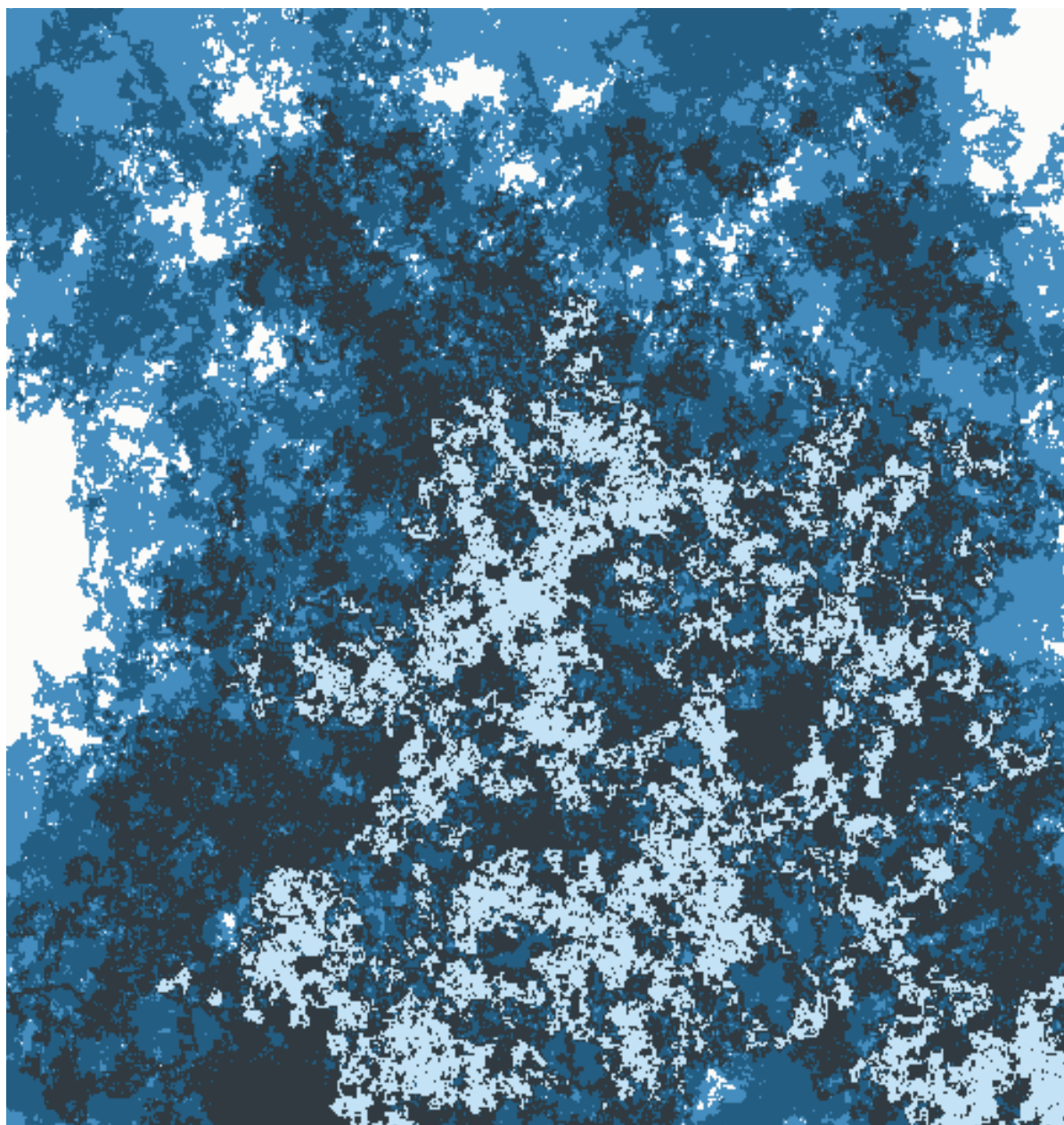


Forskningsöversikt: datakomprimering, informationsfusion och dataanalys

RONNIE JOHANSSON (RED.),
EDWARD TJÖRNHAMMAR (RED.),
JOHAN SCHUBERT, GUNNAR ERIKSSON



Ronnie Johansson (red.),
Edward Tjörnhammar (red.),
Johan Schubert, Gunnar Eriksson

Forskningsöversikt:
datakomprimering, informationsfusion
och dataanalys

Titel	Forskningsöversikt: datakomprimering, informationsfusion och dataanalys
Title	Research Overview: Data Compression, Information Fusion, and Data Analysis
Rapportnr/Report no	FOI-R--5237--SE
Månad/Month	Februari
Utgivningsår/Year	2022
Antal sidor/Pages	72
ISSN	1650-1942
Uppdragsgivare/Client	Försvarets materielverk
Forskningsområde	Ledningsteknologi
FoT-område	Inget FoT-område
Projektnr/Project no	E86218
Godkänd av/Approved by	Cecilia Dahlgren
Ansvarig avdelning	Försvarsteknik
Exportkontroll/Export control	Innehållet är granskat och omfattar ingen information som är underställd exportkontrolllagstiftningen

Bild/Cover: Skapad av Edward Tjörnhammar med aRtsy turmite (såddvärde 86218) och FOI:s palett

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

Den här rapporten omfattar en avskanning år 2021 av tre olika områden: kompression, informationsfusion och dataanalys. En uppskattning av den framtida utvecklingen av utvalda delområden och en jämförelse med avseende på delområdenas mognadsgrad och relevans görs.

Nyckelord

Avskanning, komprimering, dataanalys, informationsfusion

Abstract

This report covers a scan of three different areas conducted in 2021: compression, information fusion and data analysis. An estimate of the future development of the selected sub-areas and a comparison with regard to the sub-areas' degree of maturity and relevance are made.

Keywords

Survey, compression, data analytics, information fusion

Innehåll

1	Introduktion	7
1.1	Uppdragslydelse	7
1.2	Metod	7
1.2.1	Komprimering	7
1.2.2	Informationsfusion	7
1.2.3	Dataanalys	8
1.3	Läsanvisning	8
2	Komprimering	9
2.1	Inledning	9
2.1.1	Generell förlustfri datakomprimering	9
2.1.2	Bild- och videokomprimering	9
2.1.3	Ljudkomprimering	10
2.1.4	Kompressiv uppfattning	10
2.2	Forskningsläget	10
2.2.1	Bild- och videokomprimering	10
2.2.2	Ljudkomprimering	12
2.3	Framtidsprognos	14
3	Informationsfusion	15
3.1	Inledning	15
3.2	Forskningsläget	18
3.2.1	Analys med Web of Sciences (WoS)	18
3.2.2	Analys med HSTOOL	20
3.2.3	Manuell analys av publikationer	23
3.3	Framtidsprognos	30
4	Dataanalys	33
4.1	Inledning	33
4.1.1	Dataanalys	33
4.1.2	Storskalig dataanalys	33
4.1.3	Maskininlärning	34
4.1.4	System	34
4.2	Forskningsläget	35
4.2.1	Neurosymboliskt lärande och resonerande	35
4.2.2	Kvantdatorstödd dataanalys	38
4.2.3	Metainlärning, fåförsöksinlärning & överföringsinlärning	41
4.2.4	Uppmärksamhetsmodeller som förtränade modeller	42

4.2.5	Multimodala modeller	44
4.2.6	Modell- & datamängdförståelse	44
4.2.7	Datavisualisering	44
4.3	Framtidsprognos	45
5	Slutsatser	47
	Referenslista	51
	Appendix A : Extramaterial	67
A.1	Matematik för kvantdatorer	67
A.2	Mjukvaror	69
A.2.1	Neurosymboliskt lärande och resonering	69
A.2.2	Kvantdatorberäkningar	69

1 Introduktion

Rapporten redovisar resultatet av ett uppdrag för FMV som omfattar en avskanning av aktuell forskningslitteratur (och i viss utsträckning äldre utvald litteratur).

1.1 Uppdragslydelse

Rapporten täcker tre forskningsområden som omfattar det aktuella forskningsläget samt gör en uppskattning av den framtida utvecklingen.

Områdena som täcks av avskanningen är:

1. Komprimering och kapacitetshöjande tekniker för trådlös överföring av data.
2. Informationsfusion och presentation för beslutsfattare i behov av realtidsbeslut.
3. Dataanalys och presentation för analytiker för reguljär underlagsframställan.

1.2 Metod

Avskanningen av områdena utfördes parallellt i ett kompetensöverskridande samarbete på FOI där angreppssätten har varierat en aning. Generellt sett har meta-sökmotorn Web of Science¹ (WoS) nyttjats. Urvalet har bland annat omfattat ämnesspecifika tidsskrifter och översiktsartiklar. En sökfråga kan exempelvis se ut så här:

```
TS = (("X" OR "Y NEAR/2 Z")AND "Q")
```

där X avser ett nyckelord, exempelvis det aktuella området, och Y NEAR\2 Z ett alternativt nyckelord Y på två ordavstånd från ett följdnyckelord Z. Slutligen måste titeln även innehålla Q. Efter urval av artiklar, inläsning och sammanfattning följde en uppskattning av områdenas tillstånd och framtidsprognos.

1.2.1 Komprimering

Vad det gäller kompression så består området traditionellt av några dominerande delområden, som video- och talkompression. I projektet har dessa följts upp tillsammans med något yngre delområde. Huvuddelen av litteratursökningen har skett med hjälp av databasen IEEE Explore som har stor täckning av tidskrifter och konferenser inom området. Denna har även kompletterats med sökningar i Google Scholar. För en bredare sökning inom komprimeringsområdet kopplat till trådlös militär kommunikation har konferenserna MILCOM (Military Communications Conference) och ICMCIS (International Conference on Military Communications and Information Systems) studerats speciellt.

1.2.2 Informationsfusion

För informationsfusionsdelen har skanningsverktyget HSTOOL[96] använts för att hitta och klustra artiklar baserat på utvalda sökord. Flera av WoS:s analysverktyg användes, bland annat för att utreda ursprunget för de militär- och försvarsrelaterade artiklar som hittats.

¹<http://https://www.webofscience.com/> (besökt februari 2022)

1.2.3 Dataanalys

Vad det gäller dataanalys har, förutom WoS, även CORE:s portal², en förteckning över artiklar för topprankade ämneskonferenser använts. Ämnestabellen 1.1 är härledd från CORE:s listning från 2021, CORE2021, och visar för vilka delområden sökningar där har begränsats till. Ämnen har också valts ut baserat på författarnas egna uppfattningar om aktuella och angelägna utvecklingar. Därtill har vi sökt artiklar från *Computer Science* (CS) kategorin i arxiv och plockat lovande resultat, till stor del från *Machine Learning* (cs.LG) som återfunnits i CORE2021 mängden.

4602	Artificial intelligence
4603	Computer vision and multimedia computation
4605	Data management and data science
4608	Human-centered computing
4611	Machine learning

Tabell 1.1. Forskningsområdeskoder för sökningar i CORE

1.3 Läsanvisning

De tre avskanningsområdena behandlas i var sitt kapitel, komprimering i kapitel 2, informationsfusion i kapitel 3, och dataanalys i kapitel 4. Kapitlena för varje avskanningsområde består av en inledning (som erbjuder en kort bakgrund till området), en del om forskningsläget (som redovisar den avskannade litteraturen), samt en avslutande del som diskuterar den uppskattade framtida utvecklingen och relevans. Rapporten avslutas med en sammanfattning av kapitlens slutsatser (kapitel 5). Appendix A innehåller extramaterial för fördjupande läsning.

²Computing Research and Education Association of Australasia,
<https://www.core.edu.au/conference-portal>

2 Komprimering

2.1 Inledning

Med *komprimering* avser vi i det här kapitlet tekniker som minskar mängden data som behöver skickas över en förbindelse för att överföra viss information. I många sammanhang används också termen *källkodning* för i stort sett samma sak. Ur ett informationsteoretiskt perspektiv innebär källkodning att redundant information elimineras från en dataström så att mängden data som behöver överföras minskas. Det senare är mycket viktigt när informationen ska skickas över kanaler med begränsad kapacitet, vilket många gånger är fallet när informationen ska överföras via radio. Datakomprimeringen kan vara såväl förlustfri som icke-förlustfri. I det första fallet kan informationen återskapas perfekt, medan det i det andra fallet tillåts viss distortion av återskapade data. För vissa tillämpningar, till exempel ljud- och bildöverföring, är viss distortion av rekonstruerat data acceptabelt, medan det för andra tillämpningar, till exempel överföring av datafiler, inte går att tillåta någon distortion eftersom mottagaren måste ha en exakt kopia av originalet.

Icke-förlustfri datakomprimering kan för vissa tillämpningar, som exempelvis bild- och ljudkomprimering, ge stora komprimeringsvinster genom att utnyttja att människans perceptionsförmåga för olika former av distortion varierar. Komprimeringsmetoderna blir därför oftast mycket olika mellan olika tillämpningsområden och utgör oftast helt separata forskningsområden.

2.1.1 Generell förlustfri datakomprimering

I de fall ingen förvanskning av överförd data kan tolereras är komprimeringsalgoritmerna hänvisade till att försöka eliminera redundant information. Graden av möjlig komprimering varierar starkt mellan olika typer av meddelanden. Textmeddelanden i klartext ger normalt en hög komprimeringsgrad på grund av att normalt språk har en hög grad av redundans. Redundansen kommer sig av att bokstäver bara kan kombineras på ett begränsat sätt för att ge en meningsfull text. Om texten däremot är krypterad ska den teoretiskt inte gå att skilja från slumpmässigt genererade tecken och inget av redundansen i det ursprungliga meddelandet finns kvar att eliminera. Vid krypterad överföring måste därför meddelandet komprimeras före kryptering.

2.1.2 Bild- och videokomprimering

Bild- och videokomprimering är nästan alltid av icke-förlustfri karaktär eftersom det i de flesta tillämpningar krävs höga komprimeringsgrader. Det finns dock även förlustfria algoritmer som genom att utnyttja typiska strukturer i bilder kan ge högre kompressionsgrader än algoritmerna för generell förlustfri datakomprimering.

Den dominerande tekniken vid bildkompression är transformkodning. Vid komprimering av stillbilder utnyttjas nästan alltid bildens tvådimensionella karaktär. Den vanligaste metoden är att projicera en dimension i taget på en uppsättning basfunktioner, till exempel med hjälp av en diskret Fouriertransform eller diskret cosinustransform. Genom att många komponenter i den därigenom erhållna spektralrepresentationen av bilden i de flesta fall är relativt svaga och kan försummas kan en komprimerad representation av bilden erhållas.

För rörliga bilder (video), vilka består av en serie stillbilder, kan varje enskild bild komprimeras som en enskild stillbild. Vid komprimering av rörliga

bilder kan dock korrelationen mellan flera på varandra följande bilder utnyttjas för att uppnå en betydligt högre komprimeringsgrad än för en serie stillbilder.

2.1.3 Ljudkomprimering

Metoderna för ljudkomprimering är, liksom metoderna för bildkomprimering, nästan alltid av icke-förlustfri karaktär, men det finns även speciella förlustfria algoritmer för ljudkomprimering. Metoder för komprimering av ljud spänner över ett mycket stort område när det gäller vilket frekvensområde och vilken ljudkvalitet som ska representeras; från en mycket god återgivning av musik till återgivning av enbart tal med olika hög kvalitet. Komprimeringsalgoritmer för högkvalitativ återgivning är i allmänhet baserade på transformkodning, där ljudet delas upp i olika spektralkomponenter och där nivåerna för de olika komponenterna sedan kvantiserar. För lägre kvalitetsnivåer på ljudåtergivning, ungefär motsvarande trådbunden telefonikvalité och lägre, utnyttjas oftast modellbaserade tekniker som baserar sig på hur ljudet genereras i människans talorgan. När det gäller den senare kategorin av komprimeringsmetoder använder man i litteraturen oftast begreppet *talkodning* (eng. *speech coding*). Vi behandlar i rapporten i huvudsak komprimering av tal och har speciellt tittat på talkomprimering med hög komprimeringsgrad för militära tillämpningar.

2.1.4 Kompressiv uppfattning

Kompressiv uppfattning (eng. *compressive sensing* eller *compressive sampling*) är en teknik som medger att en signal kan samplas glesare än vad som ges av det vanliga samplingsteoremet och ändå rekonstrueras [32, 135]. En förutsättning för att det ska fungera är att signalen som ska representeras har en gles representation i någon transformdomän. Signalen behöver dock inte överföras till den glesa transformrepresentationen vid källan (sensorn), utan det räcker med att bilda linjärkombinationer av signalen. Antalet linjärkombinationer (sampler) som krävs är beroende på signalens gleshet. Signalen rekonstrueras sedan från samplen med hjälp av någon form av optimeringsalgoritm. Rekonstruktionen kan för vissa typer av optimeringsalgoritmer vara en mycket beräkningstung del.

Kompressiv sensing har snabbt fått en stor spridning inom många områden av signalbehandling. Förutom kompressionsområdet kan nämnas signalklassificering och signaldetektion. I den här studien har vi enbart beaktat tillämpningar av kompressiv uppfattning inom bild- och ljudkomprimeringsområdet.

2.2 Forskningsläget

2.2.1 Bild- och videokomprimering

Framstegen inom forskningen kring bildkomprimering, framför allt för rörliga bilder (video), har i allmänhet snabbt fångats upp och kommit in i olika internationella standarder för audiovisuella tjänster, bland annat inom ramen för Internationella teleunionen (ITU) som är ett organ inom Förenta nationerna (FN). Den senaste i raden av standarder för videokomprimering är H.265/HEVC (eng. high efficiency video coding) [168]. Den första versionen av H.265-standard utkom 2013 [86]. Standarden har därefter uppdaterats regelbundet (ungefär årligen) och den senaste versionen (V8) utkom i augusti 2021 [87]. Föregångaren till H.265 är H.264/AVC (eng. advanced video coding), vars första utgåva kom ut 2003 [84]. H.264-standard är fortfarande den mest använda för distribution av television, både via marksändningar och via ström-

ningstjänster över internet. Standarden för H.264 har även den uppdaterats regelbundet sedan den första utgåvan. Den senaste versionen (V14) kom också den ut i augusti 2021 [85]. Tilläggen under de senare åren har dock varit ganska begränsade.

Båda dessa standarder omfattar videokomprimering för ett stort spann av bildupplösningar och komprimeringsgrader. Kraven på kodare och avkodare ges i form av olika profiler, som i sin tur har ett antal nivåer. En profil anger vilken delmängd av algoritmer som en kodare och avkodare måste stödja, medan nivåerna sätter gränser på vilka värden som måste accepteras för de olika parametrarna som kontrollerar algoritmerna. För en given profil motsvarar de olika nivåerna, i allmänhet, olika krav på en avkodares beräknings- och minneskapacitet.

Komprimeringsalgoritmerna i H.265 kan i allmänhet ge en högre komprimeringsgrad för en given bildkvalité än algoritmerna i den äldre standarden H.264. Priset för detta är dock oftast en betydligt högre komplexitet hos kodare och avkodare. När det gäller beräkningskomplexitet har avsikten varit att hålla den på en relativt måttlig nivå för avkodaren, så att även högupplöst video ska kunna avkodas på relativt enkla enheter. Beräkningskomplexiteten i kodaren är dock avsevärt mycket större för H.265 än för H.264, vilket kan leda till svårigheter att göra realtidskodning av högupplösande video med hög komprimeringsgrad om inte avsevärda beräkningsresurser finns tillgängliga. En jämförelse mellan komprimeringsgraden för H.265 och H.264 för mycket högupplöst video (3840×2160 och 4096×2048 bildpunkter) visar att H.265 ger ca 2–3 gånger högre komprimeringsgrad vid höga kvalitetsnivåer [183].

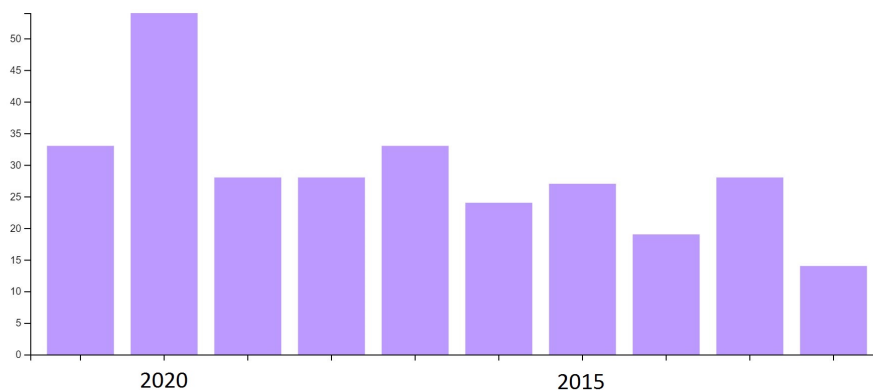
En undersökning av hur H.265 skulle kunna ersätta äldre videokomprimeringsalgoritmer i militära taktiska nät har studerats i [133] för olika användarfall. Författarna har genomfört en mappning av standardiserade videokvalitetsnivåer inom Nato till motsvarande profiler och nivåer inom H.265-standard. För de undersökta kvalitetsnivåerna erhöles i genomsnitt en reducering av bandbredden med 11% i förhållande till H.264. Författarna rapporterar också en ökad beräkningskomplexitet för vissa av kvalitetsnivåerna.

Även om de moderna videokomprimeringsstandarderna är mycket flexibla och konfigureringsbara finns det användarfall där andra lösningar kan vara att föredra. Det kanske framförallt gäller fall där en lägre kodningskomplexitet än vad standarderna medger är önskvärd för en viss komprimeringsnivå. Ett sådant användarfall i form av videoöverföring från en obemannad flygmaskin (UAV) diskuteras i [19]. Den föreslagna lösningen bygger på att rörelsen i videosekvensen är ett resultat av UAV:ns rörelse som är känd. Komprimeringsalgoritmerna i standarderna har en blockbaserad rörelseestimering för att effektivt kunna hantera komplicerade rörelser i bilden, vilket är överflödigt för detta användarfall.

Under de senaste femton åren har metoder baserade på kompressiv uppfattning applicerats inom bild- och videokomprimeringsområdet. Intresset har främst varit inriktat på tillämpningar som trådlösa sensornätverk och mindre UAV:er, där en låg beräkningskomplexitet och energiförbrukning är nödvändig på kodarsidan medan komplexiteten på avkodningssidan är mindre kritisk. Exempel på arbeten inom området är [16, 130, 152, 180, 182]. Trenden inom videokomprimering för tillfället är att applicera olika former av maskininlärningsmetoder på komprimeringsproblemet, se t. ex. [21, 40, 89, 121]. Begreppet maskininläring beskrivs närmare i avsnitt 4.1.3 i denna rapport. Maskininlärningsmetoder har även tillämpats på avkodarsidan för videokomprimeringsmetoder baserade på kompressiv uppfattning. I [198] beskrivs en sådan hybridmetod där den tänkta tillämpningen är en liten UAV för spaning.

En sökning efter publicerade artiklar med Web of Science innehållandes

orden “image coding” och “compressive sensing” i titel, sammanfattning eller nyckelord gav 288 träffar och den årliga fördelningen (2012–2021) framgår av figur 2.1.¹



Figur 2.1. Staplarna visar antal publicerade artiklar, år för år mellan 2012 och 2021, som nämner “image coding” och “compressive sensing” i titel, sammanfattning eller nyckelord.

2.2.2 Ljudkomprimering

Ljudkomprimering, eller ljudkodning som det ofta också kallas, spänner, som vi noterade i avsnitt 2.1.3, över ett brett kvalitetsregister och komprimeringsteknikerna för de olika kvalitetsnivåerna är i många fall olika. Det allmänna begreppet *ljudkomprimering* används oftast för de högre kvalitetsnivåerna för komprimering av till exempel musik, medan begreppen *talkomprimering* och *talkodning* oftast används i sammanhang där ljudet enbart består av tal. En speciell nisch inom talkomprimeringsområdet är *lågtaletalkomprimering* (eng. *low bit rate speech coding*). Kvalitetsnivån på det återskapade talet är i allmänhet ganska låg för dessa talkodare. Tillämpningsområdet för den senare typen av talkodare är i första hand krypterad militär trådlös kommunikation över smalbandiga radiokanaler.

Forskningen inom ljudkomprimering får anses ha startat inom delområdet lågtaletalkodningen vid tiden strax före andra världskriget. En tidig användning av talkodare för överföring av krypterat tal över kortvågsradio var systemet SIGSALY² [23]. De första stationerna installerades 1943 för säker talkommunikation mellan Washington och London. Vid den tiden skedde all annan överföring av ljud i analog form och det kom att dröja ganska länge innan ljudkomprimering fick ett mer allmänt intresse utanför det militära området.

Forskningen inom ljudkomprimeringsområdet för högre kvalitetsnivåer har i stor utsträckning varit ett parallellt spår till forskningen inom videokomprimeringen för distribution av audiovisuella tjänster. För området talkodning

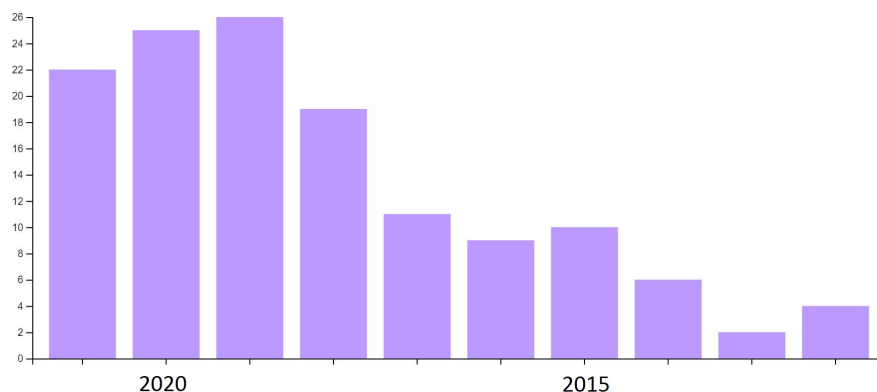
¹Fullständig sökförfråga i WoS: TS=(((image OR video)NEAR/0 coding)OR ((image OR video)NEAR/0 compression))AND (("compressive sensing")OR ("compressive sampling"))AND (SU = (Computer Science)OR SU = (Mathematics)OR SU = (Automation & Control Systems)OR SU = (Engineering)OR SU = (Remote Sensing)OR SU = (Robotics)OR SU = (Science & Technology Other Topics)OR SU = (Telecommunications))

²<https://www.cryptomuseum.com/crypto/usa/sigsaly/index.htm> (besökt februari 2022)

har mycket forskning skett inom ramen för olika standarder för mobiltelefoni. Inom området lågtaktstalkodning för militära tillämpningar gjordes ganska stora framsteg i slutet på 1990- och början av 2000-talet [35, 46, 77, 102, 128, 150], vilket ledde fram till en ny natostandard [134]. Forskningsaktiviteten inom lågtaktstalkodning för civila tillämpningar under motsvarande tidsperiod förefaller ha varit relativt låg, men bland relevanta arbeten kan nämnas [65, 148, 154]. Talkodaren som presenteras i [65] ger en datahastighet på 450 b/s och är en vidareutveckling av de tekniker som presenterats i [154].

