontrastere symbolsk og konneksjonistisk AI

Symbolisk AI (GOFAI):

- “Good Old-Fashioned AI”
- Eksplisitte regler og symboler
- Logisk resonnering
- Kunnskapsrepresentasjon
- Ekspert-systemer

Styrker:

- Forklarbar
- Presise resonnementer
- Kodifiserer ekspertkunnskap

Konneksjonistisk AI:

- Nevrale nettverk, dyplæring
- Lærer fra data
- Subsymbolsk representasjon
- Mønstergjenkjenning
- “Black box”

Styrker:

- Lærer komplekse mønstre
- Håndterer støy og usikkerhet
- Skalerer med data

Historisk perspektiv

AI har svingt mellom disse paradigmene – se [AI-vintrene](#). Nå: Kan vi kombinere dem?

S02: Forklare konseptet nevrosymbolisk integrasjon

Nevrosymbolisk AI = det beste fra to verdener

Hovedidé:

- Kombiner **nevralt** (læring fra data, mønstergjenkjenning)
- med **symbolisk** (resonnering, kunnskap, forklarbarhet)

Integrasjonsmønstre:

- 1 **Neural** → **Symbolic**: Nevralt nettverk produserer symboler for resonnering
- 2 **Symbolic** → **Neural**: Symbolsk kunnskap guider nevralt nettverk
- 3 **Hybrid**: Tett integrasjon der begge informerer hverandre

Eksempel: Medisinsk diagnose (hjernetumor)

Nevral: CNN analyserer MR-bilde, identifiserer tumor-features

Symbolisk: Regelbasert klassifikasjon iht. [WHO-kriterier](#)

Integrert: Kombinerer bildefunn med kliniske regler for diagnose

S03: Beskrive hva en kunnskapsgraf er

Kunnskapsgraf (Knowledge Graph):

- Strukturert representasjon av kunnskap som en **graf**
- Noder = **entiteter** (personer, sykdommer, legemidler)
- Kanter = **relasjoner** (“behandles med”, “forårsaker”)

Trippel-representasjon:

(Subjekt, Predikat, Objekt) – (*Diabetes, behandles_med, Metformin*)

Anvendelser:

- [Google Knowledge Graph](#)
- Medisinsk beslutningsstøtte
- Legemiddelinteraksjoner

Fordeler:

- Maskinlesbar kunnskap
- Resonnering over relasjoner
- [Linked Data](#)

Teknologier

[RDF](#), [OWL](#), [SPARQL](#), [SHACL](#) – standarder for kunnskapsrepresentasjon, spørring og validering

S04: Kjenne til medisinske ontologier

Ontologi = formell spesifikasjon av begreper og relasjoner i et domene

Viktige medisinske ontologier/terminologier:

Navn	Beskrivelse	Anvendelse
SNOMED CT	Kliniske termer, 350 000+ begreper	EPJ, klassifisering
ICD-10/11	Internasjonal sykdomsklassifikasjon	Rapportering, statistikk
LOINC	Laboratorieundersøkelser	Labresultater
RxNorm	Legemidler	Forskrivning
WHO CNS 2021	CNS-tumorklassifikasjon	Gliomdiagnose

WHO 2021 CNS-klassifikasjon

Integrerer histologiske og molekylære kriterier. Eksempel: IDH-mutant astrocytom krever både histologi og IDH-mutasjonsstatus.

S05: Diskutere fordeler med nevrosymbolisk AI i medisin

Potensielle fordeler:

- 1 **Forklarbarhet:** Symbolsk komponent gir tolkbare resonnementer
- 2 **Dataholdighet:** Symbolsk kunnskap kompenserer for lite data
- 3 **Robusthet:** Regelbaserte begrensninger forhindrer absurde prediksjoner
- 4 **Oppdaterbarhet:** Ny medisinsk kunnskap legges til som regler/fakta
- 5 **Samsvar med retningslinjer:** Kode WHO-kriterier, guidelines direkte
- 6 **Validering:** Sjekk om prediksjoner er konsistente med kjent kunnskap

Hvorfor dette er viktig i medisin

Medisin er et domene der ekspertkunnskap er rik, forklarbarhet er kritisk, og feil kan være fatale. Ren dyplæring er ofte utilstrekkelig.

S06: Nevrosymbolsk AI for gliomdiagnostikk

Case study fra Lab 3, Notebook 08:

Problemet:

- Gliom-klassifikasjon krever både **bildeanalyse** (MRI) og **molekylære markører** (IDH, MGMT, 1p/19q)
- WHO 2021 krever integrasjon av flere informasjonskilder

Nevrosymbolsk løsning:

- 1 **Nevral komponent:** CNN segmenterer tumor i MRI (BraTS)
- 2 **Symbolsk komponent:** WHO-kriterier kodet som kunnskapsgraf
- 3 **Integrasjon:** Regelbasert klassifikator bruker CNN-output + molekylære markører

Resultat

Diagnose som er både **bildebasert** (CNN) og **kriteriekonform** (WHO) – med forklarbar resonnering.

