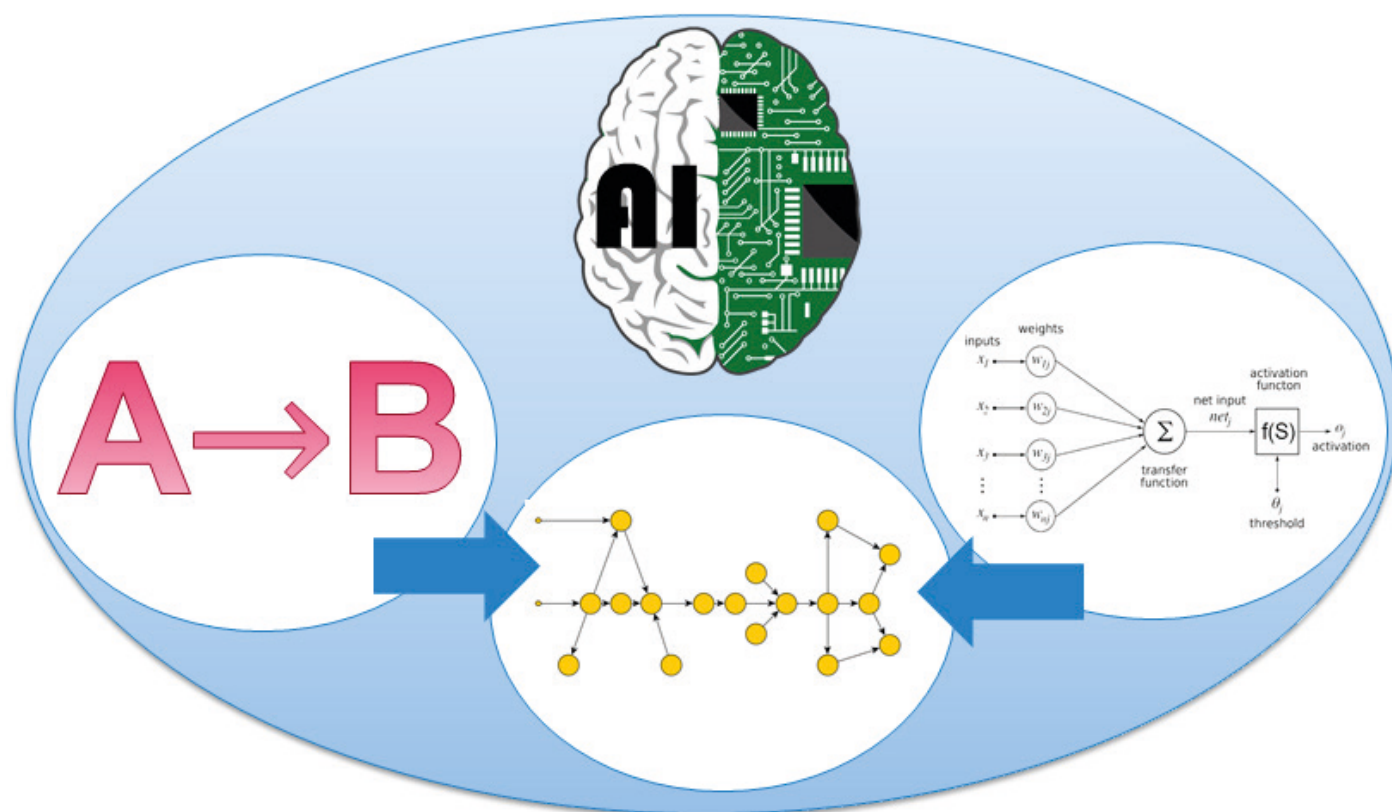


Avskanning av AI-området neuro-symboliskt resonerande

Resultat och kortfattat om metod

ANNA LINDBERG, CECILIA DURING,
RONNIE JOHANSSON, JOHAN SCHUBERT,
GÖRAN KINDVALL, EVA DALBERG



Anna Lindberg, Cecilia During, Ronnie Johansson,
Johan Schubert, Göran Kindvall, Eva Dalberg

Avskanning av AI-området neuro-symboliskt resonerande

Resultat och kortfattat om metod

Titel	Avskanning av AI-området neuro-symboliskt resonerande – Resultat och kortfattat om metod
Title	Horizon scanning of the AI domain Neuro-Symbolic Reasoning
Rapportnr	FOI-R--5224--SE
Månad	December
Utgivningsår	2021
Antal sidor	64
ISSN	1650-1942
Uppdragsgivare	Försvarsmakten
Forskningsområde	Övrigt
FoT-område	Inget FoT-område
Projektnr	E12123
Godkänd av	Malek Finn Khan
Ansvarig avdelning	Försvarsanalys

Bild/Cover: Geetika Saini (Wikimedia commons), Gordon Dylan (Open clipart), Ronnie Johansson

Detta verk är skyddat enligt lagen (1960:729) om upphovsrätt till litterära och konstnärliga verk, vilket bl.a. innebär att citering är tillåten i enlighet med vad som anges i 22 § i nämnd lag. För att använda verket på ett sätt som inte medges direkt av svensk lag krävs särskild överenskommelse.

This work is protected by the Swedish Act on Copyright in Literary and Artistic Works (1960:729). Citation is permitted in accordance with article 22 in said act. Any form of use that goes beyond what is permitted by Swedish copyright law, requires the written permission of FOI.

Sammanfattning

Neuro-symboliskt resonerande (NSR) är ett fält inom AI-forskningen som på senare år rönt allt större intresse. En avskanning av forskningsfronten inom området har utförts som visar att:

- NSR är en del av 3:e vågens AI och tros kunna bidra till att lösa kritiska utmaningar såsom förklarbarhet. Den 3:e vågens AI kopplar samman den första och andra AI-vågen, dvs. resonerande, symbolisk AI (1:a vågen) och lärande AI, exempelvis konnektionistisk AI genom neurala nätverk (2:a vågen).
- Flera olika typer av aktörer bedriver forskning och utveckling inom området. Mängden forskning och antalet aktörer ser ut att ha ökat under senare år. Stora internationella företag bedriver nu verksamhet inom området tillsammans med akademi.
- NSR är ett relevant teknikområde för försvar och säkerhet. Det finns många olika tillämpningsområden och en relativt stor del av forskningen har varit/är finansierad av försvarsorganisationer eller organisationer med försvarsanknytning.

Sammanfattningsvis observerar vi att avskanningsresultaten påvisar potentiella tillämpningsområden av NSR med intresse för försvar och säkerhet, såsom klassificering, resonerande AI/beslutsstöd, robotik/autonomi, lära av expert/simulering samt textanalys. FOI bedriver redan AI-relaterad forskning och vissa av NSR-teknikerna ligger i nära anslutning till FOI:s befintliga forskning. Vi har funnit att NSR är militärt relevant, kanske särskilt för tillämpningar där det idag saknas eller endast finns otillräckliga tekniska lösningar.

Vår slutsats efter denna avskanning är att en låg, men rimlig, ambition är att fortsatt följa utvecklingen av NSR-området inom berörda forskningsområden, men vi har inte närmare analyserat var inom FOI som NSR kan vara av störst intresse.

För underlag om genomförande av processen eller fördjupade resultat från arbetet hänvisas till FoT-projektet Avskanning av forskningsfronten.

Nyckelord: Neuro-symboliskt resonerande, NSR, artificiell intelligens, AI, avskanning, forskningsfront, horizon scanning, robotik, autonomi, textanalys

Summary

Neuro-symbolic reasoning (NSR) is a technology area within the AI domain that recently has received growing attention. Our horizon scanning of NSR shows that

- NSR belongs to the so-called third wave of AI, and may provide solutions to critical challenges such as explainable AI (XAI). The third wave of AI connects the first and second AI waves, i.e., reasoning and symbolic AI (1st wave) with learning AI such as connectionist AI through neural networks (2nd wave).
- Different actors pursue research and development within this field. Over the past few years the amount of research and the number of actors has increased. Major international commercial enterprises are among these actors together with prestigious academia.
- NSR is a relevant technology area for defence and security. There are several potential applications. A relatively large fraction of the research is or has been funded by defence organisations or organisations with an interest in defence.

To summarize, our observation is that NSR may provide applications with relevance for defence and security, such as classification, decision support, robotics/autonomy, learning from experts/simulation and text analysis. FOI already pursues AI-related research in these areas and some of the NSR technologies are related to this ongoing research. We have found that NSR is relevant for defence, in particular for applications where technologies today are missing or are insufficient.

Our conclusion is that a low, but reasonable, ambition is to follow the continuing development of NSR within related research projects, but we have not analysed where the potential interest or relevance for NSR is largest within FOI.

For further details about the results and the execution of this horizon scanning activity, we refer to the Horizon scanning project, which is part of the Swedish Armed Forces joint programme for Horizon Scanning and Technology Watch.

Keywords: Neuro-symbolic reasoning, NSR, artificial intelligence, AI, horizon scanning, robotics, autonomy, text analysis

Innehållsförteckning

1	Rapporten i sammandrag.....	7
1.1	Sammanfattande översikt av resultat	7
1.2	Rekommendationer för AI-området NSR	8
2	Inledning	9
2.1	Syfte och mål för avskanning av NSR.....	9
3	Avskanningsprocessen.....	10
4	Neuro-symboliskt resonerande – en fördjupande introduktion	13
4.1	Bakgrund.....	13
4.2	Egenskaper.....	14
4.3	Konstruktioner: hybrid till integrerad.....	14
4.4	Uttrycksfullhet	15
4.5	Relaterat arbete	16
5	Resultat	17
5.1	Den inledande navigeringen - domänkartan	17
5.2	Resultat från den breda skanningen	18
5.3	Resultat från den djupa skanningen.....	22
5.4	Resultat från värderingen	22
5.5	Analys och grunder för rekommendationer	27
6	Erfarenheter av processen.....	29
7	Referenser.....	30
	Bilaga 1 - Översikt av innehåll i kluster	32
	Bilaga 2 - Teknik- och tillämpningskort	35
	Kluster 1 - Tracking	36
	Kluster 2 - Textanalys	37
	Kluster 3 - Lågnivå (läges)bild	38
	Kluster 4 - Högnivå (läges)bild.....	41
	Kluster 5 - Beslutsstöd och delfunktioner	45
	Kluster 6 - Lära av expert/Simulering	47
	Kluster 7 - Stöd för förstärkt inläring	48
	Kluster 8 - Övrigt	49
	Bilaga 3 – Metod för genomförande	54
	Steg 1, problemformulering och utformning av avskanningen	54
	Steg 2, bred och djup skanning	57
	Steg 2.5, hantering av skanningsunderlag	58
	Steg 3, värdering och analys	59
	Steg 4, resultat och kommunikation.....	61
	Bilaga 4 - Anteckningar från brainstorm av tillämpningsområden.....	62

1 Rapporten i sammandrag

1.1 Sammanfattande översikt av resultat

Vi presenterar i detta avsnitt en översikt av de resultat som skapats genom hela avskanningsprocessen. För att underlätta spårbarhet hänvisas till de kapitel i rapporten där respektive resultat redovisas.

Ett inledande arbete, som tjänade som utgångspunkt för den fortsatta avskanningen, är den så kallade domänkartan som visar en översiktlig kartläggning av området NSR. Den är också ett viktigt resultat från studien. Domänkartan skapades tidigt i arbetet av samtliga medarbetare i projektet och visar på en bredd av aktörer, tillämpningar av potentiellt militärt intresse samt områden där forskningen är särskilt aktiv och där det finns stora behov av tekniska lösningar. Domänkartan beskrivs i avsnitt 5.1.

En större mängd litteratur har lästs av de två tekniska/vetenskapliga experter som deltagit i avskanningen. Litteratur som samlats in genom en manuell bred skanning sammanfattas i den inledande översikten över området i kapitel 4.

En bred avskanning baserad på en automatiserad sökning av publicerad vetenskaplig litteratur har bland annat möjliggjort att studera vem som publicerat peer-review-material om olika delområden, författarnas geografiska hemvist samt vilka finansiärerna varit över tid. Genom att studera publikationer från de senaste dryga tre åren och jämföra med utfall från en längre tidsperiod har de aktuella trenderna kunnat studeras. Dessa resultat presenteras i avsnitt 5.2.

Med hjälp av kriterier¹ valde experterna sedan ut särskilt intressanta artiklar ur de breda sökningarna för att fördjupa skanningen (se avsnitt 5.3). Teknikkort togs fram baserade på de utvalda artiklarna och finns i sin helhet i bilaga 2. Korten användes sedan vid en workshop med syfte att värdera vad som skulle kunna vara av särskilt intresse för det militära området. I bilaga 1 redovisas klustrad information om NSR som användes vid genomförandet av värderingsworkshopen. Vid denna workshop deltog ett antal av FOI:s forskningsområdesföreträdare samt teknisk/vetenskaplig expertis med fokus på framförallt AI-områden. Ett mindre antal representanter från FMV och FHS representerade såväl specifik teknisk som bred kompetens inom den militära domänen. Resultat från workshopen återfinns i avsnitt 5.4. Bilaga 4 redogör för de diskussioner som fördes under värderingsworkshopen.

Sammanfattningsvis observerar vi att:

- NSR förmodas kunna bidra till att nå den 3:e vågens AI och bidra till att lösa kritiska utmaningar såsom förklarbarhet.
- Flera olika typer av aktörer bedriver forskning och utveckling inom området. Mängden forskning och antalet aktörer ser ut att ha ökat under senare år. Stora internationella företag bedriver nu verksamhet inom området tillsammans med akademi.
- NSR är ett relevant teknikområde för försvar och säkerhet. Det finns många olika tillämpningsområden och en relativt stor del av forskningen har varit/är finansierad av försvarsorganisationer eller organisationer med försvarsanknytning.

Avskanningen påvisar potentiella tillämpningsområden av NSR för försvar och säkerhet, såsom klassificering, resonerande AI/beslutsstöd, robotik/autonomi, lära av expert/simulering och textanalys. Inom de flesta områdena bedriver FOI redan AI-relaterad forskning. Analys och slutsatser presenteras i avsnitt 5.5.

¹ Mognad, tillämpbarhet och relevans för militär AI användning

Genomförandet av avskanningen följde FOI:s avskanningsprocess. En redovisning av avskanningsprocessens fyra steg finns i kapitel 3. En sammanställning av metoder för genomförande återfinns i bilaga 3 och erfarenheter från arbetet redovisas i kapitel 6.

1.2 Rekommendationer för AI-området NSR

Vår slutsats är att FOI bör gå vidare med arbete med NSR. Vi tror att kontakter med samarbetspartners är en väg att gå liksom små fortsatta arbeten inom vissa av FOI:s egna forskningsområden. En låg, men rimlig, ambition är att fortsatt följa utvecklingen av NSR-området.

Fokus för fortsatt arbete bör i första hand vara att utreda tillämpning av NSR inom de forskningsområden som nyttjar AI-tekniker och har behov av möjliga tillämpningar NSR specifikt skulle kunna bidra med. För att underlätta omhändertagande föreslås muntlig presentation av resultat för intressenter.

Ett alternativ som vi är kluvna till är att fortsatt omhänderta NSR inom avskanningsprojektet, bland annat för att varje fördjupning behöver vägas mot möjligheten att avskanna ytterligare områden. En ytterligare avskanning skulle kunna samla information om var NSR bedöms ge, respektive inte bedöms ge, stort bidrag till försvarsspecifika AI-behov och -tillämpningar. Samma typ av arbete skulle emellertid kunna genomföras inom andra projekt med pågående forskning inom AI eller där det finns stora behov och potential för nyttiggörande av NSR.

2 Inledning

Neuro-symboliskt resonerande (NSR) är ett område inom AI som på senare år tillskrivits allt större intresse. Projektet Avskanning av forskningsfronten har under 2021 genomfört såväl en bred som en djup avskanning av området genom att använda den process som projektet tidigare arbetat fram (Lindberg o.a., 2021).

FOI:s arbete med avskanning genomförs inom projektet Avskanning av forskningsfronten 2021 (FM beställning AT.9221603 Avskanning FOI 21) som är en del av Försvarmaktens (FM) program för Forskningsskanning och teknisk prognos där Totalförsvarets Forskningsinstitut (FOI), Försvarets Materielverk (FMV) och Förvarshögskolan (FHS) samverkar under FM ledning. Projektet har i arbetet, särskilt vid den värderingsworkshop som genomfördes i slutet av arbetet, involverat såväl teknisk expertis som experter med mer övergripande inriktning vid FOI samt företrädare från FM, FHS och FMV.

I rapporten redovisar vi en övergripande bild av aktörer, inriktningar och syften med forskning inom området neuro-symboliskt resonerande. Vi beskriver även vad NSR är och redovisar NSR-tekniker som kan vara intressanta för försvar och säkerhet och hur vi värderat dessas relevans. Begreppen neuro och neural används synonymt, i likhet med hur de används i peer-review-artiklar. I rapporten fokuserar vi på resultaten från avskanningen av NSR, men beskriver även övergripande de metoder vi använt. Vi har utgått från FOI:s process för avskanning. Metoder som projektgruppen sedan tidigare är väl förtrogna med har valts vid genomförandet. Slutligen ger vi förslag för fortsatt hantering av området NSR för FOI. Vi ger råd om hur FOI kan gå vidare med området NSR, utgående från skanning, värdering av NSR-tekniker och efterföljande analys.

2.1 Syfte och mål för avskanning av NSR

Avskanning ska hitta det nya. NSR är ett område som inte är helt okänt för FOI, men väl ett område FOI inte bedriver egen forskning inom. Bakgrunden till att NSR föreslogs som ett ämne för avskanning är dels att forskare vid FOI över tid funnit området intressant, dels att NSR tillskrivs allt större intresse internationellt av en mängd aktörer. I projektets årsrapport (Dalberg o.a., 2021), beskrivs varför NSR valts för avskanning.

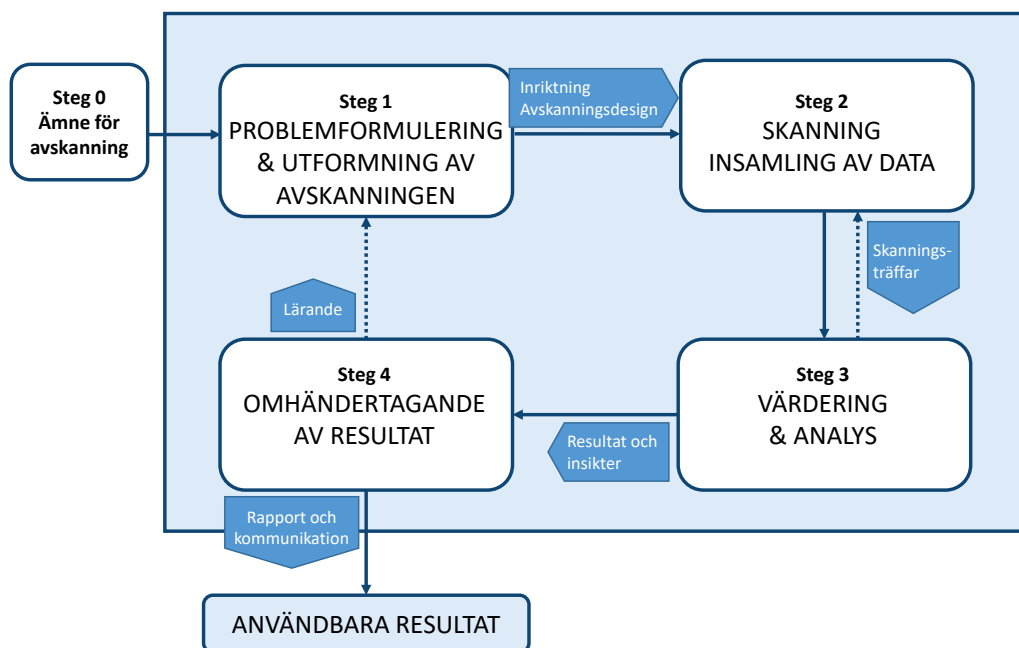
Syftet och målet för avskanningen är att ge ett underlag för att kunna avgöra om området är relevant för FOI att bedriva forskning inom och i så fall varför. Vi vill även orientera personal verksamma såväl inom som utom FOI, exempelvis inom FoT-programmet för forskningsskanning och teknisk prognos, om områdets status och möjliga tillämpningar.

Primära intressenter och målgrupp för föreliggande rapport är FOI:s forskningsområdesföreträdare (FoF) och FOI:s forskningsdirektör. Även FM, FMV och FHS är målgrupp för rapporten.

3 Avskanningsprocessen

En utmaning med den genomförda avskanningen var att skrapa på ytan, det vill säga att bara skrapa tillräckligt väl, för att kunna ge rekommendationer om det avskannande ämnets fortsatta omhändertagande utan att fastna i enskilda detaljer. Avskanningen syftade till att värdera området NSR som helhet. Detta har genomförts och FOI:s process för avskanning ligger till grund för genomförandet.²

Avskanningsprocessen består av fyra steg och varje steg innehåller ett antal aktiviteter och principer för att öka användbarheten av såväl processen som resultaten.



Figur 1 Avskanningsprocessens fyra steg.

Steg 0: Val av Ämne för avskanning

Arbetet inleddes med att identifiera ämnena att skanna. Detta ligger utanför processen för avskanning och vi beskriver det därför som steg 0. Vid genomförandet deltog all personal i projektet och val av ämnen för avskanning beslutades i samråd med FOI:s forskningsdirektör. Utöver NSR valdes vid samma tillfälle även Framtidens energi som ett avskanningsämne. Steg 0 beskrivs i Dalberg o.a. (2021).

Sammanfattningsvis valdes NSR då området fick högre poäng vid urval av ämnena i steg 0 än andra AI-tekniker. NSR rankades högst på grund av att det a) upplevs som hett forskningsmässigt och b) är ett sätt att koppla ihop 2:a vågens AI med 1:a vågens AI, vilket de andra teknikerna inte möjliggör. NSR tros kunna bidra till att fylla behov som förklarbarhet och en, vad vi i brist på bättre benämning i rapporten kallar, smartare AI.

Vi använder genomgående begreppet neuro-symboliskt resonerande. Vid sökningar i litteraturen har begreppen neuro och neural används synonymt, i likhet med hur de används i peer-review-artiklar.

² Vi redovisar i detta kapitel processens arbetsflöde. I Bilaga 3 redovisar vi övergripande hur respektive steg genomförts avseende metod, inriktning och avgränsningar. Erfarenheter från arbetet sammanfattas i kapitel 6. Avskanningsprocessen beskrivs i Lindberg o.a. (2021).

Steg 1: Problemformulering och utformning av avskanningen

Steg 1 innebär att:

- Tydliggöra mål och syfte med arbetet,
- Skapa ledstång för genomförandet.