Även inom talkodningsområdet förekommer det viss forskning kring metoder baserade på kompressiv uppfattning, men aktiviteten är mycket låg. Som exempel kan nämnas [72, 137, 158]. Resultaten visar inte på några större genombrott i förhållande till klassiska komprimeringsmetoder. Inga arbeten inom lågtaktstalkodning har hittats när det gäller kompressiv uppfattning. Under de senaste åren kan vi notera ett starkt ökat intresse för metoder baserade på maskininlärning inom talkodningsområdet, där [34, 80, 93, 192] utgör några exempel på arbeten inom området. Inom delområdet lågtaktstalkodning kan vi som exempel nämna [33, 68, 99, 119]. I [68] presenteras en jämförelse av talkvalité för en maskininlärningsbaserad talkodare och ett antal traditionella talkodare. Resultaten från jämförelsen visar att den maskininlärningsbaserade talkodaren gav en signifikant högre talkvalité vid 1,6 kb/s än vad de traditionella talkodarna gav vid 2,4 kb/s, och att den maskininlärningsbaserade talkodaren vid 800 b/s fortfarande gav en högre talkvalité än de traditionella talkodarna vid 2,4 kb/s.

En sökning efter publicerade artiklar med Web of Science innehållandes orden "speech coding" och "machine learning" i titel, sammanfattning eller nyckelord gav 134 träffar och den årliga fördelningen (2012–2021) framgår av figur 2.2.³



Figur 2.2. Staplarna visar antal publicerade artiklar, år för år mellan 2012 och 2021, som nämner "speech coding" och "machine learning" i titel, sammanfattning eller nyckelord.

³Fullständig sökning i WoS: TS=(((('speech coding')OR ('speech compression'))AND ("machine learning" OR "variational autoencoder"))AND (SU = (Computer Science)OR SU = (Mathematics)OR SU = (Automation & Control Systems)OR SU = (Engineering)OR SU = (Remote Sensing)OR SU = (Robotics)OR SU = (Science & Technology Other Topics)OR SU = (Telecommunications))

2.3 Framtidsprognos

- Forskningen inom bild- och ljudoptimering under den kommande tioårsperioden bedöms även fortsättningsvis vara driven av behoven inom de audiovisuella strömmningstjänsterna via internet och den fortsatt förväntade ökningen av trådlös anslutning via kommande mobiltelefonistandarder.
- Forskningen inom kompressionsområdet generellt kommer troligen att följa den pågående trenden och i stor utsträckning kretsa kring tillämpning av olika maskininlärningsmetoder, men också till viss del kring metoder baserade på kompressiv uppfattning.
- Det är troligt att vi i kommande standarder för audiovisuell komprimering kommer att finna algoritmer baserade på maskininläring och kompressiv uppfattning. Komprimeringsalgoritmerna baserade på kompressiv uppfattning kommer då att vara inriktade på tillämpningar med krav på låg beräkningskomplexitet hos kodaren, men där det finns möjlighet att flytta över komplexiteten till avkodaren.
- De flesta militära audiovisuella tillämpningar för mera högkvalitativ återgivning kommer troligen att även i fortsättningen baseras på civila standardiserade komprimeringsmetoder. Troligen kan ingen större forskningsaktivitet med specifik inriktning på militära system förväntas för dessa tillämpningar.
- Inom forskningen kring talkodare för låg dataakt (≤ 2400 b/s) är det rimligt att anta att vi kommer att se arbeten baserade på maskininlärningsmetoder som leder till nya militära standarder inom detta område.

3 Informationsfusion

I detta kapitel undersöker vi state-of-the-art inom forskningsområdet *information fusion*¹ (eng. *information fusion*) med speciellt fokus på dess militära användning. Kopplingar till *artificiell intelligens* (eng. *artificial intelligence* - AI) mellan informationsfusion och maskininlärningsområdet² (eng. *machine learning*), djupinläring³ (eng. *deep learning*) [74], samt mellan AI och metodiker för visualisering, undersöks också.

Informationsfusion⁴ kan beskrivas som nyttjandet av information (eller data) från flera källor (eller en källa vid flera tidpunkter) för att skapa en bättre förståelse för den situation som källorna reflekterar (observerar). Denna förståelse kan exempelvis uttryckas i form av en lägesbild som blir mindre osäker eller mer innehållsrik. En vidare diskussion om definitionen av informationsfusion finns i [25]. En samling av algoritmer och idéer finns i [111].

Informationsfusion är ett forskningsområde med dubbla rötter. Det härstammar dels från området *målföljning* (eng. *target tracking*) [101], speciellt ifrån området målföljning med radar, och dels från det område inom AI som benämns osäker information (eng. *management of uncertainty*) med metoder som Bayesiansk härledning,⁵ Dempster-Shaferteori⁶ (eng. *belief functions*) och oskarp logik⁷ (eng. *fuzzy logic*).

3.1 Inledning

Låt oss börja med en bred avskanning av fusionsområdet innan vi fokuserar på militära tillämpningar. Antalet forskningspublikationer inom fusionsområdet har vuxit kraftigt under de senaste decennierna. Mot Web of Science⁸ (WoS) ställer vi följande sökfråga⁹:

```
TS = ("information fusion" OR "data fusion" OR "sensor fusion")
```

och finner 36 610 publikationer mellan åren 1997–2021. Notera att vi använder söktermen "information fusion" OR "data fusion" OR "sensor fusion" då dessa i viss mån används synonymt av en del författare. I figur 3.1 visas hur dessa artiklar fördelar sig över åren.

Vi noterar i figur 3.1 en ökning av forskningslitteraturens omfattning från slutet av 1990-talet till idag om cirka tio gånger, från drygt 300 publikationer per år till drygt 3 000 per år.

Ofta används dock termerna i sökfrågan något olika, där datafusion och särskilt sensorfusion beskriver forskningen som rör fusion från sensorer, medan informationsfusion beskriver fusionsmetodiker högre upp i beslutskedjan i ett ledningssystem som därmed typiskt ligger närmare beslutsfattaren. I figur 3.2 visar resultatet av en sökning i WoS med sökfrågan

¹<https://isif.org/> (besökt februari 2022) ger en översikt över aktiviteter inom fusionsområdet.

²https://en.wikipedia.org/wiki/Machine_learning (besökt februari 2022). Se även avsnitt 4.1.3 i denna rapport.

³https://en.wikipedia.org/wiki/Deep_learning (besökt februari 2022).

⁴Relaterade termer som förekommer i litteraturen är data fusion och sensor data fusion.

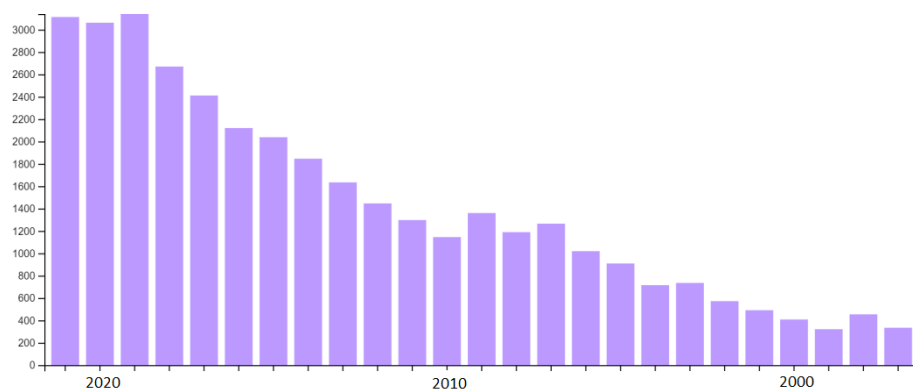
⁵https://en.wikipedia.org/wiki/Bayesian_statistics (besökt februari 2022).

⁶https://en.wikipedia.org/wiki/Dempster-Shafer_theory (besökt februari 2022).

⁷https://en.wikipedia.org/wiki/Fuzzy_logic (besökt februari 2022).

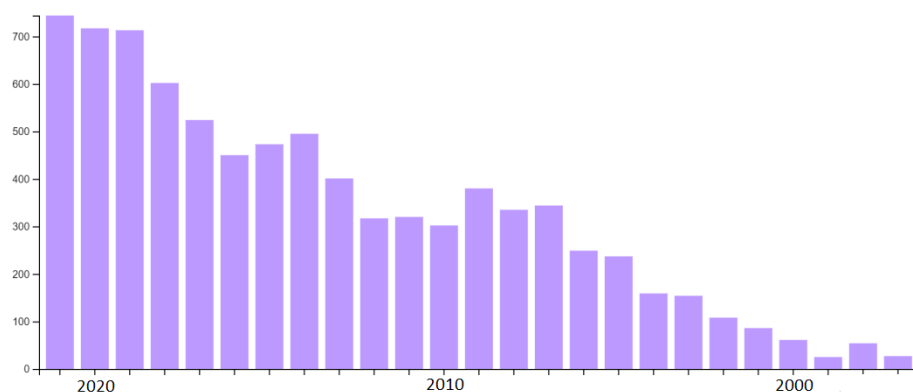
⁸<https://www.webofknowledge.com> (besökt februari 2022)

⁹TS är en fältetikett som betyder *topic*.



Figur 3.1. Antalet forskningspublikationer per år inom hela fusionsområdet 1997–2021.

TS = ("information fusion")



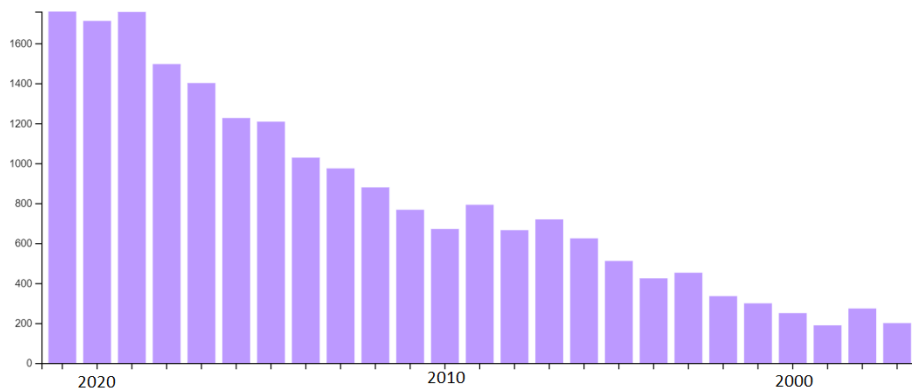
Figur 3.2. Antalet forskningspublikationer per år med sökning med söktermen "information fusion" 1997–2021.

Det totala antalet publikationer under perioden är 8 344. Mätningen uppvisar en kraftig uppgång i antalet publikationer per år under perioden med en faktor om cirka 25 gånger. Det gör informationsfusion till den snabbast stigande söktermen av de tre i den första sökfrågan. Detta speglar väl en utveckling inom fusionsforskningen som vi noterat sedan mitten av 2000-talet med en vidgning av forskningsområdet mot frågeställningar på en högre ledningsnivå närmare beslutsfattande.¹⁰

Datafusion som begrepp är bredare än både sensorfusion och informationsfusion och kan omfatta problemställningar högre upp i informationskedjan än sensorfusion. Datafusion som begrepp har en längre tradition än informationsfusion och då den är bredare i sitt omfång ger den fler träffar vid sökning i WoS. I figur 3.3 presenterar vi resultatet från en sökning i WoS med sökfrågan

TS = ("data fusion").

¹⁰FOI anordnade 2004 den sjunde internationella informationsfusionskonferensen med ett uttalat mål om bredda området från sin traditionella bas i sensor- och datafusion mot de forskningsfrågor som ligger högre upp i beslutskedjan från sensor till beslutsfattare, se <https://isif.org/events/conference/fusion-2004> (besökt februari 2022).



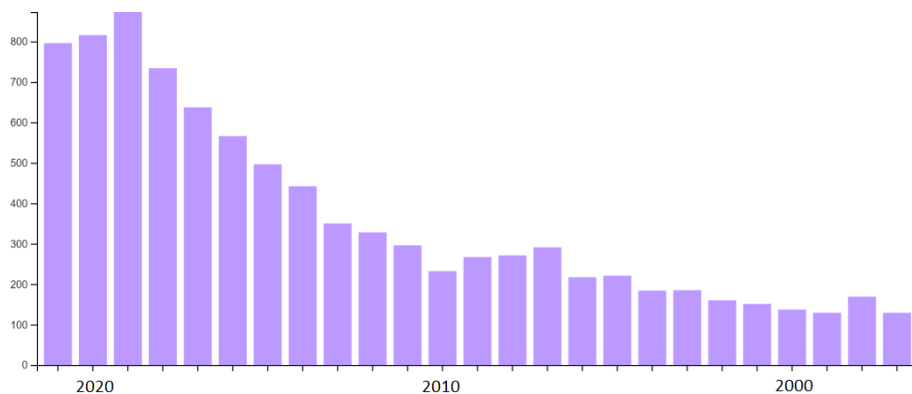
Figur 3.3. Antalet forskningspublikationer per år med sökning med söktermen "data fusion" 1997–2021.

Vi finner här 21 151 forskningspublikationer under perioden från 1997 fram till 2021. Området visar på en niofaldig uppgång sedan början på perioden till nu.

Avslutningsvis undersöker vi den tredje termen i den första sökfrågan. Vi söker i WoS med

```
TS=("sensor fusion").
```

Antalet funna publikationer är nästan det samma som för informationsfusion. Vi har 9 413 träffar med denna sökfråga. Ökningen över perioden är dock lägre med en ökning av antalet publikationer under perioden om cirka sju gånger, figur 3.4.



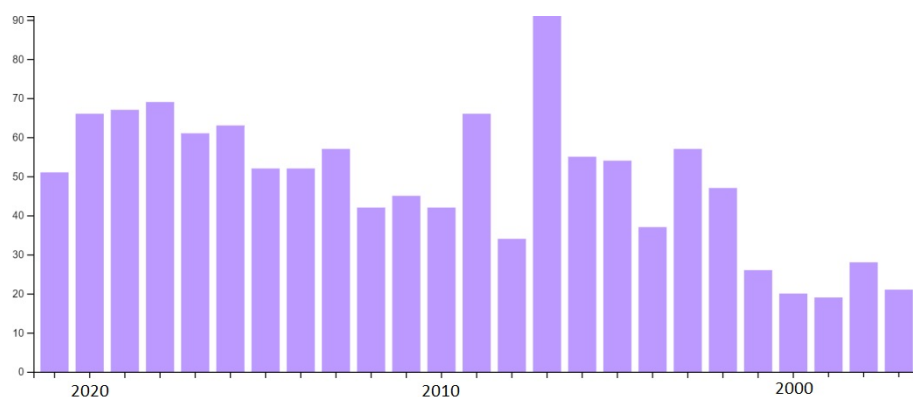
Figur 3.4. Antalet forskningspublikationer per år med sökning med söktermen "sensor fusion" 1997–2021.

Det sammanlagda antalet träffar av de senaste tre sökningarna är 38 908, vilket demonstrerar ett oväntat lågt överlapp då alla publikationerna från de tre senaste sökfrågorna (figur 3.2–3.4) är dragna från den mängd om 36 610 träffar som erhöles med den första sökfrågan. Det låga överlappet stärker analysen i figur 3.2–3.4.

Låt oss nu fokusera på militära tillämpningar. Vi utgår från den första sökfrågan och kompletterar den med några militära termer. Vi söker i WoS med

```
TS = ("information fusion" OR "data fusion" OR "sensor fusion")AND (military OR defence OR defense OR "decision support"OR "command and control"))
```

och finner 1 247 publikationer under hela perioden 1997–2021. Vi noterar en ökning under perioden med en faktor om cirka tre, från runt 20 publikationer per år till drygt 60 publikationer per år,¹¹ figur 3.5.



Figur 3.5. Antalet militära forskningspublikationer per år inom fusionsområdet 1997–2021.

Flera av publikationerna från den första sökfrågan med vetenskapligt eller tekniskt fokus kan dock ofta ha militära tillämpningar utan att det omnämns i titeln eller sammanfattningen som sökningen riktas mot. I den efterföljande analysen i nästa sektion fokuserar vi på de publikationer i den sista sökfrågan som är från de tre senaste åren 2019–2021. Vi söker i WoS med sökfrågan

```
TS = ("information fusion" OR "data fusion" OR "sensor
fusion")AND (military OR defence OR defense OR "decision
support" OR "command and control"))AND PY = (2019-2021)
```

där PY är en fältterm som betyder *publication year*. Antalet publikationer under dessa tre år är 184.

3.2 Forskningsläget

Innan vi börjar analysera forskningsinriktningarna i de 184 publikationerna om militär fusion från de senaste tre åren, låt oss med WoS och HSTOOL[95, 96, 162] analysera sökresultatet vidare.

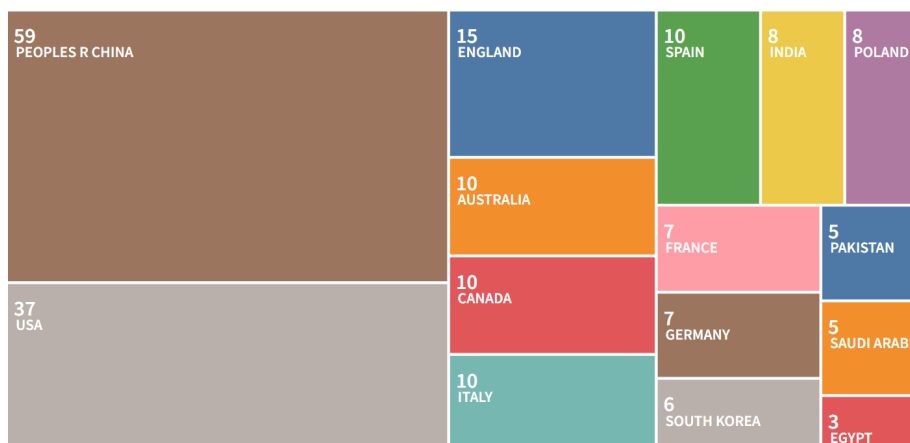
3.2.1 Analys med Web of Sciences (WoS)

Med hjälp av WoS kan vi ta fram samlad statistik om vilka forskningsområden som dominerar, var forskningen är producerad, vilka som har finansierat forskningen etc. I figur 3.6 noterar vi att Kina dominerar inom den militära fusionsforskningen under de senaste tre åren, men att Nordamerika och Europa tillsammans har en nästan dubbelt så stor volym.

I figur 3.7 visas forskarnas organisatoriska tillhörighet. Vi noterar en blandning av flera försvarsforskningsinstitut och några universitet. Amerikanska försvarsmakten, kanadensiska DRDC,¹² Nato och några kinesiska universitet finns bland de mest produktiva utförarna av forskningen i de 184 publikationerna. Ingen svensk organisation finns bland de 15 främsta organisationerna bland

¹¹Den nedgång som syns för 2021 beror sannolikt på en eftersläpning i WoS då sökningen genomfördes under januari månad 2022.

¹²Defence Research and Development Canada



Figur 3.6. Forskningens ursprungsländer för militära forskningspublikationer inom fusionsområdet 2019–2021.

publikationerna under perioden 2019–2021. FOI hade tidigare en stark närvaro inom forskningsområdet men saknas i denna sökning eftersom forskningen lades om i riktning mot datorlingvistik och AI.



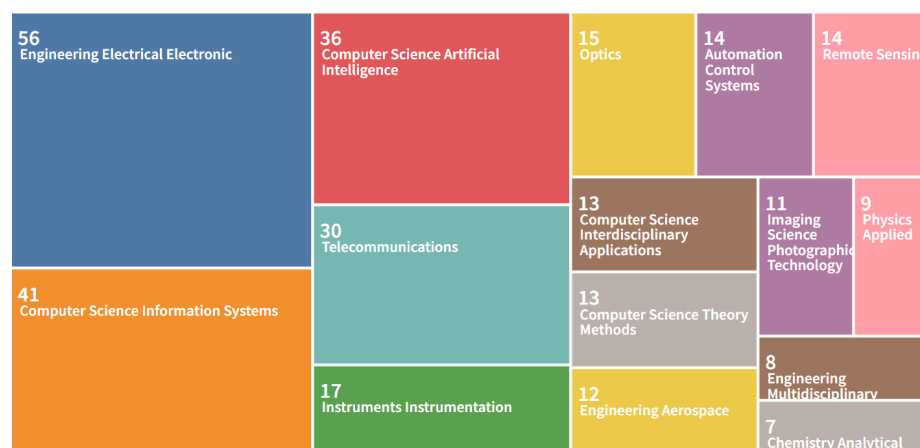
Figur 3.7. Forskarnas organisatoriska tillhörighet för militära forskningspublikationer inom fusionsområdet 2019–2021.

I figur 3.8 ser vi en översikt över de viktigaste externa forskningsfinansiärerna. Vi noterar ett antal kinesiska forskningsfinansiärer med det kinesiska naturvetenskapliga forskningsrådet främst, följt av den europeiska kommissionen. Den bristande överensstämmelse som tycks finnas mellan forskningsutförarna (figur 3.7) och finansiärerna (figur 3.8) beror sannolikt på att mycket forskning genomförs utan externt angivna finansiärer hos statliga forskningsinstitut.

Avslutningsvis gör vi en betraktelse över de 15 största forskningsområdena enligt den kategorisering som är gjord inom WoS (figur 3.9). Vi noterar att fusionsforskning i tillämpningar som klassas som elektroniktillämpningar dominerar före användning av fusion i informationssystem och därefter inom AI-tillämpningar.



Figur 3.8. Forskningsfinansiärer för militära forskningspublikationer inom fusionsområdet 2019–2021.



Figur 3.9. Tillämpningsområden för militära forskningspublikationer inom fusionsområdet 2019–2021.

3.2.2 Analys med HSTOOL

De 184 vetenskapliga publikationerna som erhöles med den sist nämnda sökningen i WoS har klustrats och rankats med FOI:s datorverktyg HSTOOL[95, 96]. Denna klustring bygger dels på samstämmighet mellan orden i alla titlar och sammanfattningar hos publikationerna inom ett och samma kluster och dels på skillnad mellan alla ords frekvens hos publikationerna i ett visst kluster gentemot alla andra publikationer som placerats i andra kluster.

Genom denna klustringen erhöles vi sex kluster vars innehåll beskrivs i tabell 3.1 utifrån de viktigaste orden i varje kluster.¹³

¹³De ingående klusternumren har ingen betydelse utan är enbart ett löpnummer för att särskilja klustren åt.

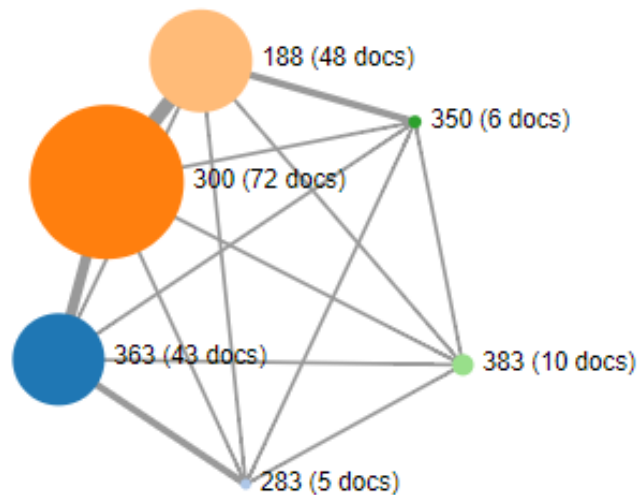
Tabell 3.1. Topp tio ämnen, nyckelord, beskrivande ord och urskiljande ord för de sex klustrena.

Cluster	Top Subjects	Top Keywords	Top Descriptive Words	Top Distinguishing Words
383	computer science, artificial intelligence, information systems, interdisciplinary applications, theory & methods	andness, information fusion, aggregation operators, banzhaf function, banzhaf value, characterization, choquet integral, copula, customer engagement, customer relationship management	information, decision, science, computer, fusion, method, operator, technology, model, review	text, banzhaf, rung, orthopair, rofs, fm, aggregator, preference, business, reasoning
283	engineering, instruments & instrumentation, telecommunications, aerospace, electrical & electronic, automation & control systems	sensor fusion, fuzzy logic system (fls), game theory, unmanned aircraft systems (uas), cyber-physical system, decision support systems (dss), decision support systems (dss), defence in depth, drone fleets, edge computing	system, fleet, detection, engineering, vehicle, space, mem, technology, design, mission	iae, assistant, fl, vhd1, intel, altera, openvino, gop, fleet, interface
350	geography, agriculture, agronomy, chemistry, engineering, forestry, instruments & instrumentation, meteorology & atmospheric sciences, plant sciences, public administration	data fusion, assemblage, athabasca river watershed, ceres-maize, crop model, danger visualisation, decision making, decision support system, ecology, geopolitics	data, science, yield, crop, irrigation, map, technology, area, palm, method	season, grain, scientist, irrigation, yield, palm, wireless, dempster, navy, simulator
363	engineering, instruments & instrumentation, electrical & electronic, optics, computer science, imaging science & photographic technology, physics, remote sensing, robotics, telecommunications	sensor fusion, data fusion, estimation, imu, sensors, additive manufacturing, autonomous flight, bundle adjustment, flight mechanics, image fusion	technology, system, sensor, fusion, science, data, detection, method, engineering, information	compensation, indoor, office, odometry, pipe, readout, in, pitot, tube, inspection
188	computer science, engineering, telecommunications, information systems, electrical & electronic, artificial intelligence, theory & methods, imaging science & photographic technology, multidisciplinary, automation & control systems	data fusion, machine learning, deep learning, sensors, edge computing, feature extraction, information fusion, security, artificial neural networks, cloud computing & science, data, technology, network, sensor, fusion, computer, system, attack, information	question, biometrics, authentication, prevalence, ad, wearable, revolution, cryptography, head, interconnection	
300	computer science, engineering, automation & control systems, electrical & electronic, information systems, artificial intelligence, remote sensing, optics, interdisciplinary applications, operations research & management science	data fusion, decision support, decision support system, information fusion, machine learning, artificial intelligence, big data, business analytics, decision support systems, situation awareness	system, data, science, technology, decision, fusion, information, support, computer, sensor	community, warfare, joint, analytics, stream, actor, lesson, provision, tbm, implication

Kolumnen *Top Descriptive Words* beskriver klustret med hjälp av de tio vanligast förekommande orden från titlar och sammanfattningar hos alla i klustret ingående publikationerna. Kolumnen *Top Distinguishing Words* listar de mest särskiljande orden mellan det aktuella klustret gentemot alla övriga kluster. Detta sker genom en beräkning av entropin hos en fördelning för varje ord över alla kluster. De två första kolumnerna innehåller de vanligaste klassificeringarna i WoS för publikationerna i respektive kluster samt motsvarande för nyckelorden hos publikationerna i respektive kluster.