S07: Forklare hvordan kunnskapsgrafer kan validere nevro-onkologiske prediksjoner

Validering av **nevro-onkologiske** prediksjoner med symbolsk kunnskap:

Scenario:

- CNN predikerer: "Glioblastom med IDH-mutasjon"
- Kunnskapsgraf inneholder: "Glioblastom er per definisjon IDH-villtype (WHO 2021)"
- **Konflikt detektert!**

Valideringsmekanismer:

- 1 **Konsistenssjekk:** Er prediksjonen logisk konsistent?
- 2 **Regelvalidering:** Oppfyller prediksjonen nødvendige kriterier?
- 3 **Plausibilitetssjekk:** Er kombinasjonen av funn realistisk?

Verdi

Fanger opp potensielle CNN-feil før de når kliniker. Øker tilliten til systemet.

S08: Definere agentisk AI og dens kjerneegenskaper

Agentisk AI = AI som handler autonomt mot mål

Kjerneegenskaper:

- 1 **Autonomi:** Handler (gjærne iterativt) uten kontinuerlig menneskelig input
- 2 **Målrettet:** Jobber mot definerte mål
- 3 **Planlegging:** Bryter ned komplekse mål i deloppgaver
- 4 **Verktøybruk:** Kan bruke eksterne verktøy (API, databaser)
- 5 **Refleksjon:** Evaluerer egen fremgang, justerer strategi
- 6 **Minne:** Husker kontekst over flere interaksjoner og iterasjoner

LLM-baserte agenter:

- LLM som “hjerne” og “orkestrator” – resonnerer og planlegger
- Verktøy-kall ([function calling](#)) for handlinger
- Eksempler: [AutoGPT](#), [LangChain](#), [Claude tool use](#), [OpenAI Operator](#), [CrewAI](#)

S09: Beskrive hvordan en AI-agent kan orkestrere klinisk arbeidsflyt

Eksempel: Agentisk AI for gliomutredning

Scenario: Lege ber om “Utred ny MR for mulig gliom”

Agenten kan:

- 1 🔍 **Hent data:** Koble til [PACS](#), laste ned MR-bilder
- 2 🧠 **Analyser:** Kjør [BraTS-segmentering](#), beregn tumorvolum
- 3 📖 **Litteratur:** Søk [PubMed](#) for relevante studier
- 4 🗄️ **Biobank:** Sjekk om molekulær analyse er tilgjengelig
- 5 📄 **Klassifiser:** Bruk WHO-kriterier for tentativ diagnose
- 6 📋 **Rapporter:** Generer strukturert rapport for [MDT-møte](#)

Orkestrering

Agenten **koordinerer** flere AI-systemer og datakilder mot ett klinisk mål – uten at legen må gjøre hvert steg manuelt.

S10: Forklare konseptet human-in-the-loop i agentiske systemer

Human-in-the-loop (HITL) i agentisk AI:

- Selv autonome agenter trenger **menneskelig tilsyn**
- Spesielt viktig for beslutninger med store konsekvenser

Implementeringsmønstre:

- 1 **Godkjenningspunkter:** Agent pauser for godkjenning før kritiske handlinger
- 2 **Konfidensterskel:** Menneske involveres når agent er usikker
- 3 **Audit trail:** Alle handlinger logges for review
- 4 **Overstyring (Overriding):** Menneske kan avbryte eller korrigere agent

I medisinsk kontekst

Agent kan samle data og foreslå diagnose, men **legen** tar endelig beslutning. Agenten er en “superkraftig assistent”, ikke en autonom beslutningstaker.

S11: Diskutere etiske utfordringer med autonome AI-agenter i helsevesenet

Etiske utfordringer:

- ❶ **Ansvar:** Hvem er ansvarlig når en agent gjør feil?
 - Utvikler? Sykehus? Lege som “slapp løs” agenten?
- ❷ **Autonomi og kontroll:** Hvor mye autonomi er for mye?
 - Risiko for utilsiktet handling
- ❸ **Transparens:** Kan vi forstå hvorfor agenten tok beslutningene?
- ❹ **Sikkerhet:** Agenter med tilgang til systemer = angrepsflate
- ❺ **Dehumanisering:** Risiko for å redusere menneskelig kontakt i helse
- ❻ **Overtillit:** “Automation bias” – blindt stole på agenten

Prinsipper for ansvarlig agentisk AI

Gradvis autonomi, robust HITL, klar ansvarsfordeling, omfattende testing, kontinuerlig overvåking

Oppsummering: Nevrosymbolisk AI og Agentisk AI

Nevrosymbolisk AI:

- S01–S02: Kombinerer nevralt (læring) og symbolisk (resonnering)
- S03–S04: Kunnskapsgrafer og medisinske ontologier
- S05–S07: Fordeler i medisin, case gliomdiagnose, validering

Agentisk AI:

- S08: Autonome, målrettede AI-systemer med planlegging
- S09: Orkestrering av kliniske arbeidsflyter
- S10–S11: HITL og etiske utfordringer

Lab 3, Notebook 08

Utforsk en detaljert case study som demonstrerer nevrosymbolisk gliomklassifisering og agentisk orkestrering.