Steg 1 syftar till att skapa stadga för fortsatt arbete, något som är viktigt vid avskanning som allt som oftast innebär att navigera i materia som till stor del är okänd eller osäker. En så kallad domänkarta är ett sätt att navigera.

Steg 2: Skanning och insamling av data

Vi valde att göra både en bred och en djup skanning. Syfte och mål för den breda skanningen var att:

- Övergripande kartlägga området, hitta termer till den djupa skanningen och förbättra förståelse samt förbättra domänkartan från steg 1.
- Redogöra för aktörer som bedriver FoU och möjliga tillämpningar.

Syfte och mål för den djupa skanningen var att:

- Identifiera resultat att ta vidare i processen, så kallade skanningsträffar. Ta fram underlag i form av särskilda tekniker eller utvecklingar som visar på teknikforskning och forskningshorisonten inom NSR.
- Lista skanningsträffar och skapa användbara underlag för värdering.
- Identifiera och med hjälp av kriterier välja ut teknikforskning för fortsatt analys
- Klustra de utvalda underlagen inför värdering.

Steg 3: Värdering och analys

Steg 3 innebar att redovisa, diskutera och kort värdera underlag från steg 2. Inför, och under inledningen av, steg 3 arbetade vi relativt mycket med att omhänderta och omforma underlagen från steg 2. Denna databearbetning benämner vi steg 2,5. Arbetet krävdes för att analysera redovisade skanningsträffar med syfte att integrera dessa till områden, dvs. att skapa kluster av skanningsträffarna. Arbetet behövdes även för att skapa underlag inför en värderingsworkshop i syfte att ge deltagarna bättre möjlighet att förstå området och genomföra värderingen. Syfte och mål för steg 3, värdering, var att:

- Presentera möjliga tillämpningar som kan ge lösningar på tekniska utmaningar som identifierats hos FOI, FM, FMV, FHS.
- Värdera relevans, exempelvis för:
 - FOI:s nuvarande eller framtida forskning,
 - FM:s behov och utmaningar.
- Värdera hur NSR fortsatt ska behandlas.

Steg 4: Omhändertagande av resultat

Syfte och mål med steg 4 är att rapportera resultat, kommunicera dessa med mottagare och följa upp användning av resultaten från avskanningen.

I övergången mellan steg 3 och steg 4 gjorde vi den slutliga analysen och syntesen av skanningen av NSR, skrev rapport och identifierade rekommendationer. Vi drog lärdom av erfarenheter av arbetet i processen. Vid skrivande av rapporten återstår ännu att:

- i. Förankra och kommunicera resultaten med intressenter,
- ii. Lära av erfarenheter avseende processteg och användande av underlag från avskanningen.

Punkterna i och ii planeras att omhändertas vid möte med FM, FMV och FHS. Erfarenhetsanalys görs vid internt projektmöte när rapporten är färdig.

Arbetsgrupper och möten

Avskanningen har genomförts i olika konstellationer. En *större avskanningsgrupp* bestående av all personal i avskanningsprojektet (7 personer) engagerades framförallt i problemformuleringen och framtagande av domänkartan, i värderingsworkshopen och vid redo-

visning av resultat. En *mindre avskanningsgrupp* (4 personer) bestående av avskanningsledare/metodexperter arbetade kontinuerligt genom projektet. Metodexperter för HSTOOL³ och tillika ämnesexperter på AI lade stor kraft på genomförandet av själva skanningen (såväl den breda som den djupa).

Avstämningsmöten med projektledare genomfördes kontinuerligt, både formellt och informellt.

FOI:s forskningsområdesföreträdare (FoF) och tekniska experter från FOI engagerades inför och under värderingen. Vid den värderingsworkshop som genomfördes i slutfasen av arbetet deltog även representanter från FMV, FHS och Inspektionen för strategiska produkter (ISP).

På grund av Covid var alla möten digitala. Vi hade inte tillgång till något särskilt verktyg anpassat för kreativt arbete på distans utan använde PowerPoint-presentationer, anteckningar i PowerPoint och Word som hjälpmedel.

Avskanningen har resulterat i en stor mängd underlag vilka sparas av projektet som arbetsmaterial.

³ HSTOOL är en datorbaserad automatisk metod utvecklad av FOI för avskanning av forskningsfronten genom sökningar i publicerad vetenskaplig litteratur, som används för att upptäcka vetenskapliga trender (Karasalo, Schubert, 2019).

4 Neuro-symboliskt resonerande – en fördjupande introduktion

Kapitlet ger en översikt över området neuro-symboliskt resonerande. Översikten är ett resultat av såväl den breda som den djupa skanningen. Sammanställningen ger en inledning till AI-delområdet NSR och vi har därför valt att redovisa underlaget inledningsvis i denna rapport.

4.1 Bakgrund

Under det senaste decenniet har det datavetenskapliga området artificiell intelligens (AI) erhållit ett enormt intresse, både för civila och militära tillämpningar, främst till följd av framgångarna inom delområdet djupinlärning (*eng.* deep learning) (Bengio, 2016).⁴ Resultaten, som i litteraturen ofta beskrivs som övermänniskliga (*eng.* super human), återfinns främst inom tillämpningsområdena bild- och textanalys, och inom olika former av spel som schack och go (Vincent, 2019). Framgångarna till trots har vissa begränsningar och utmaningar med djupinlärning identifierats, även av upphovsmännen (Bengio, Hinton och LeCun) själva (Dickson, 2021).

Den amerikanska försvarsforskningsfinansiären Defense Advanced Research Projects Agency (DARPA) som med sin projektportfölj bland annat finansierar AI-relaterad forskning, delar in AI-utvecklingen i tre vågor (Wang, 2017).

Första vågen (med start på 1960-talet) omfattar automatiserat resonerande (logisk deduktion) med mänskligt hårdkodad expertkunskap (symbolisk AI), vilket var den dominerande inriktningen under AI:s barndom. Förhoppningarna var stora men uteblivna resultat ledde fram till den så kallade "AI-vintern". Ett problem var att första vågens AI-system saknade förmåga att lära från data och resonera med osäker information.⁵

Den andra vågen (från 1990-talet) omfattar enligt DARPA metoder för att lära modeller från (eventuellt osäker) data. Ett exempel på metodområde är artificiella neuronnät (ANN). Detta metodområde, som även kallas konnektionistisk AI, har använts mycket framgångsrikt för djupinlärning som nämnts ovan, men metoderna uppvisar vissa avgörande brister, bland annat att de kan manipuleras och dra felaktiga slutsatser med potentiellt katastrofala följder.

DARPA ser behov av större tillförlitlighet, bättre förklarbarhet, och har ett behov av metoder med mindre krav på tillgång till träningsdata. De ser framför sig den tredje vågens AI som sammanför erfarenheterna från de båda AI-vågorna, dvs. både resonerande (symbolisk AI) och lärande AI (exempelvis konnektionistisk ANN).⁶

Olika förslag att sammanföra lärande och resonerande inom AI har presenterats. Ett förslag handlar om att använda och sammanföra (metoderna) logiskt resonerande med probabilistiska grafiska modeller (de Raedt, Dumancic, Manhaeve, & Marra, 2020). Ett annat förslag är att sammanföra logiskt resonerande med ANN och metoderna från djupinlärning, vilket är det område, neuro-symboliskt resonerande (NSR⁷), som avskannas i det här projektet.

⁴ Djupinlärning är en klass av maskininlärningsmetoder som har varit väldigt framgångsrik under det senaste decenniet och bygger på att man använder artificiella neuronnät med betydligt fler neuroner än i tidigare generationer. En introduktion och översikt över djupinlärning finns exempelvis på https://en.wikipedia.org/wiki/Deep_learning.

⁵ I detta sammanhang används begreppen "lära" och "resonera" på det sätt dessa används inom AI-litteraturen, dvs. att AI-metoderna ges en aktiv roll. Här avses med lärandet ett maskinbaserat lärande, ett artificiellt automatiskt resonerande etc. Detta gäller även liknande begrepp, såsom "förledas" eller att de "drar" slutsatser. Vi skriver i texten inte alltid ut kopplingen till det maskinella eller algoritmiska utan använder begrepp som lära, resonera och så vidare.

⁶ Notera att upphovsmännen och förgrundspersonerna för deep learning inte delar denna uppfattning utan jobbar på att utveckla djupinlärningens ANN för att hantera de problem som identifierats. (Dickson, 2021)

⁷ Alternativa namn är neural-symbolic reasoning and learning, neural-symbolic integration, och neurosymbolic AI.

4.2 Egenskaper

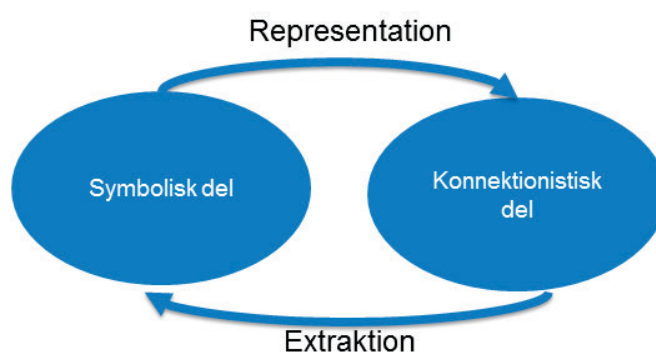
Den neuro-symboliska forskningen baseras på ursprungsidéer om mänskligt tänkande. Ett övergripande mål med NSR är att skapa AI med förmåga till både lärande (maskininläring/neurala nätverk) och resonerande (symbolisk logik). Denna kombination av två olika AI-metoder är ett sätt att bidra till en smartare och mer generell AI, det som ibland benämns ”smart” AI. NSR kombinerar de båda AI-teknikerna neurala nätverk och symbolisk logik (logisk och symbolisk AI). För att förstå NSR är det nyttigt att titta närmare på egenskaperna hos symbolisk och konnektionistisk AI.

Fördelar med första vågens symbolisk AI är att den stödjer logisk deduktion och abstraktion, samt att den expertkunskap som nyttjas i systemet lätt kan uttryckas och förstås av mänskliga användare, vilket skapar förtroende för resultaten. Nackdelarna uppstår i kontakten med verkligheten och dess komplexa och osäkra data. De data som experter inledningsvis matar in eller representerar är i normalfallet delmängder av all data som modellen därefter kommer behöva hantera. Det betyder att den symboliskt representerade expertkunskapen (i vissa fall) är otillräcklig för att fullständigt representera och slutligen lösa det aktuella problemet. Den tar inte hänsyn till de utfallsrum eller osäkerheter som kan komma att uppstå. Detta beror både på att det är svårt att uttrycka vissa relationer samt att information kan vara svåråtkomlig i stora datamängder och därmed kommer att förbises när experter används för att skapa den symboliska representationen.

Den andra vågens konnektionistiska AI, å andra sidan, kan effektivt hantera mycket stora mängder data, och extrahera mindre explicit information och samband, men hur resultatet från tillämpningen av dessa metoder framkommit blir svårförklarligt och resultaten kan därför misstros av människan.

4.3 Konstruktioner: hybrid till integrerad

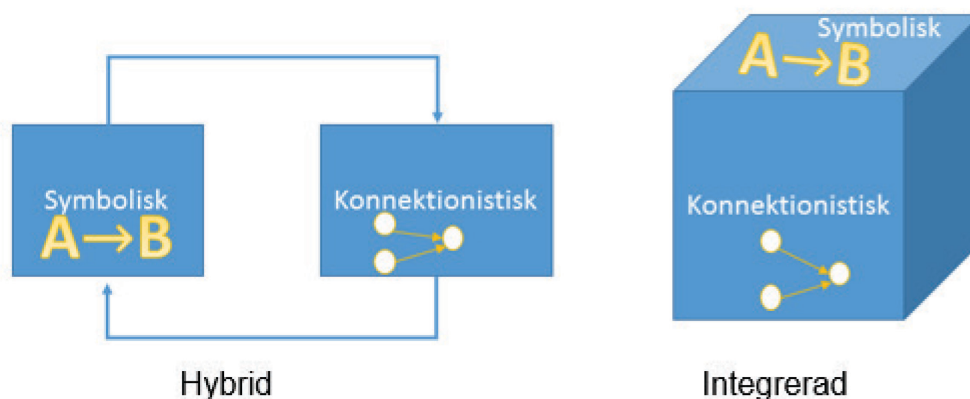
Ett sätt att åstadkomma neuro-symboliskt resonerande är att sammanföra en komponent för logiskt resonerande (symbolisk) och en konnektionistisk (ANN) komponent till en hybrid⁸ av sammanlänkade system. En hybrid i denna mening är en pragmatisk lösning som bygger på redan välutvecklade metoder inom symbolisk respektive konnektionistisk AI. En hybridmetod förutsätter typiskt att satser i den logiska komponenten kan användas för att skapa ett neuronnät, samt att logiska regler kan extraheras ur neuronnätet. Det här beskrivs i den så kallade *neural-symbolic learning cycle* (Svatoš, Šourek, & Železný, 2019) som finns avbildad i Figur 2.



Figur 2 Hybridmetoder kan till exempel bestå av en symbolisk del och en konnektionistisk del. Dessa samverkar genom "Neural-symbolic learning cycle" där symbolisk kunskap överförs (representation) till en ANN-struktur (i bästa fall utan approximationer) som kan anpassas via inläring från tillgänglig data. Det uppdaterade ANN:et uttrycks sedan återigen i symboler (extraktion) och behandlas av den symboliska delen (exempelvis genom att lösa upp motstridigheter som kan ha uppstått genom inläringen). Därefter kan cykeln upprepas.

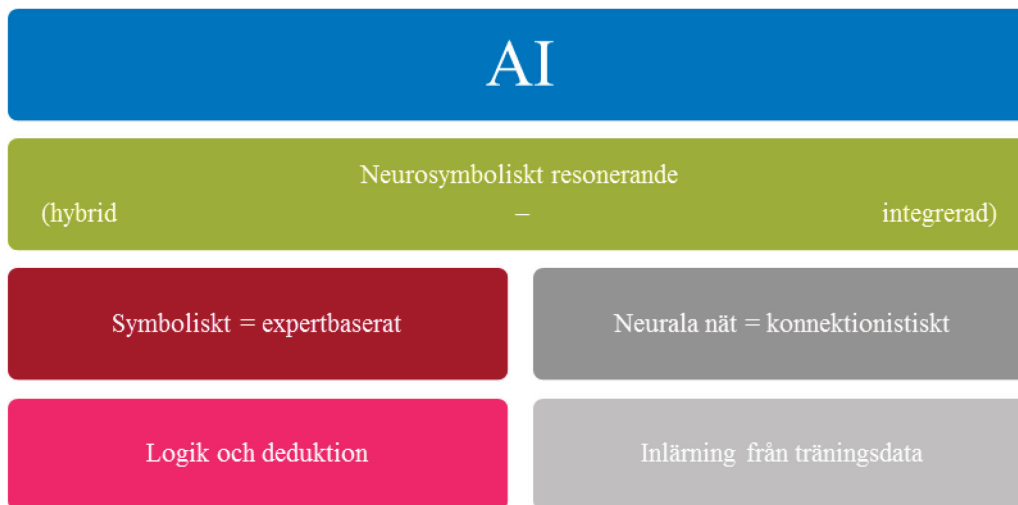
⁸ *Hybrid intelligent systems* (HIS) är ett samlingsnamn för sammansättningar av AI-metoder för att lösa något visst problem eller imitera någon egenskap av generell intelligens. HIS är ett generellt begrepp och ett sådant system behöver inte bestå av just symboliska och konnektionistiska AI-komponenter. Ett HIS behöver inte heller bestå av just två separata men kommunicerande metoder som i NSR utan en AI-metod kan använda en annan AI-metod för lösning av ett delproblem (exempelvis kan en AI-metod användas för att välja styrparametrar till en annan AI-metod).

Om hybridmetoder befinner sig i ena änden av en skala, så återfinns så kallade integrerade metoder i den andra änden. En integrerad metod fokuserar på ett neuronnät och vilken typ av logik metoden kan uttrycka. För en integrerad metod behövs normalt ingen extraktion av (approximativa) logiska satser för det aktuella neuronnätets logiska egenskaper, utan det logiska resonerandet uttrycks snarare med hjälp av ANN:er. Den principiella skillnaden mellan dessa två extremer illustreras också i Figur 3.



Figur 3 Illustration av skillnaden mellan en hybridmetod och en integrerad metod. En hybridmetod kan till exempel bestå av två separata delar, en symbolisk och en konnektionistisk del. Dessa delar samverkar på olika sätt för att sammanföra de två delarnas fördelar. En rent integrerad metod, å andra sidan, kan samtidigt uttryckas både symboliskt och konnektionistiskt.

En viss NSR-metod kan anses befinna sig någonstans mellan dessa ändpunkter, mer eller mindre hybrid eller integrerad (d'Avila Garcez & Lamb, 2020). Figur 4 sammanfattar en del av det som presenterats här om NSR.



Figur 4 NSR är ett delområde inom AI. En viss NSR-metod befinner sig någonstans på en skala mellan att vara en hybrid mellan en symbolisk och en konnektionistisk komponent, eller en metod där både den symboliska och konnektionistiska delen är integrerade. Den symboliska, eller expertbaserade, dimensionen bygger på logik och deduktion. Den konnektionistiska dimensionen bygger på neurala nätverk tränade med hjälp av utvalda data.

4.4 Uttrycksfullhet

Ett problem som man hoppas NSR ska kunna bidra till att lösa är ANN:ers oförmåga att uttrycka och resonera med hjälp av logik, vilket är en central utmaning för NSR-området. En av AI-områdets grundare, John McCarthy, betecknade ANN:er som att ha en "propositional fixation" syftandes på dess oförmåga att representera något mer uttrycksfullt än satslogik (*eng.* propositional logic), det vill säga en begränsad form av logik (McCarthy, 1988).

Sedan McCarthys kritik år 1988 har uttrycksfullheten utökats bland annat genom att kombinera ett flertal ANN på olika sätt (d'Avila Garcez, o.a., 2015).

En viktig form av logik som ännu inte kan uttryckas på ett integrerat sätt är vad som kallas första ordningens logik (ofta förkortat FOL) (d'Avila Garcez & Lamb, 2020; d'Avila Garcez, 2005). Första ordningens logik är exempelvis användbart för att representera den information som kan uttryckas i naturligt språk.⁹ En hybridmetod som sammanför första ordningens logik och djupinlärning är *Logic tensor networks* (Badreddine, d'Avila Garcez, Serafini, & Spranger, 2022).

4.5 Relaterat arbete

Distinktionen mellan hybrida och integrerade lösningar till NSR har redan nämnts och är användbar för att skilja mellan olika sätt att få till stånd NSR. Inom robotikområdet talar man om hybrider av symboliska och undersymboliska (*eng.* symbolic and sub-symbolic) system där, analogt med NSR, den symboliska delen typiskt omfattar logiskt resonerande, och där den undersymboliska delen kan innehålla ANN men där även andra maskininlärningsmetoder är möjliga (Kelley, Avery, Long, & Dimperio, 2009). Man kan likaledes tänka sig en minimalistisk undersymbolisk nivå som enbart består av mönster av ”rå” sensordata som motsvarar symboler på den högre nivån.

Kritik mot djupinlärning kommer också från professor Judea Pearl¹⁰. För ytterligare resonemang om djupinlärningens oförmåga att hantera kausala samband, se (Schubert & Johansson, 2020). Ett eget delområde har skapats kring detta behov, men forskarna inom NSR hävdar att NSR-metoder kan hantera de brister som pekats ut av Pearl (d'Avila Garcez & Lamb, 2020).

Som redan nämnts ovan så finns olika ansatser till att sammanföra första och andra vågens AI, dvs. resonerande och lärande. Till exempel jämför de Raedt o.a. (2020) SRL (*eng.* stochastic relational learning) med NSR och hittar vissa likheter.

⁹ *First-order logic in artificial intelligence*. Hämtat från <https://www.javatpoint.com/first-order-logic-in-artificial-intelligence> den 13 augusti 2021.

¹⁰ Pearl, verksam vid University of California, Los Angeles, är den ledande forskaren inom AI-området kausalitet. För mer information, se Pearl (2000).

5 Resultat

Resultat från olika steg i processen presenteras nedan på följande sätt. I avsnitt 5.1 presenteras domänkartan, som i huvudsak är ett resultat från steg 1, men som kompletterats efterhand under de efterföljande stegen. Andra resultat från steg 1 presenteras framförallt i bilaga 3. Steg 2, skanningen där insamling av data sker, gjordes på två sätt. I avsnitt 5.2 presenteras resultat från en bred skanning. Här har olika typer av källor använts för att brett belysa NSR. Insamlingen skedde såväl automatiserat som manuellt. Resultat från detta arbete utgör även den huvudsakliga grunden till den inledande beskrivningen av NSR-området som presenterades i kapitel 4. I avsnitt 5.3 presenteras sedan resultat från den andra delen av steg 2 där en automatiserad sökning i vetenskaplig litteratur kombinerad med manuell sökning av experterna ger ett smalare, men djupare, perspektiv på området. Detta då experterna läst utvald litteratur och med detta som grund skrivit ett antal teknik- och tillämpningskort vilka klustrats i ett fåtal delområden. Resultat från den värdering av dessa teknik- och tillämpningskort som gjordes vid en workshop i steg 3 presenteras sedan i avsnitt 5.4. Avsnitt 5.5 innehåller analys och slutsatser i form av rekommendationer (steg 4).

5.1 Den inledande navigeringen - domänkartan

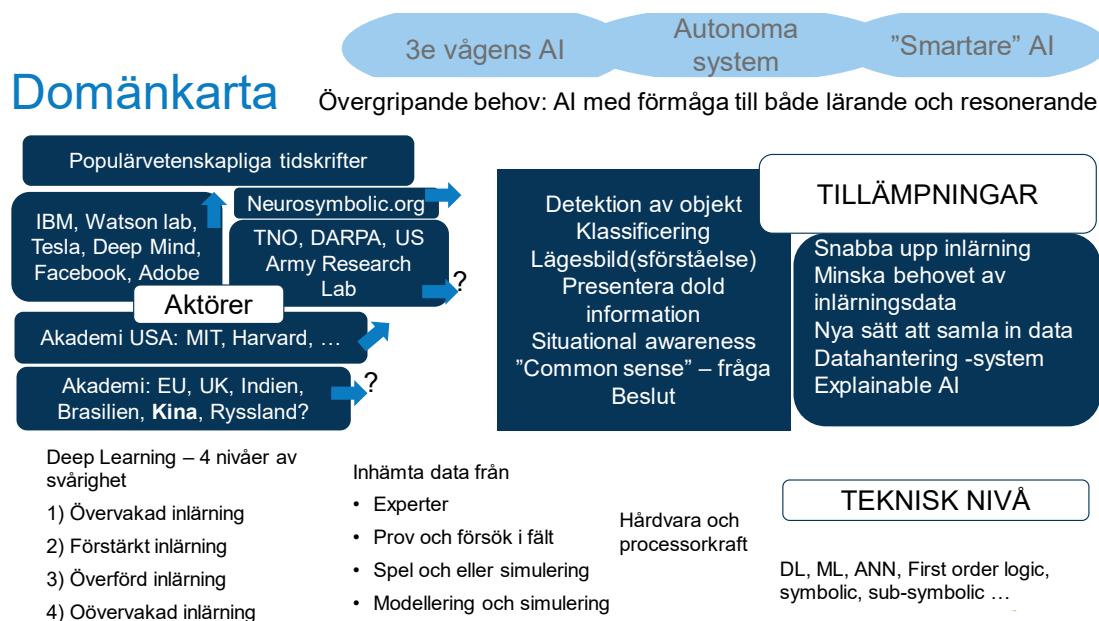
Redan tidigt i avskanningen (steg 1) utvecklade vi en domänkarta (se Figur 5) för att skapa översikt och få en bättre förståelse för området NSR. Domänkartan baserades inledningsvis på den information deltagarna i den större arbetsgruppen individuellt hittade under maximalt en timmes sökning. Denna tydliga avgränsning i söktid i detta inledande skede har vi funnit vara en bra balans mellan önskad effekt och använda resurser. Domänkartan kompletterades därefter med aktörer under steg 2 och vissa tillämpningar lades till.