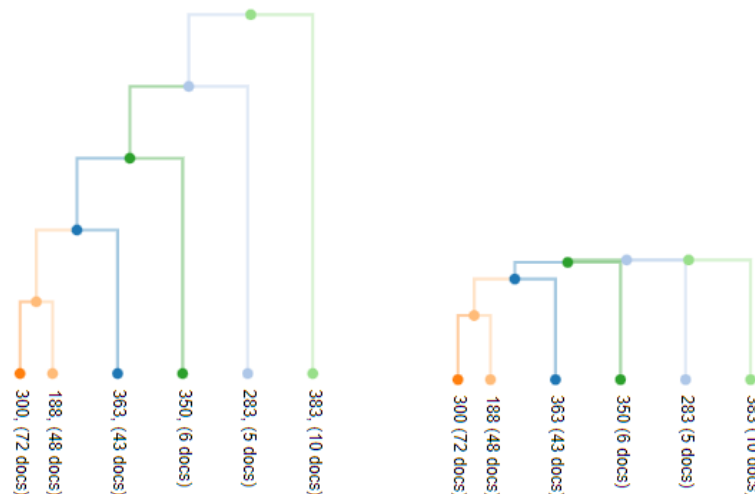
Vi kan därefter ytterligare analysera klustren med hjälp av några nytillkomna analysfunktioner i den senaste versionen av HSTOOL[162]. I figur 3.10 visas en klustergraf där varje klusters storlek är proportionell mot antalet pub-

likationer i klustret och tjockleken på bågarna i grafen mellan kluster står i proportion till ett beräknat avstånd mellan de två klustren baserat på beräkning av entropier för alla ingående ord [162].



Figur 3.10. En klustergraf över de sex funna klustren.

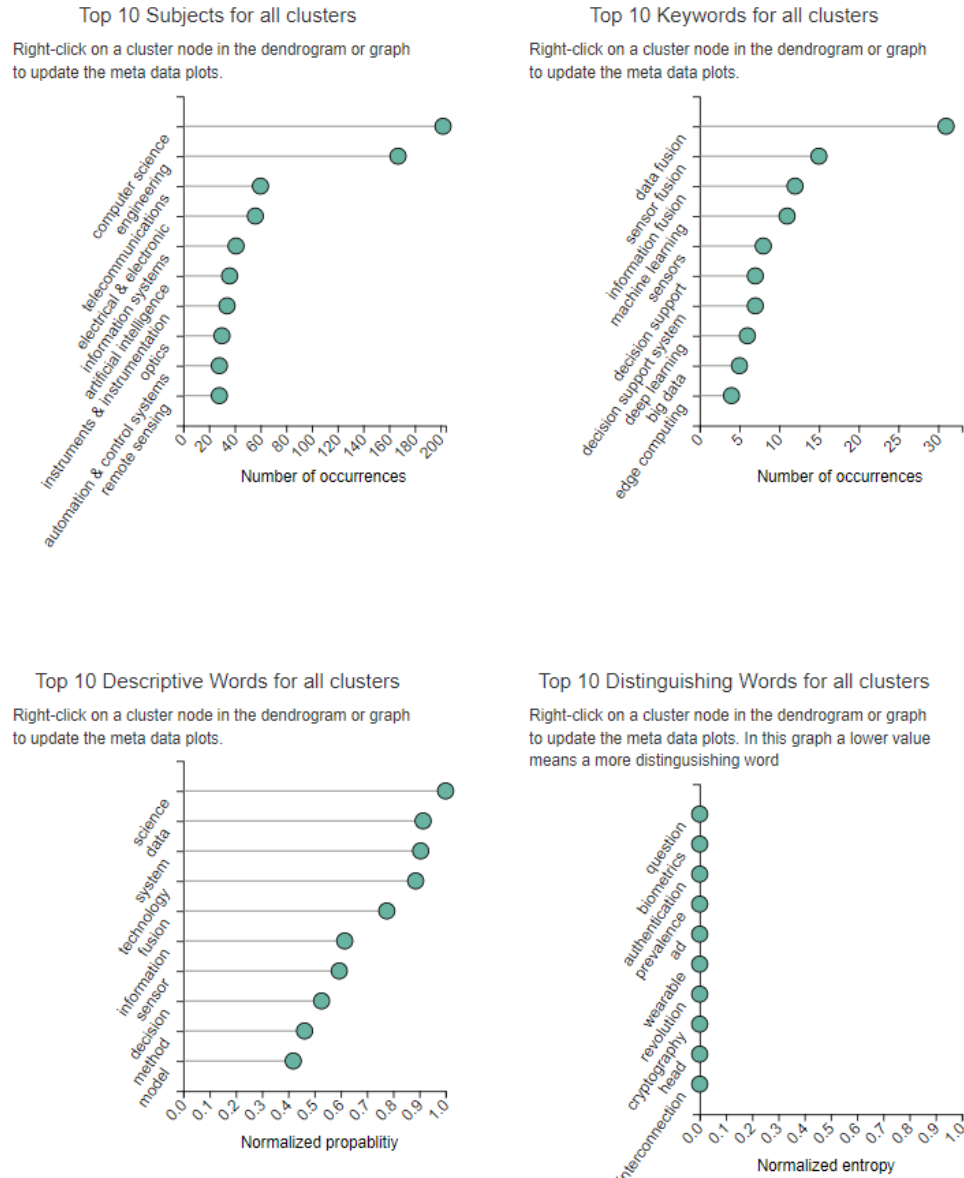
Ett kompletterande sätt att beskriva de sex klustrens inbördes förhållande är med hjälp av ett dendritträd. I figur 3.11 visas två dendritträd över de sex klustren. Till vänster i figuren visas vilka kluster som ligger närmast varandra enligt det beräknade entropiavståndet. Här ser vi att kluster 188 och 300 ligger närmast varandra och därefter kluster 363, osv. I den vänstra delen av figuren är grenarna i dendritträdet inte proportionella mot avståndet mellan klustren. I den högra delen visas samma dendritträd med proportionella grenar.



Figur 3.11. Dendritträd över de sex funna klustren.

Slutligen kan man med HSTOOL få en presentation av förekomsten av orden i de fyra kolumnerna i tabell 3.1. I figur 3.12 presenteras förekomsterna för ämnesområden, nyckelord, beskrivande ord och särskiljande ord för hela mängden

med de 184 publikationerna. Det går även bra att få dessa beskrivningar individuellt för var och en av de sex klustren.



Figur 3.12. Förekomst av ämnesområden och ord bland de 184 publikationerna.

3.2.3 Manuell analys av publikationer

Analysen av de 184 publikationerna sker i flera steg. Initialt använder vi citeringsstatistik som hämtas från WoS av HSTOOL. Statistiken presenteras i kalkylblad med en fil för varje kluster.

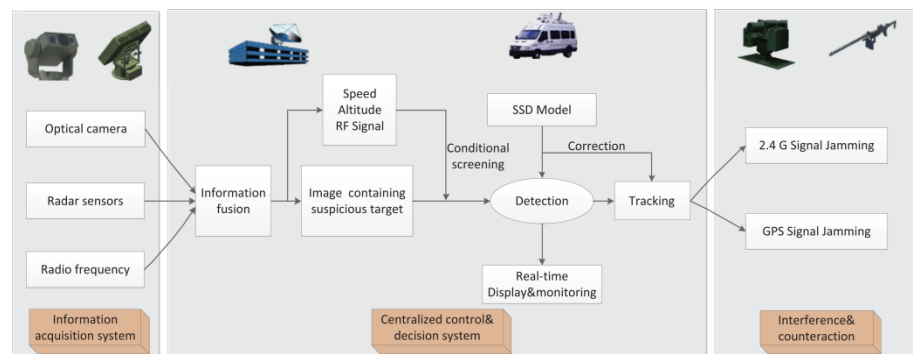
Eftersom antalet publikationer är relativt hanterbart gör vi endast ett urval av publikationerna och väljer dem som har minst en (1) citering. Därtill adderar vi ett fåtal publikationer som saknar citeringar men vars titel eller samfattning inkluderar termerna *visualisering* eller *deep learning*, då vi är intresserade av att se hur dessa områden kan kombineras med fusion. Denna selektion reducerade antalet publikationer till 75.

För de kvarvarande 75 publikationerna läser vi alla sammanfattningar och väljer att fortsätta med de publikationer som har någon militär relevans eller där vi finner något tekniskt speciellt av intresse. Detta ger oss 30 kvarvarande publikationer som studeras i sin helhet. Av dessa sammanfattar vi nedan under olika rubriker de 16 artiklar som är mest intressanta och militärt relevanta.

3.2.3.1 UAV-spaning

In en artikel av Thakoor m.fl. [176] studeras routingproblemet för flera obemannade flygfarkoster (UAV) där en grupp av okoordinerade UAV:er behöver samla in information över ett stort område. Författarna formulerar problemet som ett spelteoretiskt problem där UAV:er är spelare och deras strategier är de olika vägarna de kan flyga. Modellen inkluderar informationsfusion där hänsyn tas till om flera UAV:er samtidigt besöker ett delområde och den kvarvarande osäkerhet som finns kvar för nästa UAV efter att den förra avspanat området. På detta sätt styr fusionsberäkningarna den kvarvarande osäkerhet som finns kvar att minimera genom att spana på samma plats som en annan UAV redan spanat vid och om det är värt att göra. Artikeln påvisar möjligheten att kombinera informationsfusion i dynamisk uppdragsplanering vid spaning med flockar (eng. swarms) av UAV:er.

Chen m.fl. [39] adresserar å andra sidan problemet med att upptäcka UAV:er genom användandet av multisensorfusionsmetoder för detektering. De föreslår en arkitektur för ett intelligent anti-UAV försvarssystem för låga flyghöjder. För att lösa problemet med att undersöka misstänkta mål har de designat en fusionsmetod som kombinerar radar med en detektionsmodell för att identifiera UAV:er i bilder för att kunna upptäcka, särskilja och målfölja inkräktande UAV:er. Olika egenskaper hos olika typer av sensorer kombineras för att förbättra effekten för att upptäcka och detektera UAV:er. Ett övergripande schematiskt diagram visas i figur 3.13. Författarna har även konstruerat en experimentell plattform för att demonstrera metodens effektivitet. Resultaten påvisar god prestanda i räckvidd, noggrannhet och framgång för övervakning.

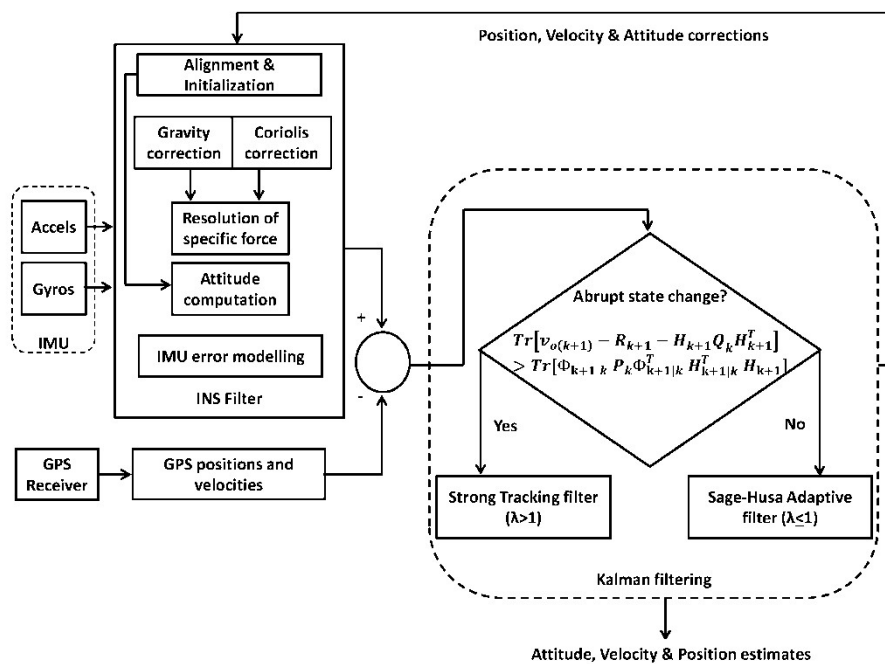


Figur 3.13. Övergripande arkitektur för systemet för skytt mot fientliga UAV:er [39] (Creative Commons Attribution 4.0 International License, <http://creativecommons.org/licenses/by/4.0/>).

3.2.3.2 Målföljning

Ahmed m.fl. [4] föreslår ett ramverk och en algoritm som kombinerar ett adaptivt Sage-Husa Kalmanfilter [156] med en stark målföljande Kalman-filtreringsalgoritm [69]. Metodiken är implementerad för målföljning av markfordon

baserad på GPS¹⁴ och tröghetsnavigering. Sage-Husas adaptiva Kalmanfilter är baserat på maximal *a priori* estimering. Eftersom gamla mätningar påverkar lägesestimeringen kan filtrets prestanda optimeras genom att använda ett antal av de senaste mätningarna. Å andra sidan, när man hanterar icke-linjära dynamiska modeller där tillståndsförändringar kan ske plötsligt kan ett starkt målföljande Kalmanfilter vara mer robust. Författarna presenterar ett hybridfilter som erbjuder balans mellan noggrannhet och robusthet i situationer med plötsliga förändringar i systemtillståndet. När tillståndsändringen inte är abrupta kommer filtret att prioritera uppskattning av Sage-Husa Kalmanalgoritmen, annars kommer uppskattningen att utföras genom det starka målföljande Kalmanfiltret. Figur 3.14 visar filterfunktionen. Den föreslagna algoritmen har i experiment genererat en överlägsen navigering för uppskattning av ett markfordons position och hastighet. Algoritmen erbjuder en god balans mellan noggrannhet och stark konvergens för optimala uppskattningar för dynamiska system. Samma hybridfilter kommer att producera ännu bättre navigationsskattningar för målspårning av flyg och UAV:er, där ett konventionellt Kalman-filter kommer att släpa efter det verkliga läget på grund av sitt beroende av a-priori information.



Figur 3.14. Ett hybridfilter baserat på Sage-Husa Kalmanfilter och starka målföljande Kalmanfilter [4] (Creative Commons Attribution 4.0 International License, <http://creativecommons.org/licenses/by/4.0/>).

I en artikel av Sherer m.fl. [164] används radar- och GPS-mätningar i ett fusionssystem med flera sensorer som tillsammans möjliggör robusta avståndsbedömningar genom att utnyttja noggrannheten hos en differentiell GPS (DGPS) med den högre uppdateringshastigheten hos en radar i ett Kalmanfilter. En fusionsalgoritm som kombinerar dynamisk realtidskinematik (DRTK) med ett probabilistiskt radardata-associationsfilter (PDAF) utvecklas av författarna, där DRTK- och PDAF-lösningarna är integrerade i ett enda Kalman-

¹⁴Global Positioning System, https://sv.wikipedia.org/wiki/Global_Positioning_System (besökt februari 2022)

filter. Målet är en GPS- och radarfusionsalgoritm som är både robust, exakt och kostnadseffektiv för relativ positionering mellan fordon för konvojkörning i realtid. Analys av algoritmen visar att multisensorkonfigurationen fungerar med högre noggrannhet och stabilitet än de fristående GPS- och PDFA-lösningarna.

I en tredje artikel med Tektonidis & Monnin [175] som författare ligger fokus på att upptäcka hot från improviserade sprängladdningar (IED) mot en militär fordonskonvoj. Författarna presenterar ett tillvägagångssätt för detektering av bildförändringar som uppskattar risker utmed vägar, med hjälp av ett kamerasystem. Kamerasystemet är en del av ett multisensorsystem monterat på det ledande fordonet i fordonskonvojen. Uppskattade sannolikheter för IED:er från varje sensor kombineras av en datafusionsalgoritm. För experimentell utvärdering har man använt bildsekvenser inhämtade från vägavsnitt som inkluderar olika typer av förändringar utmed vägen och med stora skillnader i belysning. Metoden har gett ett bra resultat.

3.2.3.3 Operatörsstöd

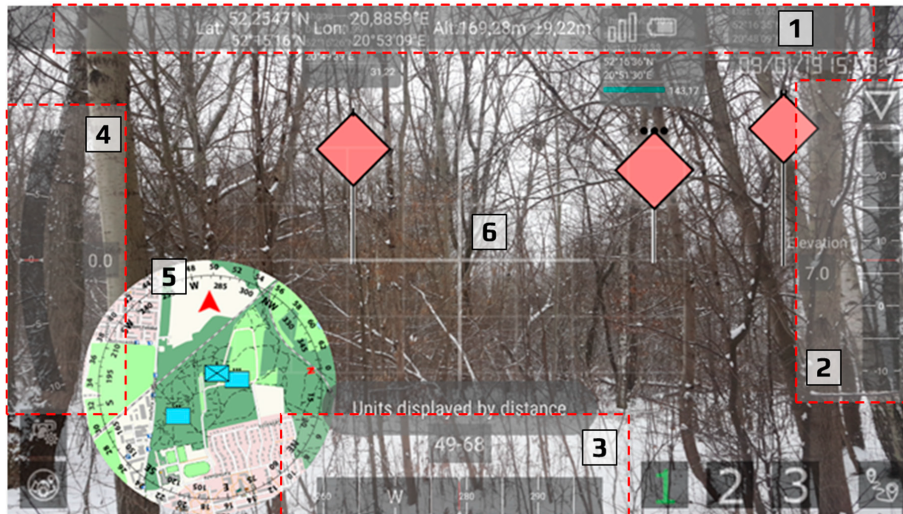
Huvudmonterade displayer kan vara användbara för syntetisk träning och förstärkning vid militärt beslutsfattande i ledningssystem. Användning av sådana displayer kan dock resultera i *åksjuka* och minskad arbetsprestation. Dennison m.fl. [54] använder maskininlärning på data som samlats in från heterogena, icke-invasiva sensorer som bärs av användare som navigerat i en virtuell miljö för att bedöma graden av åksjuka. För att göra detta, jämförde man den maximala noggrannheten för en uppsättning maskininlärningsmodeller. Särskilt jämförde de noggrannheten hos modeller som tränats på funktioner som härrör från enstaka sensorer med modeller som hade tillgång till funktioner från en uppsättning heterogena mått. De upptäckte att två alternativa modeller där en är tränad på heterogena sensordata och en är tränad endast på EEG¹⁵-data så kunde båda klassificera åksjukans svårighetsgrad med hög noggrannhet. Författarna noterar dock att modeller som tränats på funktioner från individuella sensorer visade mycket varierande klassificeringsnoggrannhet. Bäst prestanda uppnåddes från en bagging-modell¹⁶ bestående av 13 egenskaper härledda från EEG-data. De resultat som presenteras indikerar hur militären bäst kan utrusta den framtida soldaten med sensorer i en stridsmiljö.

Chmielewski m.fl. [41] vid den Polska militärhögskolan presenterar i en artikel utvecklingen av mobilapplikationer för taktiskt beslutsstöd som använder förstärkt verklighet (eng. augmented reality) vid produktion av kontextuell data för presentation i en smartphone. Författarna presenterar mobila stödverktyg som används för att stödja militär personal under strid och krishantering. Verktøyen ger nya möjligheter för visualisering och identifiering av militära hot, uppdragsplanering och uppgifter, vilket hjälper användarna att överblicka en aktuell situation. Den utvecklade mjukvaruplattformen kallad *mobil gemensam operativ bild* (mCOP) erbjuder ett individuellt distribuerat mobilsystem för enskilda soldater och befälhavare fokuserat på strid och stöder *blue-force*-följning, fusion av spaningsdata samt utvärderingar av hotnivå för militära- och krishanteringsscenarier. Författarna framhåller att information som levereras i rätt tid och medger väl förberedda beslut är avgörande för att uppnå beslutsöverlägsenhet. För att uppnå den nödvändiga situationsmedvetenheten måste information om stridsområdet vara tillgänglig i realtid. Därför har mCOP utvecklats till att återge taktisk information med avseende på varje användares egen geografiska plats med hjälp av förstärkt verklighet av den omgivande stridsmiljön, figur 3.15. Alla inhämtade data fusioneras och används för att

¹⁵<https://sv.wikipedia.org/wiki/Elektroencefalografi> (besökt februari 2022)

¹⁶https://en.wikipedia.org/wiki/Bootstrap_aggregating (besökt februari 2022)

uppskatta de viktigaste egenskaperna hos stridssituationen.



Figur 3.15. Taktisk vy av objekt i stridsrummet presenterad med förstärkt verklighet [41] (Creative Commons Attribution 4.0 International License, <http://creativecommons.org/licenses/by/4.0/>).

3.2.3.4 Beslutsstöd

Multi-sensor informationsfusion (IF) kan förbättra noggrannheten i modellering genom att kombinera information från flera sensorer. För att ge ett heltäckande beslutsstöd bör både kvantitativ data och experters domänkunskap inkluderas i fusionsprocessen. I många praktiska tillämpningar finns det dock otillräckligt med data vilket gör att konventionella metoder inte kan användas. En tilltros-regelbas (eng. belief rule base - BRB) [193] har fördelar i icke-linjär modellering med otillräckliga data och experters domänkunskap. BRB är sammansatt av flera regler som finns i samma struktur. Zhou m.fl.¹⁷ [202] föreslår en kombinerad BRB-IF-modell för multisensorfusions där både kvantitativ multisensordata och kvalitativa experters domänkunskap effektivt kan inkluderas till ett enhetligt resultat för beslutsfattande. För att effektivt kunna tillämpa BRB-IF-modellen i multisensorfusion måste modellens parametrar optimeras. Författarna konstruerar en optimeringsmodell och en motsvarande optimeringsalgorithm där *belief function*-algoritmen används för slutledning, och partikelsvärmoptimering¹⁸ och differential-evolutionäroptimering¹⁹ används för optimeringen.

Att använda attribut för att karakterisera och mäta informationskvalitet är nödvändigt för att säkerställa framgångsrik mjuk och hård datafusion av öppna data. Sghaier m.fl. [136] introducerar naturliga språkbehandlingsmetoder (eng. natural language processing - NLP) med djupinlärning för att klassificera specifik domänrelevant information och kunskap från öppna data i ett mjukt och hårt informationsfusionssystem (eng. soft and hard information fusion) för beslutsstöd. Metoderna inkluderar utöver informationskvalitetsattribut också relevans och ett nytt attribut för informationens vikt för att förbättra kontextuell kunskap vid informationsfusion. Metoderna tilläm-

¹⁷Det kan noteras att författarna citerar FOI:s Informationsfusionsdemonstrator [3] som ett exempel på användningen av informationsfusion i ledningssystem.

¹⁸https://en.wikipedia.org/wiki/Particle_swarm_optimization (besökt februari 2022)

¹⁹https://en.wikipedia.org/wiki/Differential_evolution (besökt februari 2022)

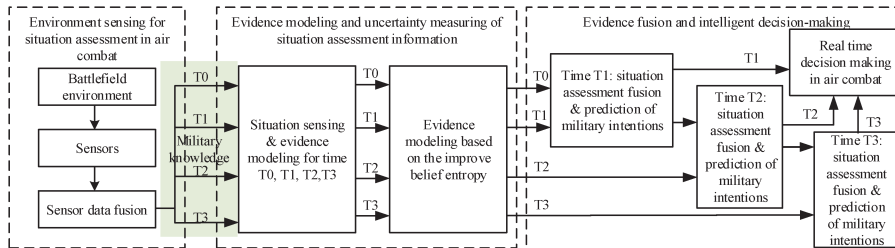
pas och utvärderas i denna artikel i samband med situationsförståelse hos en maritim situation. Den utvecklade metoden kombinerar ett relevansfilter med maritim meningstagning för att extrahera användbar information ur text med maritimt innehåll. Syftet med detta relevansfilter är att identifiera maritima relevanta meningar för att underlätta taggning i syfte att extrahera information som sedan ska användas av en datamodell i ett system för mjuk och hård informationsfusion. För den maritima domänen finns det ingen allmänt tillgänglig databas, därför var det nödvändigt att bygga upp en databas med manuella noteringar (eng. manual annotations). På så sätt är det möjligt att samtidigt skapa två separata databaser för två olika uppgifter: (i) för sjöfartsrelaterad meningsklassificering, och (ii) för maritim meningstagning. Metoden har utvärderats och visade sig vara effektiv i klassificeringen och valet av viktiga taggar och var robust avseende sanningshalten hos öppen data. Metoderna är utvecklade hos *Defence Research and Development Canada* och är tänkta att användas i kvalitetsdrivna hårda och mjuka fusionsprocesser för att underlätta dataassociering, särskilt i tvekydiga situationer i syfte att förbättra extraheringen av relevant information om fartyg, vilket gör det möjligt att få noggranna fusionsresultat.

3.2.3.5 Flygstrid

Osäker information finns inom varje procedur för bedömning av en luftstridssituation. Vid situationsbedömningen är beslutsprocessen för piloter och liksom för UAV:er beroende av intelligent informationsbehandling. I bedömning av en luftstridssituation inkluderas ofta en hotbedömning (eng. threat assessment) såsom förutsägelse av fiendens militära avsikter. Zhou m.fl. [201] studerar den osäkra informationen i informationsfusion av en luftstridssituation inom ramen för Dempster-Shaferteori. Författarna framhåller att särskild uppmärksamhet bör riktas mot fusionsprocesser där heterogen information kommer från flera olika källor. De föreslår en utvecklad metod för att hantera osäkerheten inom informationsfusion vid bedömning av aktuell situation i luftstrid. Genom att uppnå ett förbättrat fusionsresultat avseende förutsägelse om motståndarens militära avsikt förbättras möjligheten till beslutsfattande i luftstriden. Författarna utvecklar en metod som syftar till att erhålla ett mer exakt fusionsresultat i situationsbedömningen baserad på en förbättrad belief function-entropi [200] genom att förbehandla osäkerheten i all inkommande information innan fusionering sker. Fusion av information efter denna förbearbetning baseras på den standardiserade kombinationen inom Dempster-Shaferteori. Ett ramverk för en entropi-baserad metod för situationsbedömning vid luftstrid presenteras i figur 3.16. Metoden består av tre huvudprocesser: (i) *omvärldsöversikt* för bedömning av luftstridssituationen, (ii) *osäkerhetsmodellering* och utveckling av funktioner för bedömning av situationen, och (iii) *bedömning* av luftstriden med ett fusionsbaserat intelligent beslutfattande. I det sista steget sker en realtidsklassificering av den aktuella situationen som ett underlag till beslutsstöd för piloter eller alternativt till beslutsprocessen för en UAV.