Från arbetet med domänkartan drog vi slutsatsen att en övergripande drivkraft för forskning inom NSR är en önskan från de som tillämpar AI på praktiska problem om att stärka AI:s förmåga till att resonera, det vill säga att bli mer generell och i allmänt språk smartare. Ofta förekommande tillämpningsområden, och därmed drivkrafter för utvecklingen, är autonoma system, såsom självkörande bilar, samt tekniskt efterfrågade funktionaliteter för generell AI, såsom förklarbar AI (*eng. explainable AI, XAI*).

Forskning om NSR har pågått under flera decennier, men på senare år har antalet och framförallt typen av aktörer förändrats. Utgående från domänkartan och inledande sökningar fann vi att NSR är ett teknikområde som antas kunna övervinna centrala utmaningar med djupinlärning (Deep Learning).

Baserat på resultat från den inledande sökningen och efter den breda skanningens analys uppdaterade vi domänkartan. Pilarna i Figur 5 visar hur vi bedömer utvecklingen för respektive aktör baserat på framförallt den breda skanningen. Industri/kommersiella aktörer och samarbete mellan industri och akademi har ökat. Noterbart är kopplingen till militärt knuten finansiering och samarbeten mellan stora IT-företag och internationella toppuniversitet.

Det finns många tillämpningar för NSR-tekniker. Till tillämpningarna hör detektion av objekt, klassificering, skapande av lägesbilder, lägesbildsförståelse och beslutsstöd. NSR-tekniker kan användas för att förbättra prestanda jämfört med dagens AI, såsom snabbare inlärning, bättre nyttjande av inlärningsdata eller minskat behov av träningsdata, ökad förklarbarhet med mera (se rutan Tillämpningar i Figur 5).



Figur 5 Domänkarta över området neuro-symboliskt resonerande. Den är inte alls fullständig men används för att ge en överblick av en del av innehållet i området. DL – Deep Learning, ML – Machine Learning, ANN – Artificial Neural Network.

Tekniskt finns en stor spännvidd av information och några få exempel på områden syns längst ner i bilden under rubriken "Teknisk nivå".

Ett av dessa områden är "djupinlärning" (eng. Deep Learning, ofta förkortat DL) där NSR kan bidra till vissa lösningar. Vi har resonerat om fyra nivåer av komplexitet:

1. Övervakad inlärning (eng. supervised learning) – inlärning utifrån stora mängder av uppmärkt data (exempel självkörande bilar, spelet go),
2. Förstärkt inlärning (eng. reinforcement learning) – som ovan men inkluderar dessutom återkoppling (feedback),
3. Överförd inlärning (eng. transfer learning) – inlärning sker genom att använda data genererade av en simulator eller genom att använda data från närliggande problem (används vid avsaknad av stora datamängder),
4. Oövervakad inlärning (eng. unsupervised learning) – inget behov av feedback eller taggade (symboliskt märkta) data vid inlärning.

5.2 Resultat från den breda skanningen

Från de breda sökningarna i steg 2 framkom att NSR är ganska litet som forskningsområde betraktat.

Själva begreppet NSR verkar ha använts som term sedan 1990-talet, och är ett relativt litet område publiceringsmässigt. Det finns närliggande forskningstermer som inte klassas som "ren" NSR, men som vi valde att inkludera i de första breda sökningarna. Dessa närliggande termer är *sub-symbolic reasoning*, *cognitive computational reasoning*, *statistical-relational-learning* samt *hybrid-system*. Till den fortsatta breda och djupa skanningen togs begreppen neuro/neural symbolic och sub-symbolic reasoning vidare.

Utifrån resultaten kunde vi dra slutsatsen att termen sub-symbolic ligger nära begreppet neuro/neural. Begreppet sub-symbolic verkar användas mer inom sensor- och robotik-områdena.

Hybrid-system är en term som vi finner utmanande. Den används både väldigt brett, och som delmängd till neuro/neural-symbolic.

Sökning i Web of Science

Den breda automatiserade skanningen genomfördes som sökningar i Web of Science (WoS). Härvid observerades att det är ett fåtal (10-tal) grupper som publicerar peer-review-artiklar om NSR.

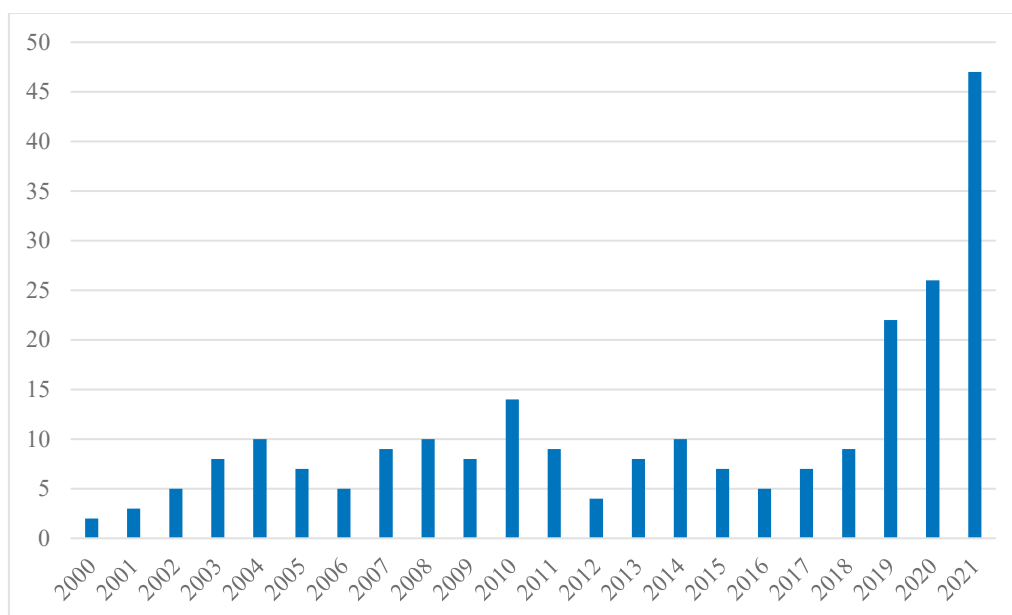
Efter en bred inledande skanning och läsning av resultat från en Nato-studie¹¹ genomfördes en sökning med termerna (neural OR neuro, alla år). Därefter genomfördes en ytterligare bred automatiserad sökning.¹²

I den breda skanningen identifierade vi aktörer som publicerar och forskar inom NSR, men som inte alltid publicerar i artikelform. Vi valde att inkludera dessa i den manuella expert-baserade sökningen.

Analys av neuro-symbolic reasoning i Web of Science

En analys av underlagen från två av WoS sökningens söksträngar genomfördes av den mindre arbetsgruppen. Analysen avsåg utvecklingen avseende de av WoS fördefinierade parametrarna Forskningsområde, Länder, Organisationer och Finansiärer för de båda tidsperioderna. Analysen presenteras separat för söksträngarna neuro/neural symbolic respektive sub-symbolic. I detta stycke presenteras resultat för neuro-symbolic reasoning.

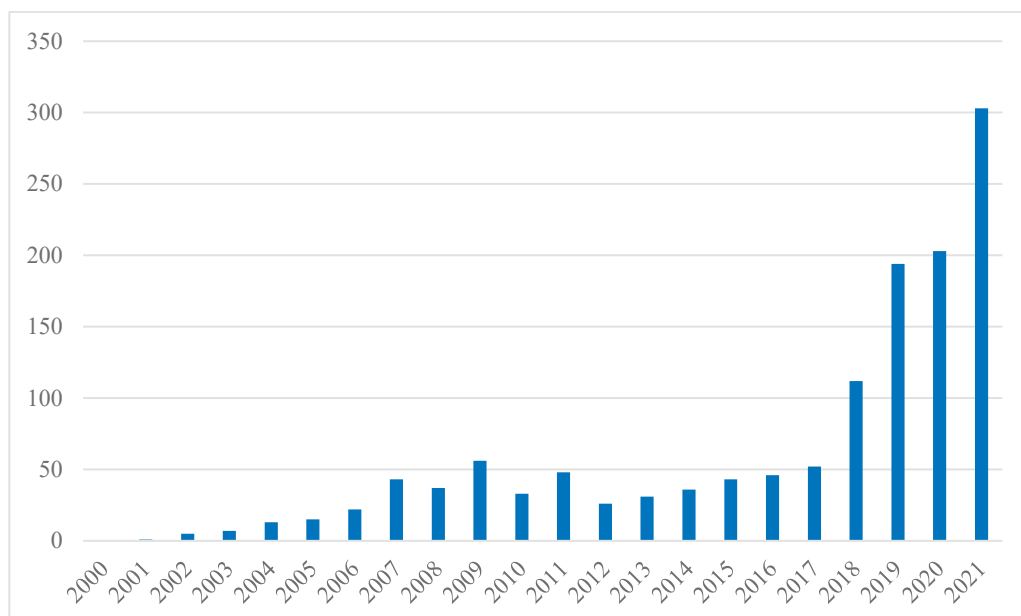
I den breda automatiserade sökningen ser vi att antalet publikationer (Figur 6) fluktuerar över tid men ökar på senare år. Antalet citeringar (Figur 7) ökar tydligt under senare år.



Figur 6 Antal publikationer per år för söksträngen (neural OR neuro) NEAR/0 symbolic.

¹¹ Nato (2021), Nato technical activity. IST-ET-118, Complementary use of symbolic and sub-symbolic AI methods (<https://www.sto.nato.int/Pages/activitieslisting.aspx?FilterField1=Panel&FilterValue1=Information%20Systems%20Technology>)

¹² Sökningen genomfördes med söksträngarna: (neural OR neuro) NEAR/0 symbolic och sub NEAR/0 symbolic. Sökresultaten analyserades för två tidsintervall: samtliga år samt avgränsat till åren 2018-2021. Att inkludera reasoning i söktermen visade sig vara onödigt.



Figur 7 Antal citeringar per år för söksträngen (neural OR neuro) NEAR/0 symbolic.

Forskningsområde

Datavetenskaplig forskning (eng. *computer science*) av olika slag dominerar över tid. Där- emot ändras de tio största områdena 2018-2021 jämfört med skanning över alla år i och med att *chemistry* och *biomedicine engineering* kommer in istället för *neurosciences*, *automation control systems* och *applied mathematics*. Sammantaget är det få publikationer så det är svårt att dra några säkra slutsatser, men vi ser att den datalogiska forskningen är kontinuerligt förekommande och att ett ytterligare antal andra forskningsområden har intresse av NSR.

Land

De länder som bedriver forskning har inte förändrats när man fokuserar på 2018-2021 jämfört med hela tidsperioden, med skillnaden att länder som USA och Kina på senare tid kommer ifatt Storbritannien som initialt dominerade området. Därutöver finns det forskning inom området i Tyskland, Brasilien och Italien. Det är dock ett så pass begränsat underlag att enskilda individer och forskargrupper får genomslag i statistiken.

Organisation

Beträffande aktiva organisationer har det skett en relativt stor förändring över tid. Flera företag, till exempel IBM, Google och Microsoft, finns med under perioden 2018-2021 som nya forskande organisationer och därtill förekommer nya toppuniversitet i USA. Aktörer från Storbritannien, till exempel brittiska universitet, minskar sin relativa betydelse gentemot aktörer från andra länder. Från Kina ser vi exempelvis att Shanghai Jiao Tong University som första kinesiska universitet har tagit sig in på topp tio-listan över forskningsorganisationer med flest publikationer i perioden 2018-2021.

Storbritannien som land tappar således mark i listan som rangordnar organisationer aktiva inom NSR-forskningen men behåller platsen på listan över antal publikationer. USA som land toppar organisationslistan och står även för en stor andel av alla publikationer. Bland organisationerna finns tre stora företag på topp tio-listan och därutöver MIT, Harvard och Kansas State University. På topp 20-listan förekommer även Facebook. Vår analys är att NSR antagligen är ett område där man hoppas på (snar) kommersialisering och tillämpning.

Finansiär

EU är en ledande finansiär, men ligger något efter USA, där försvaret är en viktig aktör. Cirka 9 % av NSR-forskningen är finansierad av USA:s försvarsdepartement (US Department of Defense, DoD) summerat över alla hela tidsperioden (de senaste 20 åren). Under de senaste fyra åren kompletterar de amerikanska företagen US DoD på topplistan, de finansierar nu på samma nivå som försvaret. Den amerikanska dominansen är särskilt tydlig de senaste fyra åren, där USA rycker ifrån kraftigt beroende på kombinationen av försvarsforskning och industriell forskning. Inom EU förekommer (sett över hela tidsperioden) både kommissionen och det tyska vetenskapsrådet som stora finansiärer. Vid en jämförelse mellan de senaste fyra åren och hela tidsperioden är EU-kommissionen kvar som finansiär, Kina behåller sin position samtidigt som länder som Malaysia och Indien kommer in som nya finansiärer.

Analys av sub-symbolic reasoning (SSR) i Web of Science

Forskningsområde

Här finner vi delvis samma forskningsområden som för NSR. Det finns dock en del skillnader. Vi hittar här forskningsområdena *robotics* och *automation control systems* i området *sub-symbolic reasoning* vilka inte återfinns inom NSR. Sökning med sub-symbolic finner även artiklar inom områdena *psychology* och *neuroscience*. De stora klossarna av olika typer av *computer science* är dock i stort desamma. Även publikationer inom *electrical/electronic engineering* återfinns i denna sökning.

Till följd av att termen sub-symbolic är vanligt förekommande inom autonomiområdet och robotik, vilka båda verkar vara drivkrafter till utvecklingen av NSR, valde vi att inkludera söktermen *sub NEAR/0 symbolic* i fortsatta sökningar.

Land

Här finns det skillnader jämfört med NSR. Europa som helhet är relativt starkare i SSR än NSR, och flera länder, såsom Tyskland, Italien, Nederländerna, Spanien, Belgien och Sverige är aktiva inom SSR. USA:s relativa betydelse minskar under den senare tidsperioden. Utöver europeiska länder och USA är Singapore en aktör. Storbritannien har markant mycket mindre betydelse inom SSR än NSR och Kina saknas helt.

Organisation

Forskningen inom området SSR drivs framförallt av universitet, men det finns även några kommersiella aktörer. Det kan noteras att Örebro Universitet finns med på topp tio-listan.

Finansiär

De tongivande finansiärerna har oftast sin hemvist inom EU och USA, vilket är stabilt över hela den undersökta tidsperioden.

Resultat från manuell systematisk bred sökning

Den manuella expertbaserade breda sökningen i steg 2 fokuserade framförallt på att söka efter tillämpningar och metoder i översiktsartiklar, samt på forskning och tillämpningar hos de företag som sedan tidigare identifierats. Även en bok från 2009¹³ användes för att skapa övergripande förståelse om området och dess bakgrund.

Resultatet från den manuella systematiska sökningen redovisas framförallt i den inledande beskrivande texten om NSR i kapitel 4. En stor del av de resultat som identifierades i den breda manuella sökningen överlappar med de resultat som redovisas i domänkartan, men ger något djupare beskrivningar. Vi har därför infogat delar av resultaten i redovisningen av domänkartan.

¹³ d'Avila Garcez, A. S., Lamb, L. C., & Gabbay, D. M. (2009). *Neural-symbolic cognitive reasoning*. Springer.

5.3 Resultat från den djupa skanningen

I steg 2, Skanning, har vi med automatiska sökningar (med HSTOOL) och manuella expertbaserade sökningar omhändertagit underlag om NSR som vi bedömt vara särskilt intressanta. Kriterier som användes för att värdera litteraturen var *nyhetsvärde, i framkant av FoU och av relevans för försvar och säkerhet*. De skanningsträffar som värderades högst sammanställdes i teknik- och tillämpningskort. Dessa presenteras i sin helhet i Bilaga 2.

Korten klustrades vid två tillfällen, se avsnitt 5.4. Vid det första tillfället skapades de områden som redovisas i Bilaga 2. Detta underlag innehåller samtliga kort. De flesta kort bedömdes som relevanta för värdering och då den första klustringen enbart utgick från rubriker klustrades korten på nytt med avseende på tillämpning. Dessa områden och inplaceringen av korten redovisas i Bilaga 1.

5.4 Resultat från värderingen

Värderingen genomfördes vid en workshop som del av steg 3. Värderingen utgick från klustringen baserad på tillämpning enligt nedan:

- Klassificera objekt,
- Fråga AI:n om råd,
- Robotik/Autonoma system,
- Lära av expert/simulering,
- Textanalys.

Dessa tillämpningsområden (kluster) för värdering redovisas i Bilaga 1 och samtliga kort redovisas i Bilaga 2. Vissa kort bedömdes vara mindre prioriterade och omhändertogs inte vid workshopen.

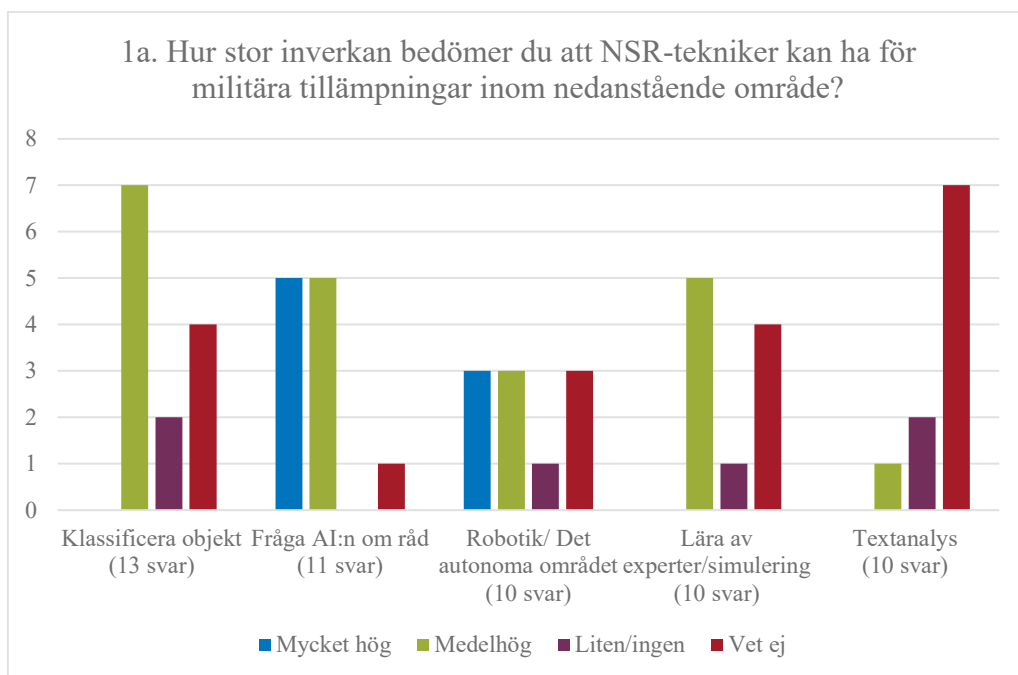
Samtliga kluster värderades, men *Klassificera objekt* och *Fråga AI:n om råd* gavs mest tid. Dessa två områden var inledningsvis ett enda stort område vid klustringen. Vi delade detta i två för att få en hanterbar mängd underlag vid värderingen, och redovisas delvis separerat nedan, förutom efteranalysen som gjordes samlat. Övriga områden värderades mer översiktligt. Därtill ställdes övergripande frågor till deltagarna i workshopen för att möjliggöra en sammanfattande värdering av NSR-området i stort.

Vi redovisar i detta avsnitt resultat med följande struktur:

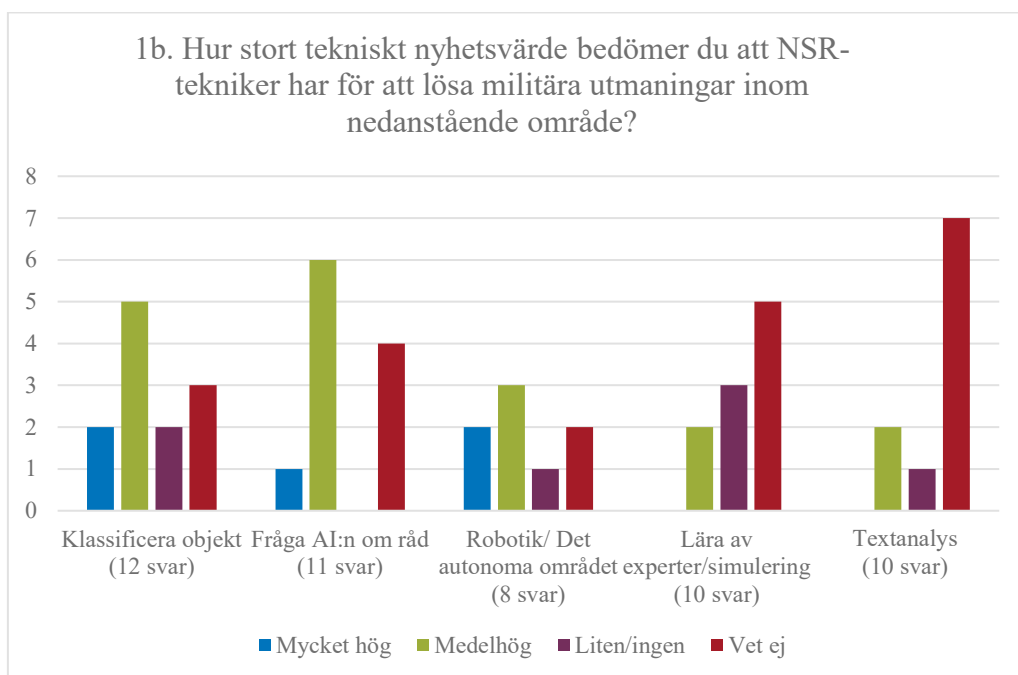
- Värdering av individuella områden (kluster):
 - Sammanställd värdering av respektive område (kluster), dvs. en sammanställd bild som redovisar resultat från workshopen.
 - Redogörelse för de inledande diskussioner som genomfördes av deltagarna vid workshopen efter presentation av respektive område (kluster). Redogörelse för den större avskanningsgruppens analys av underlagen från värderings-workshopen, genomförd efter workshopen.
- En sammanfattande värdering av NSR gjord vid WS.