3.2.3.6 Minröjning

Minröjning är ett viktigt område för informationsfusion. Krtalić & Bajić [105] beskriver i en artikel *TIRAMISU Advanced Intelligence Decision Support System* vilket är ett system för humanitär minröjning som kombinerar fjärranalys och datafusionsmetoder med avancerad övervakning och spaning i ett operativt system. Systemet stödjer processen att fatta beslut om misstänkta farliga områden genom att genererar kartor över områden där det kan finnas minor. Författarna beskriver systemets struktur och dess potentiella tillämpning som



Figur 3.16. Ett ramverk för bedömning av luftstridssituationen baserad på den förbättrade entropin av sensordata [201] (Creative Commons Attribution 4.0 International License, <http://creativecommons.org/licenses/by/4.0/>)

en del av humanitär minröjning.

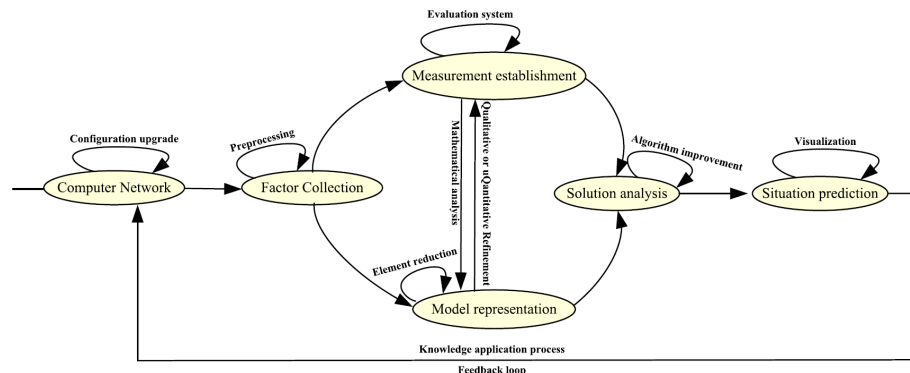
Krtalić m.fl. [104] presenterar ett koncept för att ta fram kartor över minfara för ett misstänkt riskområde. Indata omfattar information lagrad i informationssystem om mineringar och annan data som samlats in om det misstänkta området. De resulterande kartorna förbättrar identifieringen av misstänkt farliga områden för minröjning. Metoden använder en *indikator på minnärvaro* (eng. *indicators of mine presence - IMP*) som exempelvis ett militärt befäst objekt eller övergiven jordbruksmark i en terräng som experter antar försvaras av minor. Experter på humanitär minröjning avgör vikten av varje IMP baserat på sin kunskap om militära doktriner i det specifika geografiska området. Vikterna för alla IMPs beräknas med hjälp av den analytiska hierarkiprocessen (eng. *analytic hierarchy process - AHP*) vilket är en metodik för beslutsfattande baserad på flera kriterier som hjälper till att strukturera problemet med att fatta beslut. AHP är baserat på ömsesidiga jämförelser av alternativ och expertbeslut genom att omvandla experters subjektiva bedömningar om IMP:s relativa betydelse till en uppsättning vikter. Datafusion används för att visualisera det samlade inflytandet av alla IMP:s på miljön och till att förse experter med information om det misstänkt farliga området för att hjälpa dem att fatta beslut om att definiera områdets storlek. Slutresultatet av datafusionen är kartor över minfara.

3.2.3.7 Cyberförsvar

Avancerade ihållande hot (eng. *advanced persistent threats - APT*) utgör ett allvarligt hot i cyberrymden på grund av deras dolda verkan. I en artikel föreslår Lv m.fl. [120] en hybridstrategi baserad på ett dynamiskt spelteoretiskt försvar för att optimalt allokera begränsade säkra resurser i ett nätverk. Nyttovärden för spelares vinster i spelet beräknas med en ny fusionsmetod som kallas **NetF**. Med hjälp av **NetF** kan man kombinera data som observerats från trådlösa sensornätverk med data från internet. Fusionsmetoden bygger på att värdera varje nod med en funktion F som utgör en summa av flera nyttofunktioner där varje nods beräknade nyttovärde associeras med dess förmåga att lämna och ta emot meddelanden. Baserat på beräkningarna kan vi allokera säkerhetsresurser till dessa noder. Med hjälp av metoden kan data som observerats från internet och trådlösa sensornätverk fusioneras för att exakt beräkna förmågan att fungera för varje nod i nätverket. Metoden tillåter en dynamisk optimal strategi att allokeras till varje nod vid varje tid.

I en artikel lägger Li m.fl. [110] fram ett logiskt analysramverk rörande situationsmedvetenhet om nätverkssäkerhet ur datavärdekedjans (eng. *data value chain*) perspektiv. Ramverket delar upp processen för situationsmedvetenhet om nätverkssäkerhet i fem kontinuerliga bearbetningssteg: elementdefi-

inition, modellrepresentation, mätetablering, lösningsanalys och situationsprediktion, figur 3.17. Författarna sammanfattar forskningens framstegen i vardera av de fem stegen och diskuterar metoder och resultat.



Figur 3.17. Process för situationsmedvetenhet för nätverkssäkerhet [110]
(Creative Commons Attribution 4.0 International License,
<http://creativecommons.org/licenses/by/4.0/>)

I definitionsstadiet av element klassificeras och sammanfattas perceptuella data och designen och implementeringen av en databas beskrivs. I modellpresentationsstadiet diskuteras koncept, teknologier och modelleringsresultat för varje modell. I mätetableringsstadiet kvantifieras modellelementen. I lösningsanalyssteget diskuteras förutsättningarna i den aktuella tillämpningen och algoritmer analyseras. I situationsprediktionen betonas vikten av återkoppling och metoder för visualisering av analysresultat och val av försvarsåtgärder diskuteras.

I en annan artikel studeras s.k. exfiltrerande²⁰ hot. Exfiltrerande avancerade ihållande hot (eng. exfiltration advanced persistent threats - XAPT) står i hög grad för incidenter som rör kritisk informationsexfiltrering från mål med högt värde.

Befintliga ramverk för cyberförsvar och datafusionsmodeller klarar inte av XAPT:er på grund av bristande möjligheter att hantera flerfasattacker som kännetecknas av osäkerhet och motstridig information. Ioannou m.fl. [88] utvecklar en Markov Multi-Phase Transferable Belief Model²¹ (MM-TBM) som utvidgar TBM till att hantera cyberattacker som karakteriseras av flera faser för att uppnå svårbedömd cybersituationsmedvetenhet (eng. situational awareness). MM-TBM använder informationskonflikt i fusionsprocessen som en indikator till att identifiera paradoxer för att förbättra detekteringsförmågan. Författarna utvecklar nya funktioner för cybersituationsmedvetenhet och en algoritmisk metod för att upptäcka, spåra och förutsäga XAPT:er baserat på grafstrukturer som visar den potentiella sekvensen av åtgärder som en angripare kan vidta. MM-TBM erbjuder funktioner som spårning och förutsägelse av XAPT:er, multipel och samtidig attackdetektion med en diagnostisk funktionalitet.

3.3 Framtidsprognos

- Som framgår av figur 3.1 har antalet forskningspublikationer per år inom fusionsområdet ökat med en faktor tio under 24 år medan antalet forskningspublikationer per år inom fusion med militär inriktning ökat med

²⁰Motsatsen till infiltrerande

²¹https://en.wikipedia.org/wiki/Transferable_belief_model (besökt februari 2022)

en faktor tre (figur 3.5). Bedömningen är att dessa trender fortsätter. Tidigare dominerades forskningen av institutioner från EU och USA. Idag ligger Europa (inkl. Storbritannien) och Kina på samma nivå med USA strax därefter (figur 3.6).

- Fusionsområdet har flera delområden med mycket olika mognadsgrad. Området härstammar från radarutvecklingen där området målföljning har långa rötter med fundamental utveckling av algoritmer under 1970-till 1990-talet. Trots dess mognad så sker det ytterligare nyutveckling av flera slag, t.ex. inom målföljning av markmål. Även spaning är ett område med lång historik och även här sker utveckling inom flera delområden, t.ex. inom spaning med flockar av UAV:er. Vi bedömer att forskningen breddas inom spaning och målföljning till ytterligare nya tillämpningar.
- Områden som datafusion för operatörsstöd, informationsfusion för beslutsstöd och stöd vid flygstrid är betydligt yngre forskningsområden med ett mycket lägre antal forskningspublikationer. Operatörsstöd kan gälla stöd till piloter och beslutsstöd kan vara stöd från stridsteknisk nivå upp till stöd till staber och operativ nivå. Det är högst sannolikt att dessa metodiker kommer att samverka med många andra metoder främst inom artificiell intelligens, t.ex. med metoder som automatiskt och probabilistiskt resonerande, maskininlärning men också med metoder för optimering av processer och taktiskt uppträdande.
- Datafusion för minröjning har ökat i betydelse sedan inbördeskriget i Jugoslavien. Vi är osäkra om områdets framtida utvecklingstrend.
- Fusion för cyberförsvar är ett område under stark utveckling. Området utvecklas starkt både som ett civilt och som ett militärt forskningsområde. Vi tror att båda dessa utvecklingar fortsätter under lång tid med stark inbördes påverkan.

4 Dataanalys

4.1 Inledning

I den engelskspråkiga litteraturen förekommer två närbesläktade begrepp som berör *dataanalys*, nämligen *data analysis* och *data analytics*. De båda är lätt betraktade som utbytbara, men det finns en åtskillnad. *Data analytics* definieras brett, så tillvida att det betecknar användandet av både data och verktyg för att skapa beslutsunderlag eller förståelse. *Data analysis* används för att beskriva den delmängd av totalbegreppet *data analytics* som enbart rör förädlingen av data.

För denna rapport låter vi dataanalys omfatta både *data analytics* och *data analysis*.

4.1.1 Dataanalys

Den klasiska dataanalysen syftar till fyra stycken delfakulteter, vilka bildar en analyshierarki:

Beskrivande (eng. descriptive), Förklarar *vad* i data. Den del av analysen vilken berör visualisering som en platt företeelse, eller den icke aktiva visualiseringen.

Diagnosticerande (eng. diagnostic), Förklarar *varför* i data. Den del av analysen vilken strukturerar och relaterar informationen givet den beskrivande analysen. Denna inkluderar interaktiv visualisering.

Förutsäggande (eng. predictive), Förklarar *trolighet* i data. Vilken, givet den diagnosticerande analysen, kan ge svar på vad det troligaste alternativet, givet det vi vet, är.

Föreskrivande (eng. prescriptive), Förklarar *agerandebehov* i data. Vilket, den givna förutsäggande analysen, kan ge bra och konkret beslutsstödsunderlag.

I den retrospektiva studien[48] gör man ett rimligt försök till att etablera en rollmodell för datanalytiker. Kortfattat delas analytikerrollen in i fyra högre processteg med specifika delsteg:

Förberedelser, med delstegen definiera behov, datainhämtning, datageniering, profilering och databearbetning

Analys, med delstegen experiment, utforskande, modellering, verifiering och tolkning

Utplacering, med delstegen övervakning och förädling

Kommunicering, med delstegen disseminering och dokumentation

där visualisering, eller presentation, kommer in i ett av delstegen till varje övergripande processteg, nämligen: *profilering*, *tolkning*, *övervakning* och *disseminering*. I denna rapport ämnar vi att beskriva tekniker vilka kan nyttjas för en eller flera av information- och presentationsindelningarna, men vi kommer ej återkoppla till de specifika kategorierna direkt.

4.1.2 Storskalig dataanalys

Med storskalig dataanalys åsyftas metoder, och till viss del tekniker, för analys av datamängder där det inte går att få en totalbild över hela datamängden. Detta brukar exemplifieras med att en enskild dators lagringsplats inte räcker

till. Med lagringsplats åsyftas här hela systemets minneshierarki, dvs. inkluderar snabbt primärminne såväl som långsammare enheter. Det som gör storskalig dataanalys till just detta är dock inte den enskilda datorns begränsningar, utan just frånvaron av en totalbild över datamängden och att metoder anpassade till sådana situationer används. Detta eftersom högprestandaberäkningar inkluderar datamängder vilka partitionerats över flera maskiners minneshierarkier, men där man till skillnad från den storskaliga dataanalysen, har en totalbild över datamängden för ens problem.

Specifikt för storskalig dataanalys är att hela analyskedjan måste anpassas till en ofullständig vy, eller en approximation, av datan. Det har skett många framsteg inom området, såsom approximativa algoritmer för de statistiska momenten, men på grund av datalogistiken inskränks analysen alltid för:

- *inkommande strömmande data*, till olika sofistikerade aggregationsfönster.
- *redan lagrad data*, till datapartitionering.

4.1.3 Maskininlärning

På grund av framstegen inom djup maskininlärning är domänen värd att ta upp i sin egen rätt, även om den är en del av den förutsäggande dataanalysen. Traditionellt har maskininlärning använts för problem där det funnits god tillgång till data men med en sämre kvalitativ problemlösningsförmåga, eller problemlösningsförståelse. Man har använt träningsdata med syfte att träna en effektmodell för den domän datan beskriver. Modellerna kan således användas för problemlösning genom extrapolering från känd data, då dessa är en icke-förlustfri komprimering av den lärda datamängden. De första modellerna var av enkel karaktär där modellkoefficienterna tränades manuellt men när datorerna blev kraftfullare kunde dessa tränas automatiskt med bakåtpropagering. Efter denna tidpunkt gällde det istället att manuellt beskriva och välja ut de mest effektfyllda särdragen i ens data. När maskininlärningsforskarna lärde sig att använda SPMD¹-egenskaperna hos grafikprocessorerna inleddes övergången till djupa maskininlärningsmodeller. I skiftet från grunda till djupa maskininlärningsmodeller gick man från manuell särdragsidentifiering till automatisk sådan. Modellerna gick från få enkla lager till många intrikata lager och växte med avsevärda magnitudordningar.

4.1.4 System

Under inte minst det senare decenniet har mängden data ökat drastiskt, vilket ställer livcykelkrav på hanteringen av modeller samt generella resultat från dataanalysen. Detta får konsekvenser på den genomförbara analysen samt vilka analysmodeller som kan användas. Ofta används speciell hårdvara eller distributionsmodeller för att minska eller sprida ut mängden data.

Exempelvis kan kvantisering[55, 83] drastiskt minska modellstorleken, typiskt 32 gånger, men kräver istället ett mer gediget kvantitativt underlag. Kvantiseringen kräver även speciellt hårdvarustöd eftersom få processorer stöder drastiskt minskad mantissa av flyttal eller under 8 bitars heltal. Populära implementeringar av detta är BFLOAT eller INT8 för Nvidia korten, eller Xilinx PYNQ platform där man enkodar modellen med en FPGA bitström och således kan stödja lägre precisioner av parametrar[71].

¹En demultiplexering av SIMD för internaliserade GPU-processer

Man försöker även aktivt finna vägar för effektivare och maximalt modelnutnyttjande för djupa nät med hjälp av utvecklandet av parameterservrar². Exempel på sådana är Petuum[190] (vilken var först men nu under namnet sailing-pmls), hopsworks³ och Triton⁴. Tyvärr kräver dessa ofta ett annat programmeringsparadigm än vad en typisk maskininlärningsingenjör är bekant med.

Ett nyare inslag är att reducera parameterrymden genom att exploatera kompletterande världsinformation[75, 98]. Modeller som är anpassade för att förlita sig på externa kunskapsbaser levererar jämförbar prestanda[195] med stora modeller, men kräver drastiskt mycket mindre parametrar[24]. Man har helt enkelt gått från att baka in världsinformation i modellerna till att externalisa denna överflödiga och naturligt externa del av kunskapen. Detta förhållningssätt kräver ett tätare integrationsarbete då modellerna blir mindre men får externa beroenden.

Sammantaget kan man säga att datamängderna kräver aggregerade, partitionerade eller approximativa analysmetoder. Det finns dock nedslagspunkter av samsyn. Vi ser att både t.ex. PYNQ och Triton förlitar sig på användandet av Jupyter's cellutvärderade anteckningsblock⁵ för programmeraråtkomst samt visualisering. Man går mer mot återanvändandet av modeller och analysarbetsflöden, men instansierar dessa lokalt. Det finns klara skäl att anta att vi kommer gå mer åt deklarativa, parametriserade arbetsflöden med tillhörande specifikt programmeringsparadigm. Ett analysarbetsflöde kommer att fungera i en högprestandaberäkningsmiljö, likväl som på ett litet FPGA-kort där data- och modellförsörjningen är en implementationsdetalj.

4.2 Forskningsläget

4.2.1 Neurosymboliskt lärande och resonerande

Under det senaste decenniet har maskininlärningsområdet erhållit ett enormt intresse främst till följd av framgångarna inom delforskningsområdet djupinläring⁶ (eng. deep learning) [18]. De stora framgångarna till trots har vissa begränsningar och utmaningar med djupinläring identifierats, till och med av upphovsmännen (Yoshua Bengio, Geoffrey Hinton och Yann LeCun) själva [57].

Den amerikanska försvarsforskningsfinansiären Darpa betraktar området med stort intresse, men är samtidigt medvetet om bristerna. Vad man hoppas på är att djupinläringens lärande funktion kan kombineras med en resonerande funktion.

Olika förslag för att sammanföra lärande och resonerande inom AI har föreslagits. Ett förslag (som översiktligt illustreras i figur 4.1) är att sammanföra logiskt resonerande (symbolisk AI) med ANN (konnektionistisk AI) och metoderna från djupinläring, ett område som kallas *neurosymboliskt lärande och resonerande* (NSR).

För att förstå NSR är det nyttigt att titta närmare på egenskaperna hos symbolisk och konnektionistisk AI. Fördelar med 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

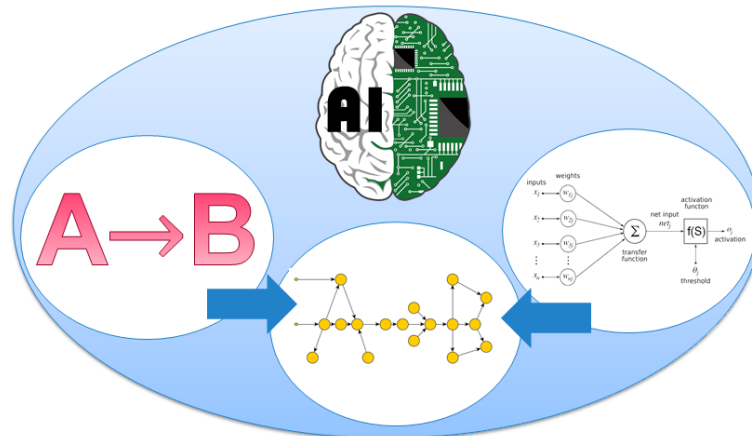
²Initiativ för lagring och delning av modellparametrar, <https://www.featurestore.org> (besökt februari 2022)

³HDFS kompatibel och fri mjukvara, <https://docs.hopsworks.ai/feature-store-api/latest/> (besökt februari 2022)

⁴Mer fokuserad på härledningsförmågedelning än parameterdelning <https://developer.nvidia.com/nvidia-triton-inference-server> (besökt februari 2022)

⁵Cellutvärderade anteckningsblock, med stöd för flera typer av programmeringsspråk-kärnor, och är fri mjukvara, <https://jupyter.org> (besökt februari 2022)

⁶Mer om djupinläring finns att läsa i avsnitt 4.1.3



Figur 4.1. Neurosymboliskt lärande och resonerande (NSR) är en form av artificiell intelligens som består av en kombination av en symbolisk del (vänster) och konnektionistisk del (höger). Resultatet (mitten) har egenskaper från båda delarna.

förtroende för resultaten. Nackdelarna uppstår i kontakten med verklighetens komplexitet och osäkra data. De data som experter inledningsvis matar in eller representerar är en delmängd av all data som modellen därefter kommer att 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 experter är oförmögna att uttrycka vissa relationer samt att information som är svåråtkomlig och som "gömmar" sig i stora datamängder kommer att förbises när experter används för symbolisk representation.

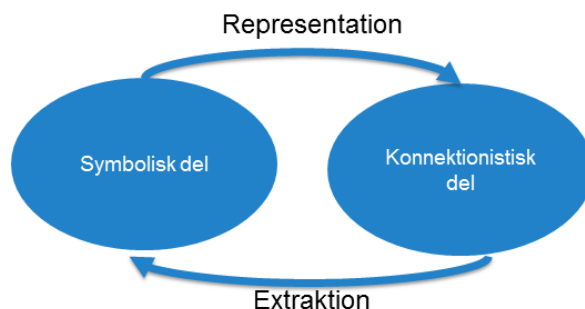
Konnektionistisk AI, å andra sidan, kan effektivt hantera stora mängder data och extrahera outtalad underliggande information och samband, men nackdelen med att använda dessa metoder är istället att resultatet är svåröklart och kan misstros av en mänsklig användare.

Ett direkt sätt att åstadkomma neurosymboliskt resonerande är att sammanföra en komponent för logiskt resonerande (symbolisk) och en konnektionistisk (ANN) komponent till en hybrid⁷ 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 neuronät, samt att logiska regler kan extraheras ur neuronätet. Det här beskrivs i den så kallade *Neural-symbolic learning cycle* (NSLC) [169] som finns avbildad i figur 4.2.

Hybridmetodernas symboliska del och konnektionistiska del samverkar genom NSLC 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.

Ett problem som man hoppas att NSR ska kunna bidra till att lösa är

⁷HIS (eng. Hybrid intelligent systems) är ett samlingsnamn för sammansättningar av AI-metoder för att lösa något visst problem eller imitera någon egenskap hos generell intelligens. HIS är ett generellt begrepp och ett sådant system behöver inte bestå av just symboliska och konnektionistiska AI-komponenter.



Figur 4.2. Hybridmetoder består av en symbolisk del och en konnektionistisk del. Dessa samverkar genom Neural-symbolic learning cycle där symbolisk kunskap överförs (representation) till en ANN-struktur som kan anpassas via inlärning 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 inlärningen).

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” syftande på dess oförmåga att representera något mer uttrycksfullt än satslogik (eng. propositional logic), det vill säga en begränsad form av logik [127]. Sedan McCarthys kritik år 1988 har uttrycksfullheten utökats bland annat genom att kombinera ett flertal ANN på olika sätt [51].

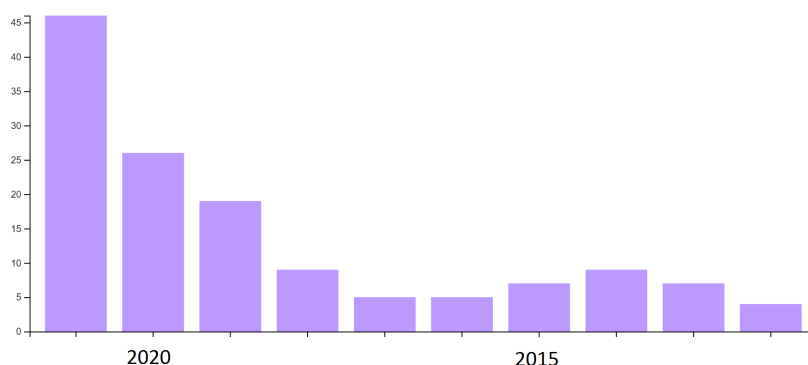
En viktig form av logik är vad som kallas första ordningens logik [49, 50]. Första ordningens logik är exempelvis användbart för att representera den information som kan uttryckas i naturligt språk. En hybridmetod som sammanför första ordningens logik och djupinlärning är “Logic tensor networks” [12].

Under de senaste åren har IBM tillsammans med MIT (MIT-IBM Watson AI Lab, förkortat Watson) satsat på att utveckla NSR-system för praktiska tillämpningar [194]. Medan den “vanliga” djupinlärningen är skicklig på att hitta och klassificera förmål i bilder, så fokuserar Watson AI Lab på en tillämpning som möjliggör att en användare kan ställa frågor till de bilder som analyseras, exempelvis “Vilken form har det röda objektet” där svaret exempelvis kan vara “sfär”. Watsons system kallas NS-VQA (eng. Neuro-symbolic visual querying and answering) och kombinerar en djupinlärningsbaserad bildanalysdel (som kan extrahera och klassificera objekt i bilder och uttrycka resultatet med logiska fakta) med en textanalysdel (också baserat på djupinlärning) som omvandlar frågor till en logisk representation. Därefter försöker en resonerande funktion lösa den ställda frågan, med hjälp av de extraherade objekten. Förutom att man kan uppvisa bättre resultat än andra metoder, så behöver man betydligt mindre data för inlärning, och som bieffekt får man en textuell beskrivning av hur NS-VQA-systemet nått sitt resultat (vilket kan användas för felsökning och för att skapa förtroende för systemet hos användarna). Watson har gått vidare med arbetet för att även kunna ställa frågor kring video.

En sökning efter publicerade artiklar med Web of Science innehållande orden “neuro symbolic” eller “neural symbolic” i titel, sammanfattning eller nyckelord gav den årliga fördelning som framgår av figur 4.3.⁸ Det absoluta antalet artiklar är litet men växande. NSR-området stöds också av en workshop

⁸Sökfråga: TS=((neural OR neuro)NEAR/0 symbolic)

som har hållits årligen sedan 2005.⁹



Figur 4.3. Staplarna visar antal publicerade artiklar, år för år mellan 2012 och 2021, som nämner "neuro symbolic" eller "neural symbolic" i titel, sammanfattning eller nyckelord.

Mer om NSR finns att läsa i en FOI-rapport [112] som bland annat presenterar och diskuterar försvarstillämpningar. En lista med länkar till olika mjukvarupaket för NSR finns i Appendix A.2.1.