Värdering av individuella kluster

Varje kluster värderades individuellt. Värderingen inleddes med en kort redovisning av innehållet i klustret vilken därefter följdes av en gemensam diskussion. Därefter ombads deltagarna att bedöma hur stor inverkan NSR-tekniker kan ha för militära tillämpningar samt vilket tekniskt nyhetsvärde NSR har för dessa. Svaren redovisas i Figur 8 och 9 nedan. Därefter återges diskussionerna vid workshopen och vår efterföljande analys av respektive kluster.



Figur 8 Svar på frågan om NSR:s inverkan för militära tillämpningar. Deltagarnas svar redovisas för samtliga värderade kluster.



Figur 9 Svar på frågan om NSR:s tekniska nyhetsvärde. Deltagarnas svar redovisas för samtliga värderade kluster.

Klassificera objekt

Diskussionen inleddes med en fråga angående varför NSR behövs inom klassificering, där djupinlärning sedan länge är ett tillämpat AI-område. Det framkom från diskussionen att dagens system för klassificering inte är fullt tillräckliga för dagens tillämpningar inom försvars- och säkerhetsområdet. Förbättringspotential sågs för:

- Klassificering av objekt i komplexa miljöer där en kontext måste vägas in i scenen. För sådana fall är dagens AI-lösningar långt ifrån tillräckliga.
- Situationer med begränsad tillgång till inlärningsdata, till exempel vid övning inför ett framtida krig.
- Upptäckt av svaga signaler i en kontext, exempelvis händelser i fysisk miljö, som människor lätt kan missa.

Vissa typer av klassificering kräver absolut korrekthet och en fråga ställdes om NSR skulle kunna bidra till denna tekniska utmaning. Detta kunde vi inte svara på.

Svaren på värderingsfrågan (se Figur 8) visar att NSR bedöms tillämpligt för att klassificera objekt, men ingen deltagare har röstat på mycket hög inverkan. Vid röstning om tekniskt nyhetsvärde är svaren nästintill desamma.

Fråga AI:n om råd

Det finns en stor efterfrågan på de tillämpningar som klustret redovisar. Utmaningarna är potentiellt stora. Förklarbar AI (XAI) och validerbar AI diskuterades, liksom algoritmers konfidensgrad, produktions- och systemkrav och tillhörande rättsliga krav (legala ramverk för produkter). NSR:s möjlighet att bidra till att lösa vissa av de tekniska utmaningar som en mer generell AI innebär gjorde att klustret värderades högt vad gäller frågan om tillämpning (Figur 8). Behovet av symboliskt resonerande är mer uppenbart för detta kluster än för att klassificera objekt.

Tillämpningar som diskuterades var

- Möjligheten att upptäcka anomalier och urskilja/särskilja objekt i komplexa miljöer eller i komplexa datamängder.
- Potential för bättre precision och mer tillförlitlig information (beslutsunderlag) i högnivålägesbilder.

Analys av Klassificera objekt och Fråga AI:n om råd

Klustren *Klassificera objekt* och *Fråga AI:n om råd* går in i varandra. Vid den innehållsmässiga klustringen bildade dessa kluster inledningsvis ett gemensamt, mycket stort kluster. För att minska mängden information delade vi klustret till de två som värderats.

Flera behov och tillämpningar som idag inte tillgodoses lyftes under endera eller båda klustrens diskussioner.

Diskussionen om klassificering berörde framförallt tillämpningar där dagens tekniker är otillräckliga. Däremot kom diskussionen om *Fråga AI:n om råd* att fokusera på centrala utmaningar för en mer generell AI, tekniska lösningar och om NSR kan bidra till att lösa dessa. Exempel är förklarbarhet, transparens, validering och verifiering.

En reflektion är att korten för klustren beskrev två olika sätt att hantera och använda träningsdata. Klustret *Klassificera objekt* redovisade möjligheter med kontinuerlig insamling och inlärning, medan klustret *Fråga AI:n om råd* diskuterade hur utmaningen med för lite träningsdata och XAI skulle överkommas. Under värderingen diskuterades den mycket viktiga frågeställningen tillgång till och nyttjande av träningsdata. Denna frågeställning vore intressant att fördjupa i en eventuell fortsatt analys. Frågan kopplar an till utmaningar med djupinlärning (se domänkartan), metoder för att samla in och nyttja data, möjligheter med simulering och träning samt övervakad inlärning.

Diskussionen gled under värderingen diffust mellan resonemang om autonom offensiv verkan, lägesbild och underlag för beslutsstöd. Vid eventuella fortsatta arbeten bör frågan

om tillförlitlighet och förutsättningar för användning förfinas och diskuteras separat för olika typer av tillämpningar och system.

Tillämpningsdiskussionen fördes framförallt utifrån ett perspektiv sett från blå sida. Vi upplevde en diskrepans av bedömningen av NSR:s relevans mellan forskare som skapar lösningar för en fysisk miljö och de som gör det för en helt digital miljö. Utmaningen med klassificering och nyttan med NSR verkar vara större för komplexa fysiska miljöer. Inom området *Fråga AI:n om råd* finns många stora utmaningar vilka rent tekniskt inte kan tillgodoses med befintliga AI-metoder. Detta öppnar för tillämpningar av NSR.

Robotik/Autonoma system

Diskussionen baserades på endast en artikel. Den presenterade NSR-tekniken ansågs relevant för alla tillämpningar som inbegriper sensorer på farkoster eller andra rörliga objekt och farkoster. Tillämpningar som diskuterades var:

- Kontinuerligt uppdaterad och delad lägesbild för sensorer/autonoma system.
- Förutsättningar för robotars/autonoma systems egen planering.
- Internet of Things (IoT) i allmänhet.

Värderingen visar att fler deltagare bedömer att NSR kan ha hög snarare än låg inverkan på militära tillämpningar (se Figur 8). Det tekniska nyhetsvärdet bedöms på liknande sätt.

Analys av Robotik/Autonoma system

Robotik och autonoma system är en av de drivkrafter som vi funnit för användning av NSR. Annorlunda typer av begrepp för AI-tekniker verkar användas inom detta område än inom andra, varför det är möjligt att vi missat underlag som ändå skulle kunna klassas som NSR. Försvarsmakten bedriver utvecklingsarbete och beställer forskning inom obemannade och autonoma system och tillämpningen av NSR bör därför vara relevant att undersöka vidare.

Lära av expert/Simulering

Diskussionen om klustret lära av expert/simulering var kort, men en värdering genomfördes som för övriga kluster.

Det finns ett tekniskt behov av att skapa regler baserat på observationer av människors beteenden i komplexa situationer. Det kan vara svårt att nå/skriva fungerande regler för komplexa situationer. Tillämpningar finns exempelvis för både självkörande fordon och simulering.

Analys av Lära av expert/Simulering

Under workshopen diskuterades omvänd förstärkt inlärning medan den funktionalitet som presenterats med bas i teknikkort inte lika tydligt diskuterades. Vi kan därför inte lita på värderingsresultatet då röstningar kan avse en annan sakfråga än avsett. Vi tror dock fortsatt att teknikerna/tillämpningarna kan vara militärt relevanta, exempelvis för stridsflygssimulering.

Textanalys

FOI bedriver tillämpad forskning inom området textanalys. Mycket kort diskuterades huruvida de NSR-tekniker som presenterades överhuvudtaget skulle kunna tillämpas och, om så blir fallet, när detta kan komma att ske.

Analys av Textanalys

Textanalys är ett mycket nischat område såväl tekniskt som tillämpningsmässigt. NSR kan vara relevant för textanalys och de uppdrag som FOI har. Vi uppfattar att FOI-experten kanske kommer att undersöka området vidare. Värderingen innehåller en särskilt stor andel *vet ej*, vilket gör det svårt att dra slutsatser.

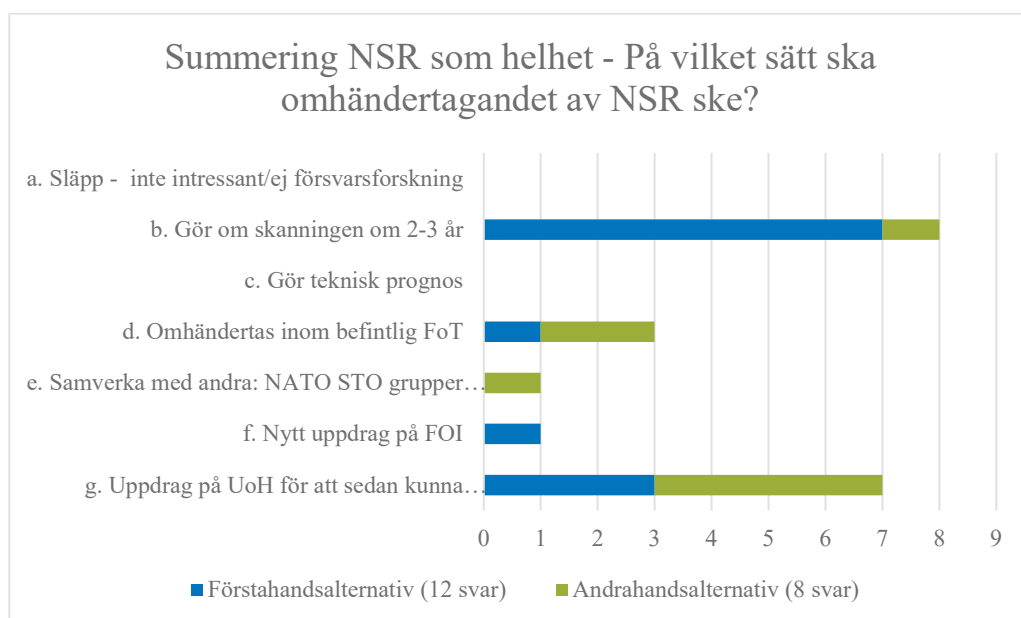
Sammanfattande värdering av NSR

Den sammanfattande värderingen gjordes tillsammans med deltagarna som avslutning av workshopen. Syftet var att omhänderta deltagarnas åsikter beträffande området NSR, för att kunna dra slutsatser om det är försvarsintressant och borde undersökas vidare eller inte. Värderingen visar att NSR-tekniker bedöms vara relevanta för militära tillämpningar och intressanta för forskning inom försvarsområdet. Specifikt frågade vi deltagarna om de ansåg att området bör finansieras med medel från Försvarsmakten. Vi frågade även hur de ansåg att NSR-området bör tas vidare.

Det kan först noteras att några deltagare valde att inte rösta alls, medan andra valde att rösta vet ej.

På frågan om man ansåg att Försvarsmakten bör finansiera forskning inom NSR svarade sju deltagare ja. Två deltagare tyckte inte att Försvarsmakten bör finansiera och fyra deltagare svarade vet ej.

Vi frågade även hur de ansåg att NSR-området bör tas vidare. Deltagarna kunde ge såväl ett förstahandsval som ett andrahandsval. Resultatet visas i Figur 10.



Figur 10 Summerande fråga om hur NSR-området kan tas vidare. Utöver de givna svarsalternativen i bilden fanns möjlighet att ge ett fritt svar. Ett sådant "övrigt" förslag gavs, samma vid båda omröstningarna, i form av möjligheten att göra en studie inom FOI:s anslagspost för egeninitierad forskning.

Vid den första omröstningen (deltagarna angav sin prioritet 1) ansåg en majoritet av deltagarna att en ny avskanning om 2-3 år var bästa alternativet (7 st röster), medan andra deltagare föredrog alternativet att lägga uppdrag på universitet och högskolor som sedan skulle omhändertas (3 st). Vid den andra omröstningen på samma frågor (prioritet 2) hamnade istället uppdrag på universitet och högskolor i topp (4 st röster), med alternativet omhändertas inom befintlig FoT på andra plats (2 st). Det kan noteras att alternativet att helt släppa området inte fick några röster alls.

Analys av Sammanfattande värdering av NSR

Svaren på de summerande frågorna visar att området vid workshopen bedömdes vara relevant att undersöka vidare. Vi anser att experterna vid värderingen pekat på flera riktningar där NSR kan vara viktigt för FOI att fortsatt bevaka.

Utfallet synliggör samtidigt att många av deltagarna kände osäkerhet, vilket också framkom under diskussionerna. Vi uppfattar att denna osäkerhet rör såväl förståelse av NSR som teknik- och forskningsområde som innebörden av tillämpningsområdena (klustren).

Slutligen noterar vi att resultaten innehåller bias då värderingen tar avstamp i våra tekniska redovisningar. Urvalet av medverkande experter innebär också en bias, vilken alltid kommer att förekomma i någon grad vid denna typ av värdering. Samtidigt konstaterar vi att värderingsworkshopen bidrog till kunskapsspridning mellan deltagarna.

5.5 Analys och grunder för rekommendationer

En utmaning med avskanning är att begränsa sig till att skrapa på ytan, det vill säga att bara skrapa tillräckligt djupt för att kunna ge rekommendationer om det avskannande ämnets fortsatta omhändertagande utan att fastna i enskilda detaljer. Avskanningen som beskrivs i denna rapport syftade till att värdera området NSR som helhet. En allmän slutsats är att vi, utifrån resultatet från den genomförda avskanningen, inte kan avfärda NSR som ett intressant forsknings- eller tillämpningsområde.

Samtidigt har vi inte tillräckligt djupa underlag för att koppla specifika NSR-tekniker till särskilda tillämpningsområden, eller bestämda tidshorisonter för utveckling. Vi kan inte kartlägga hur olika NSR-tekniker kopplas till de alternativ för omhändertagande som deltagarna värderade. Därtill kan vi konstatera att den genomförda avskanningen har luckor. I domänkartan identifierade vi exempelvis namngivna militära aktörer som utvecklar tillämpningar inom området, men som vi inte har omhändertagit i själva skanningen (steg 2). Vi har inkluderat, men inte grävt djupare i, den utveckling som drivs av civila aktörer, bland annat i form av samarbeten mellan industri och universitet. Vår bedömning är att just industri-akademisamarbetena i framförallt USA gör NSR till ett relevant område för FOI att fortsatt omhänderta. Till detta kommer att den amerikanska forskningen till stor del och under längre tid finansierats av amerikanska försvarsdepartementet. Det verkar som att det är först under senare år som NSR har fångats upp av civila aktörer. Drivkrafterna verkar dock vara gemensamma i form av behov av ”smartare” AI, XAI och effektivare nyttjande av data och träningsdata för att nämna de mest uppenbara. NSR är en del i utvecklingen mot 3:e vågens AI. NSR kombinerar symbolisk logik, det vill säga 1:a vågens AI, med 2:a vågens AI i form av exempelvis djupinlärning (Deep Learning, DL) och maskininlärning (Machine Learning, ML). Kombinationen används för att dra nytta av respektive vågs styrkor och för att överkomma de stora utmaningar som finns när exempelvis DL/ML fallerar. Vi skriver ”smartare”, då NSR kan vara en teknik på vägen mot det som kallas Artificiell Generell Intelligens, något förenklat liknat vid en helt självtänkande AI.

Kombinationen av behoven av en teknisk lösning för utmaningarna inom DL/ML, mängden finansiering och intresset hos kompetenta aktörer kan påverka utvecklingshastigheten på oförutsedda sätt. Utgående från våra underlag kan vi inte avgöra hur snabbt utvecklingen går eller hur hög mognaden egentligen är. Våra bedömningar avseende TRL för teknik- och tillämpningskort varierar. De flesta teknikerna bedöms befinna sig på låga TRL-nivåer (TRL 2), men det finns ett fåtal som redan bedöms befinna sig på högre TRL-nivåer (upp till TRL 7). Bland de tekniker som bedöms befinna sig på låg TRL-nivå finns en stor spännvidd i våra bedömningar avseende hur snabb den framtida utvecklingen kan komma att vara för de olika teknikerna. Det finns tekniker som vi bedömer har mycket stora osäkerheter vad gäller möjligheterna att överhuvudtaget nå höga TRL-nivåer någon gång i framtiden och andra där vi bedömer en sådan utveckling som ganska säker och snabb.

Vi har inte gjort en fördjupad analys av resultaten med FOI:s forskningsområdesföreträdare då vi bedömde att vi, givet de resurser projektet kunde lägga på skanningen, fått tillräckliga underlag vid den workshop som genomfördes.¹⁴

¹⁴ Vi hade planerat att göra en uppföljande värdering av resultaten från workshopen med forskningsområdesföreträdare med inriktning mot AI. Frågor vi skulle ställt om så kunnat ske är i) om NSR-teknikerna/kluster bedömdes ligga

Diskussion

Givet underlag från den genomförda skanningen drar vi slutsatserna att:

- Vissa av NSR-teknikerna ligger i anslutning till FOI:s befintliga forskning. De omhändertas däremot inte inom befintliga projekt.
- Vissa av NSR-teknikerna ligger längre ifrån FOI:s befintliga forskning, men kan potentiellt vara militärt relevanta.
- Flera av NSR-teknikerna är potentiellt relevanta för militära tillämpningar där det idag saknas eller finns otillräckliga tekniska lösningar.

Just de militära tillämpningarna och den militära finansieringen av forskning internationellt finner vi särskilt intressant.

Vid värderingen identifierade vi tillämpningar i relation till de av oss identifierade klustren, men frågor kom också från deltagare om ytterligare tillämpningar. Frågorna visar att det kan finnas behov som idag inte omhändertas av befintlig AI-forskning och -teknik och att NSR-tekniker här skulle kunna innebära nya möjligheter. Detta gäller t.ex. NSR-tekniker till stöd för klassificering och en ”smartare” AI (beslutsstöd och autonomi).

Skanningen visar att NSR till del är sprunget ur de behov och utmaningar som finns inom AI för att nå mer effektiva/användbara och smarta AI-system. Tekniskt finns utmaningar som berör video- och textanalys, tillgång till och effektivt nyttjande av data för träning, samt minskat beroende av människan för inläring och formulering av regler.

Alternativ för eventuellt framtida arbete

Vid värderingsworkshopen röstades *avvakta och gör om avskanning om 2-3 år* och *lägga uppdrag på universitet och högskola*, fram som prioriterade alternativ för fortsatt arbete. Den breda spännvidden på TRL, försvarsrelevanta tillämpningar och nya aktörers verksamhet inom området gör oss osäkra på om alternativet avvakta är rätt väg att gå. Vi befarar att lägga området på hyllan och göra om avskanningen om 2-3 år kan vara ett onödigt stort risktagande. Tre års forskning, om det visar sig att NSR är viktigt för militära tillämpningar och utvecklingen tar ytterligare fart, kan inom det snabbväxande AI-området innebära att steg hinner tas mot höga TRL-nivåer. En grund för den farhågan är den starka tillväxt vi noterat i samverkan mellan akademi och företag. Takten i AI-teknikens utveckling behöver inte nödvändigtvis förhålla sig till linjära TRL-skalar vare sig gällande mognad eller utvecklingstakt (tid).

Inledningsvis i arbetet diskuterade vi att kontakta universitet för informationsinhämtning. Vi har i arbetet träffat på litteratur med koppling till vissa lärosäten i Sverige. Vi ser att sådana kontakter fortsatt är bra men att det primära intresset vid fortsatt arbete är militära tekniska behov och tillämpningar. Vi bedömer därför att det i första hand är andra aktörer än svenska universitet och högskolor som blir intressanta att kontakta. Kontakter med utvalda samarbetspartners är således en möjlig väg att gå, liksom små fortsatta arbeten inom vissa av FOI:s egna forskningsområden.

Ett sätt att gå vidare är att utreda NSR inom de forskningsområden som nyttjar AI-tekniker och har behov av de tillämpningar NSR kan bidra med. Ett annat sätt är att närmare utreda andra möjliga militära tillämpningar och vilka typer av NSR-tekniker som specifikt adresserar dessa behov. En låg, men rimlig, ambition är att följa utvecklingen av NSR-området.

Vi har inte analyserat var inom FOI som NSR kan vara av störst intresse, eller möjliga samarbeten med andra forskningsutövare. Detta var dock inte heller syftet med arbetet. För att underlätta omhändertagande av resultaten av FOI:s tekniska experter i berörda forskningsprojekt föreslås muntlig presentation av resultat.

nära eller långt ifrån befintlig forskning vid FOI? ii) vara en möjligt kritisk teknologi eller tillämpning mot framtiden? iii) prioriterade att gå vidare med?

6 Erfarenheter av processen

Under arbetet har vi lärt oss mycket angående det praktiska utövandet av själva avskanningsprocessen.

Något som påverkat genomförandet av värderingen är bredden av tekniker och tillämpningar hos området NSR. Vi gick in i arbetet med ett antagande om att NSR är ett smalt, nischat område. Antagandet kan sägas stämma sett till antal publikationer. Antagandet stämmer inte alls för teknikområdet i sig eller möjligheterna till tillämpning. Teknikområdet är brett; en kombination av neurala nätverk med symbolisk logik på många olika sätt och olika nivåer, och det finns många potentiella tillämpningar.

Avskanningen hade två syften: skannande och kunskapsbyggande. Avskanning innebär att skrapa på ytan för att ge underlag för beslut om fortsatt omhändertagande. Kunskapsbyggande innebär fördjupning för att förstå bättre. Förslagsvis blandas inte en bred avskanning med ett kunskapsbyggande syfte. Det har varit en utmaning att kombinera dessa. Kunskapsbyggande är förmodligen enklare att förena med en djupare avskanning.

Erfarenheter med fokus på genomförande och metod dokumenteras vid sidan av denna rapport.

7 Referenser

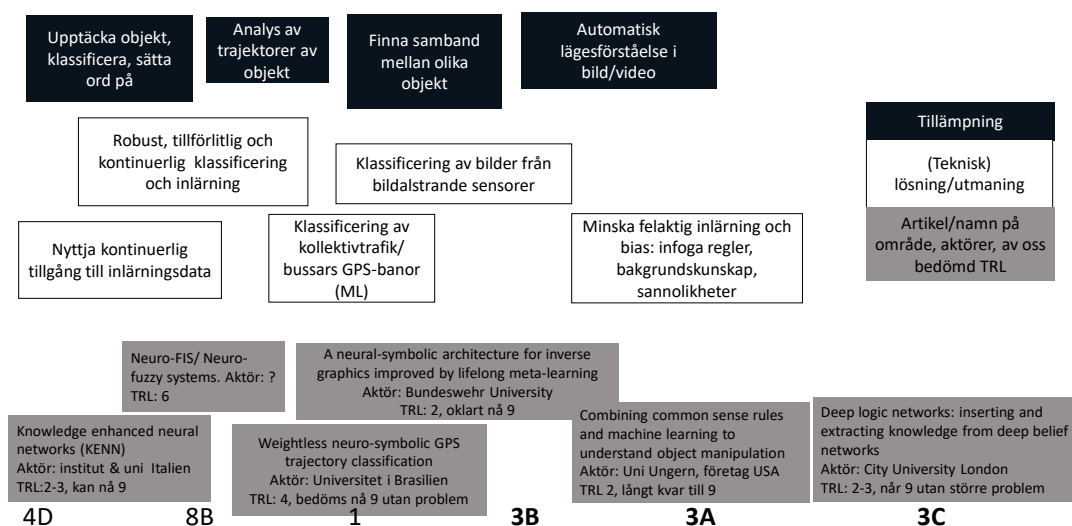
- Badreddine, S., d'Avila Garcez, A., Serafini, L., & Spranger, M. (2022). *Logic Tensor Networks. Artificial Intelligence* 303:103649. doi:10.1016/j.artint.2021.103649
- Bengio, Y. (2016). The rise of AI. *Scientific American* (June), 45-51.
- d'Avila Garcez, A. S. (2005). Fewer epistemological challenges for connectionism. *First Conference on Computability in Europe, CiE 2005, 3526*, ss. 289-325. Amsterdam. doi:10.1007/11494645_18
- d'Avila Garcez, A. S., & Lamb, L. C. (2020). *Neurosymbolic AI: The 3rd Wave. arXiv preprint arXiv:2012.05876v2 [cs.AI]*.
- d'Avila Garcez, A. S., Besold, T. R., de Raedt, L., Földiák, P., Hitzler, P., Icard, T., Silver, D. L. (2015). Neural-Symbolic Learning and Reasoning: Contributions and Challenges. *2015 AAAI Spring Symposium*, (ss. 18-21).
- d'Avila Garcez, A. S., Lamb, L. C., & Gabbay, D. M. (2009). *Neural-symbolic cognitive reasoning*. Springer.
- Dalberg, E., During, C., Kindvall, G., Lindberg, A., Schubert, J., Wickenberg Bolin, U. (2021). *Avskanning av forskningsfronten 2021*. FOI-Memo 7702.
- de Raedt, L., Dumancic, S., Manhaeve, R., & Marra, G. (2020). From Statistical Relational to Neural-Symbolic Artificial Intelligence. *Proceedings of the Twenty-Ninth International Joint Conference on Artificial Intelligence (IJCAI-20)*, (ss. 4943-4950).
- Dickson, B. (2021). *The future of deep learning, according to its pioneers* 1 juli 2021. Hämtat från TechTalks: <https://bdtechtalks.com/2021/07/01/deep-learning-future-bengio-hinton-lecun/> den 11 augusti 2021
- Karasalo, M., Schubert, J. (2019). HSTOOL for horizon scanning of scientific literature, FOI-R--4760--SE.
- Kelley, T. D., Avery, E., Long, L. N., & Dimperio, E. (2009). A Hybrid Symbolic and Sub-Symbolic Intelligent System. *AAAI InfoTech@Aerospace Conference*. Seattle.
- Lindberg, A. Grönwall, C., Lång, E., Dalberg, E. (2021). *A utility focused process for enhancing the use of horizon scanning in defence research*. FOI-R--5149--SE.
- McCarthy, J. (1988). Epistemological challenges for connectionism. *Behavioral and Brain Sciences*, 11(1), 44-44.
- Nato (2021), Nato technical activity, IST-ET-118, *Complementary use of symbolic and sub-symbolic AI methods*. Återfinns på <https://www.sto.nato.int/Pages/activitieslisting.aspx?FilterField1=Panel&FilterValue1=Information%20Systems%20Technology>.
- Pearl, J. (2000). *Causality: Models, Reasoning and Inference*. Cambridge University Press, New York, NY

- Schubert, J., & Johansson, R. (2020). *Learning causal structures from data*. FOI-R--4882-SE, Totalförsvarets forskningsinstitut (FOI).
- Svatoš, M., Šourek, G., & Železný, F. (2019). Revisiting Neural-Symbolic Learning Cycle. *Proceedings of the 2019 International Workshop on Neural-Symbolic Learning and Reasoning (NeSy'19)*.
- Wang, B. (2017). *Three waves of AI*. 22 februari 2017. Hämtat från <https://www.nextbigfuture.com/2017/02/darpa-sees-future-third-wave-of.html> den 11 augusti 2021
- Vincent, J. (2019). *Former Go champion beaten by DeepMind retires after declaring AI invincible*. 27 november 2019. Hämtat från The Verge: <https://www.theverge.com/2019/11/27/20985260/ai-go-alphago-lee-se-dol-retired-deepmind-defeat> den 11 augusti 2021

Bilaga 1 - Översikt av innehåll i kluster

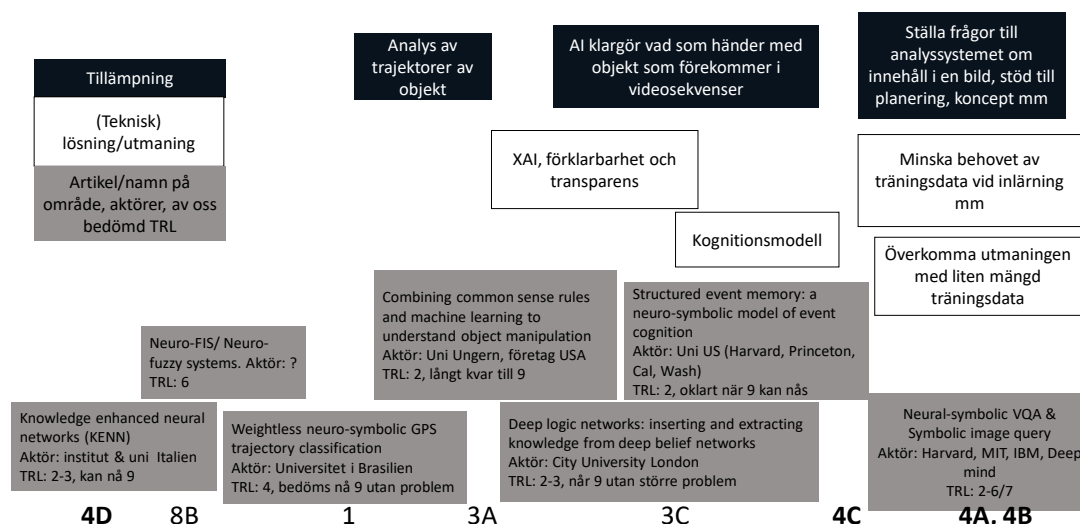
Bilderna nedan redovisar de kluster, dvs. områden för tillämpning, som användes som inspel för diskussionen vid värderingen.¹⁵ Det var inte individuella kort utan klustren i bilderna som diskuterades och värderades. Bilderna visar idéer om tillämpningar, tekniska lösningar/utmaningar, hänvisningar till teknikkort, källor och bedömd TRL.

Klassificera objekt



Figur 11 Figuren visar, med bas i teknikkort och egen sammanställning, olika tillämpningar inom området Klassificera objekt, vilka tekniska lösningar som diskuteras i litteraturen och vilka källor som använts vid klustring, aktör som bedriver forskning och teknikernas bedömda TRL nivå.

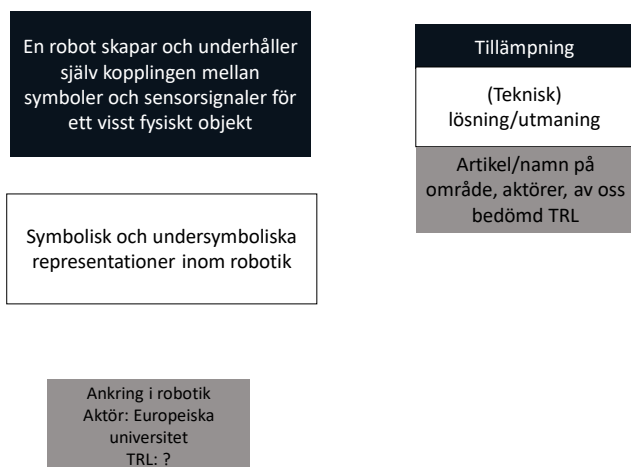
Fråga AI:n om råd



Figur 12 Figuren visar, med bas i teknikkort och egen sammanställning, olika tillämpningar inom området Fråga AI:n om råd, tekniska lösningar och vilka källor som använts vid klustring, aktör som bedriver forskning och teknikernas bedömda TRL nivå.

¹⁵ Vid värderings-workshopen användes ett bildspel som innehöll såväl information om workshopen, NSR och bilder för att stödja värderingen. Detta bildspel finns sparad och presenteras delvis nedan. Diskussionerna från brainstorming-momenten finns sparade i såväl anteckningar som i PowerPoint-filen som visades vid workshopen.

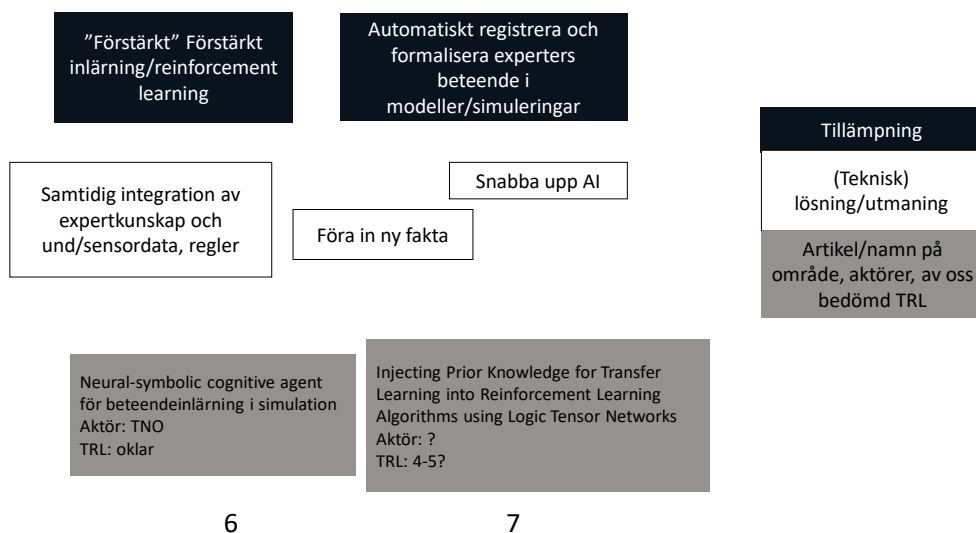
Robotik/autonoma system



8A

Figur 13 Figuren visar, med bas i teknikkort och egen sammanställning, olika tillämpningar inom området Robotik/autonoma system, tekniska lösningar och vilka källor som använts vid klustring, aktör som bedriver forskning och teknikernas bedömda TRL nivå.

Lära av experter/simulering

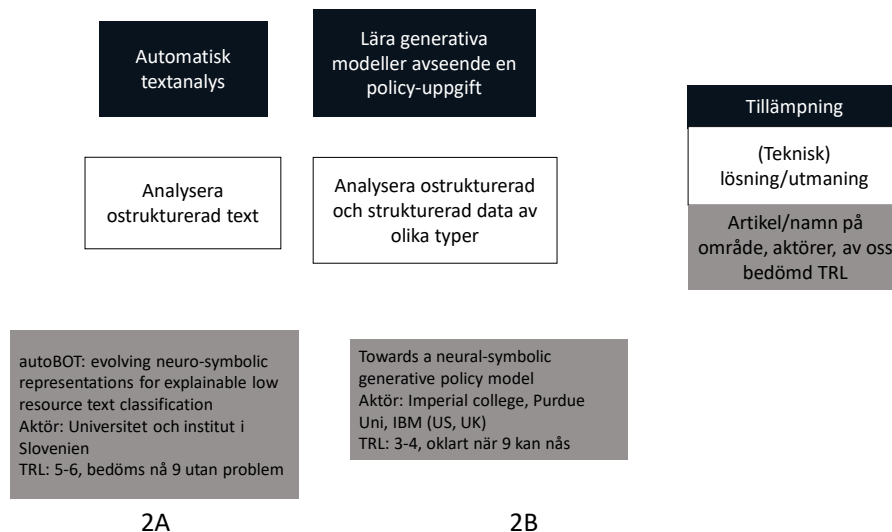


6

7

Figur 14 Figuren visar, med bas i teknikkort och egen sammanställning, olika tillämpningar inom området Lära av expert/Simulering, tekniska lösningar och vilka källor som använts vid klustring, aktör som bedriver forskning och teknikernas bedömda TRL nivå.

Textanalys



Figur 15 Figuren visar, med bas i teknikkort och egen sammanställning, olika tillämpningar inom området Textanalys, tekniska lösningar och vilka källor som använts vid klustring, aktör som bedriver forskning och teknikernas bedömda TRL nivå.

Bilaga 2 - Teknik- och tillämpningskort

Teknik- och tillämpningskort sammanställdes av de två experterna och skannarna i den mindre avskanningsgruppen. I stort har samma mallar och rubriker använts, men kriterier för bedömning av relevans skiljer något. Korten från den automatiserade skanningen redovisar relevans avseende nyvärde, för FOI och för FM i en skala 1-3, där 1 är lägst och 3 är högst. Värderingen användes för att välja ut urval från den automatiserade sökningens långa listor med artiklar, och diskuterades vid klustring. Den manuella sökningen redovisar kortets/områdets bedömda relevans i en skala 1-5 efter läsning av litteratur.

Alla kort är inte fullständiga. Vi har fyllt i den information som vi funnit vid läsning, men inte letat information när sådan saknas i syfte att få fullständiga kort. Teknikkorten är tekniska till sin natur och innehåller därmed en del facktermer.

Kluster	Teknik- och tillämpningskort samt nummer från första klustring
1. Tracking	1. Weightless neuro-symbolic GPS trajectory classification
2. Textanalys	2a. autoBOT: evolving neuro-symbolic representations for explainable low resource text classification
	2b. Towards a neural-symbolic generative policy model
3. Lågnivå (läges)bild	3a. Combining common sense rules and machine learning to understand object manipulation
	3b. A neural-symbolic architecture for inverse graphics improved by lifelong meta-learning / neuro-symbolisk kapsel
	3c. Deep logic networks: inserting and extracting knowledge from deep belief networks
4. Högnivå (läges)bild	4a. Neural-symbolic VQA: disentangling reasoning from vision and language understanding
	4b. Symbolic image querying
	4c. Structured event memory: a neuro-symbolic model of event cognition
	4d. Knowledge enhanced neural networks (KENN)
5. Beslutsstöd och delfunktioner	5a. Neuro-Symbolic integration with uncertainty
	5b. Regelextraktion
	5c. Conceptual space representation
6. Lära av expert/simulering	6. Neural-symbolic cognitive agent för beteendeläring i simulation/Lära av instruktör NSCA i simulering
7. Stöd för förstärkt inläring	7. Logic tensor network för att förbättra reinforcement learning
8. Övrigt	8a. Ankring i robotik
	8b. Neuro-FIS/Neuro-fuzzy systems
	8c. Neural-symbolic learning cycle (NSLC)
	8d. Generative art
	8e. T-PRISM

Kluster 1 - Tracking

1. Weightless neuro-symbolic GPS trajectory classification

Kort beskrivning av tekniken – vad den är och hur den kategoriseras

Tekniken är en användning av maskininlärning för klassificering. Tekniken består av ett ramverk för att hantera problemet med klassificering av GPS-banor. Ett sådant ramverk kombinerar en mångsidig WiSARD¹⁶-klassificerare med en i förväg definierad uppsättning regler som tillsammans utgör ett neural-symboliskt inlärningssystem.

Tillämpning av tekniken – till vad används den

Fokus: Informationsfusion.

Tekniken används av Rio de Janeiros kollektivtrafiksystem för klassificering av bussars GPS-banor. Tekniken har testats mot ett stort antal bussar i Rio de Janeiro. Resultaten bekräftar den praktiska tillämpningen av dessa, eftersom det uppnådda resultatet var lika bra som andra moderna rivaliserande metoder i de flesta scenarier.

Bedömd mognad och tekniska beroenden, hinder eller möjligheter för att nå TRL 9

TRL 4. Bör kunna nå TRL 9 utan problem.

Aktörer

Centro Federal de Educação Tecnológica Celso Suckow da Fonseca, Brasilien.

Universidade Federal do Rio de Janeiro, Brasilien.

Datum för publicering (år)

2017

Bedömning av relevans (skala 1-3, med motivering)

Nyvärde	För FOI	För FM
3	3	3

Tekniken är relevant för både FOI:s forskning och FM:s utveckling av tekniker för situationsanalys i militära beslutsstödssystem där ett beslutsstödssystem försöker förstå olika objekts avsikter.

Referens

Barbosa, R., Cardoso, D. O., Carvalho, D., Franca, F. M. (2018). Weightless neuro-symbolic GPS trajectory classification. *Neurocomputing* 298:100–108.
doi:10.1016/j.neucom.2017.11.075

Egen analys/tankar

Tekniken bör kunna ingå som en av flera komponenter i ett beslutsstödsystem som försöker förstå motståndarens avsikter genom att klassificera motståndarens rörelser som tillhörande viss kategori av beteenden med alternativa avsikter. Militär tillämpning bedöms finnas främst i mark- och sjöstrid.

Intressant för områden som vill analysera trajektorer av objekt.

¹⁶ Wilkes, Stonham and Aleksander Recognition Device, 1984.

Kluster 2 - Textanalys

2a. autoBOT: evolving neuro-symbolic representations for explainable low resource text classification

Kort beskrivning av tekniken – vad den är och hur den kategoriseras

Språkmodeller har visat goda resultat för textklassificering, men de är omständliga att träna och kräver stora mängder data för att ställa in hundratals miljoner parametrar. Ett alternativ är automatiskt utvecklade textrepresentationer till modeller med automatiserad hyperparameterinställning. autoBOT (automatic Bags-Of-Tokens) är en teknik med tillämpning vid inlärningsscenarier där både hårdvara och mängden data som krävs är begränsad. autoBOT använder en evolutionär algoritm som optimerar så kallade glesa representationer. Metoden erbjuder konkurrenskraftig klassificeringsprestanda på fjorton verkliga klassificeringsuppgifter i en genomförd utvärdering jämfört med ett konkurrenskraftigt autoML-tillvägagångssätt som utvecklar ensemblemodeller, liksom gentemot moderna neurala språkmodeller som BERT och RoBERTa.

Tillämpning av tekniken – till vad används den

Fokus: Textanalys.

Automatisk textanalys vid Open Source Intelligence, t.ex. av sociala medier.

Bedömd mognad och tekniska beroenden, hinder eller möjligheter för att nå TRL 9

TRL 5-6. Troligen inga stora problem att nå TRL 9.

Aktörer

Jožef Stefan Institute, Slovenien.

University of Nova Gorica, Slovenien.

Datum för publicering (år)

2021

Bedömning av relevans (skala 1-3, med motivering)

Nyvärde	För FOI	För FM
3	2	2

Tekniken är nydanande. Dock begränsad till ett militärt delområde inom underrättelseinhämtning.

Referens

Škrlj, B., Martinc, M., Lavrač, N., Pollak, S. (2021). autoBOT: evolving neuro-symbolic representations for explainable low resource text classification. *Machine Learning* 110(5):989–1028. doi:10.1007/s10994-021-05968-x

Egen analys/tankar

En ny metod som bör kunna användas tillsammans med etablerade metoder inom OSINT.

Det finns FOI-projekt som forskar kring denna problemställning (MIA – Beslutsstöds-systems största FoT projekt), men som inte använder denna metod.

2b. Towards a neural-symbolic generative policy model

Kort beskrivning av tekniken – vad den är och hur den kategoriseras

I en distribuerad miljö måste man kunna hantera ostrukturerad data från olika sensorer. Det är viktigt att ostrukturerad kontextuell information kan analyseras för att möjliggöra att generativa policymodeller kan tillämpas på komplexa uppgifter. En ny neuro-symbolisk maskininlärningsteknik som kan analysera både strukturerade och ostrukturerade data utvärderas i referensen. Det som utvärderas är förmåga att lära generativa modeller en policy-inlärningsuppgift utifrån ostrukturerade data.

Tillämpning av tekniken – till vad används den

Fokus: Textanalys.

Tekniken kan användas för att analysera ostrukturerad text.

Bedömd mognad och tekniska beroenden, hinder eller möjligheter för att nå TRL 9

TRL 3-4. Oklart hur mycket arbete som krävs för att nå TRL 9.

Aktörer

IBM Research, Storbritannien.

IBM Research, USA.

Imperial College, Storbritannien.

Purdue University, USA.

Datum för publicering (år)

2019

Bedömning av relevans (skala 1-3, med motivering)

Nyvärde	För FOI	För FM
2	2	2

Referens

Cunnington, D., Law, M., Russo, A., Bertino, E., & Calo, S. (2019). Towards a neural-symbolic generative policy model. In *2019 IEEE International Conference on Big Data (Big Data)*, pp. 4008–4016. doi:10.1109/BigData47090.2019.9005569

Egen analys/tankar

Ytterligare en metodik med tillämpning på textanalys. Antagligen tillämpning inom OSINT.

Kluster 3 - Lågnivå (läges)bild

3a. Combining common sense rules and machine learning to understand object manipulation

Kort beskrivning av tekniken – vad den är och hur den kategoriseras

Automatisk lägesförståelse i video har förbättrats under de senaste åren. State-of-the-art bildbehandlingsmetoder har dock fortfarande brister: de kräver vanligtvis inlärningsdata för varje objektsklass och kan ha höga falska positiva eller falska negativa klassificeringar/förklaringar. En alternativ teknik är att inkludera sunt förnufts-regler och att utnyttja olika

state-of-the-art djupa neurala nätverk (DNN) som detektorer för villkoren i dessa regler. Tekniken testas på problem med att upptäcka: *att plocka upp ett objekt från bordet och lägga ett objekt på bordet*. Tekniken utvärderas med analys av video. En slutsats är att den föreslagna neural-symboliska metoden a) minskar den nödvändiga storleken på data för inlärning och b) möjliggör nya applikationer där märkta data är svåra att få.

Tillämpning av tekniken – till vad används den

Fokus: Bildanalys.

Att förstå vad som händer med objekt som förekommer i videosekvenser.

Bedömd mognad och tekniska beroenden, hinder eller möjligheter för att nå TRL 9

TRL 2. Testas i begränsade bildbehandlingsscenarier. Långt kvar till TRL 9 och generiska tillämpningar.