4.2.2 Kvantdatorstödd dataanalys

Åtskilliga möjliga tillämpningar av kvantmekaniska fenomen har föreslagits inklusive tillämpningar för datorer, simuleringar, kommunikation och sensorer. Det är viktigt att fastställa att *kvantdatorer* (eng. *quantum computer*) inte är en ersättare till dagens digitala datorer. Kvantdatorerna kommer att kunna lösa vissa speciella uppgifter effektivare än sin digitala motsvarighet, men inte minst på grund av kvantdatorernas instabila och (i klassisk mening) mycket begränsade arbetsminne kommer de flesta uppgifter som digitala datorer utför idag inte att utföras av kvantdatorer. Istället lär klassiska digitala datorer och kvantdatorer komma att samverka. Vidare har kvantdatorer flera tillämpningar där den förmodligen mest omtalade är kvantkryptering och, inte minst, kvantdekryptering som väntas få stora effekter på hur känslig information hanteras på ett säkert sätt. I den här rapporten fokuseras dock på de kvantdatorberäkningar som berör dataanalys.

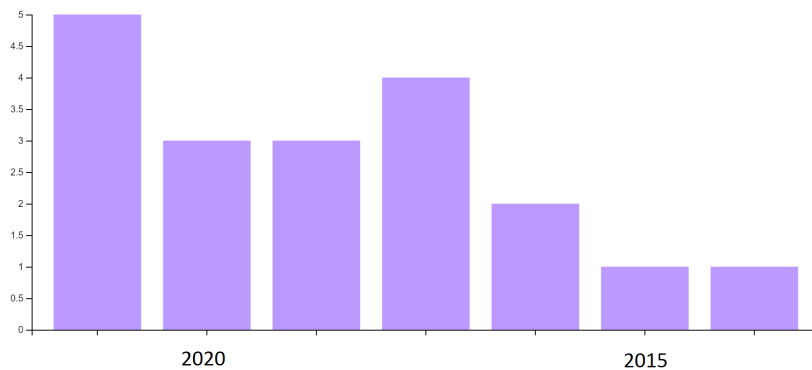
På senare tid har den potentiella nyttan av kvantdatorer för att stödja dataanalys rapporterats allt flitigare i takt med att satsningarna på kvantdatorer lett till vissa framgångar. Utvecklingen av publiceringarna om kvantbaserad AI och maskininlärning framgår av sökningar i WoS figur 4.4¹⁰ respektive 4.5.¹¹

Fysikern Richard Feynman (1918-1988) brukar tillskrivas idén att använda kvantmekaniska fenomen för att göra beräkningar [143]. Feynman ville simulera kvanttillstånd hos en stor mängd partiklar men insåg att tidens klassiska datorer inte maktade med att skapa en sån simulering. Istället föreslog han i början av 1980-talet [66] att man skulle kunna använda kontrollerade kvantmekaniska tillstånd, en kvantdator, för åstadkomma denna simulering. Trots att det ännu inte fanns kvantdatorer började det under 1990-talet fors-

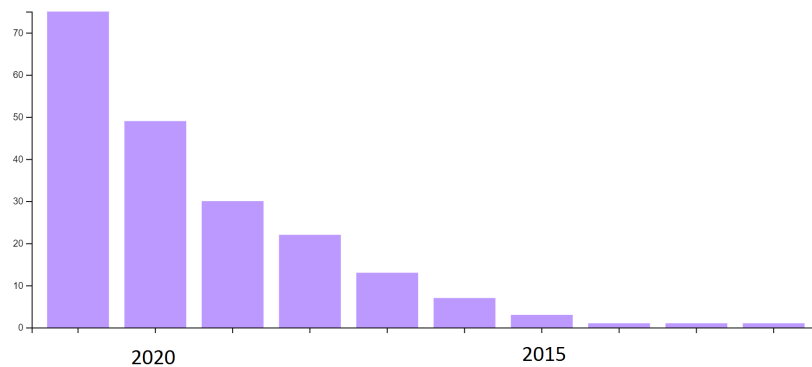
⁹Workshop series on Neural-Symbolic Learning and Reasoning, <https://www.neural-symbolic.org> (besökt februari 2022)

¹⁰Sökfråga: TS=("quantum artificial intelligence" OR "quantum AI")

¹¹Sökfråga: TS=("quantum" NEAR/0 ("artificial intelligence" OR "AI" OR "machine learning"))



Figur 4.4. Staplarna visar antal publicerade artiklar, år för år mellan 2012 och 2021, som nämner “quantum artificial intelligence” i titel, sammanfattning eller nyckelord.



Figur 4.5. Staplarna visar antal publicerade artiklar, år för år mellan 2012 och 2021, som nämner “quantum artificial intelligence” eller “quantum AI” i titel, sammanfattning eller nyckelord.

kas fram algoritmer för dessa ännu ej existerande maskiner. Det stora allmänna intresset för kvantdatorer kom i och med Shors arbete [166] om primtalsfaktorisering. Anledningen är att primtalsfaktorisering kan användas för att dekryptera RSA-krypterad data, och Shors kvantalgoritm kan teoretiskt göra detta avsevärt snabbare än en klassisk dator.

Grovers sökalgoritm [76] är en annan “klassisk” kvantdatoralgoritm som tydligare är användbar för dataanalys än Shors. Grovers algoritm går ut på att hitta ett visst data i en mängd data (ur en fast databas eller kanske ännu hellre ur strömmande data). Grovers sökalgoritm har en söktid som är snabbare än linjär, i det här fallet $O(\sqrt{n})$, jämfört med linjär, dvs. $O(n)$, för den digitala motsvarigheten, dvs. en kvadratisk förbättring i antal data (n). Det som är särskilt intressant med Grovers algoritm är att den demonstrerar en princip kallad *amplitudförstärkning*¹² som kan användas i flera andra problem [7, avsnitt 3.1.2.3],[26].

Den mest fundamentala dataenheten i en klassisk digital dator är en *bit* (binary digit), som kan anta värdena *ett* (1) eller *noll* (0). Logiskt sett består

¹²Dvs. en förstärkning av sannolikheten för den önskade lösningen.

digital data enbart av ettor och nollor, exempelvis en 8-bitars byte 01101111, som sedan omtolkas till instruktioner, tal, eller tecken beroende på kontext. För att använda digital data i praktiken (dvs. för att lagra, förmedla, och processa den digitala datan) behöver den manifesteras fysiskt. Det sker normalt antingen med en elektrisk spänning (hög eller låg beroende på värdet) i primärminnet eller magnetisk polarisation i sekundärminne (exempelvis hårddisk).

Kvantdatorn däremot är inte baserad på digitala bitar, utan *kvantbitar* (eng. *qubit*) [126]. En kvantbit kan representera en digital nolla och etta, uttryckt med symbolerna $|0\rangle$ respektive $|1\rangle$ (en matematisk definition av kvantbitar och -tillstånd finns i Bilaga A.1).

För att kunna realisera de kvanttillstånd och beräkningar som beskrivs ovan så behövs särskild hårdvara, dvs. kvantdatorer, och det första lyckade fysiska experimentet genomfördes 1998 [43]. Utvecklingen av fysiska kvantdatorer har sedan tagit fart, både inom akademien och i industrin, och gjort stora framsteg på den sista tiden. Det finns i dagsläget inte en dominerande teknik utan naturen erbjuder flera sätt att skapa och manipulera kvanttillstånd. Kortfattat är teknikerna bland annat baserade på *supraledning*, *fotonik*, respektive *jonfälla*. Mer om dessa tekniker finns beskrivet i exempelvis FOI-rapporten [7, avs. 3.3] och en prestandajämförelse är nyligen utförd av Gill m.fl. [73, avs. 6.4]. Teknikernas egenskaper varierar stort vad det gäller behov av kylning (supraledande och jonfällor måste kylas till nära nollpunkten), skalbarhet, och hur störningskänsliga kvantbitarna är.

IBM¹³ har en 127-bitars kvantdator baserad på supraledning (kallad Eagle) och erbjuder kunder möjligheten att använda den via en molnlösning.¹⁴ Xanadu¹⁵ använder fotonik för sin kvantdator [9].¹⁶ IonQ¹⁷ har en fotonikbaserad kvantdator på 32 bitar. En lovande uppstickare är QuEra¹⁸ som i slutet av 2021 rapporterade¹⁹ om en maskin med 256 kvantbitar, som är baserad på excitering av atomer till ett så kallat "rydbergstillstånd" [62].^{20,21}

Klassiska digitala datorer kan drabbas av störningar som gör att bitar byter värden (s.k. "bit-flip"), men detta problem är hanterbart. I kvantdatorer däremot är störningar ett mycket stort problem och flera olika typer av störningar kan ske, exempelvis motsvarigheten till bit-flip, men också att kvanttillståndet kollapsar okontrollerat (s.k. *dekoherens*) innan beräkningarna avslutas [153].

I kvantdatorns barndom var det på grund av denna störningskänslighet osäkert om man någonsin skulle kunna bygga användbara kvantdatorer. Återigen bidrog dock Peter Shor (som först konstruerade kvantalgoritmen för primtalsfaktorisering) med ett viktigt resultat [167]. Trots de framsteg som skett sedan dess kräver felkorrigering i kvantdator tusentals gånger fler faktiska kvantbitar än de som i praktiken kan användas algoritmiskt. En populär uppskattning är att det kommer att krävas "en miljon kvantbitar" för att få

¹³<https://www.ibm.com/quantum-computing/> (besökt februari 2022)

¹⁴<https://quantum-computing.ibm.com/> (besökt februari 2022)

¹⁵<https://www.xanadu.ai/hardware> (besökt februari 2022)

¹⁶<https://spectrum.ieee.org/race-to-hundreds-of-photonic-qubits-xanadu-scalable-photon> (besökt februari 2022)

¹⁷<https://ionq.com/> (besökt februari 2022)

¹⁸<https://www.quera.com/> (besökt februari 2022)

¹⁹<https://www.technologyreview.com/2021/11/17/1040243/quantum-computer-256-bit-startup/> (besökt februari 2022)

²⁰Atomens "rydbergstillstånd" är uppkallat efter den svenska fysikern Johannes (Janne) Rydberg (1854-1919).

²¹I sammanhanget måste även kvantdatorpionjären D-wave (<https://www.dwavesys.com/learn/quantum-computing/>, besökt februari 2022) nämnas som har en kvantdator som redan har flera tusen kvantbitar. D-waves kvantdator kan dock inte direkt jämföras med de övriga då den är mer specialiserad.

kvantdatordominans.

För att vidare undersöka nyttan av kvantdatorer för dataanalys undersöktes litteraturen för *quantum machine learning* (QML), vilket omfattar maskininlärningsmetoder som antingen består av kvantalgoritmer, kvantdata, eller både och. Av litteraturen [30, 163] framgår att man har tagit fram kvantalgoritmiska motsvarigheter för många maskininlärningsmetoder som utvecklats för digitala datorer, men att det fortfarande är oklart om kvantalgoritmerna erbjuder bättre lösningar.

John Preskill är fysiker och Richard P. Feynman-professor vid Caltech i Kalifornien och en av förgrundsfigurerna inom kvantdatorer. Preskills bedömning är att djupinlärningsnät skulle kunna tränas mer effektivt eller att man skulle kunna dra nytta av kvantdatorns förmåga att lagra högdimensionell data effektivt, men att det är för tidigt att dra någon slutsats [144].²²

Trots att riktigt användbara kvantdatorer ännu inte finns, så hjälper mjukvarorna till att underlätta utvecklingen av nya kvantalgoritmer. En lista med länkar till olika mjukvarupaket och algoritmer för kvantdatorberäkningar finns i Appendix A.2.2. Mer att läsa om kvantdatorer finns i andra FOI-rapporter [7, 100].

4.2.3 Metainläring, fåförsöksinläring & överföringsinläring

För detta avsnitt avskannas maskininlärningsområdena:

Metainläring (eng. meta learning), eller lära sig att lära. Syftar till gradientbaserade inlärningsmetoder vilka optimerar över uppgiftspecifika körningar med syfte att lära sig att träna modellparametrar. En vanlig tillämpning av metainläring är federativ inläring, alltså medelvärdestagning och delning av modellparametrar över distribuerade inlärningsuppgifter. Typiskt för metainläringen är en yttre och en inre slinga av modellparameteruppdatering där dessa propagerar fram och tillbaka över modellansvarsgränser.

Fåförsöksinläring (eng. few shot learning), används typiskt vid metainläring där mängden exempel för varje distribuerad uppgift uppgår till tiotals men inte mer. Notera att enförsöksinläring kommer från den bayesianska statistiken där visuell klassificering görs efter mindre än tio exempel. Fåförsöksinläring kallas ibland även k -försöks n -fall då modellen skall lära sig klassificera n klasser efter träning på k delexempel, eller i metainlärningsfallet då k -försöksinläring syftar till k distribuerade uppgifter.

Överföringsinläring (eng. transfer learning), syftar till de tekniker där modellparametrar inte uppdateras avsevärt mycket i sig, som i metainlärningsfallet, utan återanvänds tillsammans med någon finjustering (eng. fine tuning) av modellen. Ett exempel är att nyttja ett sista lager, exempelvis differentierbart maximum (eng: softmax), att använda ett avsevärt mindre inlärningssteg, eller att bara träna om senare lager i modellen.

och förutsätter en fördjupad kunskap inom ämnesområdena. Ordvalet "inläring" istället för "lärande" syftar till att belysa det faktum att bara en mängd modellparametrar uppdateras och att det inte handlar om traditionellt lärande.

Metainläring är ett fortsatt aktivt och aktuellt[197] område där man oftare använder sig av Fisheriansk statistik²³ och hierarkisk klustering för att

²²<https://www.cornell.edu/video/john-preskill-quantum-computing-nisq-era-beyond> (besökt februari 2022)

²³klassisk beskrivande statistik

grannskapsbestämma inlärningsuppgifter[2] eller för att syntetisera dataset genom metainläring och simulering[94]. Ett nytt inslag är att lära sig avståndsmetriken för nätets målfunktion med hjälp av metainläring[37], samt att använda evolutionära algoritmer i syfte att reducera antalet modellvikter[118]. Man ser fortfarande federativa tillämpningar förekomma[108]. Att just federativt genomsnitt överutnyttjas för identifiering av mänsklig aktivitet har vi inte någon förklaring till.

I en jämförelse mellan metainläring och överföringsinläring har överföringsinläring visat sig prestera bättre än metainläring som en fåförsökslärandemetod[61]. Vidare har en metod för fåförsöksinläring, FLUTE[178], visat sig prestera bra för generell partiell paramaterinitialisering för FiLM-lager[139] (bildigenkänning).

Det sker fortsatt mycket arbete inom noll- och fåförsöksinläring. Nästan uteslutande exploateras någon känd isomorfi för att generativt svälla kända etiketter. Detta kan exempelvis ske genom generativa modellutökningar eller datamängder som etiketteras med hjälp av kunskapsgrafer[70, 78]. Vi ser även att man nyttjar metainläring för att lösa isomorfin[58].

Framsteg har gjorts i analysen kring massiv uppgiftsframställan vid övervakad inläring[14] av språkmodeller där man har identifierat viktiga prestandaförändrande parametrar. Flertalet kontextkänsliga mixturmodeller har framtagits[8, 157, 184], specifikt vid förträning och för en ansenlig mängd olika typer av uppgifter[129]. Dessa modeller har visat sig ge bra utväxling på nedströmsprestanda i specifika domänområden. Därtill ger de förbättrad prestanda över fleruppgifts fåförsöksinläring för språkmodeller.

4.2.4 Uppmärksamhetsmodeller som förtränade modeller

För detta avsnitt avskannas:

Förträning (eng. pre-training), syftar till tekniker och modeller för nära på optimal modellparameterinitialisering istället för slumpmässig sådan. Naivt uttryckt kan man säga att det är bättre om modellen redan från början förstår sig på vår eftersökta datamängd och kan göra förutsägelser utan någon träning, relativt detta till de tidigare koncepten i sektion 4.2.3.

Uppmärksamhetsmodeller (eng. attention based models), även kända som transformerare (eng. transformers) syftar till modeller innehållande en uppmärksamhetsavkodare, ett typ av korrelationslager, kopplat till en förbättringsmekanism (eng. boosting) av signalen från uppmärksamhetsavkodaren. För denna rapport används uppmärksamhetsmodell istället för transformerare då det är uppmärksamhetsfunktionens icke sekvensiella och associativa natur som är av fundamental betydelse.

Vi ser en fortsatt tillförlitlighet och aktivitet runt stora förtränade modeller[22]. Språkmodeller[181], har arkitekturmässigt vidareutvecklats till funktionsdugliga högre dimensionsmodeller. Dessa uppmärksamhetsbaserade modeller, vilka tidigare rönt stor framgång inom naturligt språk, har framgångsrikt anpassats och använts inom bildigenkänning[60] samt skalats upp[196].

Man har nyligen visat att djupa nätverk, specifikt transformerare, klarar att lära sig lösa komplexa matematiska problem[36] såsom att lösa stabilitet eller ekvilibria för system av partiella differentialekvationer. Tidigare ansatser på området har sammanställts i en översiktsstudie där man resonerar kring modellernas komplexitet, effektivitet, och etablerar en taxonomi för dessa[173]. Man kan där finna 7 huvudklasser av arkitekturella knep för att nå högre modelleffektivitet. Exempelvis kan man ta *kärnisering* (eng. *kernelization*) av

uppmärksamhetslager för att undvika beräkningen av $N \times N$ matrisen för sekvenslängden[97]

På grund av att uppmärksamhetsvikterna är kvadratiska med avseende på sekvenslängden har det även framställts andra approximativa, förutom kärnisering, självuppmärksamhetsansatser[122, 138]. Vidare har transformerare analyserats utifrån ett skalningsperspektiv där man ser empiriskt att uppgiftseffektivitet uppström står i direkt konflikt med nedströmprestanda[1, 174]. Man har försökt att komma bort från tokeniseringen av text genom nyttjandet av k -shinglar, även kallat n -gram, för transformerare[172, 191] där man visat på ökad robusthet och förbättrad prestanda. Vi kan även konstatera att text-till-text transformeraren T5[146] är väldigt populär att utöka eller jämföra med.

Andra använda knep är att använda en intern latent bas och korskorrelerad uppmärksamhet[90, 91], här alterneras själv- och kors-uppmärksamhetsvikterna internt mellan beräkningsstegen. Det finns därtill en uppsjö av försök att ersätta uppmärksamhetslager med alternativ[106, 114, 177]. Det har även visat sig att konvolutionella nät kan förtränas på samma sätt som transformerare och erhålla liknande prestanda[171].

Som tidigare nämnts har man tagit fram metoder för att externalisera omvärldsinformation, och på så vis krympa modellstorlekarna[24]. Man har dock inte kommit längre än att nyttja närmsta-granne-sökningar, eller sämre[195], för sin omvärldsinformationsinhämtning.

Det är även populärt att bygga modeller för frågor och svar, t.ex. genom maximerat trolighetsvärde[155] eller i vissa fall genom användandet av webbsökningar[103, 131].

4.2.4.1 Frågeställningsframställan

För detta avsnitt avskannas:

Frågeställningsframställan (eng. prompting), syftar till användandet av fyller-i framställan för att plocka fram associativer. Exempelvis kan en frågeställningsframställan ta formen “bilen har fyra _” där “_” fylls i eller används i tränings syfte, svaren “hjul” eller “säten” kan då ges med olika tillförlitlighet. Frågeställningsframställan kan även mindre lyckat användas för att svara på snäppet mer avancerade frågor som “_⁴ = e”

Frågeställningsframställan är en relativt ny teknik för klassprediktion och inlärning via fyller-i uppgifter[28]. Själva området har övervakats i en systematisk litteraturstudie[115] men övergripande kan man säga att modellen tränas genom att inmatningsrelationer ersätts med hål, eller tillfälliga symboler, vilket ställer nya ingenjörskrav på frågeframställningssemantik. Själva formen för denna är prestandakänslig [140] och trots att man rönt framgång inom fåförsöksinlärning [160, 161, 170] finns det ännu inget riktigt bra evalueringskriterium för dessa[199]. Frågeställningsframställan visar dock på modellträningspotential då en väl vald frågeställningsframställanstrategi reducerar mängden nödvändig träningsdata[159]. Frågeställningsframställan som en domän kopplar i teorin ihop programspråksemantik med datautvidgning[151]. Det finns även ett initiativ att dela frågeställningsframställandeinstanser[11]. Man har börjat göra framsteg med förklarbarhet inom naturligt språk tack vare frågeställningsframställan[189]

Då stora förtränade modeller dominerar så totalt har metoder för att finjustera vikter sett framgångar både inom frågeställningsframställan[107, 117] samt genom nyttjande av variationella autoenkodrar som arkitekturella översättnoder[63], eller så kallade adaptrar, för nedströmsanpassning. Det ska näm-

nas att dessa adaptrar för uppmärksamhetsmodeller visar sig inte bara vara prestandahållbara[141] utan är även i sig delbara för olika uppgifter nedströms²⁴. Metodutveckling för finjusterandet av färre modellvikter har sett framgång, exempelvis genom injektion av uppgiftsberoende vikter[123] eller ramverk för uppgiftsberoende modellviktselektion[81, 125].

4.2.5 Multimodala modeller

För detta avsnitt avskannas:

Multimodala modeller(eng. multimodal models), bygger förenklat en korrelationsmatris mellan de olika modaliteternas särdragsvektorer och kan således associera de textuella särdragen för en utter med de grafiska särdragen för en utter.

Multimodala modeller har överlag blivit populära och prestandamässigt intressanta att titta på[31, 82], men även rönt framgångar inom förstärkningsinlärning[38] samt fåförsöksinlärning[179]. Inom talsyntes har man kombinerat kända modeller för naturligt språk med ljudinbäddningar[13] och i självövervakat träningsläge uppnått snabb konvergens med förtränade modeller[10, 44]. Talsyntes är bara ett exempel på mångsidigheten i senare djupa modeller där man undersökt koppling mellan modaliteter genom analys av videotransformrare och språkinbäddning[67], eller talenkodning genom att kombinera tal och text vid förträning[15]. Det är populärt att generera bilder från text[149] men det verkligt nya är att nyttja diffusionstekniker, alternerade framåt-bakåt pass till konvergens, och multimodala inbäddningar[145] för att generera än mer trovärdiga bilder[132] än vad som är möjligt med generativa nätverk[56]. Det ska dock nämnas att även generativa korrektionsmetoder, för t.ex. ansiktsgenerering, genom generativ inversion blivit snabbare och klarar av att producera än mer tillförlitliga resultat[42].

4.2.6 Modell- & datamängdförståelse

Inom teoretiskt modellförståelse ser vi ett nytt grundläggande resultat gällande behovet av överparametrisering för djupa neurala nät för att nå robusthet. Beviset berör högre dimensionaliteters geometrier och redogör för en universell robusthetslag[29]. Det var tidigare känt, tack vare Håstads Lemma[79], att ett neuralt nät med k -lager och en värdesräckvidd på n krävde minst en parameterstorlek på $2^{n^{\frac{1}{k-1}}}$ för god approximering, men nu är alltså kunskapen om överparametriseringens påverkan för robustheten känd samt varför överparametriseringen är nödvändig.

För växande modellstorlekar krävs verktyg för att hantera datakvalitet[20] och dataprovenans[59]. Därtill visar multi-kulturell modellprestanda på försvagad generell robusthet när man använder andra språk än engelska[113].

Vidare har man lagt energi på att problematisera förtränade modellers partiskhet (eng. *bias*) från olika genusperspektiv (i bred bemärkelse) samt vilka risker det medför[5, 17, 116, 185, 186]. Det finns anledning att anta att även modeller producerade i Sverige innehåller, eller kodifierar, oönskad partiskhet.

4.2.7 Datavisualisering

Cellutvärderade digitala anteckningsblock²⁵, vilka tidigare omnämnts i 4.1.4, är populära för återanvändning i visualiseringssyfte. Anteckningsblocken skiljer

²⁴<https://adapterhub.ml> (besökt februari 2022)

²⁵I stort sett enbart jupyter-notebook (<https://jupyter.org>, besökt februari 2022) blad men det finns andra implementationer, t.ex. zeppelin (<https://zeppelin.apache.org/>, besökt februari 2022)

sig från klassiska kalkylblad i att man får stöd för fler programmeringsspråk direkt i en webbläsare. Denna popularitet har lett fram till frågor gällande provenans, komposition och inte minst övergripande visualisering för anteckningsblocken[188]. Återanvändbarhet av dessa anteckningsblock börjar ta fart i form av visualiserings- och programcellsrekommendationsystem[147].

t -SNE²⁶ är fortsatt intressant för att visa dimensionsreducerade processbilder, exempelvis skulle processprestandavisualisering[109] kunna specialiseras på modellprestanda och modelldriv över tid och inte bara visa på generell processprestanda. Vidare har man utvärderat ett symbolsammanfattande verktyg för klustringsresultat av dekompositioner[92] och med en enkätundersökning fått bra resultat från respondenter med avseende på deras förmåga att upptäcka utstickare och utföra metodjämförelser. Inom social nätverksanalys har man tagit fram visuella hjälpmedel, baserade på hierarkisk klustring av dekompositioner, för att analysera dominanshierarkier[45]. Just dimensionsreduktion kopplat till klustring är väldigt populärt[187].

Darpa har finansierat ett verktyg för framtagandet av visuella grafjämförelser med fokus på maskininlärningsarbetsflöden och dess produkter[47]. Man har visat på verktygsnytta vid modelljämförelser, och specifikt jämförelseverktyg för hyperparameterstämmande[53]. Det har tillkommit en helt ny metod för värmekartsgeneration av kontinuerlig tidsseriedata, vilken är snabbare än befintliga metoder[142]. Inom fordonsindustrin används kontinuerligt upptränade modeller, tillsammans med sensorinformatiksdekompositioner till 3-dimensioner, lyckats få lågutbildade operatörer att explorativt upptäcka maskinfel i elmotorer[64]. För att hjälpa manuell visuell analys av feletiketterad kontextuell data har man tagit fram en korrigeringsprototyp med kvalitativt bra resultat[124], det har även tagits fram ett interaktivt verktyg för utvärderandet av olika objektivfunktioner[52].

Strategiskt spatial lägesdominans, eller områdestvist, i realtids har erhållit en visualiseringsmetod[6]. Vid explorativ analys har man upptäckt en ny metod för grafisk sekvensanalys. Detta genom användandet av förstärkningsinlärning för att finna optimala grafiska historiker vid korrelation-, anomali- och klusteranalys[165].