Aktörer

Eötvös Loránd University, Ungern.

Argus Cognitive Inc., USA.

Datum för publicering (år)

2019

Bedömning av relevans (skala 1-3, med motivering)

Nyvärde	För FOI	För FM
3	2	2

Denna nya metodik kan eventuellt bli viktig i framtiden i militära sammanhang då den inte kräver stora mängder träningsdata.

Referens

Sárkány, A., Csákvári, M., Olasz, M. (2019). Combining common sense rules and machine learning to understand object manipulation. *Acta Cybernetica* 24(1):157–172. doi:10.14232/actacyb.24.1.2019.11

Egen analys/tankar

Tekniken kan kanske utvecklas till en teknik som har betydelse för scenarioförståelse i bildanalyserande sensorer. Möjlig tillämpning i markstrid.

3b. A neural-symbolic architecture for inverse graphics improved by lifelong meta-learning / neuro-symbolisk kapsel

Kort beskrivning av tekniken – vad den är och hur den kategoriseras

Denna teknik består av en "neural-symbolisk kapsel" för att fånga semantisk information i bilder. Utgående från ett nätverk med en initial uppsättning neural-symboliska kapslar förbättrar kontinuerlig inlärning nätverkets detektionsförmåga genom att lägga till kapslar för nya och mer komplexa objekt som nätverket upptäcker. Preliminära resultat visar potential i den nya tekniken.

Tillämpning av tekniken – till vad används den

Fokus: Bildanalys.

Upptäckt av objekt i bilder hos bildalstrande sensorer.

Bedömd mognad och tekniska beroenden, hinder eller möjligheter för att nå TRL 9

TRL 2-3. Oklart vad som krävs för att nå TRL 9.

Aktörer

Bundeswehr University Munich, Tyskland.

Datum för publicering (år)

2019

Bedömning av relevans (skala 1-3, med motivering)

Nyvärde	För FOI	För FM
---------	---------	--------

3	2	2
---	---	---

Metodik för analys kopplad till bildalstrande sensorer. Relevant för alla militära arenor, men högst nytta där nya okända objekt kan upptäckas (ev. mest inom mark- och sjöstrid).

Referens

Kissner M., Mayer H. (2019) A Neural-Symbolic Architecture for Inverse Graphics Improved by Lifelong Meta-learning. In *Pattern Recognition* (DAGM GCPR 2019). Springer (LNCS 11824). doi:10.1007/978-3-030-33676-9_33

Egen analys/tankar

Kan behöva kommunikation genom ett nätverksbaserat ledningssystem för att uppnå maximal potential då tekniken innehåller kontinuerlig inlärning.

3c. Deep logic networks: inserting and extracting knowledge from deep belief networks

Kort beskrivning av tekniken – vad den är och hur den kategoriseras

Forskningen inom djup inlärning har utvecklats mot att kombinera oövervakad inlärning med övervakad inlärning för finjustering. Djup inlärning kan ses som ett modulärt system för inlärning av representationer. En modulär teknik för att infoga bakgrundskunskap i djupa nätverk testas för att se om det kan förbättra inlärningsprestanda och för att utvinna kunskap från djupa nätverk om de representationer som lärs av sådana nätverk. Tekniken använder ett enkelt symboliskt språk för kvantitativa resonemang i djupa nätverk. Med användning av denna metod föreslås och utvärderas ett djupt neuralt-symboliskt system, med de experimentella resultaten som indikerar att modularitet genom användning av konfidsensregler och kunskapsinförande kan vara till nytta för nätverkets klassificeringsprestanda.

Tillämpning av tekniken – till vad används den

Fokus: Bildanalys.

Tillämpning inom klassificering av bilder från bildalstrande sensorer.

Bedömd mognad och tekniska beroenden, hinder eller möjligheter för att nå TRL 9

TRL 2-3. Bör kunna utvecklas till att ingå i system på TRL 9-nivå utan större problem.

Aktörer

City University London, Storbritannien.

Datum för publicering (år)

2018

Bedömning av relevans (skala 1-3, med motivering)

Nyvärde	För FOI	För FM
---------	---------	--------

2	3	2
---	---	---

Referens

Tran, S. N., Garcez, A. S. D. A. (2018). Deep logic networks: Inserting and extracting knowledge from deep belief networks. *IEEE Transactions on Neural Networks and Learning Systems* 29(2):246–258. doi:10.1109/TNNLS.2016.2603784

Egen analys/tankar

Metodik för klassificering inom sensordatafusion av bilder. Militär tillämpning inom objekt-klassificering inom alla militära vapengrenar.

Kluster 4 - Högnivå (läges)bild

4a. Neural-symbolic VQA: disentangling reasoning from vision and language understanding

Kort beskrivning av tekniken – vad den är och hur den kategoriseras

Tekniken kombinerar djup representationsinlärning för visuellt igenkännande och symboliskt resonerande. Ett neural-symboliskt system för visuella frågesvar (NS-VQA) inhämtar en strukturell representation från en bild. Det kör sedan ett program på representationen för att få ett svar på frågan. Att systemet utnyttjar symbolisk struktur som förkunskap erbjuder tre fördelar: 1. Symboliskt resonerande är mer robust, 2. Modellen är mer data- och minnes-effektiv: den fungerar bra efter att ha lärt sig ett litet antal träningsdata, 3. Symboliskt resonerande ger fullständig XAI-transparens för resonemangsprocessen.

Tillämpning av tekniken – till vad används den

Fokus: Bildanalys.

Tekniken kan användas för att människor ska kunna interagera med bilder på ett effektivt sätt och ställa frågor mot ett analysystem om vad som förekommer i bilderna.

Bedömd mognad och tekniska beroenden, hinder eller möjligheter för att nå TRL 9

TRL 2-4. Tekniken behöver antagligen mer forskning och utveckling, men vissa tillämpningar kan eventuellt nå operativ användning tidigare än andra.

Aktörer

Harvard University, USA.

Massachusetts Institute of Technology, USA.

DeepMind, Storbritannien.

Datum för publicering (år)

2018

Bedömning av relevans (skala 1-3, med motivering)

Nyvärde	För FOI	För FM
3	2	2

En ny intressant teknik att följa i skärningen bildanalys/underrättelseanalys, analys av spaningsrapporter och fotoanalys. Tillämpningar möjliga inom flygspaning, satellitspaning mm.

Referens

Yi, K., Wu, J., Gan, C., Torralba, A., Kohli, P., Tenenbaum, J. B. (2018). Neural-symbolic VQA: Disentangling reasoning from vision and language understanding. In *Advances in Neural Information Processing Systems* 31 (NeurIPS 2018). arXiv preprint arXiv:1810.02338

Egen analys/tankar

Tekniken bör kunna finna tillämpning inom framtidens spaning genom att ställa frågor till ett analysystem om vad som förekommer inom bilder.

4b. Symbolic image querying**Kort beskrivning av tekniken - vad den är och hur den kategoriseras**

Det här är ett hybridsystem som består av en bildbehandlingsdel och en textbehandlingsdel. Bildbehandlingen, som är den neurala delen, analyserar bilder och extraherar objekt och deras egenskaper (färg, form, material, etc.). Språkdelen, den symboliska delen, hanterar högnivåfrågor om bilderna (objekten i bilderna), exempelvis "Vilken form har objektet bakom den röda cylindern?" Systemet kan också användas för konceptinläring (*eng. concept learning*), förstärkt inläring (*eng. reinforcement learning*), planering (*eng. planning*) och mycket annat.

Tillämpning av tekniken – till vad används den

Tekniken används för att upptäcka objekt i filmer, att analysera objekts rörelser och att resonera kring deras uppträdande. I den litteratur vi läst strävar eller uppnår (vi är osäkra på vilket) man med labbdatabaserna CLEVR och CLEVRER (CoLLision Events for Video Representation and Reasoning) efter denna funktionalitet.

Bedömd mognad och tekniska beroenden, hinder eller möjligheter för att nå TRL 9

TRL 6. Tekniken verkar stabil och komplett, men verkar främst ha utvärderats på labbdatabaserna CLEVR och CLEVRER.

Aktörer

MIT och IBM

Datum för publicering (år)

2020

Bedömning av relevans (skala 1-5, med motivering)

4, man vill förmodligen kunna ställa underrättelsefrågor till stillbilder och video. Fördelar med tekniken är:

1. Korrekt svar
2. Behöver inte mycket träningsdata som deep learning
3. Förklaringsbarheten (XAI) är högre än för enbart ANN.

Referens

Neurosymbolic VQA, Yi, Neurips 2018

Egen analys/tankar

<https://slideslive.com/38946767/neurosymbolic-visual-concept-learning>

Samma resonemang som för Neurosymbolic VQA ovan men bredare. Mycket är nytt, med stort nyhetsvärde för FOI. Kan ge nya tillämpningar.

4c. Structured event memory: a neuro-symbolic model of event cognition

Kort beskrivning av tekniken – vad den är och hur den kategoriseras

Människor organiserar spontant en kontinuerlig upplevelse till diskreta händelser och använder den inlärd strukturen för dessa händelser för att generalisera och organisera minne. Structured Event Memory (SEM) är en kognitionsmodell som står för mänsklig förmåga att segmentera händelser i minnet. SEM härrör från en sannolikhetsgenerativ modell av händelsedynamik definierad över strukturerade symboliska scener. Genom att bädda in symboliska scenrepresentationer kombinerar SEM fördelarna med strukturerade och neurala nätverksmetoder för kognition på hög nivå. Genom att använda sannolikhetsresonemang över denna generativa modell kan SEM dra slutsatser från händelser, lära sig händelsescheman och använda händelsekunskap för att rekonstruera tidigare erfarenheter. Vi visar att SEM kan producera människoliknande förståelse för naturalistiska videodata.

Tillämpning av tekniken – till vad används den

Fokus: Informationsfusion & Bildanalys.

Teknik för automatisk förståelse på hög kognitiv nivå av vad som sker i en videoström från en bildsensor.

Bedömd mognad och tekniska beroenden, hinder eller möjligheter för att nå TRL 9

TRL 2. Kan kräva mycket forskning för att klättra till högre TRL-nivåer. Oklart om/när TRL 9 kan/kommer att nås.

Aktörer

Harvard University, USA.

Princeton University, USA.

University of California, Davis, USA.

Washington University in St. Louis, USA.

Datum för publicering (år)

2020

Bedömning av relevans (skala 1-3, med motivering)

Nyvärde	För FOI	För FM
3	3	2

Kan vara en mycket viktig teknik för framtiden i automatisk förståelse av video.

Referens

Franklin, N. T., Norman, K. A., Ranganath, C., Zacks, J. M., Gershman, S. J. (2020). Structured event memory: A neuro-symbolic model of event cognition. *Psychological Review* 127(3): 327–361. doi:10.1037/rev0000177

Egen analys/tankar

Mycket forskning återstår. Tycks gå lägre i förståelse än andra tekniker.

4d. Knowledge enhanced neural networks (KENN)

Kort beskrivning av tekniken – vad den är och hur den kategoriseras

En ny teknik kallad Knowledge Enhanced Neural Networks (KENN) har utvecklats. Det är en arkitektur för att inkludera förkunskap i form av en uppsättning logiska klausuler i ett neuralt nätverk. Dessa klausuler ingår i nätverket som ett extra lager. Som en konsekvens kan KENN ignorera träningsdata som strider mot logiska bivillkor, vilket gör systemet robust mot förekomsten av felaktig förkunskap. KENN har utvärderats på två standarddata-uppsättningar för klassificering som visar att inkluderingen av logiska klausuler avsevärt förbättrar klassificeringen. KENN har testats för att lösa problemet med att hitta samband mellan upptäckta objekt i bilder. Utvärderingen visar att KENN överträffar state-of-the-art för denna uppgift.

Tillämpning av tekniken – till vad används den

Fokus: Informationsfusion & Bildanalys.

Tekniken KENN har tillämpningar inom bildanalys specifikt för att finna samband mellan olika objekt. Militär tillämpning bör finnas inom markstrid med flera samarbetande fientliga objekt där man behöver förstå vad som sker.

Bedömd mognad och tekniska beroenden, hinder eller möjligheter för att nå TRL 9

TRL 2-3. Bör kunna nå hög TRL med sedvanlig forskning och utveckling.

Aktörer

Fondazione Bruno Kessler, Italien.

University of Florence, Italien.

Datum för publicering (år)

2019

Bedömning av relevans (skala 1-3, med motivering)

Nyvärde	För FOI	För FM
3	3	3

Potentiellt viktig metodik i situations- och hotanalys i ledningssystem för att få god förståelse av situationen.

Referens

Daniele, A., Serafini, L. (2019). Knowledge enhanced neural networks. In *Pacific Rim International Conference on Artificial Intelligence (PRICAI 2019)*, pp. 542–554. Springer (LNCS 11670). doi:10.1007/978-3-030-29908-8_43

Egen analys/tankar

Kan vara en viktig teknik för att förstå vad som sker på ett komplicerat slagfält. Även robustheten är viktig och kan kanske motverka vilseledning.

Kluster 5 - Beslutsstöd och delfunktioner

5a. Neuro-Symbolic integration with uncertainty

Kort beskrivning av tekniken – vad den är och hur den kategoriseras

De flesta beslut behöver hantera osäkerhet i samband med dem. Förekomsten av osäkerhet kan bero på flera orsaker, såsom ofullständig domänkunskap, opålitlig eller tvetydig data på grund av mätfel, inkonsekvent datarepresentation. Här beskrivs en teknik för att hantera osäkerheten i s.k. neuruler, ett neural-symboliskt integrationskoncept. Neurala nätverk används inom ett symboliskt ramverk för att förbättra prestandan för symboliska regler. Osäkerheten, dvs. troligheten att en osäker händelse kan inträffa hanteras genom att beräkna de sammanvägda troligheterna för ofullständiga eller motstridiga data. Med implementeringen av osäkerhetshantering i neuruler kan noggrannheten slutledningen och prestanda förbättras.

Tillämpning av tekniken – till vad används den

Fokus: Informationsfusion, hantering av osäker information i beslutsunderlag/ beslutsfattande.

Tekniken inkluderar neurala nätverk inom ramen för symboliskt-baserat beslutsfattande för att hantera osäker information.

Bedömd mognad och tekniska beroenden, hinder eller möjligheter för att nå TRL 9

TRL 3-4. Bör kunna utvecklas till en komponent som kan ingå i större system på TRL 9-nivå.

Aktörer

Indian Institute of Technology Bombay, Indien.

Datum för publicering (år)

2018

Bedömning av relevans (skala 1-3, med motivering)

Nyvärde	För FOI	För FM
3	3	3

Referens

Sreelekha, S. (2018). NeuroSymbolic integration with uncertainty. *Annals of Mathematics and Artificial Intelligence* 84(3):201–220. doi:10.1007/s10472-018-9605-y

Egen analys/tankar

Tillämpning i beslutssystem inom ledningssystem för att fatta beslut under osäkerhet. Tillämpning främst i mark- och sjöstrid.

5b. Regelextraktion

Regelextraktion är en teknisk basfunktion som är nödvändig i många NSR-tillämpningar. Beskrivningen av kortet skiljer sig därför från andra kort.

Teknisk beskrivning

Att extrahera logiska regler, exempelvis hornsatser (dvs. $\text{Premiss1} \wedge \text{Premiss2} \rightarrow \text{Konsekvens}$) från ANN. Detta är en basfunktion i många NSR-tillämpningar.

Tillämpning av tekniken – till vad används den

För att förklara vad ett ANN har lärt sig. Och för att – i NSR – erbjuda en ”återkoppling” till den kunskapsbas som (i NSLC – neuro-symbolic learning cycle) initialt har beskrivits.

Bedömd mognad och tekniska beroenden, hinder eller möjligheter för att nå TRL 9

TRL 7-9. Algoritmer för regelextraktion för standard ANN har funnits länge. Däremot är regelextraktion för djupa neuronnät inte tillräckligt utrett.

Aktörer

Okänt

Datum för publicering (år)

1990-talet kanske tidigare

Bedömning av relevans (skala 1-5, med motivering)

5. Regelextraktion är en viktig funktion i NSR för att knyta ihop konnektionistiska och symboliska komponenter i hybridsystem.

Referens

R. Andrews, J. Diederich and A.B. Tickle, "Survey and critique of techniques for extracting rules from trained artificial neural networks", *Knowledge-Based Systems*, vol. 8, no. 6, pp. 373-384, 1995. T. Hailesilassie, "Rule Extraction Algorithm for Deep Neural Networks: A Review", *International Journal of Computer Science and Information Security (IJCSIS)*, Vol. 14, No. 7, July 2016, page 371-381.

Egen analys/tankar

Regelextraktion är en AI-funktion som måste finnas för att få NSR att fungera. Det är en del av learning cycle och har inget egenvärde, det är en pusselbit som måste finnas med.

Om man exempelvis vill extrahera regler från neurala nätverk eller nå förklarbar AI (XAI) måste denna typ av tillämpning/teknik finnas på plats. För FM är XAI viktigt, exempelvis inom beslutsstödsystem och ledningssystem.

Bedömningen är att FOI måste ha kompetens inom området, rent tillämpningsmässigt. DL & NN forskar FOI inom och då måste man inkludera regelextraktion. FOI bör forska på området tillämpningsmässigt för försvarslösningar (högre systemnivå). Grundforskning – universitet och högskola.

5c. Conceptual space representation

Kort beskrivning av tekniken - vad den är och hur den kategoriseras

Fokuserar på nyttjandet av representationen Conceptual spaces (Gärdenfors, 2000) för att sammanföra symbolisk information med sub-symbolisk. Representationen är bland annat intressant för att den, enligt Gärdenfors kan förklara biologisk representation av koncept och den har inspirerat en hel del forskning.

Ett konceptrum är ett mellanlager där symbolisk och undersymbolisk information ”möts”. En punkt i konceptrummet är undersymbolisk och regioner motsvarar koncept (eller symboler). Enligt litteraturen verkar representationen främst vara lämplig för oövervakad inlärning och för att upptäcka koncept (symboler).

Tillämpning av tekniken – till vad används den

Oklart var den gör som mest nytta.

Bedömd mognad och tekniska beroenden, hinder eller möjligheter för att nå TRL 9

TRL 2 – arbetet med det här verkar vara i ett tidigt skede

Aktörer

Främst akademin, exempelvis Lunds universitet och universitetet i Osnabrück

Datum för publicering (år)

P. Gärdenfors, ”Conceptual spaces: The geometry of thought”, The MIT Press ISBN: 9780262273558, 2000.

Bedömning av relevans (skala 1-5, med motivering)

3 – med tiden kan det visa sig att det conceptual spaces är en lämplig representation för att sammanföra symboler och undersymbolisk data, men just nu är det oklart.

Referens

Towards Grounding Conceptual Spaces in Neural Representations, Bechberger, Kühnberger, 2017

Egen analys/tankar

Forskningsnyhetsvärdet av conceptual space representation kan vara stort. Publikationen presenterar en ny nivå – ett conceptual space – mellan symbolisk och subsymbolisk nivå.

Det är dock ett omoget område och bedöms vara ett bevakningsområde. Svensk AI-forskning vid universitet bedrivs.

Tillämpas inom olika områden men inte NSR. Kan finnas vissa egenskaper som kan leda till hybridlösning. Måste integreras i system.

Kluster 6 - Lära av expert/Simulering

6. Neural-symbolic cognitive agent för beteendelinlärning i simulation/Lära av instruktör NSCA i simulering

Kort beskrivning av tekniken - vad den är och hur den kategoriseras

En neural-symbolic cognitive agent (NSCA) är i princip en modul som observerar data från en expert och tillsammans med viss förkunskap uppdaterar förkunskaperna med nya regler. Tekniken bygger i det här fallet på en så kallad Recurrent Temporal Restricted Boltzmann Machine (RTRBM), som erbjuder att man specificerar ungefärliga regler (som trafikregler, se nedan) och att nya regler kan läras från simulerad data.

Tillämpning av tekniken – till vad används den

Utgår från behovet av att automatiskt registrera och formalisera experters beteende. I ett exempel tittar man på elever som övningskör bil i en simulator och där instruktörer bedömer

deras körning med avseende på fordonskontroll, säker körning, tydlighet och ekonomisk körning. Sensordata som samlas in från simuleringen omfattar exempelvis bilens position i relation till övrig trafik, hastighet, växel, m.m. Grova körregler som exempelvis hur man bör bete sig vid en korsning för att inte bryta mot "högerregeln" finns kodade i temporal logik. Vissa fria parametrar finns i reglerna som kan läras in under träning.

Bedömd mognad och tekniska beroenden, hinder eller möjligheter för att nå TRL 9

2011 var det här arbetet i ett tidigt skede. Extraherade regler från körskolesimulatorens såg krångliga ut. Oklart vad som hänt sedan 2011.

Aktörer

TNO

Datum för publicering (år)

2011

Bedömning av relevans (skala 1-5, med motivering)

4+, extrahera regler från träningsdata (exempelvis vid flygsimulator) för att symboliskt beskriva en pilot resonerande, att användas för rapportering av pilotens agerande och för utbildning

Referens

A Neural-Symbolic Cognitive Agent for Online Learning and Reasoning, de Pennig mfl, IJCAI 2011.

Egen analys/tankar

AI-tekniken NSCA kan lära sig parametrar i trafikregler genom elevens körning och bedömarens bedömning. Tekniken passar in i en learning cycle.

Publikationen beskriver hur en agent, dvs. simulerad bil, samlar in data från fysisk förare (elev). Ytterligare data ges genom en expert (fysisk person) som bedömer körningen. Dessutom läggs trafikregler (regler) till AI-modellen. Det finns fria parametrar (ex avstånd till ankommande bil) som systemet kan lära sig.

Ett tillämpningsområde är att extrahera expertkunskap/beteende, eller skapa ett system som blir mer autonomt baserat på expertdata.

Möjliga tillämpningar: flygsimulatorer: extrahera regler genom motsvarande ramverk. (regler kan vara svåra för piloter/expertter att uttrycka men kan fångas genom data).

Kluster 7 - Stöd för förstärkt inlärning

7. Logic tensor network för att förbättra reinforcement learning

Kort beskrivning av tekniken – vad den är och hur den kategoriseras

Förstärkt inlärning (Reinforcement learning, RL) är en maskininlärningsform från data som resulterar i ett beslutssystem, en matematisk funktion som föreslår ett beslut baserat på aktuellt läge. Deep RL (DRL) har de senaste åren visat fantastiska resultat vad gäller olika former av spel som schack och go. Inlärningen är dock långsam och saknar mänsklig förmåga att snabbt lära från ett fåtal exempel.

Logic tensor networks är en NSR som möjliggör specificering av kunskap i first-order logic och här appliceras för att snabba upp DRL genom att göra det möjligt att införa vissa fakta i inlärningsprocessen.

Vad driver utvecklingen

Begränsningen i nuvarande RL-metoder

Bedömd mognad och tekniska beroenden, hinder eller möjligheter för att nå TRL 9

Medelhög TRL, uttestad på leksaksexempel, men kanske inte passar till alla RL problem

Aktörer

Fondazione Bruno Kessler, City university of London

Datum för publicering (år)

2016

Bedömning av relevans (skala 1-5, med motivering)

3, oklart hur viktigt förstärkt inläring är för FM

Referens

Injecting Prior Knowledge for Transfer Learning into Reinforcement Learning Algorithms using Logic Tensor Networks, Badreddine and Spranger, NeSy 2019

Egen analys/tankar

Logic Tensor Networks har stort nyhetsvärde. Det är ett hybridsystem som kan användas för att förbättra och snabba upp tiden som krävs för förstärkt inläring. FOI forskar inom området förstärkt inläring.

Tillämpningsområden är beslutsstödssystem och flyg-AI-forskning.

Kluster 8 - Övrigt

8a. Ankring i robotik

Kort beskrivning av tillämpningen

”Ankring” handlar om att en robot skapar och underhåller kopplingen mellan symboler och sensorsignaler för ett visst fysiskt objekt. Utvecklingsvägar/-potential är bland annat att kunna utbyta förankrade symboler mellan robotar/agenter/beslutsstödssystem.

Vad driver utvecklingen

Behovet för autonoma system att resonera kring den miljö som de agerar i och att kunna kommunicera sin kunskap (exempelvis en karta med innehåll) till andra agenter/robotar. Inspirerat av idéer från det övergripande AI-området (IJCAI-16) diskuteras symboliska och undersymboliska representationer inom robotik. Behovet av både symbolisk och undersymbolisk nivå har stöd även om vissa anser att undersymbolisk nivå är tillräckligt för intelligent beteende (Rodney Brooks).

Kritiska möjliggörande tekniker

Ej identifierade

Mognad –TRL, färdig produkt eller prototyp

Området har forskats på under lång tid men det finns fortfarande utmaningar

Aktörer

Örebro universitet, City university of London, Universität Osnabrück

Datum för publicering (år)

Från omkring år 2000

Bedömning av relevans (skala 1-5, med motivering)

3 – inriktat på mobil robotik. Bör även vara relevant för exempelvis drönare och annan typ av autonoma fordon. Det borde också vara relevant för lägesbilsarbete.

Referens

Besold, T. R., Kuehnberger, K-U., Garcez, A., Saffiotti, A., Fischer, M. H. and Bundy, A. (2015). Anchoring Knowledge in Interaction: Towards a Harmonic Subsymbolic/Symbolic Framework and Architecture of Computational Cognition. Lecture Notes in Computer Science, 9205, pp. 35-45. doi: 10.1007/978-3-319-21365-1_4

Egen analys/tankar

Ankring i robotik bedöms vara av intresse för robotik och autonoma system. Robotik och autonoma system har "ett eget" AI-område som delvis är skilt från NSR. Det finns en tydlig koppling till neural-symboliska metoder, men symboliska/undersymboliska metoder fokuserar inte på ANN (det är bara en delmängd).

Forskning bedrivs i Sverige, litteratur från Örebro universitet.

8b. Neuro-FIS/Neuro-fuzzy systems

Teknisk beskrivning

Neuro-fuzzy inference systems är en integrationsmetod som utgår från expertkunskap i form av ett så kallat *fuzzy inference system* (If A then B-regler). Med hjälp av reglerna specificeras ett ANN. Genom träning med data (övervakad inlärning) kan vissa parametrar i nätverket justeras för att förbättra prestandan.

Ett exempel på Neuro-FIS är ANFIS (Jang, 1993) som tillåter uppdatering av medlemsfunktioner och hur utdata från reglerna (alternativa tolkningar) viktas i slutändan. ANFIS kan därför bara uppdatera medlemsmängderna. ANFIS är baserat på Sugeno FIS.

Tillämpning av tekniken – till vad används den

Den typiska tillämpningen är för styrning av något ställdon baserat på sensorvärden. Klassificering baserat på sensordata är också en möjlighet.

Aktörer

National Taiwan University

Datum för publicering (år)

1993

Mognad –TRL, färdig produkt eller prototyp

6 (se relevans nedan)

Bedömning av relevans (skala 1-5, med motivering)

3. Tekniken är utvecklad och användbar men begränsad i sin uttrycksfullhet.

Referens

Jyh-Shing R Jang, ” ANFIS Adaptive-Network-based Fuzzy Inference System”, IEEE Transactions on Systems Man and Cybernetics, 1993.

Egen analys/tankar

Neuro-FIS/Neuro-fuzzy systems ger förståelse för vad som händer i systemet. Det är en logisk del kombinerad med en neural del och är inte någon svart låda-lösning. Fuzzy rules är en välanvänd lösning och enkel teknik. Fördelen med kombinationen med neutrala nätverk är att vid träning uppdateras fuzzy rules direkt. Det blir således ingen diskrepans vid konvertering, då nätverk och regler är samma sak. Teknikkortet ger högre nivå av transparens än övriga exempel. Det är ett integrerat system med främsta fokus på klassificering [av objekt].

Teknikkortet är relevant för personer som håller på med fuzzy rules och AI-lösningar med hög TRL/mognad.

Det är en generell metod och underlagen vi läst är abstrakta/teoretiska. Vid värdering kan möjliga tillämpningar undersökas. Deltagare kan fråga om deras problem kan lösas mha metoden.

8c. Neural-symbolic learning cycle (NSLC)

Kort beskrivning av tillämpningen – vad är den och hur används den

En hybridprocedur som beskriver samverkan mellan en symbolisk och en konnektionistisk komponent. Proceduren utgår från symbolisk expertkunskap som kodas som en ANN-modell (nämnt representation). ANN-modellen tränas med observationer. Symboler extraheras och kan återföras till den symboliska databasen (benämnt extraktion). Cykeln kan därefter startas om, ifall mer träningsdata blir tillgänglig. Observera att cykeln är användbar även i de fall där det inte blir fler cykler än en. I tidigare underlag finns en figur över resonemanget som är hjälpsam (se referens med publikations år 2019).

Vad driver utvecklingen

Generellt tillämpbar användning av NSR, som i sig bidrar till att förklara och motivera nyttan med NSR (genom att möjliggöra samtidig integration av expertkunskap och underlättelserapporter och sensordata).

Kritiska möjliggörande tekniker

Representationsmetod för att skapa ANN och Extraktionsmetod för att skapa symboler från ANN. Dessa metoder kan vara approximativa och störa metodens prestanda. Observera att om man vill försöka välja en ny konnektionistisk metod så kan man behöva utveckla en ny passande extraktionsmetod.

Mognad –TRL, färdig produkt eller prototyp

Skapad 1993 (Towell, Shavlik) men en ”ordentlig experimentell utvärdering saknas” (NeSy’19)

Aktörer

NSLC är en generell metod och har inga särskilda intressenter.

Datum för publicering (år)

Från 1993

Bedömning av relevans (skala 1-5, med motivering)

3 – typisk metod, kan vara användbart för stöd för lägesbild

Referens

Martin Svatos, Gustav Sourek and Filip Zelezny, "Revisiting Neural-Symbolic Learning Cycle", NeSy'19, 2019.

Geoffrey G Towell and Jude W Shavlik. "Extracting refined rules from knowledge-based neural networks", Machine learning, 13:71–101, 1993.

Egen analys/tankar

Generell metodik, ett hybridsystem som dock inte är så väl utrett.

Teknikkortet kan användas vid presentation i utbildande syfte för att visuellt visa vad NSR är. Inte intressant för värdering.

8d. Generative art

Kort beskrivning av tillämpningen – vad är den och hur används den

En hybridmetod som kombinerar ett GAN (Generative Adversarial Network) för att generera konst med en symbolisk metod för att skapa träningsdata. Resultatet har bedömts av mänskliga betraktare med avseende på "bättre", "ovanligare", "mer kreativ", m.m.

Vad driver utvecklingen

Nyfikenhet

Kritiska möjliggörande tekniker

Deep learning/GAN

Mognad –TRL, färdig produkt eller prototyp

Färdig prototyp

Aktörer

Facebook, Adobe

Datum för publicering (år)

2020

Bedömning av relevans (skala 1-5, med motivering)

2 – kan möjligtvis användas för att generera bra syntetiskt data eller syntetiska scenarier

Referens

Neuro-symbolic generative art: a preliminary study, Aggarwal, Parikh, 2020

Egen analys/tankar

Vi resonerar om möjliga tillämpningar. Metoden beskrivs ge resultat som är mer kreativa, mer annorlunda och så vidare. Kreativitet och osäkerhet är något som är vanligt förekommande vid exempelvis användning av mjuka operationsanalytiska metoder. Skannarna är mycket osäkra på nytta för militär verksamhet och kortet värderades inte.

8e. T-PRISM

Kort beskrivning av tekniken -vad den är och hur den kategoriseras

T-PRISM erbjuder ett logikprogrammeringsspråk som möjliggör konstruktion av både grafiska strukturer och neurala nätverk.

Tillämpning av tekniken – till vad används den

Används när man inuti ett program vill styra konstruktionen av, exempelvis, ett neuralt nätverk.

Bedömd mognad och tekniska beroenden, hinder eller möjligheter för att nå TRL 9

Förefaller kunna användas direkt (<https://github.com/kojima-r/prism>)

Aktörer**Datum för publicering (år)**

Språket PRISM (1995)

Bedömning av relevans (skala 1-5, med motivering)

2 – inte så viktigt i NeSy-sammanhang verkar det som.

Referens

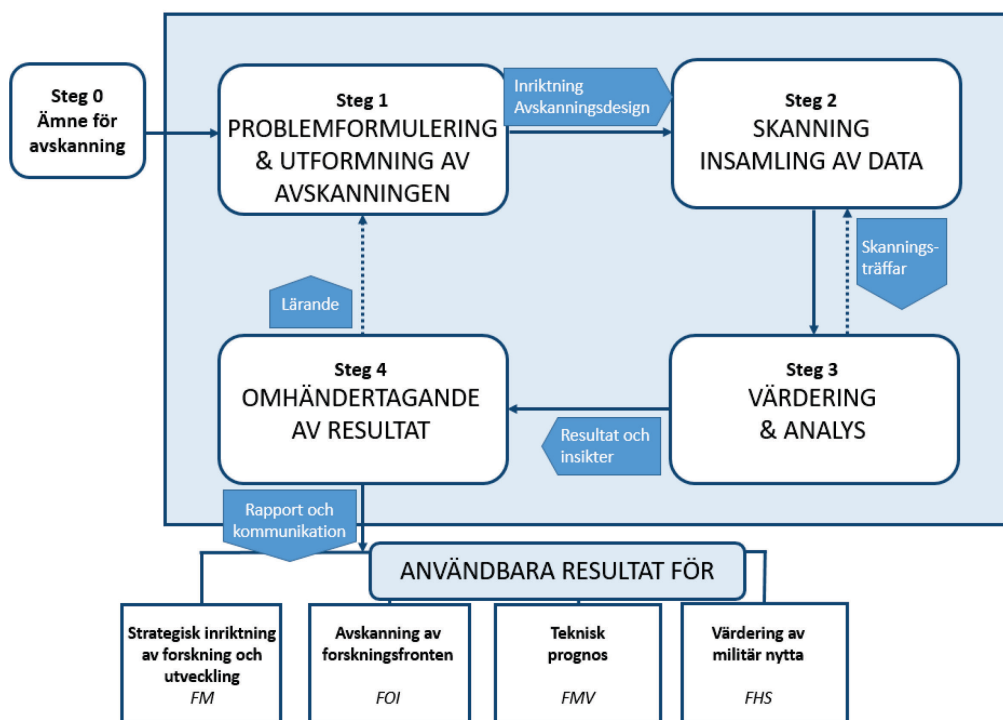
T-PRISM: A tensorized logic programming language for data modelling, NeSy '19, 2019

Egen analys/tankar

Det känns inte som att T-PRISM erbjuder något unikt och spännande för NSR-området. Har inte värderats eller omhändertagits

Bilaga 3 – Metod för genomförande

I denna bilaga görs en genomgång av arbetet med NSR. Vi redovisar de fyra steg i den process som tagits fram (Lindberg o.a., 2021). Processen framgår övergripande i Figur 16 nedan.



Figur 16 Avskanningsprocessens fyra steg, inklusive steg 0 där ämne för avskanningen bestäms. Dessutom visas de primära mottagarna av resultaten.

Under arbetet använde vi principer för att öka användbarheten av processen och resultaten från avskanningen (Lindberg o.a., 2021).

Steg 1, problemformulering och utformning av avskanningen

Steg 1 innebär att planera och lägga grunderna för genomförandet av avskanningen.

Resultaten från steget innebär att:

- Mål och syfte med arbetet har tydliggjorts, det vill säga att arbetet inriktats.
- Ledstång för genomförandet har skapats genom att en avskanningsdesign utformats.

Vi genomförde steg 1 som ett flöde av aktiviteter:

1. Bestämma ämne att skanna och inleda steget.
2. Skissa en grov aktivitetsplan.
3. Skapa en bild av omfattning för området och arbetet (scope).
4. Genomföra intressentanalys.
5. Formulera problemet, bestämma området för skanning (frame).
6. Utforma designen för genomförande av avskanningsaktiviteten.
7. Färdigställa aktivitetsplanen i samråd med projektledaren.
8. Rapportera resultat avseende problemformulering och problemdefinition samt design för genomförande av avskanning.

Inom avskanning är steg 1 viktigt. Själva kärnfrågan är att utforska något som är okänt. För att navigera i det okända krävs en ledstång, en tydlighet i vad som ska uppnås och hur det ska gå till. Detta avser till exempel val av metoder, typ av resultat mellan steg och utformning av arbetet i stort. En stor del av detta material redovisas inte i föreliggande rapport men finns sparad i projektet Avskanning av forskningsfronten.

Vi genomförde fyra arbetsmöten med den större avskanningsgruppen. I den större arbetsgruppen diskuterades ledning inklusive personalplanering, arbetsfördelning och förankring med chefer. Vi tog även fram tillämpningsområde ("scope") för skanningen i form av en domänkarta, se avsnitt 4.1. Dessutom gjordes en intressentanalys. Intressentanalysen är central avseende deltagande i processen, mottagare av resultat och andra experter som kan vara intresserade av underlag. Sist men inte minst förankrades skrivningarna om syfte och mål med avskanningen, utformningen av processen och metodval inom den större arbetsgruppen som till stor del deltagit vid framtagning av underlaget. Mellan arbetsmöten i den större arbetsgruppen omhändertog, analyserade och förberedde avskanningsledarna/metodexperterna för nästa aktivitet.

Planering av avskanningen vid steg 1

Arbetet med avskanning inleds med att problemstrukturera, definiera problemet och utforma genomförandet i en design. Steg 1 resulterade i en inriktning och en avskanningsdesign. Vi redovisar dessa i den form, punktat, som de använts under arbetets gång.

Inriktning från problemformulering

Bakgrund

- Inom FOI har området neuro-symboliskt resonerande varit känt en längre tid. Det är ett forskningsområde vi kastat ett öga på. Det är ett område som nu börjar tillskrivas allt större intresse både forskningsmässigt och för tillämpningar.

Syfte och mål med arbetet

- En frågeställning, och möjligt övergripande mål, är att genom värdering¹⁷ avgöra om området är relevant för FOI att bedriva forskning inom och i så fall varför. Dessutom orientera om områdets status och möjliga tillämpningar för FOI:s forskningsorganisation och övriga aktörer inom FoT-programmet forsknings-skanning och teknisk prognos.
- För skannarna är arbetet kunskapsbyggande avseende teknikerna, forskningsstatus och potentiella tillämpningar. Man vill reda ut vad områdets status, visa på vad som är nytt och svårt.

Intressenter

- Primära intressenter och användare av resultat: FOI:s forskningsområdesföreträdare (FoF), FOI:s forskningsdirektör.
- Sekundära intressenter: Materielområdesföreträdare (MoF) vid FMV, samt berörda vid FM och FHS, vilka bjuds in till steg 3 och steg 4.

Avgränsning

- Samlar inte evidens som talar för eller emot hypoteser vid djupare skanning.
- Militärt försvar i fokus.

¹⁷ Utgående från kriterierna mognad, tillämpbarhet och relevans för militär användning

Process och metodval

- Utgår från FOI process för avskanning. Metoder som projektgruppen sedan tidigare är väl förtrogna med väljs för genomförande.

Möjliga rapporteringspunkter

Frågeställningar som ska undersökas och besvaras¹⁸

- Vad är oväntat/nytt inom området.
- Vad är nyttan med NSR för militär verksamhet.
- Varför investerar andra aktörer, militära (primärt) och civila (sekundärt), inom NSR.
- Finns det områden som FOI eller andra försvarsanslutna aktörer bör bevaka eller bedriva forskning inom.
- Förslag för fortsatt hantering av området NSR.

Avskanningsdesign: Inriktningar för metod och genomförande av processen

- Metoder som projektgruppen sedan tidigare är väl förtrogna med har valts för genomförande. (Domänkarta, HSTOOL, web-sökningar, nyttjande av kriterier, värdering vid workshop)
- Anpassning av steg 2. För att tillåta AI och tillika skanningsexperter att läsa in sig görs en inledande bred skanning, som följs av en djup skanning på ett urval av områden. För att välja sökord/områden för den djupa skanningen så tillämpas värderings-/urvalskriterier (se metodbeskrivning för steg 2). Arbetet innebär en sväng in i steg 3, värdering, innan den djupa skanningen påbörjas. Ämnena (skanningsträffar) som hittats i den djupa skanningen ställs samman och förs vidare till steg 3, värdering.
- Värderingen i steg 3 är expertbaserad. Inledningsvis avsåg vi värdera:
 - Vad inom området/vilken teknikutveckling som ryms eller inte ryms inom befintlig FoT,
 - Om civila aktörer kan bedriva forskning inom området som FOI och försvarsmyndigheter kan merutnyttja, kommer NSR då att bli ett säkerhets- eller integritetskritiskt område?
 - Vad som är försvarsintressant och varför – vilka tillämpningar NSR kan ge/vilka militära AI-problem NSR kan lösa.

Ovanstående uppdaterades kontinuerligt och använda värderingskriterier återges i metodbeskrivning av steg 3, värdering.

- Vi diskuterade bias och tillförlitlighet och såg medverkan av experter från annan försvarsmyndighet eller universitet/högskola till steg 3 som ett sätt att hantera detta. Alternativt genom intervjuer.

En övergripande bild över vägval i genomförandet av avskanningen syns i Tabell 1 Vägval i genomförandet av avskanningen.

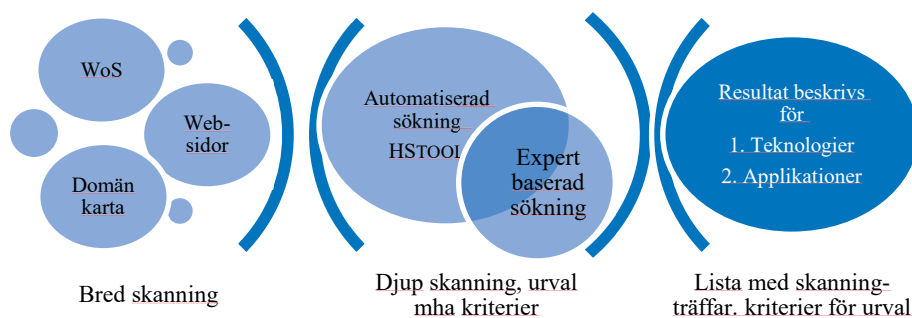
¹⁸ Frågeställningarna skapades av arbetsgruppen, men frågan ställdes även till FOI:s FoF:ar.

Tabell 1 Vägval i genomförandet av avskanningen.

Avskanning och analys				
Typ av avskanning	Utforskande		Ämnesspecifik	
Angreppssätt	Bred och förut-sättningslös	Bred men inriktad	Fördjupande	Söka svar/evidens
Metod för insamling av data	Automatiserade metoder	Systematisk manuell undersökning		Dialog med experter
Metod för urval avskanningsträffar	Välja ut träffar		Ta bort irrelevanta träffar	
Grad av bevis/säkerhet	Låg	(Medel)		Hög
Värdering och analys av	Individuella träffar	Relationer mellan träffar		Sammanställda träffar
Ambitionsnivå för analys och värdering	Prioritera skanningsträffar	Snabb värdering med kriterier	Djupare analys med kriterier	Värdering av påverkan/konsekvens
Metod för analys och värdering	Workshop	Intervju, enkät etc.		Kvanitativ

Steg 2, bred och djup skanning

Steg 2, skanning, omfattade en bred och en djup skanning. Vi utgick från domänkartan som skapades i steg 1.



Figur 17 Arbetsprocess och metodval i steg 2 skanning.

Vi valde att göra både en bred och en djup skanning. Syfte och mål för bred skanning var att:

- Övergripande kartlägga området, hitta termer för djup skanning och förbättra förståelse samt förbättra domänkartan från steg 1,
- Redogöra för aktörer som bedriver FoU och tillämpningar.

Syfte och mål för djup skanning var att:

- Redogöra för teknikforskning och forskningshorisonten inom NSR, att använda vid värdering av underlag i steg 3,
- Identifiera teknikforskning med särskilt intresse för försvarsfrågor,
- Testa klustring och rankning för analys och urval med hjälp av kriterier (steg 3 – första iteration).

Vid den breda sökningen användes Web of Science (WoS) och sökmotorer. Arbetet inleddes brett och vi valde att inkludera termer som återfinns inom robotik och autonoma system. Sökning och omhändertagande av resultat genomfördes enligt en enhetlig systematik. Den automatiserade sökningen innebar att skannaren först sökte på WoS och sammanställde underlag för ett antal söktermer. Detta underlag användes sedan för att välja söktermer för automatiserad sökning med hjälp av HSTOOL.

Den manuella skanningen innebar manuella sökningar på sökmotorer och olika webbplatser, läsning i böcker och konferenspaper. Arbetet resulterade i den översikt av NSR som redovisas i Kapitel 4, Neuro-symboliskt resonerande – en fördjupande introduktion. Resultaten från den breda skanningen redovisades, diskuterades och användes som ingångsvärde till den djupa skanningen.

I likhet med den breda skanningen valde vi två metoder för att genomföra den djupa skanningen, en automatiserad med hjälp av HSTOOL och en expertbaserad manuell sökning.

Från sökningarna valdes underlag för fortsatt omhändertagande med hjälp av kriterierna: *stort nyhetsvärde, i framkant av forskningen/tillämpningen och av relevans för försvar och säkerhet*. Respektive skannare värderade sina egna underlag.