4.3 Framtidsprognos

- För ansiktsgenerering återkommer fundamentala problem som höger/vänster irisassymmetri eller artefaktbildning för många generade bilder.
- Vi ser att fåförsöksinlärning trumfar metainlärning för fleruppgiftsfall.
- Metainlärning gör promptinlärning effektivare genom promptindelning på uppgiftsnivå.
- Transformerare, eller uppmärksamhetsmodeller, kommer att fortsätta effektiviseras och troligtvis användas som ett intermediärt språklager för en mängd modaliteter.
- Delning av organisationsöverskridande maskininlärningresultat kommer fortsätta, tidigare har man delat hela modeller²⁷ och nu ser vi att man delar med sig av uppgiftspecifika finjusterade lager²⁸ (besökt februari 2022).
- Vi ser att modellstorlekarna, trots framgångar i diverse reduktionstekniker, kommer att fortsätta växa från biljonparametermodeller och kräva

²⁶ t -distribuerad stokastisk granninbäddning är en statistisk visualiseringsmetod för högdimensionell data

²⁷<https://pytorch.org/hub> (besökt februari 2022) och <https://huggingface.co> (besökt februari 2022)

²⁸<https://adapterhub.ml>

datamängder i motsvarande magnitudordningar.

- Neurosymboliskt lärande och resonerande var, tills för några år sedan, inte tillämpat utan mer av ett teoretiskt område. Den potentiella nyttan med att koppla NSR till utvecklingen inom djupinlärning har dock lett till stort intresse för området och resulterat i intressanta praktiska tillämpningar. Systemet NS-VQA av Watson, som presenteras i avsnitt 4.2.1 är av försvarsintresse, exempelvis för att stödja underrättelseanalys (interaktiv bildanalys beskrivs i avsnitt 4.2.1). NSR-systemens förmåga att förklara slutsatser är, som bland annat noterats av Darpa, också av intresse i försvarssammanhang.
- Framstegen med utvecklingen av fysiska kvantdatorer de senaste åren har väckt ett allmänt intresse och hopp om att kvantdatortekniken ska börja bidra på olika områden, exempelvis som stöd till AI, kvant-AI och kvant-ML. Förutsättningarna för att lära mer om kvantdatorns praktiska möjligheter och begränsningar är nu stor, och likaså utveckling av nya algoritmer (till följd av tillgången till mjukvarustöd). För närvarande krävs dock vidare utveckling av stabila kvantdatorer med fler tillgängliga kvantbitar samt lämpliga kvantalgoritmer. Även om hårdvara och algoritmer snart kan börja leverera kvantdominans, så kan FM förmodligen inte använda en kvantdator i en molntjänst (exempelvis IBM:s) utan behöver skaffa och underhålla en själv, vilket är ett betydligt större och mer avlägset mål.

5 Slutsatser

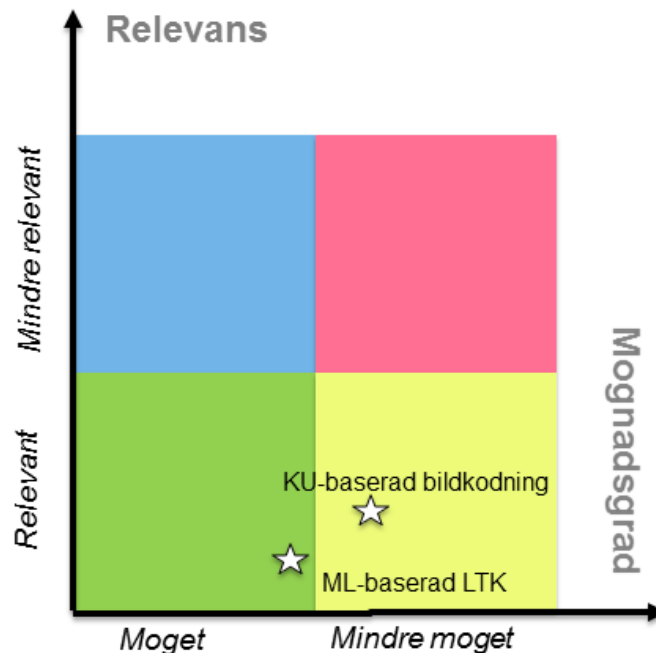
Rapporten omfattar en avskanning av områdena *komprimering* (kap. 2), *informationsfusion* (kap. 3), och *dataanalys* (kap. 4). Avskanningen är utförd under årskiftet 2021/2022 och den använda metodiken för de tre områdena redovisas i avsnitt 1.2. Skanningen har resulterat i att en mängd delområden valts ut och studerats närmare, vilket redovisas i respektive kapitel. Nedan följer en kort sammanfattning av de viktigaste slutsatserna för de studerade områdena och jämförande visualisering.

Bild- och talkomprimering bedöms som två viktiga områden när det gäller trådlös överföring av data inom Försvarsmakten. Båda forskningsområdena är sedan lång tid väletablerade. Den tillgängliga datatakten i mobila taktiska radionät är, i förhållande till fasta nät och andra radionät baserade på fast infrastruktur, relativt begränsad. Även om den tekniska utvecklingen över tid sannolikt leder till en ökning av de möjliga datatakterna kommer denna typ av radionät att vara begränsade till relativt låga datatakt även i framtiden. En stor del av de data som överförs i dessa nät består av digitaliserat krypterat tal. En effektivare talkomprimering (med bibehållen eller ökad talkvalité) medger att fler talförbindelser kan överföras vid en given kapacitet hos nätet. Forskningen inom lågtaktstalkodare baserade på nyare tekniker som maskininläring bedöms på sikt kunna ge en signifikant ökad komprimeringsgrad av talet.

Den begränsade datatakten i mobila taktiska radionät medger inte överföring av högupplöst video, men utvecklingen inom bildkomprimeringsområdet bör kunna förbättra möjligheten att överföra stillbilder och långsam video. Ett annat tillämpningsområde, där forskningen inom videokomprimering kan få stor betydelse, är videolänkar för högupplöst video från små obemannade flygande farkoster (UAV:er). Radioförbindelsen utgörs i dessa fall oftast av en frisiktsförbindelse som medger en relativt hög datatakt. Befintliga videokomprimeringstekniker för högupplöst video kräver stora beräkningsresurser vid realtidskomprimering, vilket kan vara svårt att realisera på små UAV:er. Med videokomprimeringstekniker baserade på kompressiv uppfattning finns en möjlighet att flytta beräkningskomplexitet från kodaren till avkodaren.

I figur 5.1 görs en grov jämförelse mellan metoder för komprimering med avseende på uppskattad *mognad* och *relevans* för Försvarsmakten. Med egenskapen *mognad* avses hur välutvecklat området är i termer av möjligheten att bidra till praktiska tillämpningar. Försvarande i sammanhanget kan vara att algoritmer mest kan hantera förenklade problem, eller att tillämpningen kräver svåråtkomlig data eller expertkunskaper. Egenskapen *relevans* handlar om i vilken utsträckning området kan tänkas bidra till försvarstillämpningar. De avskannade områdenas placering i diagrammet indikeras med en ☆-symbol. Ju närmare origo ett område placerats desto mer intressant är det ur ett försvarstillämpningsperspektiv. I figur 5.1 jämförs två områden, maskininlärningsbaserad lågtaktstalkodning (ML-baserad LTK) och bildkodning baserad på kompressiv uppfattning (KU-baserad bildkodning).

Informationsfusionsområdet består av flera delområden med mycket olika mognadsgrad. Vi bedömer att utvecklingen kommer att gå fort i de högre systemområdena där fusion kombineras med artificiell intelligens, optimering, med mera. Områden som fusion för beslutstöd och nya metoder för spaning, t.ex. spaning med flockar (eng. *emphswarms*) av UAV:er, är områden som förutspås få en snabb utveckling. Informationsfusion bör dock inte betraktas som en separat metodik som endast ska leverera ett underlag för agerande, utan



Figur 5.1. I figuren jämförs tekniker för *komprimering* som omfattas av skanningen med avseende på uppskattad mognadsgrad och relevans (nytta) för försvarstillämpningar.

som en metodik integrerad med artificiell intelligens, med mera, i system som tar fram föreslagna och motiverade alternativ för agerande på alla nivåer.

I figur 5.2 görs en jämförelse mellan de studerade delområdena inom informationsfusion. Intressanta observationer från avskanningen av dataanalysområdet är:

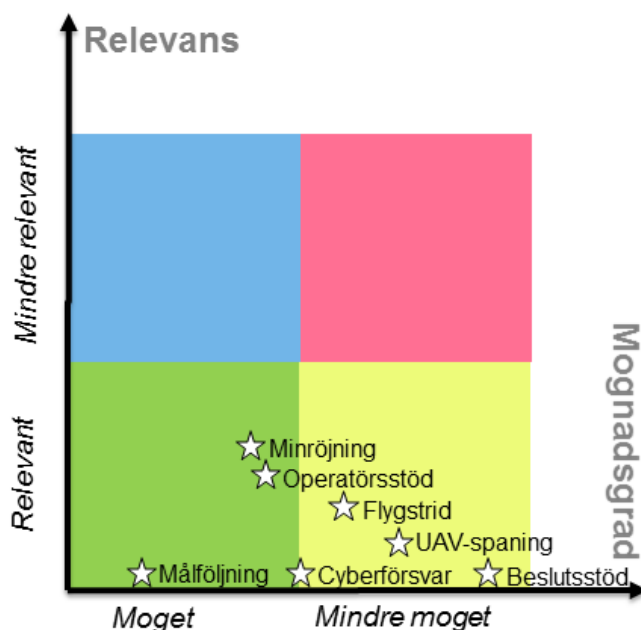
Neurosymboliskt lärande och resonerande (NSR): Har mognat rejält de senaste åren, men behöver mogna ytterligare för att passa skarpa försvarstillämpningar. Försvarsintresset för området är stort, vilket inte minst framgår av Darpas engagemang i området.

Kvantdatorstödd dataanalys (QAI): Kvantdatorer har potential att revolutionera isolerade delar av försvarstillämpningar, men både hårdvara och mjukvara har till synes fortfarande en lång väg att gå.

Kvantisering: I de fall man har en stor datamängd men begränsat med SPMD-beräkningskapacitet, t.ex. på grund av otillräckligt fysiskt utrymme eller effekttillgänglighet, är tekniken ett lovande tillskott som kommer att återfinnas i försvarstillämpningar.

Parameterservrar: Fortfarande används parameterservrar knappt trots deras utlovade potential för maskininlärningsingenjörer att dela modeller, eller parameter vikter, med varandra samt skala upp till enorma modellstorlekar. Vi ser tyvärr inte att området kommer ge frukt förrän man på ett bättre sätt löst modellogistik till SPMD-processenheterna. Det finns inget enhetligt paradigm för att effektivt använda dem och varje tillämpning tycks erbjuda sin egen lösning på modelldelning.

Cellutvärderade anteckningsblock: Används som pedagogiska verktyg i undervisning, för att dela experimentuppsättningar bland forskare, för att ge experimentkodsrekommendationer och för att programmera enheter med nedladdat programmerbar åtkomst. Vi tror att man kommer att se



Figur 5.2. I figuren jämförs tekniker för *informationfusion* som omfattas av skanningen med avseende på uppskattad mognadsgrad och relevans (nytta) för försvarstillämpningar.

en fortsatt utveckling av dessa där de blir mer integrerade, likt vanliga programpaket, i analyskedjor, dvs. parametriserade i ett återbrukssyfte.

Uppmärksamhetsbaserade modeller: Har visat på stor tillförlitlighet, möjlighet att återanvändas för många uppgifter och kommer att minska i storlek allt eftersom vi lär oss att mer aggresivt externalisera omvärldsinformation från dem. Vi kommer att se dessa modeller i bruk i många olika domänområden då grunddatamängden, ostrukturerade textuella beskrivningar av verkligheten, är det enda sättet vi människor hittills lyckats fånga semantik, och därmed förståelse av vår omvärld.

Frågeställningsframställan: Vi ser att man även i framtiden kommer att dela olika former av frågeframställaninstanser och det kommer ske fortsatt forskning för att undersöka vilka former som ger bäst resultat för olika domänområden. De militära tillämpningarna kommer komma igång efter att man lyckats koppla samman området med sökningar för externaliserad omvärldsinformation. Vi ser att finstämda komponenter kommer fortsätta att delas genom repositorer, som idag, men att det även kommer utvecklas mot mer deklarativa delningsformalismen.

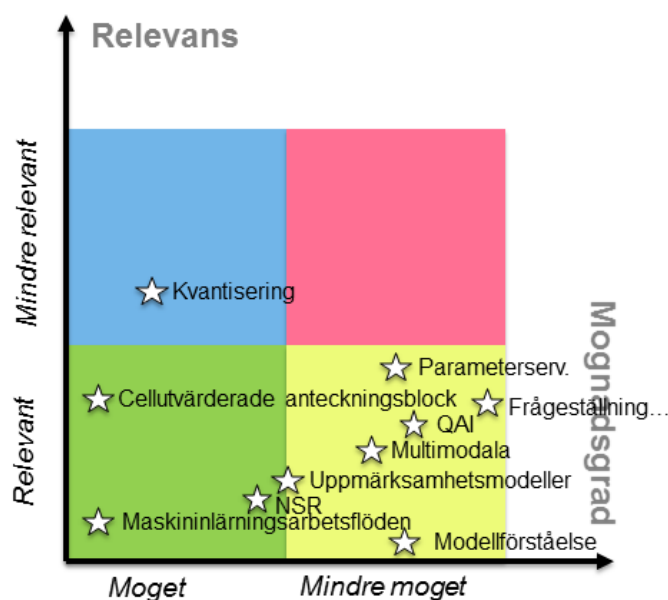
Multimodala modeller: Multimodala modeller har länge varit av intresse inom informationsfusionen på grund av sin förmåga att hjälpa målföljning. Vi ser att man kommer fortsätta utvecklingen med att använda transformerare som ett intermediärt format för multimodal syntes.

Maskininlärningsflöden: Som en logisk fortsättning på cellutvärderade anteckningsblock och dataflödesprogrammering kommer arbetsflöden att ta en mer central plats i analysarbetet, dessa kommer även att utvecklas mot mer deklarativa formalismer, där strävan i utvecklingen av formalismerna är finkornighet för att uppnå reproducerbarhet. Dessa kommer att ha stor betydelse för Försvarsmakten och andra svenska myndigheter eftersom de kommer att ha möjliggöra delning av modellframställan utan att nödvändigtvis behöva dela parametrar eller grunddata.

Modellförståelse: Vi har sett en imponerande teoretisk föreståelseframgång detta år, som kommer att fortsätta att ske. De militära tillämpningarna kommer dock att handla mer om att värdera partiskhet i beslutsunderlag och skifta fokus mot datakvalitetsgarantier och dataprovenans. Detta är alltså utöver all verksamhet som redan sker inom prediktionsförståelse, där den mesta energin tycks läggas och där det redan finns bra tillämpade system för utvärdering.

Från erfarenheter i skilda verksamheter inom totalförsvaret och de framtida ökande kraven på dataanalysen, från lagring och (åter-)användning, är det även tydligt att det existerar ett överskridande säkert och ackrediterat behov att, för myndigheter under försvarsdepartementet, dela datamängder, maskinlärningsresultat samt hårdvara för utvärderande av analysmodeller samt träningsresurser.

I figur 5.3 görs en jämförelse mellan dessa resultat.



Figur 5.3. I figuren jämförs tekniker för *dataanalys* som omfattas av skanningen med avseende på uppskattad mognadsgrad och relevans (nytta) för försvarstillämpningar.

Referenslista

- [1] Samira Abnar m.fl. "Exploring the Limits of Large Scale Pre-training". 5 okt. 2021. arXiv: 2110.02095 [cs, stat]. URL: <http://arxiv.org/abs/2110.02095> (hämtad 2022-01-28) (se s. 43).
- [2] Alessandro Achille m.fl. "Task2Vec: Task Embedding for Meta-Learning". I: *2019 IEEE/CVF International Conference on Computer Vision, ICCV 2019, Seoul, Korea (South), October 27 - November 2, 2019*. IEEE, 2019, s. 6429–6438. DOI: 10.1109/ICCV.2019.00653 (se s. 42).
- [3] Simon Ahlberg m.fl. "An Information Fusion Demonstrator for Tactical Intelligence Processing in Network-Based Defense". I: *Information Fusion*. Special Issue on the Seventh International Conference on Information Fusion-Part II 8.1 (1 jan. 2007), s. 84–107. ISSN: 1566-2535. DOI: 10.1016/j.inffus.2005.11.002. URL: <https://www.sciencedirect.com/science/article/pii/S1566253505000928> (hämtad 2022-02-26) (se s. 27).
- [4] Haseeb Ahmed m.fl. "Adaptive Filtering on GPS-Aided MEMS-IMU for Optimal Estimation of Ground Vehicle Trajectory". I: *Sensors* 19.24 (24 jan. 2019), s. 5357. ISSN: 1424-8220. DOI: 10.3390/s19245357. URL: <https://www.mdpi.com/1424-8220/19/24/5357> (hämtad 2022-02-26) (se s. 24, 25).
- [5] Jaimeen Ahn och Alice Oh. "Mitigating Language-Dependent Ethnic Bias in BERT". 14 sept. 2021. arXiv: 2109.05704 [cs]. URL: <http://arxiv.org/abs/2109.05704> (hämtad 2022-02-10) (se s. 44).
- [6] Jules Allegrè och Romain Vuillemot. "Visualizing and Analyzing Disputed Areas in Soccer". I: *2020 Visualization in Data Science (VDS)*. 2020 IEEE Visualization in Data Science (VDS). Salt Lake City, UT, USA: IEEE, okt. 2020, s. 52–60. ISBN: 978-1-72819-284-0. DOI: 10.1109/VDS51726.2020.00010. URL: <https://ieeexplore.ieee.org/document/9355217/> (hämtad 2022-02-23) (se s. 45).
- [7] Elias Amselem m.fl. *Kvantinformatik - teknisk prognos*. R FOI-R-3920-SE. FOI, 2014 (se s. 39–41, 68).
- [8] Vamsi Aribandi m.fl. "ExT5: Towards Extreme Multi-Task Scaling for Transfer Learning". 21 nov. 2021. arXiv: 2111.10952 [cs]. URL: <http://arxiv.org/abs/2111.10952> (hämtad 2022-01-29) (se s. 42).
- [9] J. M. Arrazola m.fl. "Quantum Circuits with Many Photons on a Programmable Nanophotonic Chip". I: *Nature* 591.7848 (7848 mars 2021), s. 54–60. ISSN: 1476-4687. DOI: 10.1038/s41586-021-03202-1. URL: <https://www.nature.com/articles/s41586-021-03202-1> (hämtad 2022-02-25) (se s. 40).
- [10] Arun Babu m.fl. "XLS-R: Self-supervised Cross-lingual Speech Representation Learning at Scale". 16 dec. 2021. arXiv: 2111.09296 [cs, eess]. URL: <http://arxiv.org/abs/2111.09296> (hämtad 2022-01-28) (se s. 44).
- [11] Stephen H. Bach m.fl. "PromptSource: An Integrated Development Environment and Repository for Natural Language Prompts". 2 febr. 2022. arXiv: 2202.01279 [cs]. URL: <http://arxiv.org/abs/2202.01279> (hämtad 2022-02-15) (se s. 43).

- [12] Samy Badreddine m.fl. *Logic Tensor Networks*. 2021. URL: <https://arxiv.org/abs/2012.13635> (hämtad 2022-01-28) (se s. 37).
- [13] Alexei Baevski m.fl. "Wav2vec 2.0: A Framework for Self-Supervised Learning of Speech Representations". 22 okt. 2020. arXiv: 2006.11477 [cs, eess]. URL: <http://arxiv.org/abs/2006.11477> (hämtad 2022-01-28) (se s. 44).
- [14] Trapit Bansal m.fl. "Diverse Distributions of Self-Supervised Tasks for Meta-Learning in NLP". I: *Proceedings of the 2021 Conference on Empirical Methods in Natural Language Processing*. Online and Punta Cana, Dominican Republic: Association for Computational Linguistics, nov. 2021, s. 5812–5824. DOI: 10.18653/v1/2021.emnlp-main.469. URL: <https://aclanthology.org/2021.emnlp-main.469> (hämtad 2022-02-14) (se s. 42).
- [15] Ankur Bapna m.fl. "SLAM: A Unified Encoder for Speech and Language Modeling via Speech-Text Joint Pre-Training". 19 okt. 2021. arXiv: 2110.10329 [cs]. URL: <http://arxiv.org/abs/2110.10329> (hämtad 2022-01-28) (se s. 44).
- [16] Evgeny Belyaev. "Compressive Sensed Video Coding Having JPEG Compatibility". I: *IEEE International Conference on Image Processing (ICIP)*. 2020, s. 1128–1132. DOI: 10.1109/ICIP40778.2020.9190857 (se s. 11).
- [17] Emily M. Bender m.fl. "On the Dangers of Stochastic Parrots: Can Language Models Be Too Big? U+1F99C". I: *Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency*. FAccT '21. New York, NY, USA: Association for Computing Machinery, 2021, s. 610–623. ISBN: 978-1-4503-8309-7. DOI: 10.1145/3442188.3445922. URL: <https://doi.org/10.1145/3442188.3445922> (hämtad 2022-02-14) (se s. 44).
- [18] Yoshua Bengio. "The Rise of AI". I: *Scientific American* 2016 (June 2016), s. 45–51 (se s. 35).
- [19] Malavika Bhaskaranand och Jerry D. Gibson. "Low-Complexity Video Encoding for UAV Reconnaissance and Surveillance". I: *Proc. Military Communications Conference (MILCOM)*. 2011, s. 1633–1638. DOI: 10.1109/MILCOM.2011.6127543 (se s. 11).
- [20] Abeba Birhane m.fl. "Multimodal Datasets: Misogyny, Pornography, and Malignant Stereotypes". 5 okt. 2021. arXiv: 2110.01963 [cs]. URL: <http://arxiv.org/abs/2110.01963> (hämtad 2022-02-10) (se s. 44).
- [21] Raz Birman m.fl. "Overview of Research in the Field of Video Compression Using Deep Neural Networks". I: *Multimedia tools and applications* 79 (2020), s. 11699–11722. DOI: 10.1007/s11042-019-08572-3 (se s. 11).
- [22] Rishi Bommasani m.fl. "On the Opportunities and Risks of Foundation Models". 18 aug. 2021. arXiv: 2108.07258 [cs]. URL: <http://arxiv.org/abs/2108.07258> (hämtad 2021-08-23) (se s. 42).
- [23] J. V. Boone och R.R. Peterson. *The Start of the Digital Revolution: SIGSALY Secure Digital Voice Communications in World War II*. Juli 2000. URL: <https://www.nsa.gov/portals/75/documents/about/cryptologic-heritage/historical-figures-publications/publications/wwii/sigsaly.pdf> (hämtad 2022-03-19) (se s. 12).

- [24] Sebastian Borgeaud m.fl. "Improving Language Models by Retrieving from Trillions of Tokens". 7 febr. 2022. arXiv: 2112.04426 [cs]. URL: <http://arxiv.org/abs/2112.04426> (hämtad 2022-02-10) (se s. 35, 43).
- [25] Henrik Boström m.fl. *On the Definition of Information Fusion as a Field of Research*. HS-IKI-TR-07-006. Institutionen för kommunikation och information, 2007. URL: <http://urn.kb.se/resolve?urn=urn:nbn:se:his:diva-1256> (hämtad 2022-02-28) (se s. 15).
- [26] Gilles Brassard m.fl. "Quantum Amplitude Amplification and Estimation". I: *AMS Contemporary Mathematics* 305 (2002), s. 53–74. DOI: 10.1090/conm/305/05215. arXiv: quant-ph/0005055. URL: <http://arxiv.org/abs/quant-ph/0005055> (hämtad 2022-02-25) (se s. 39).
- [27] Michael Broughton m.fl. "TensorFlow Quantum: A Software Framework for Quantum Machine Learning". 26 aug. 2021. arXiv: 2003.02989 [cond-mat, physics:quant-ph]. URL: <http://arxiv.org/abs/2003.02989> (hämtad 2022-02-25) (se s. 69).
- [28] Tom B. Brown m.fl. "Language Models Are Few-Shot Learners". 22 juli 2020. arXiv: 2005.14165 [cs]. URL: <http://arxiv.org/abs/2005.14165> (hämtad 2022-02-09) (se s. 43).
- [29] Sébastien Bubeck och Mark Sellke. "A Universal Law of Robustness via Isoperimetry". 21 okt. 2021. arXiv: 2105.12806 [cs, stat]. URL: <http://arxiv.org/abs/2105.12806> (hämtad 2022-02-11) (se s. 44).
- [30] Lorenzo Buffoni och Filippo Caruso. "New Trends in Quantum Machine Learning (a)". I: *Europhysics Letters* 132.6 (dec. 2020), s. 60004. ISSN: 0295-5075. DOI: 10.1209/0295-5075/132/60004. URL: <https://doi.org/10.1209/0295-5075/132/60004> (hämtad 2022-02-28) (se s. 41).
- [31] Emanuele Bugliarello m.fl. "Multimodal Pretraining Unmasked: A Meta-Analysis and a Unified Framework of Vision-and-Language BERTs". I: *Transactions of the Association for Computational Linguistics* 9 (8 sept. 2021), s. 978–994. ISSN: 2307-387X. DOI: 10.1162/tac1_a_00408. URL: https://doi.org/10.1162/tac1_a_00408 (hämtad 2022-01-28) (se s. 44).
- [32] Emmanuel J. Candes och Michael B. Wakin. "An Introduction to Compressive Sampling". I: *IEEE Signal Processing Magazine* 25.2 (2008), s. 21–30. DOI: 10.1109/MSP.2007.914731 (se s. 10).
- [33] Milos Cernak m.fl. "Composition of Deep and Spiking Neural Networks for Very Low Bit Rate Speech Coding". I: *IEEE/ACM Transactions on Audio, Speech, and Language Processing* 24.12 (2016), s. 2301–2312. DOI: 10.1109/TASLP.2016.2604566 (se s. 13).
- [34] Milos Cernak m.fl. "Phonological Vocoding Using Artificial Neural Networks". I: *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*. 2015, s. 4844–4848. DOI: 10.1109/ICASSP.2015.7178891 (se s. 13).
- [35] M.W. Chamberlain. "A 600 Bps MELP Vocoder for Use on HF Channels". I: *Proc. Military Communications Conference (MILCOM)*. Vol. 1. 2001, s. 447–453. DOI: 10.1109/MILCOM.2001.985836 (se s. 13).
- [36] François Charton m.fl. "Learning Advanced Mathematical Computations from Examples". 19 mars 2021. arXiv: 2006.06462 [cs]. URL: <http://arxiv.org/abs/2006.06462> (hämtad 2022-02-10) (se s. 42).

- [37] Guangyi Chen m.fl. "Deep Meta Metric Learning". I: *2019 IEEE/CVF International Conference on Computer Vision, ICCV 2019, Seoul, Korea (South), October 27 - November 2, 2019*. IEEE, 2019, s. 9546–9555. DOI: 10.1109/ICCV.2019.00964 (se s. 42).
- [38] Lili Chen m.fl. "Decision Transformer: Reinforcement Learning via Sequence Modeling". 24 juni 2021. arXiv: 2106.01345 [cs]. URL: <http://arxiv.org/abs/2106.01345> (hämtad 2022-01-28) (se s. 44).
- [39] Shuai Chen m.fl. "Low-Altitude Protection Technology of Anti-UAVs Based on Multisource Detection Information Fusion". I: *International Journal of Advanced Robotic Systems* 17.5 (1 sept. 2020). ISSN: 1729-8806. DOI: 10.1177/1729881420962907. URL: <https://doi.org/10.1177/1729881420962907> (hämtad 2022-02-26) (se s. 24).
- [40] Tong Chen m.fl. "DeepCoder: A Deep Neural Network Based Video Compression". I: *IEEE Visual Communications and Image Processing (VCIP)*. 2017, s. 1–4. DOI: 10.1109/VCIP.2017.8305033 (se s. 11).
- [41] Mariusz Chmielewski m.fl. "Application of Augmented Reality, Mobile Devices, and Sensors for a Combat Entity Quantitative Assessment Supporting Decisions and Situational Awareness Development". I: *Applied Sciences* 9.21 (21 jan. 2019), s. 4577. ISSN: 2076-3417. DOI: 10.3390/app9214577. URL: <https://www.mdpi.com/2076-3417/9/21/4577> (hämtad 2022-02-26) (se s. 26, 27).
- [42] Min Jin Chong och David Forsyth. "JoJoGAN: One Shot Face Stylization". 2 febr. 2022. arXiv: 2112.11641 [cs]. URL: <http://arxiv.org/abs/2112.11641> (hämtad 2022-02-18) (se s. 44).
- [43] Isaac L. Chuang m.fl. "Experimental Implementation of Fast Quantum Searching". I: *Physical Review Letters* 80.15 (13 april 1998), s. 3408–3411. DOI: 10.1103/PhysRevLett.80.3408. URL: <https://link.aps.org/doi/10.1103/PhysRevLett.80.3408> (hämtad 2022-02-20) (se s. 40).
- [44] Yu-An Chung m.fl. "W2v-BERT: Combining Contrastive Learning and Masked Language Modeling for Self-Supervised Speech Pre-Training". 13 sept. 2021. arXiv: 2108.06209 [cs, eess]. URL: <http://arxiv.org/abs/2108.06209> (hämtad 2022-01-28) (se s. 44).
- [45] Darius Coelho m.fl. "PeckVis: A Visual Analytics Tool to Analyze Dominance Hierarchies in Small Groups". I: *IEEE Transactions on Visualization and Computer Graphics* 26.4 (1 april 2020), s. 1650–1660. ISSN: 1077-2626, 1941-0506, 2160-9306. DOI: 10.1109/TVCG.2020.2969056. URL: <https://ieeexplore.ieee.org/document/8984335/> (hämtad 2022-02-23) (se s. 45).
- [46] J.S. Collura m.fl. "The 1.2 Kbps/2.4 Kbps MELP Speech Coding Suite with Integrated Noise Pre-Processing". I: *MILCOM 1999. IEEE Military Communications Conference Proceedings (Cat. No.99CH36341)*. Vol. 2. Okt. 1999, s. 1449–1453. DOI: 10.1109/MILCOM.1999.821443 (se s. 13).
- [47] Joseph Cottam m.fl. "Graph Comparison for Causal Discovery". I: *Visualization in data science* (2021), s. 10 (se s. 45).
- [48] Anamaria Crisan m.fl. "Passing the Data Baton : A Retrospective Analysis on Data Science Work and Workers". I: *IEEE Transactions on Visualization and Computer Graphics* 27.2 (febr. 2021), s. 1860–1870. ISSN: 1077-2626, 1941-0506, 2160-9306. DOI: 10.1109/TVCG.2020.3030340. URL: <https://ieeexplore.ieee.org/document/9222030/> (hämtad 2022-02-23) (se s. 33).

- [49] Artur S. d'Avila Garcez. "Fewer Epistemological Challenges for Connectionism". I: vol. 3526. Lecture Notes in Computer Science. First Conference on Computability in Europe, CiE 2005. Amsterdam, 2005, s. 289–325. DOI: 10.1007/11494645_18. URL: <https://openaccess.city.ac.uk/id/eprint/296/> (hämtad 2022-01-14) (se s. 37).
- [50] Artur S. d'Avila Garcez och Luís C. Lamb. *Neurosymbolic AI: The 3rd Wave*. 2020. URL: <https://arxiv.org/abs/2012.05876> (hämtad 2022-01-14) (se s. 37).
- [51] Artur S. d'Avila Garcez m.fl. "Neural-Symbolic Learning and Reasoning: Contributions and Challenges". I: 2015 AAAI Spring Symposium. 2015, s. 18–21 (se s. 37).
- [52] Subhajit Das och Alex Endert. "CACTUS: Detecting and Resolving Conflicts in Objective Functions". I: *Visualization in data science* (), s. 10 (se s. 45).
- [53] Subhajit Das och Alex Endert. "LEGION: Visually Compare Modeling Techniques for Regression". I: *2020 Visualization in Data Science (VDS)*. 2020 IEEE Visualization in Data Science (VDS). Salt Lake City, UT, USA: IEEE, okt. 2020, s. 12–21. ISBN: 978-1-72819-284-0. DOI: 10.1109/VDS51726.2020.00006. URL: <https://ieeexplore.ieee.org/document/9355220/> (hämtad 2022-02-23) (se s. 45).
- [54] Mark S. Dennison m.fl. "Improving Motion Sickness Severity Classification through Multi-Modal Data Fusion". I: *Artificial Intelligence and Machine Learning for Multi-Domain Operations Applications*. Artificial Intelligence and Machine Learning for Multi-Domain Operations Applications. Utg. av Tien Pham. Baltimore, United States: SPIE, 10 maj 2019, s. 27. ISBN: 978-1-5106-2677-5 978-1-5106-2678-2. DOI: 10.1117/12.2519085. URL: <https://www.spiedigitallibrary.org/conference-proceedings-of-spie/11006/2519085/Improving-motion-sickness-severity-classification-through-multi-modal-data-fusion/10.1117/12.2519085.full> (hämtad 2022-02-26) (se s. 26).
- [55] Tim Dettmers m.fl. "8-Bit Optimizers via Block-wise Quantization". 6 okt. 2021. arXiv: 2110.02861 [cs]. URL: <http://arxiv.org/abs/2110.02861> (hämtad 2022-02-10) (se s. 34).
- [56] Prafulla Dhariwal och Alex Nichol. "Diffusion Models Beat GANs on Image Synthesis". 1 juni 2021. arXiv: 2105.05233 [cs, stat]. URL: <http://arxiv.org/abs/2105.05233> (hämtad 2022-02-10) (se s. 44).
- [57] Ben Dickson. "The Future of Deep Learning, According to Its Pioneers". I: (juli 2021). URL: <https://bdtechtalks.com/2021/07/01/deep-learning-future-bengio-hinton-lecun/> (hämtad 2022-01-13) (se s. 35).
- [58] Kaize Ding m.fl. "Few-Shot Network Anomaly Detection via Cross-Network Meta-Learning". I: *WWW '21: The Web Conference 2021, Virtual Event / Ljubljana, Slovenia, April 19-23, 2021*. Utg. av Jure Leskovec m.fl. ACM / IW3C2, 2021, s. 2448–2456. DOI: 10.1145/3442381.3449922 (se s. 42).
- [59] Jesse Dodge m.fl. "Documenting Large Webtext Corpora: A Case Study on the Colossal Clean Crawled Corpus". 30 sept. 2021. arXiv: 2104.08758 [cs]. URL: <http://arxiv.org/abs/2104.08758> (hämtad 2022-02-10) (se s. 44).

- [60] Alexey Dosovitskiy m.fl. "An Image Is Worth 16x16 Words: Transformers for Image Recognition at Scale". 3 juni 2021. arXiv: 2010.11929 [cs]. URL: <http://arxiv.org/abs/2010.11929> (hämtad 2022-01-28) (se s. 42).
- [61] Vincent Dumoulin m.fl. "Comparing Transfer and Meta Learning Approaches on a Unified Few-Shot Classification Benchmark". 6 april 2021. arXiv: 2104.02638 [cs]. URL: <http://arxiv.org/abs/2104.02638> (hämtad 2022-02-10) (se s. 42).
- [62] Sepehr Ebadi m.fl. "Quantum Optimization of Maximum Independent Set Using Rydberg Atom Arrays". 18 febr. 2022. arXiv: 2202.09372 [cond-mat, physics:physics, physics:quant-ph]. URL: <http://arxiv.org/abs/2202.09372> (hämtad 2022-02-25) (se s. 40).
- [63] Constantin Eichenberg m.fl. "MAGMA – Multimodal Augmentation of Generative Models through Adapter-based Finetuning". 9 dec. 2021. arXiv: 2112.05253 [cs]. URL: <http://arxiv.org/abs/2112.05253> (hämtad 2022-02-10) (se s. 43).
- [64] Joscha Eirich m.fl. "VIMA: Modeling and Visualization of High Dimensional Machine Sensor Data Leveraging Multiple Sources of Domain Knowledge". I: *2020 Visualization in Data Science (VDS)*. 2020 IEEE Visualization in Data Science (VDS). Salt Lake City, UT, USA: IEEE, okt. 2020, s. 22–31. ISBN: 978-1-72819-284-0. DOI: 10.1109/VDS51726.2020.00007. URL: <https://ieeexplore.ieee.org/document/9355215/> (hämtad 2022-02-23) (se s. 45).
- [65] Stefan Erhardt m.fl. "An Open-Source Speech Codec at 450 Bit/s with Pseudo Wideband Mode". I: *Proc. 49th European Microwave Conference (EuMC)*. Paris, France, okt. 2019, s. 1048–1051 (se s. 13).
- [66] Richard P. Feynman. "Simulating Physics with Computers". I: *International Journal of Theoretical Physics* 21.6 (1 juni 1982), s. 467–488. ISSN: 1572-9575. DOI: 10.1007/BF02650179. URL: <https://doi.org/10.1007/BF02650179> (hämtad 2022-02-13) (se s. 38).
- [67] Tsu-Jui Fu m.fl. "VIOLET : End-to-End Video-Language Transformers with Masked Visual-token Modeling". 24 nov. 2021. arXiv: 2111.12681 [cs]. URL: <http://arxiv.org/abs/2111.12681> (hämtad 2022-01-28) (se s. 44).
- [68] Cristina Gărbacea m.fl. "Low Bit-Rate Speech Coding with VQ-VAE and a WaveNet Decoder". I: *Proc. IEEE Int. Conf. on Acoustics, Speech and Signal Processing (ICASSP)*. 2019, s. 735–739. DOI: 10.1109/ICASSP.2019.8683277. arXiv: 1910.06464 (se s. 13).
- [69] Quanbo Ge m.fl. "Analysis on Strong Tracking Filtering for Linear Dynamic Systems". I: *Mathematical Problems in Engineering* 2015 (10 sept. 2015), e648125. ISSN: 1024-123X. DOI: 10.1155/2015/648125. URL: <https://www.hindawi.com/journals/mpe/2015/648125/> (hämtad 2022-02-26) (se s. 24).
- [70] Yuxia Geng m.fl. "OntoZSL: Ontology-enhanced Zero-Shot Learning". I: *WWW '21: The Web Conference 2021, Virtual Event / Ljubljana, Slovenia, April 19-23, 2021*. Utg. av Jure Leskovec m.fl. ACM / IW3C2, 2021, s. 3325–3336. DOI: 10.1145/3442381.3450042 (se s. 42).
- [71] Mohammad Ghasemzadeh m.fl. "ReBNet: Residual Binarized Neural Network". 27 mars 2018. arXiv: 1711.01243 [cs]. URL: <http://arxiv.org/abs/1711.01243> (hämtad 2022-01-10) (se s. 34).

- [72] Daniele Giacobello m.fl. "Retrieving Sparse Patterns Using a Compressed Sensing Framework: Applications to Speech Coding Based on Sparse Linear Prediction". I: *IEEE Signal Processing Letters* 17.1 (2010), s. 103–106. DOI: 10.1109/LSP.2009.2034560 (se s. 13).
- [73] Sukhpal Singh Gill m.fl. "Quantum Computing: A Taxonomy, Systematic Review and Future Directions". I: *Software: Practice and Experience* 52.1 (2022), s. 66–114. ISSN: 1097-024X. DOI: 10.1002/spe.3039. URL: <https://onlinelibrary.wiley.com/doi/abs/10.1002/spe.3039> (hämtad 2022-02-24) (se s. 40).
- [74] Ian Goodfellow m.fl. *Deep Learning*. Utg. av Francis Bach. Adaptive Computation and Machine Learning Series. Cambridge, MA, USA: MIT Press, 2016. 800 s. ISBN: 978-0-262-03561-3. URL: <https://www.deeplearningbook.org/> (hämtad 2022-02-14) (se s. 15).
- [75] Edouard Grave m.fl. "Improving Neural Language Models with a Continuous Cache". I: (13 dec. 2016). URL: <https://arxiv.org/abs/1612.04426v1> (hämtad 2022-02-22) (se s. 35).
- [76] Lov K. Grover. "Quantum Mechanics Helps in Searching for a Needle in a Haystack". I: *Physical Review Letters* 79.2 (14 juli 1997), s. 325–328. ISSN: 0031-9007, 1079-7114. DOI: 10.1103/PhysRevLett.79.325. arXiv: quant-ph/9706033. URL: <http://arxiv.org/abs/quant-ph/9706033> (hämtad 2022-02-17) (se s. 39).
- [77] G. Guilmin m.fl. "New NATO STANAG Narrow Band Voice Coder at 600 Bits/s". I: *IEEE International Conference on Acoustics Speech and Signal Processing*. Vol. 1. Toulouse, France, 2006. DOI: 10.1109/ICASSP.2006.1660114 (se s. 13).
- [78] Zhichun Guo m.fl. "Few-Shot Graph Learning for Molecular Property Prediction". I: *WWW '21: The Web Conference 2021, Virtual Event / Ljubljana, Slovenia, April 19-23, 2021*. Utg. av Jure Leskovec m.fl. ACM / IW3C2, 2021, s. 2559–2567. DOI: 10.1145/3442381.3450112 (se s. 42).
- [79] Johan Håstad. *Computational Limitations of Small-Depth Circuits*. Cambridge, MA, USA: MIT Press, 1987. ISBN: 978-0-262-08167-2 (se s. 44).
- [80] Tomoki Hayashi m.fl. "An Investigation of Multi-Speaker Training for Wavenet Vocoder". I: *IEEE Automatic Speech Recognition and Understanding Workshop (ASRU)*. 2017, s. 712–718. DOI: 10.1109/ASRU.2017.8269007 (se s. 13).
- [81] Junxian He m.fl. "Towards a Unified View of Parameter-Efficient Transfer Learning". 2 febr. 2022. arXiv: 2110.04366 [cs]. URL: <http://arxiv.org/abs/2110.04366> (hämtad 2022-02-10) (se s. 44).
- [82] Lisa Anne Hendricks m.fl. "Decoupling the Role of Data, Attention, and Losses in Multimodal Transformers". I: *Transactions of the Association for Computational Linguistics* 9 (8 juli 2021), s. 570–585. ISSN: 2307-387X. DOI: 10.1162/tac1_a_00385. URL: https://doi.org/10.1162/tac1_a_00385 (hämtad 2022-01-28) (se s. 44).
- [83] Itay Hubara m.fl. "Quantized Neural Networks: Training Neural Networks with Low Precision Weights and Activations". 22 sept. 2016. arXiv: 1609.07061 [cs]. URL: <http://arxiv.org/abs/1609.07061> (hämtad 2022-01-10) (se s. 34).

- [84] International Telecommunication Union (ITU). *Advanced Video Coding for Generic Audiovisual Services*. Recommendation ITU-T H.264. Telecommunication standardization sector of ITU, mars 2003. DOI: 11.1002/1000/6312 (se s. 10).
- [85] International Telecommunication Union (ITU). *Advanced Video Coding for Generic Audiovisual Services*. Recommendation ITU-T H.264 (V14). Telecommunication standardization sector of ITU, aug. 2021. DOI: 11.1002/1000/14659 (se s. 11).
- [86] International Telecommunication Union (ITU). *High Efficiency Video Coding*. Recommendation ITU-T H.265. Telecommunication standardization sector of ITU, april 2013. DOI: 11.1002/1000/11885 (se s. 10).
- [87] International Telecommunication Union (ITU). *High Efficiency Video Coding*. Recommendation ITU-T H.265 (V8). Telecommunication standardization sector of ITU, aug. 2021. DOI: 11.1002/1000/14660 (se s. 10).
- [88] Georgios Ioannou m.fl. "A Markov Multi-Phase Transferable Belief Model for Cyber Situational Awareness". I: *IEEE access : practical innovations, open solutions* 7 (2019), s. 39305–39320. ISSN: 2169-3536. DOI: 10.1109/ACCESS.2019.2897923 (se s. 30).
- [89] Alan Jacob m.fl. "Deep Learning Approach to Video Compression". I: *IEEE Bombay Section Signature Conference (IBSSC)*. 2019, s. 1–5. DOI: 10.1109/IBSSC47189.2019.8973035 (se s. 11).
- [90] Andrew Jaegle m.fl. "Perceiver IO: A General Architecture for Structured Inputs & Outputs". 2 aug. 2021. arXiv: 2107.14795 [cs, eess]. URL: <http://arxiv.org/abs/2107.14795> (hämtad 2022-01-30) (se s. 43).
- [91] Andrew Jaegle m.fl. "Perceiver: General Perception with Iterative Attention". 22 juni 2021. arXiv: 2103.03206 [cs, eess]. URL: <http://arxiv.org/abs/2103.03206> (hämtad 2022-01-30) (se s. 43).
- [92] Dietrich Kammer m.fl. "Glyphboard: Visual Exploration of High-Dimensional Data Combining Glyphs with Dimensionality Reduction". I: *IEEE Transactions on Visualization and Computer Graphics* 26.4 (1 april 2020), s. 1661–1671. ISSN: 1077-2626, 1941-0506, 2160-9306. DOI: 10.1109/TVCG.2020.2969060. URL: <https://ieeexplore.ieee.org/document/8967136/> (hämtad 2022-02-23) (se s. 45).
- [93] Srihari Kankanahalli. "End-to-End Optimized Speech Coding with Deep Neural Networks". I: *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*. 2018, s. 2521–2525. DOI: 10.1109/ICASSP.2018.8461487 (se s. 13).
- [94] Amlan Kar m.fl. "Meta-Sim: Learning to Generate Synthetic Datasets". I: *2019 IEEE/CVF International Conference on Computer Vision, ICCV 2019, Seoul, Korea (South), October 27 - November 2, 2019*. IEEE, 2019, s. 4550–4559. DOI: 10.1109/ICCV.2019.00465 (se s. 42).
- [95] Maja Karasalo och Johan Schubert. "Developing Horizon Scanning Methods for the Discovery of Scientific Trends". I: *2019 International Conference on Document Analysis and Recognition (ICDAR)*. 2019 International Conference on Document Analysis and Recognition (ICDAR). Sept. 2019, s. 1055–1062. DOI: 10.1109/ICDAR.2019.00172 (se s. 18, 20).

- [96] Maja Karasalo och Johan Schubert. *HSTOOL for Horizon Scanning of Scientific Literature*. R FOI-R-4760-SE. Swedish Defence Research Agency (FOI), 2019 (se s. 7, 18, 20).
- [97] Angelos Katharopoulos m.fl. "Transformers Are RNNs: Fast Autoregressive Transformers with Linear Attention". 29 juni 2020. arXiv: 2006.16236v3 [cs]. URL: <https://arxiv.org/abs/2006.16236v3> (hämtad 2022-02-22) (se s. 43).
- [98] Urvashi Khandelwal m.fl. "Generalization through Memorization: Nearest Neighbor Language Models". 14 febr. 2020. arXiv: 1911.00172 [cs]. URL: <http://arxiv.org/abs/1911.00172> (hämtad 2022-02-10) (se s. 35).
- [99] Sergey Nikolaevich Kirillov m.fl. "Machine Learning Algorithms Based on Hidden Markov Models in Low-Speed Speech Codecs for Assessing Speech Quality". I: *2nd International Conference on Control Systems, Mathematical Modeling, Automation and Energy Efficiency (SUMMA)*. 2020, s. 408–412. DOI: 10.1109/SUMMA50634.2020.9280754 (se s. 13).
- [100] Jonas Kjäll m.fl. *Kvantteknologier 2020*. D FOI-D-1070-SE. FOI, 2021, s. 30 (se s. 41).
- [101] Wolfgang Koch, utg. *Tracking and Sensor Data Fusion*. 1. utg. Mathematical Engineering. 2014. ISBN: 978-3-642-39270-2. URL: <https://link.springer.com/book/10.1007/978-3-642-39271-9#about> (hämtad 2022-02-14) (se s. 15).
- [102] M.A. Kohler. "A Comparison of the New 2400 Bps MELP Federal Standard with Other Standard Coders". I: *IEEE International Conference on Acoustics, Speech, and Signal Processing*. Vol. 2. 1997, s. 1587–1590. DOI: 10.1109/ICASSP.1997.596256 (se s. 13).
- [103] Mojtaba Komeili m.fl. "Internet-Augmented Dialogue Generation". 15 juli 2021. arXiv: 2107.07566 [cs]. URL: <http://arxiv.org/abs/2107.07566> (hämtad 2022-02-10) (se s. 43).
- [104] A. Krtalić m.fl. "Visualization of Land Mine Danger, Svilaja Region (Croatia)". I: *Journal of Maps* (2018). DOI: 10.1080/17445647.2018.1552209 (se s. 29).
- [105] Andrija Krtalić och Milan Bajić. "Development of the TIRAMISU Advanced Intelligence Decision Support System". I: *European Journal of Remote Sensing* 52.1 (1 jan. 2019), s. 40–55. ISSN: null. DOI: 10.1080/22797254.2018.1550351. URL: <https://doi.org/10.1080/22797254.2018.1550351> (hämtad 2022-02-26) (se s. 28).
- [106] James Lee-Thorp m.fl. "FNet: Mixing Tokens with Fourier Transforms". 9 sept. 2021. arXiv: 2105.03824 [cs]. URL: <http://arxiv.org/abs/2105.03824> (hämtad 2022-02-09) (se s. 43).
- [107] Brian Lester m.fl. "The Power of Scale for Parameter-Efficient Prompt Tuning". 2 sept. 2021. arXiv: 2104.08691 [cs]. URL: <http://arxiv.org/abs/2104.08691> (hämtad 2022-02-10) (se s. 43).
- [108] Chenglin Li m.fl. "Meta-Har: Federated Representation Learning for Human Activity Recognition". I: *WWW '21: The Web Conference 2021, Virtual Event / Ljubljana, Slovenia, April 19-23, 2021*. Utg. av Jure Leskovec m.fl. ACM / IW3C2, 2021, s. 912–922. DOI: 10.1145/3442381.3450006 (se s. 42).

- [109] Jianping Kelvin Li m.fl. "A Visual Analytics Framework for Analyzing Parallel and Distributed Computing Applications". I: *2019 IEEE Visualization in Data Science (VDS)*. 2019 IEEE Visualization in Data Science (VDS). Vancouver, BC, Canada: IEEE, okt. 2019, s. 1–9. ISBN: 978-1-72815-047-5. DOI: 10.1109/VDS48975.2019.8973380. URL: <https://ieeexplore.ieee.org/document/8973380/> (hämtad 2022-02-23) (se s. 45).
- [110] Yan Li m.fl. "Analysis Framework of Network Security Situational Awareness and Comparison of Implementation Methods". I: *EURASIP Journal on Wireless Communications and Networking* 2019.1 (13 aug. 2019), s. 205. ISSN: 1687-1499. DOI: 10.1186/s13638-019-1506-1. URL: <https://doi.org/10.1186/s13638-019-1506-1> (hämtad 2022-02-26) (se s. 29, 30).
- [111] Martin Liggins II m.fl. *Handbook of Multisensor Data Fusion: Theory and Practice*. CRC press, 2017 (se s. 15).
- [112] Anna Lindberg m.fl. *Auskanning av AI-området neuro-symboliskt resonande - Resultat och kortfattat om metod*. R FOI-R-5224—SE. 2021, s. 66 (se s. 38).
- [113] Fangyu Liu m.fl. "Visually Grounded Reasoning across Languages and Cultures". 21 okt. 2021. arXiv: 2109.13238 [cs]. URL: <http://arxiv.org/abs/2109.13238> (hämtad 2022-02-10) (se s. 44).
- [114] Hanxiao Liu m.fl. "Pay Attention to MLPs". 1 juni 2021. arXiv: 2105.08050 [cs]. URL: <http://arxiv.org/abs/2105.08050> (hämtad 2022-02-09) (se s. 43).
- [115] Pengfei Liu m.fl. "Pre-Train, Prompt, and Predict: A Systematic Survey of Prompting Methods in Natural Language Processing". 28 juli 2021. arXiv: 2107.13586 [cs]. URL: <http://arxiv.org/abs/2107.13586> (hämtad 2022-01-30) (se s. 43).
- [116] Ruibo Liu m.fl. "Mitigating Political Bias in Language Models Through Reinforced Calibration". 30 april 2021. arXiv: 2104.14795 [cs]. URL: <http://arxiv.org/abs/2104.14795> (hämtad 2022-02-10) (se s. 44).
- [117] Xiao Liu m.fl. "GPT Understands, Too". 18 mars 2021. arXiv: 2103.10385 [cs]. URL: <http://arxiv.org/abs/2103.10385> (hämtad 2022-02-10) (se s. 43).
- [118] Zechun Liu m.fl. "MetaPruning: Meta Learning for Automatic Neural Network Channel Pruning". I: *2019 IEEE/CVF International Conference on Computer Vision, ICCV 2019, Seoul, Korea (South), October 27 - November 2, 2019*. IEEE, 2019, s. 3295–3304. DOI: 10.1109/ICCV.2019.00339 (se s. 42).
- [119] Ekaterina A. Lopukhova m.fl. "A Codec Simulation for Low-Rate Speech Coding with Radial Neural Networks". I: *Ural Symposium on Biomedical Engineering, Radioelectronics and Information Technology (US-BEREIT)*. 2021, s. 0285–0288. DOI: 10.1109/USBEREIT51232.2021.9455048 (se s. 13).
- [120] Kun Lv m.fl. "Dynamic Defense Strategy against Advanced Persistent Threat under Heterogeneous Networks". I: *Information Fusion* 49 (1 sept. 2019), s. 216–226. ISSN: 1566-2535. DOI: 10.1016/j.inffus.2019.01.001. URL: <https://www.sciencedirect.com/science/article/pii/S1566253518303774> (hämtad 2022-02-26) (se s. 29).