Den automatiserade sökningen med hjälp av HSTOOL använde enbart söktermer för neuro/neural-symbolic reasoning och resulterade i listor på artiklar och kluster baserade på textanalys. Teknikexperten som också utförde skanningen läste abstracts på flera artiklar (antal) och värderade dessa med hjälp av kriterierna ovan. De högst rankade artiklarna valdes ut för läsning och sammanfattades i teknik- och tillämpningskort.

Den manuella, expertbaserade sökningen utgick från resultaten i den breda skanningen, och vid fortsatta sökningar användes både termerna neuro/neural och sub-symbolic reasoning. Skannaren använde något som vi kan kalla för en snöbollsprincip och skanningen resulterade i ett antal intressanta träffar. Kriterierna ovan användes för att välja vilka källor som skannaren skulle läsa och sammanfatta i teknik- och tillämpningskort. Vi diskuterade även underlaget i den mindre arbetsgruppen inför urval.

När samtliga kort skrivits av skannarna och tillika sakexperterna diskuterade och analyserade vi samtliga teknik- och tillämpningskort i den mindre arbetsgruppen. De egna reflektionerna antecknades och korten klustrades.

Steg 2.5, hantering av skanningsunderlag

Vid val av metoder i steg 1 inriktade vi arbetet till att inte omhänderta resultaten från skanningen individuellt. I stället tillämpade vi ett angreppssätt där vi integrerade skanningsträffarna (klustringen) och dessa kombinerade underlag värderades i steg 3. Den expertbaserade klustringen/integreringen gjordes efter analys av korten av den mindre avskanningsgruppen.

Den första klustringen var dock inte tillräcklig. Inför värderings-WS sammanfattades underlagen i exceltabeller och en ytterligare, informationsbaserad klustring gjordes. Det vill säga vi utgick från de tekniska beskrivningar och tillämpningar som beskrivs i korten. Dessa kluster representerades i PowerPoint bilder (se Bilaga 1).

De två olika metoderna för att skapa kluster jämfördes och de överensstämde i stort. Den andra klustringen bidrog till att välja bort underlag som inte skulle värderas. Det visade sig vid den innehållsmässiga klustringen att vare sig kluster eller korten har tydliga gränser utan att de glider in i varandra. Tekniker, metoder och syften glider samman, det är svårt att sätta tydliga gränser.

Steg 3, värdering och analys

Teknik- och tillämpningskortet i steg 2 krävde analys och utveckling inför steg 3, värdering. Deltagarna delgavs underlag från den djupa skanningen inför värderingen: uppdaterad domänkarta, sammanställning från bred skanning och teknik- och tillämpningskort. Kluster som användes vid värdering presenteras i Bilaga 1.

Avskanningsledare och metodexpert utvecklade underlagen från skanningsteget ytterligare för att dels underlätta planering och genomförandet av värderingsworkshopen, dels ge underlag som deltagarna kunde orientera kring vid värderingen. Underlagen togs fram i nära dialog med saks experter. Vi skapade bland annat:

- En uppdaterad domänkarta,
- En teknisk introduktion till NSR,
- Bilder av kluster till värderingen.

Syfte och mål för steg 3, värdering var att:

- Identifiera möjliga tillämpningar som kan ge lösningar på tekniska utmaningar som finns hos FOI, FM, FMV, FHS.
- Värdera relevans, exempelvis för:
 - FOI:s forskning,
 - FM behov och utmaningar,
- Värdera hur NSR fortsatt ska behandlas.

Värderingen genomfördes i digitalt workshopformat under tre timmar. Deltagare bjöds in tre veckor innan värderingen och delgavs vissa underlag. Cirka två veckor innan värderingsworkshopen bjöds även fler tekniska experter vid FOI in till mötet. Syftet var både att få fler deltagare och att sprida underlagen inom organisationen.

Deltagare som bjöds in till workshopen hade roll som expert. Från de olika organisationerna medverkade tekniska experter och forskningsområdesföreträdare (FOI), tekniska och militära experter och ansvariga i FoT programmet för forskningsskanning och teknisk prognos från FMV och FHS. Mottagare av resultat, såsom FOI:s forskningsdirektör och Försvarmaktens FoU-detalj, bjöds in men medverkade inte vid värderingen. Någon vecka innan workshopen tillkom deltagare genom att vissa inbjudna delade kalenderlänk. På detta sätt tillkom en deltagare från ISP och flera tekniska experter från FOI.

Värderingsworkshopen genomfördes i fyra steg: inledning, värdering av kluster, sammanfattande värdering och avslutning. Agendan för värderingen var:

Agendapunkt	Cirka tider
Introduktion och upplägg, inklusive introduktion till området NSR	13.00 - 13.30
Värdering	
<i>Klassificera objekt(lågnivå sammanställning)</i>	13.30 – 14.00
<i>Fråga AI:n om råd (högnivå sammanställning)</i>	14.10 – 14.40
<i>Robotik och autonoma system</i>	14.45 -15.10
<i>Lära av experter / simulering</i>	
<i>Textanalys</i>	15.15 – 15.30
Summerande värdering	15.30 – 15.50
Avslutning	15.50 – 16.00

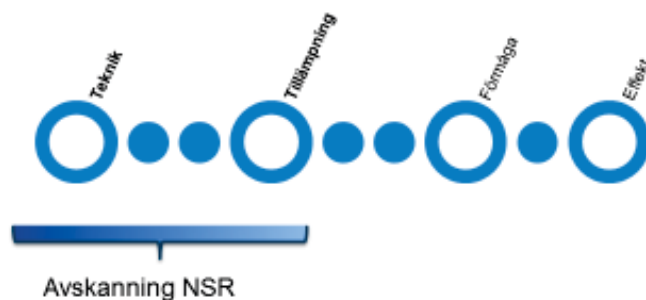
För varje område genomfördes en kortare presentation av området följt av en enklare brainstorming. Syftet var att ge deltagarna tillfälle att navigera i området. Instruktionen som ställdes till deltagarna var att identifiera militära tillämpningar för specifika NSR-tekniker och tillämpningsområdet. Därefter värderades området avseende:

- Hur stor **inverkan** bedömer du att NSR-tekniker kan ha för militära tillämpningar inom ”det specifika området” (exempelvis klassificera objekt...)?
- Hur stort **tekniskt nyhetsvärde** bedömer du att NSR-tekniker har för att lösa militära utmaningar inom ”det specifika området”?

Värderingarna genomfördes via en Skypeomröstning och fokus för värdering var på teknik och tillämpning, snarare än de mer Försvarsmaktsinriktade frågorna förmåga och effekt, se figur 18. Vid Skypeomröstningen användes skalan 5= mycket hög, 3= medelhög, 1= liten/ingen, ?= vet ej.

Metod för värdering

1. Presentation av tekniken
2. Diskutera möjliga militära tillämpningar (Brainstorm)
3. Värdera tillämpning och teknik (Skypeomröstning)



Figur 18 Upplägg för värdering av klustrade kort i tillämpningsområden.

Mest tid lades på de två inledande områdena *Klassificera objekt*, till *Fråga AI:n om råd*.

Mindre tid lades på *Robotik och autonoma system*, *Lära av experter/simulering* och *Text-analys*.

Avslutningsvis genomfördes en avslutande värdering för NSR som helhet. Två frågor ställdes: ska FM finansiera och hur, i vilka former, ska arbetet tas vidare. Frågeställningar och metod framgår vid redovisningen av resultaten.

Till grund för utformningen av arbetet i steg 3 använde vi ett antal frågor som vi ansåg vara relevanta att besvara för NSR, se Tabell 2.

Tabell 2 Grund för de frågor som ställdes vid värderingsworkshop och syntes av resultat.

Kriterium	Utkast frågor vid värdering	Värdering genom
Inverkan (Impact)	Hur kan NSR komma att påverka försvarsområdet? Har det relevans, för vad?	WS
Tillämpning (Impact)	Vilka behov som NSR skulle kunna lösa är viktiga för försvar? Varför?	WS
Nyhetsvärde (Novelty)	Innehåller underlaget från avskanningen nyheter/är nytt eller har du läst om dem förut? Är det något som du anser att FOI bör kunna/lära sig mer om?	WS
Genomförbarhet (Feasibility)	Hur bedömer du teknologisk genomförbarhet av de underlag du läst? Bedömer du att området NSR/delmängder i NSR är kost-effektivt att bedriva verksamhet inom (kostnad i relation till relevans och påverkan)?	AI expert Syntes <i>Genomfördes ej på grund av tidsbrist</i>
Rätt i tid (Timeliness)	Hur snabb bedömer du att utvecklingen inom NSR är? Behöver FOI tid att förbereda, finns det behov av att agera snabbt? Varför?	Större avskanningsgrupp, Syntes
Summering och fortsatt omhändertagande	Se alternativ i metodkapitel	WS och större avskanningsgruppens syntes
Förhållande i relation till FOI befintliga forskning Säkerhetsaspekter	Resonemang kring kriterierna	Större avskanningsgruppens syntes

Steg 4, resultat och kommunikation

Syfte och mål med steg 4 är att omhänderta och rapportera resultat, kommunicera dessa med mottagare och följa upp användning av resultaten från avskanningen. Vi har vid skrivande av rapporten omhändertagit att rapportera resultat.

Efter värderingsworkshopen sammanställdes röstningarna och utfallen diskuterades vid ett arbetsmöte med den större avskanningsgruppen. Den större avskanningsgruppens syntes av värderingsworkshopen beskrivs i steg 3. Syntesen utgör grunden för de slutsatser och rekommendationer som redovisas från avskanningen om NSR (steg 4).

En diskussion som uppstod under steg 4 var vi skulle hantera resultaten från värderingen. Vi valde att använda vissa kriterier som inte användes under workshopen (se Tabell 2) och att föra ett resonemang baserat på något som kallas *Level of regret*¹⁹, det vill säga hur mycket man skulle ångra sig om man inte vidtar åtgärder.

Relevanta underlag från projektet finns sparade på Avskanningsprojektets gemensamma area.

Vi har vid skrivande av rapporten ännu inte kommunicerat resultaten till uppdragsgivare och andra mottagare. Vi vet inte heller hur och i vilken mån resultaten från NSR kan användas eller används. Rent metodmässigt är processen inte avslutad vid färdigställande av rapporten.

Två avrapporteringstillfällen har identifierats: inom ram för FoT Avskanning och Teknisk prognos samt vid FOI:s lunchseminarium för spridning av kunskap till myndigheten i stort.

¹⁹ För en beskrivning, se metodbeskrivning i rapporter från Nato-studien FATE. Slutrapport: NATO, STO-TR-SAS-123, *Futures Assessed alongside socio-Technical Evolutions (FATE)*. Technical Report RDP, DOI 10.14339/STO-TR-SAS-123, ISBN 978-92-837-2322-6.

Bilaga 4 - Anteckningar från brainstorming av tillämpningsområden

Klassificera objekt

Diskussionen inleddes med en reflektion över varför den symboliska/logiska metoden behövs. AI-metoden djupinlärning används redan framgångsrikt för att klassificera objekt. En respons var att djupinlärning inte alltid är tillräcklig.

Det framkom under diskussionen att de utmaningar NSR försöker lösa är militärt relevanta för klassificering av objekt. Kombinationen med symbolik och logik kan t.ex. ha potential i situationer med begränsad tillgång till data, såsom militär verksamhet där man planerar och övar inför ett krig som ännu inte inträffat. NSR kan bidra genom snabb inlärning och minskat behov av träningsdata.

Dagens system för klassificering är ofta inte tillräckliga, särskilt inte i komplexa miljöer där en kontext måste vägas in i scenen. Där kan AI-funktioner fylla ett behov. Dagens djupinlärningssystem för bildanalys beskrevs av en deltagare som alltför begränsade, ordvalet autistisk användes. Möjligheten med NSR lyftes och specifikt det som beskrivs i artikel 4D – Knowledge enhanced neural networks. Om hänsyn tas till kontexten kan svaga och tvetydiga signaler tolkas bättre. Det bedömdes av deltagarna att sådan AI-funktionalitet bör utvecklas specifikt för militära ändamål då det kräver specifik kunskap om verksamheten och dess kontext. En NSR-lösning skulle kunna bidra till bättre förståelse för vad som pågår i scenen. Idag har människor ansvar för detta. Människor är dock olika bra på att uppfatta avvikelser i de miljöer man befinner sig i.

I vissa system är toleransen för felklassificeringar låg. Förenklat uttryckt innebär det att samma typ av objekt i samtliga bilder i en videosekvens/över tid ska vara korrekt klassificerade. Frågan är om NSR kan bidra till att lösa detta behov av total korrekthet.

Ett resonemang fördes som var relevant för AI-området som helhet och inte enbart för NSR. Resonemanget rörde behov av att säkerställa bättre precision och verkan, att helt säkert undvika vådabekämpning och bättre använda simulering.

Fråga AI:n om råd

Denna diskussion inleddes med ett resonemang om förklarbarhet. Snarare än att diskutera funktionalitet och tillämpningar till följd av en mer generell AI i exempelvis beslutsstödsystem och högnivå lägesbilder lyftes tekniska behov som krävs för sådan funktionalitet. Diskussionen rörde vilka tekniska utmaningar som finns och om NSR kan bemöta dessa.

För att kunna använda NSR krävs generellt en tillräcklig konfidensgrad för algoritmerna. Militära tillämpningar kräver ofta särskilt säker AI där algoritmen inte kan vara en svart låda. Den måste vara grå eller gärna vit för att AI-lösningar ska kunna användas för exempelvis verkan. Det ansågs att djupinlärning idag är otillräckligt i det avseendet. Man måste kunna förutse konsekvenser och ha tillräcklig förlitlighet. Det finns ett uppenbart behov att nå förklarbarhet för tillämpningar inom kategorierna Fråga AI:n om råd och högnivå lägesbild/beslutsstödsystem. Sådana tillämpningar är naturliga för symboliskt resonerande. Därför bedömdes NSR mer relevant till skillnad från på lägre nivå.

Möjligheten att bättre upptäcka anomalier och urskilja/särskilja objekt, och därmed få bättre precision i en lägesbild, diskuterades. Detta är särskilt viktigt i situationer där både civila och militära objekt finns i samma lägesbild/område och det skjuts skarpt.

Komplexa system är per definition svåra att förstå och en fråga ställdes om NSR kan bidra till bättre tillförlitlighet och transparens genom att lösa upp komplexa system i mindre entiteter (modularisering). Genom att isolera djup-/maskininlärningen till vissa begränsade delar av ett system kanske man kan öka förståelsen för övriga delar och på så vis för systemet som helhet. Det förblev vid workshopen oklart hur och om NSR, eller andra hybrida system,

kan tillgodose transparens på detta vis. Det hördes skeptiska röster angående hur neuron-nätsmoduler blir mer förklarbara av att de samverkar med symbolisk representation. Tekniken måste bidra till förklarbarhet och vara möjlig att validera och verifiera. Vissa av artiklarna hävdar att man är på väg mot just dessa AI-utmaningar. NSR kan kanske bidra till förklarbar AI (XAI).

Analys kopplat till områdena Klassificera objekt och Fråga AI:n om råd

NSR erbjuder militärt relevanta lösningar för klustren Klassificera objekt till fråga AI:n om råd. Vid den korta brainstormingen lyftes flera behov och tillämpningar som idag inte tillgodoses. Fokus för klassificering av objekt var på möjligheter till förbättrad klassificering och funktionalitet med hjälp av NSR, där dagens tekniker är otillräckliga.

Behoven av förklarbarhet, transparens, validering och verifiering var i fokus under diskussionerna avseende Fråga AI:n om råd. Att lösa dessa utmaningar är kritiskt för implementering av tillämpningar som beskrivs i klustret Fråga AI:n om råd. Vid fortsatta diskussioner bör frågan om tillförlitlighet förfinas och diskuteras för olika typer av tillämpning och system. I våra resonemang blandades autonom offensiv verkan med lägesbild samt med underlag för beslutsstöd.

Diskussioner om tillämpningar fördes framförallt utifrån ett blått resonemang. Bland annat lyftes utmaningen med klassificering och hantering av dessa data i komplexa miljöer. Vi upplevde en diskrepans i bedömningar av NSR:s relevans mellan forskare som skapar lösningar för en fysisk miljö och de som framförallt nyttjar rent digitala datamängder i en digital miljö.

Klassificera objekt värderades som mindre relevant för fortsatt arbete än Fråga AI:n om råd. Vi uppfattar att det av vissa deltagare vid workshopen ses som ett överkomligt problem att nå en tillräckligt funktionell AI med befintliga AI-lösningar för klassificering, varför värdering blev lägre än för området Fråga AI:n om råd. I högnivåområdet Fråga AI:n om råd finns fler utmaningar som rent tekniskt inte kan tillgodoses med befintliga AI-metoder/-lösningar.

En reflektion är att vi inte diskuterade perspektivet blå-röd, eller tillgång till tillräckliga datamängder. Kanske hade detta förändrat resonemangen. Teknik- och tillämpningskortet om KENN (kort 4d) lyftes som intressant vid diskussion om klassificera objekt, ett kort vi tidigare hade i detta kluster (innehållsbaserat) men inför värderingen lyfte bort igen.

Teknikkortet för de två olika klustren beskrev två olika sätt att hantera och använda träningsdata. Klustret Klassificera objekt redovisade möjligheter med kontinuerlig insamling och inlärning, medan klustret Fråga AI:n om råd diskuterade hur utmaningen med för få träningsdata skulle överkommas. Under värderingen kom vi inte in på dessa frågeställningar. Just tillgång till tillräcklig och rätt träningsdata, hantering av träningsdata, otillräcklig tillgång till träningsdata och så vidare är utmaningar som diskuteras för DL/ML. De togs även upp när vi inledningsvis skapade domänkartan. Tekniker för att bättre använda eller minska behov av träningsdata bör bara intressant. Frågan kopplar till möjligheter med simulering och träning samt övervakad inlärning.

Robotik/Autonoma system

En artikel redovisades och NSR-tekniken bedömdes relevant för alla tillämpningar som inbegriper sensorer på farkoster eller andra rörliga objekt. Det behövs teknik för sensorer/robotar/autonoma system som delar lägesbild och själva måste agera i relation till en lägesbild. Bilden behöver inte vara densamma som för en människa men autonoma system som rör sig i ett rum måste kunna beskriva och hantera till exempel var och hur en stol står där. En symboliskt baserad karta med kontinuerlig uppdatering kan vara ett sätt för olika sensorer i samma interagerande systemlösning att dela och förstå varandras data. Exempel på tekniska lösningar för detta finns på olika universitetscampus i USA och Tyskland.

Den lägesbild som robotar/autonoma system använder behöver inte vara mänskligt förståbar. Däremot, om exempelvis mål- och objektigenkänning hanteras av lägesbilden, kan den även förstås av människan.

Behovet av lägesbild för icke-mänskliga system gäller även planering. Autonoma system måste kunna hantera en förändrad lägesbild, där sensorerna ger varandra information om den förändrade lägesbilden. Exempel som diskuterades var en fabrik, där helt plötsligt kartonger är i vägen för den rutt som robotarna brukar ta. Ett annat exempel var ruttplanering på väg. För rutinartade uppgifter som utförs dagligdags (köra samma vägsträcka till jobbet med bil) krävs en kanske mindre avancerad AI-lösning – den är på rutin. Däremot, när det oväntade inträffar (olycka eller ishalka), eller när man ska göra en ej rutinartad uppgift (åka på en resa där bilfärden inbegriper en längre sträcka till en okänd plats och där en uppsjö av vägalternativ är möjliga) krävs en mer nogsam planering eller mer avancerad AI-lösning.

Deltagare noterade en stark koppling till smarta hem-lösningar och IoT i litteraturen. Där finns civila publikationer och forskning finns i Sverige. Örebro universitet nämndes i samband med självgående robotar i gruvor, och för att ha myntat begreppet anchoring. Även Lunds universitet bedriver forskning. DARPA och IBM har också publicerat.

Analys kopplat till området Robotik/Autonoma system

Resonemanget och röstningen var kortfattad. Robotik och autonoma system är en av de drivkrafter som vi funnit för NSR, där även andra begrepp används. Autonoma system är vanligt förekommande i studier av framtida teknik- och systemlösningar som Försvarsmakten arbetar med. Ofta studeras det autonoma systemet och dess tilltänkta funktion och funktionalitet i relation till andra autonoma eller bemannade system. Mer sällan studeras grundläggande tekniska lösningar, som AI-tekniken trots allt är för autonoma system. Vi anser att det är lämpligt att spana vidare i detta fält för NSR-tekniker och -tillämpningar.

Lära av expert/Simulering

Diskussionen av Lära av expert/Simulering blev mycket kort. Sakexperten som inledde såg artiklarna som högst relevanta. En av deltagarna belyste omvänd förstärkt inlärning.

Det finns ett tekniskt behov av att skapa regler baserat på observationer av människors beteenden i komplexa situationer. Det kan vara svårt att nå/skriva fungerande regler för sådana situationer. Exempelvis gäller det regler för hur en bil som kör in på en motorväg ska anpassa hastighet till andra bilar och förhållanden. Här kan förstärkt inlärning/lära av expert fylla en roll.

Analys kopplat till Lära av expert/Simulering

Vi uppfattar att det exempel som presenterades kanske missförstods av deltagaren som forskar inom närallgande område. Värderingen är låg men kan ha påverkats av en deltagares uttalande om att NSR-tekniken inte är relevant. Avskanningsprojektets saks experter tror dock att det blivit ett missförstånd och att själva funktionaliteten i de artiklar som presenterats inte nått fram. Vi kan därför inte lita på värderingsresultatet. Vi tror fortsatt att teknikerna/tillämpningarna är militärt relevanta, exempelvis möjligen för stridsflygssimulering.

Textanalys

Inom området textanalys bedriver FOI relativt mycket forskning. Det rådde osäkerhet om när i tid (om hur länge) de NSR-tekniker som presenterades skulle kunna tillämpas och hur långt teknikerna skulle nå.

Analys kopplat till området Textanalys

Området kan vara relevant för textanalys. FOI-experten bedöms kunna hantera frågan.

FOI är en huvudsakligen uppdragsfinansierad myndighet under Försvarsdepartementet. Kärnverksamheten är forskning, metod- och teknikutveckling till nytta för försvar och säkerhet. Organisationen har cirka 1000 anställda varav ungefär 800 är forskare. Detta gör organisationen till Sveriges största forskningsinstitut. FOI ger kunderna tillgång till ledande expertis inom ett stort antal tillämpningsområden såsom säkerhetspolitiska studier och analyser inom försvar och säkerhet, bedömning av olika typer av hot, system för ledning och hantering av kriser, skydd mot och hantering av farliga ämnen, IT-säkerhet och nya sensorers möjligheter.



FOI
Totalförsvarets forskningsinstitut
164 90 Stockholm

Tel: 08-55 50 30 00
Fax: 08-55 50 31 00

www.foi.se