- [121] Siwei Ma m.fl. "Image and Video Compression with Neural Networks: A Review". I: *IEEE Transactions on Circuits and Systems for Video Technology* 30.6 (2020), s. 1683–1698. DOI: 10.1109/TCSVT.2019.2910119 (se s. 11).
- [122] Xuezhe Ma m.fl. "Luna: Linear Unified Nested Attention". 2 nov. 2021. arXiv: 2106.01540 [cs]. URL: <http://arxiv.org/abs/2106.01540> (hämtad 2022-02-10) (se s. 43).
- [123] Rabeeh Karimi Mahabadi m.fl. "Compacter: Efficient Low-Rank Hypercomplex Adapter Layers". 27 nov. 2021. arXiv: 2106.04647 [cs]. URL: <http://arxiv.org/abs/2106.04647> (hämtad 2022-02-10) (se s. 44).
- [124] Hamid Mansoor m.fl. "DELFI: Mislabeled Human Context Detection Using Multi-Feature Similarity Linking". I: *2019 IEEE Visualization in Data Science (VDS)*. 2019 IEEE Visualization in Data Science (VDS). Vancouver, BC, Canada: IEEE, okt. 2019, s. 11–19. ISBN: 978-1-72815-047-5. DOI: 10.1109/VDS48975.2019.8973382. URL: <https://ieeexplore.ieee.org/document/8973382/> (hämtad 2022-02-23) (se s. 45).
- [125] Yuning Mao m.fl. "UniPELT: A Unified Framework for Parameter-Efficient Language Model Tuning". 14 okt. 2021. arXiv: 2110.07577 [cs]. URL: <http://arxiv.org/abs/2110.07577> (hämtad 2022-02-10) (se s. 44).
- [126] Martin D. Martin-Guerrero och Lucas Lamata. "Quantum Machine Learning: A Tutorial - ScienceDirect". I: *Neurocomputing* 470 (2022), s. 457–461. URL: <https://www.sciencedirect.com/science/article/pii/S0925231221011000> (hämtad 2022-01-25) (se s. 40, 67).
- [127] John McCarthy. "Epistemological Challenges for Connectionism". I: *Behavioral and Brain Sciences* 11.1 (1988), s. 44–44 (se s. 37).
- [128] A. McCree. "A Scalable Phonetic Vocoder Framework Using Joint Predictive Vector Quantization of MELP Parameters". I: *IEEE International Conference on Acoustics Speech and Signal Processing*. Vol. 1. Toulouse, France, 2006. DOI: 10.1109/ICASSP.2006.1660118 (se s. 13).
- [129] Sewon Min m.fl. "MetaICL: Learning to Learn In Context". 29 okt. 2021. arXiv: 2110.15943 [cs]. URL: <http://arxiv.org/abs/2110.15943> (hämtad 2022-01-29) (se s. 42).
- [130] Mohammad Hossein Moghaddam m.fl. "A Framework for Super-Resolution of Scalable Video via Sparse Reconstruction of Residual Frames". I: *Proc. Military Communications Conference (MILCOM)*. 2017, s. 164–168. DOI: 10.1109/MILCOM.2017.8170865 (se s. 11).
- [131] Reiichiro Nakano m.fl. "WebGPT: Browser-assisted Question-Answering with Human Feedback". 17 dec. 2021. arXiv: 2112.09332 [cs]. URL: <http://arxiv.org/abs/2112.09332> (hämtad 2022-02-10) (se s. 43).
- [132] Alex Nichol m.fl. "GLIDE: Towards Photorealistic Image Generation and Editing with Text-Guided Diffusion Models". 22 dec. 2021. arXiv: 2112.10741 [cs]. URL: <http://arxiv.org/abs/2112.10741> (hämtad 2022-02-10) (se s. 44).
- [133] James Nightingale m.fl. "Enhancing Visual Communications Capabilities in Tactical Networks". I: *Proc. Int. Conf. Military Communications and Information Systems (ICMCIS)*. 2015. DOI: 10.1109/ICMCIS.2015.7158692 (se s. 11).

- [134] North Atlantic Treaty Organisation (NATO). *The 600 Bit/s, 1200 Bit/s and 2400 Bit/s NATO Interoperable Narrow Band Voice Coder*. Standardization agreement (STANAG) 4591. NATO standardisation agency, 2006 (se s. 13).
- [135] Irena Orović m.fl. "Compressive Sensing in Signal Processing: Algorithms and Transform Domain Formulation". I: *Mathematical Problems in Engineering* 2016 (2016), s. 1–16. DOI: 10.1155/2016/7616393. URL: <https://doi.org/10.1155/2016/7616393> (hämtad 2022-03-19) (se s. 10).
- [136] Moslem Ouled Sghaier m.fl. *Relevance and Importance in Deep Learning for Open Source Data Processing to Enhance Context*. 3 juli 2019 (se s. 27).
- [137] P.V Pearlsy och Deepa Sankar. "Implementation of Compressive Sensing for Speech Signals". I: *8th International Symposium on Embedded Computing and System Design (ISED)*. 2018, s. 162–166. DOI: 10.1109/ISED.2018.8704026 (se s. 13).
- [138] Hao Peng m.fl. "Random Feature Attention". 19 mars 2021. arXiv: 2103.02143 [cs]. URL: <http://arxiv.org/abs/2103.02143> (hämtad 2022-02-10) (se s. 43).
- [139] Ethan Perez m.fl. "FiLM: Visual Reasoning with a General Conditioning Layer". 18 dec. 2017. arXiv: 1709.07871 [cs, stat]. URL: <http://arxiv.org/abs/1709.07871> (hämtad 2022-02-10) (se s. 42).
- [140] Ethan Perez m.fl. "True Few-Shot Learning with Language Models". 24 maj 2021. arXiv: 2105.11447 [cs, stat]. URL: <http://arxiv.org/abs/2105.11447> (hämtad 2022-02-09) (se s. 43).
- [141] Jonas Pfeiffer m.fl. "MAD-X: An Adapter-Based Framework for Multi-Task Cross-Lingual Transfer". 6 okt. 2020. arXiv: 2005.00052 [cs]. URL: <http://arxiv.org/abs/2005.00052> (hämtad 2022-02-10) (se s. 44).
- [142] Vung Pham m.fl. "ContiMap: Continuous Heatmap for Large Time Series Data". I: *2020 Visualization in Data Science (VDS)*. 2020 IEEE Visualization in Data Science (VDS). Salt Lake City, UT, USA: IEEE, okt. 2020, s. 42–51. ISBN: 978-1-72819-284-0. DOI: 10.1109/VDS51726.2020.00009. URL: <https://ieeexplore.ieee.org/document/9355216/> (hämtad 2022-02-23) (se s. 45).
- [143] John Preskill. "Quantum Computing 40 Years Later". 25 juni 2021. arXiv: 2106.10522 [quant-ph]. URL: <http://arxiv.org/abs/2106.10522> (hämtad 2022-02-12) (se s. 38).
- [144] John Preskill. "Quantum Computing in the NISQ Era and Beyond". I: *Quantum* 2 (6 aug. 2018), s. 79. DOI: 10.22331/q-2018-08-06-79. URL: <https://quantum-journal.org/papers/q-2018-08-06-79/> (hämtad 2022-02-24) (se s. 41).
- [145] Alec Radford m.fl. "Learning Transferable Visual Models From Natural Language Supervision". 26 febr. 2021. arXiv: 2103.00020 [cs]. URL: <http://arxiv.org/abs/2103.00020> (hämtad 2022-02-10) (se s. 44).
- [146] Colin Raffel m.fl. "Exploring the Limits of Transfer Learning with a Unified Text-to-Text Transformer". 23 okt. 2019. arXiv: 1910.10683v3 [lg]. URL: <https://arxiv.org/abs/1910.10683v3> (hämtad 2022-02-22) (se s. 43).

- [147] Deepthi Raghunandan m.fl. "Lodestar: Supporting Independent Learning and Rapid Experimentation Through Data-Driven Analysis Recommendations". I: *Visualization in data science* (), s. 10 (se s. 45).
- [148] V. Ramasubramanian. "Ultra Low Bit-Rate Speech Coding: An Overview and Recent Results". I: *2012 International Conference on Signal Processing and Communications (SPCOM)*. 2012 International Conference on Signal Processing and Communications (SPCOM). Bangalore, Karnataka, India: IEEE, juli 2012, s. 1–5. ISBN: 978-1-4673-2014-6 978-1-4673-2013-9 978-1-4673-2012-2. DOI: 10.1109/SPCOM.2012.6290246. URL: <https://ieeexplore.ieee.org/document/6290246> (hämtad 2021-11-15) (se s. 13).
- [149] Aditya Ramesh m.fl. "Zero-Shot Text-to-Image Generation". 26 febr. 2021. arXiv: 2102.12092 [cs]. URL: <http://arxiv.org/abs/2102.12092> (hämtad 2022-02-10) (se s. 44).
- [150] Miguel Arnjona Ramirez och Mario Minami. "Low Bit Rate Speech Coding". I: *Wiley Encyclopedia of Telecommunications*. Utg. av J. G. Proakis. Vol. 3. New York: Wiley, 2003, s. 1299–1308 (se s. 13).
- [151] Alexander Ratner m.fl. "Data Programming: Creating Large Training Sets, Quickly". 8 jan. 2017. arXiv: 1605.07723 [cs, stat]. URL: <http://arxiv.org/abs/1605.07723> (hämtad 2022-02-09) (se s. 43).
- [152] Mengjie Ren m.fl. "Research on Missile-Borne Image Transmission Technology Based on Compressive Sensing". I: *2nd International Conference on Image, Vision and Computing (ICIVC)*. 2017, s. 773–777. DOI: 10.1109/ICIVC.2017.7984659 (se s. 11).
- [153] Joschka Roffe. "Quantum Error Correction: An Introductory Guide". I: *Contemporary Physics* 60.3 (3 juli 2019), s. 226–245. ISSN: 0010-7514, 1366-5812. DOI: 10.1080/00107514.2019.1667078. arXiv: 1907.11157. URL: <http://arxiv.org/abs/1907.11157> (hämtad 2022-02-24) (se s. 40).
- [154] David Grant Rowe. "Techniques for Harmonic Sinusoidal Coding". University of South Australia, School of Physics and Electronic Systems Engineering, juli 1997 (se s. 13).
- [155] Devendra Singh Sachan m.fl. "End-to-End Training of Multi-Document Reader and Retriever for Open-Domain Question Answering". 4 dec. 2021. arXiv: 2106.05346 [cs]. URL: <http://arxiv.org/abs/2106.05346> (hämtad 2022-02-10) (se s. 43).
- [156] Andrew P. Sage och Gary W. Husa. "Adaptive Filtering with Unknown Prior Statistics". I: *Joint Automatic Control Conference* 7 (1969), s. 760–769. DOI: 10.1109/JACC.1969.4169325. URL: <https://www.infona.pl/resource/bwmeta1.element.ieee-art-000004169325> (hämtad 2022-02-26) (se s. 24).
- [157] Victor Sanh m.fl. "Multitask Prompted Training Enables Zero-Shot Task Generalization". 12 dec. 2021. arXiv: 2110.08207 [cs]. URL: <http://arxiv.org/abs/2110.08207> (hämtad 2022-01-29) (se s. 42).
- [158] M. S. Arun Sankar och P. S. Sathidevi. "Compressive Sensing Based Scalable Speech Coder with Dynamic Selection of Basis and Vector Quantization". I: *International Conference on Wireless Communications, Signal Processing and Networking (WiSPNET)*. 2017, s. 1053–1058. DOI: 10.1109/WiSPNET.2017.8299923 (se s. 13).

- [159] Teven Le Scao och Alexander M. Rush. "How Many Data Points Is a Prompt Worth?" 6 april 2021. arXiv: 2103.08493 [cs]. URL: <http://arxiv.org/abs/2103.08493> (hämtad 2022-02-09) (se s. 43).
- [160] Timo Schick och Hinrich Schütze. "Exploiting Cloze Questions for Few Shot Text Classification and Natural Language Inference". 25 jan. 2021. arXiv: 2001.07676 [cs]. URL: <http://arxiv.org/abs/2001.07676> (hämtad 2022-02-09) (se s. 43).
- [161] Timo Schick och Hinrich Schütze. "It's Not Just Size That Matters: Small Language Models Are Also Few-Shot Learners". 12 april 2021. arXiv: 2009.07118 [cs]. URL: <http://arxiv.org/abs/2009.07118> (hämtad 2022-02-09) (se s. 43).
- [162] Johan Schubert och Ulrika Wickenberg Bolin. *Cluster Management of Scientific Literature in HSTOOL*. R FOI-R-5178-SE. Swedish Defence Research Agency (FOI), 2021 (se s. 18, 21, 22).
- [163] Maria Schuld m.fl. "An Introduction to Quantum Machine Learning". I: *Contemporary Physics* 56.2 (3 april 2015), s. 172–185. ISSN: 0010-7514. DOI: 10.1080/00107514.2014.964942. URL: <https://doi.org/10.1080/00107514.2014.964942> (hämtad 2022-02-28) (se s. 41).
- [164] Tyler P. Sherer m.fl. "Radar Probabilistic Data Association Filter/DGPS Fusion for Target Selection and Relative Range Determination for a Ground Vehicle Convoy". I: *NAVIGATION, Journal of the Institute of Navigation* 66.2 (1 juni 2019), s. 441–450. ISSN: 0028-1522. DOI: 10.1002/navi.305. URL: <http://www.ion.org/publications/abstract.cfm?jp=j&articleID=102811> (hämtad 2022-02-26) (se s. 25).
- [165] Danqing Shi m.fl. "Task-Oriented Optimal Sequencing of Visualization Charts". 7 aug. 2019. arXiv: 1908.02502 [cs]. URL: <http://arxiv.org/abs/1908.02502> (hämtad 2022-02-23) (se s. 45).
- [166] P.W. Shor. "Algorithms for Quantum Computation: Discrete Logarithms and Factoring". I: *Proceedings 35th Annual Symposium on Foundations of Computer Science*. Proceedings 35th Annual Symposium on Foundations of Computer Science. Nov. 1994, s. 124–134. DOI: 10.1109/SFCS.1994.365700 (se s. 39).
- [167] Peter W. Shor. "Scheme for Reducing Decoherence in Quantum Computer Memory". I: *Physical Review A* 52.4 (1 okt. 1995), R2493–R2496. DOI: 10.1103/PhysRevA.52.R2493. URL: <https://link.aps.org/doi/10.1103/PhysRevA.52.R2493> (hämtad 2022-02-25) (se s. 40).
- [168] Gary J. Sullivan m.fl. "Overview of the High Efficiency Video Coding (HEVC) Standard". I: *IEEE Transactions on Circuits and Systems for Video Technology* 22.12 (2012), s. 1649–1668. DOI: 10.1109/TCSVT.2012.2221191 (se s. 10).
- [169] Martin Svatoš m.fl. "Revisiting Neural-Symbolic Learning Cycle". I: *Proceedings of the 2019 International Workshop on Neural-Symbolic Learning and Reasoning (NeSy'19)*. 2019 (se s. 36).
- [170] Derek Tam m.fl. "Improving and Simplifying Pattern Exploiting Training". 28 sept. 2021. arXiv: 2103.11955 [cs]. URL: <http://arxiv.org/abs/2103.11955> (hämtad 2022-02-09) (se s. 43).
- [171] Yi Tay m.fl. "Are Pre-trained Convolutions Better than Pre-trained Transformers?" 30 jan. 2022. arXiv: 2105.03322 [cs]. URL: <http://arxiv.org/abs/2105.03322> (hämtad 2022-02-09) (se s. 43).

- [172] Yi Tay m.fl. "Charformer: Fast Character Transformers via Gradient-based Subword Tokenization". 2 juli 2021. arXiv: 2106.12672 [cs]. URL: <http://arxiv.org/abs/2106.12672> (hämtad 2022-02-10) (se s. 43).
- [173] Yi Tay m.fl. "Efficient Transformers: A Survey". 16 sept. 2020. arXiv: 2009.06732 [cs]. URL: <http://arxiv.org/abs/2009.06732> (hämtad 2022-02-10) (se s. 42).
- [174] Yi Tay m.fl. "Scale Efficiently: Insights from Pre-training and Fine-tuning Transformers". 22 sept. 2021. arXiv: 2109.10686 [cs]. URL: <http://arxiv.org/abs/2109.10686> (hämtad 2022-01-29) (se s. 43).
- [175] Marco Tektonidis och David Monnin. "Illumination-Invariant Change Detection for the Protection of Vehicle Convoys". I: *Electro-Optical Remote Sensing XIII*. Electro-Optical Remote Sensing XIII. Utg. av Gary W. Kamerman och Ove Steinvall. Strasbourg, France: SPIE, 10 okt. 2019, s. 9. ISBN: 978-1-5106-3023-9 978-1-5106-3024-6. DOI: 10.1117/12.2533134. URL: <https://www.spiedigitallibrary.org/conference-proceedings-of-spie/11160/2533134/Illumination-invariant-change-detection-for-the-protection-of-vehicle-convoys/10.1117/12.2533134.full> (hämtad 2022-02-26) (se s. 26).
- [176] Omkar Thakoor m.fl. "Multiagent UAV Routing: A Game Theory Analysis With Tight Price of Anarchy Bounds". I: *IEEE Transactions on Automation Science and Engineering* 17.1 (jan. 2020), s. 100–116. ISSN: 1558-3783. DOI: 10.1109/TASE.2019.2902360 (se s. 24).
- [177] Ilya Tolstikhin m.fl. "MLP-Mixer: An All-MLP Architecture for Vision". 11 juni 2021. arXiv: 2105.01601 [cs]. URL: <http://arxiv.org/abs/2105.01601> (hämtad 2022-01-30) (se s. 43).
- [178] Eleni Triantafillou m.fl. "Learning a Universal Template for Few-shot Dataset Generalization". 21 juni 2021. arXiv: 2105.07029 [cs]. URL: <http://arxiv.org/abs/2105.07029> (hämtad 2022-01-28) (se s. 42).
- [179] Maria Tsimpoukelli m.fl. "Multimodal Few-Shot Learning with Frozen Language Models". 3 juli 2021. arXiv: 2106.13884 [cs]. URL: <http://arxiv.org/abs/2106.13884> (hämtad 2022-02-10) (se s. 44).
- [180] Saeed Ullah m.fl. "Energy and Distortion Analysis of Video Compression Schemes for Wireless Video Sensor Networks". I: *Proc. Military Communications Conference (MILCOM)*. 2011, s. 822–827. DOI: 10.1109/MILCOM.2011.6127779 (se s. 11).
- [181] Ashish Vaswani m.fl. "Attention Is All You Need". I: *Advances in Neural Information Processing Systems*. Utg. av I. Guyon m.fl. Vol. 30. Curran Associates, Inc., 2017. URL: <https://proceedings.neurips.cc/paper/2017/file/3f5ee243547dee91fbd053c1c4a845aa-Paper.pdf> (hämtad 2022-02-14) (se s. 42).
- [182] Ashish Wani och Nazanin Rahnavard. "Compressive Sampling for Energy Efficient and Loss Resilient Camera Sensor Networks". I: *Proc. Military Communications Conference (MILCOM)*. 2011, s. 1766–1771. DOI: 10.1109/MILCOM.2011.6127567 (se s. 11).
- [183] Rajitha Weerakkody m.fl. "Verification Testing of HEVC Compression Performance for UHD Video". I: *2014 IEEE Global Conference on Signal and Information Processing (GlobalSIP)*. 2014, s. 1083–1087. DOI: 10.1109/GlobalSIP.2014.7032288 (se s. 11).

- [184] Jason Wei m.fl. "Finetuned Language Models Are Zero-Shot Learners". 1 dec. 2021. arXiv: 2109.01652 [cs]. URL: <http://arxiv.org/abs/2109.01652> (hämtad 2022-01-29) (se s. 42).
- [185] Laura Weidinger m.fl. "Ethical and Social Risks of Harm from Language Models". 8 dec. 2021. arXiv: 2112.04359 [cs]. URL: <http://arxiv.org/abs/2112.04359> (hämtad 2022-02-10) (se s. 44).
- [186] Johannes Welbl m.fl. "Challenges in Detoxifying Language Models". 15 sept. 2021. arXiv: 2109.07445 [cs]. URL: <http://arxiv.org/abs/2109.07445> (hämtad 2022-02-10) (se s. 44).
- [187] John Wenskovitch och Chris North. "Pollux: Interactive Cluster-First Projections of High-Dimensional Data". I: *2019 IEEE Visualization in Data Science (VDS)*. 2019 IEEE Visualization in Data Science (VDS). Vancouver, BC, Canada: IEEE, okt. 2019, s. 38–47. ISBN: 978-1-72815-047-5. DOI: 10.1109/VDS48975.2019.8973381. URL: <https://ieeexplore.ieee.org/document/8973381/> (hämtad 2022-02-23) (se s. 45).
- [188] John Wenskovitch m.fl. "Albireo: An Interactive Tool for Visually Summarizing Computational Notebook Structure". I: *2019 IEEE Visualization in Data Science (VDS)*. 2019 IEEE Visualization in Data Science (VDS). Vancouver, BC, Canada: IEEE, okt. 2019, s. 1–10. ISBN: 978-1-72815-047-5. DOI: 10.1109/VDS48975.2019.8973385. URL: <https://ieeexplore.ieee.org/document/8973385/> (hämtad 2022-02-23) (se s. 45).
- [189] Sarah Wiegrefe och Ana Marasović. "Teach Me to Explain: A Review of Datasets for Explainable Natural Language Processing". 7 dec. 2021. arXiv: 2102.12060 [cs]. URL: <http://arxiv.org/abs/2102.12060> (hämtad 2022-02-10) (se s. 43).
- [190] Eric P. Xing m.fl. "Petuum: A New Platform for Distributed Machine Learning on Big Data". I: *Proceedings of the 21th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*. KDD '15. New York, NY, USA: Association for Computing Machinery, 2015, s. 1335–1344. ISBN: 978-1-4503-3664-2. DOI: 10.1145/2783258.2783323. URL: <https://doi.org/10.1145/2783258.2783323> (hämtad 2022-01-10) (se s. 35).
- [191] Linting Xue m.fl. "ByT5: Towards a Token-Free Future with Pre-Trained Byte-to-Byte Models". 28 maj 2021. arXiv: 2105.13626 [cs]. URL: <http://arxiv.org/abs/2105.13626> (hämtad 2022-02-10) (se s. 43).
- [192] Haici Yang m.fl. "Source-Aware Neural Speech Coding for Noisy Speech Compression". I: *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*. 2021, s. 706–710. DOI: 10.1109/ICASSP39728.2021.9413678 (se s. 13).
- [193] Jian-Bo Yang m.fl. "Optimization Models for Training Belief-Rule-Based Systems". I: *IEEE Transactions on Systems, Man, and Cybernetics - Part A: Systems and Humans* 37.4 (juli 2007), s. 569–585. ISSN: 1558-2426. DOI: 10.1109/TSMCA.2007.897606 (se s. 27).
- [194] Kexin Yi m.fl. "Neural-Symbolic VQA: Disentangling Reasoning from Vision and Language Understanding". I: *Annual Conference on Neural Information Processing Systems* 2018. Vol. 31. Montreal, Canada, 2018 (se s. 37).
- [195] Dani Yogatama m.fl. "Adaptive Semiparametric Language Models". 4 febr. 2021. arXiv: 2102.02557 [cs]. URL: <http://arxiv.org/abs/2102.02557> (hämtad 2022-01-12) (se s. 35, 43).

- [196] Xiaohua Zhai m.fl. "Scaling Vision Transformers". 8 juni 2021. arXiv: 2106.04560 [cs]. URL: <http://arxiv.org/abs/2106.04560> (hämtad 2022-01-28) (se s. 42).
- [197] Yingxue Zhang m.fl. "cST-ML: Continuous Spatial-Temporal Meta-Learning for Traffic Dynamics Prediction". I: *20th IEEE International Conference on Data Mining, ICDM 2020, Sorrento, Italy, November 17-20, 2020*. Utg. av Claudia Plant m.fl. IEEE, 2020, s. 1418–1423. DOI: 10.1109/ICDM50108.2020.00187 (se s. 41).
- [198] Chen Zhen m.fl. "A Deep Learning Based Distributed Compressive Video Sensing Reconstruction Algorithm for Small Reconnaissance UAV". I: *3rd International Conference on Unmanned Systems (ICUS)*. 2020, s. 668–672. DOI: 10.1109/ICUS50048.2020.9274972 (se s. 11).
- [199] Yanan Zheng m.fl. "FewNLU: Benchmarking State-of-the-Art Methods for Few-Shot Natural Language Understanding". 26 sept. 2021. arXiv: 2109.12742 [cs]. URL: <http://arxiv.org/abs/2109.12742> (hämtad 2022-02-09) (se s. 43).
- [200] Deyun Zhou m.fl. "An Improved Belief Entropy and Its Application in Decision-Making". I: *Complexity* 2017 (16 mars 2017), e4359195. ISSN: 1076-2787. DOI: 10.1155/2017/4359195. URL: <https://www.hindawi.com/journals/complexity/2017/4359195/> (hämtad 2022-02-26) (se s. 28).
- [201] Ying Zhou m.fl. "A Novel Uncertainty Management Approach for Air Combat Situation Assessment Based on Improved Belief Entropy". I: *Entropy. An International and Interdisciplinary Journal of Entropy and Information Studies* 21.5 (5 maj 2019), s. 495. ISSN: 1099-4300. DOI: 10.3390/e21050495. URL: <https://www.mdpi.com/1099-4300/21/5/495> (hämtad 2022-02-26) (se s. 28, 29).
- [202] Yu Zhou m.fl. "A Belief-Rule-Based Model for Information Fusion with Insufficient Multi-Sensor Data and Domain Knowledge Using Evolutionary Algorithms with Operator Recommendations". I: *Soft Computing - A Fusion of Foundations, Methodologies and Applications* 23.13 (1 juli 2019), s. 5129–5142. ISSN: 1432-7643. DOI: 10.1007/s00500-018-3179-9. URL: <https://doi.org/10.1007/s00500-018-3179-9> (hämtad 2022-02-26) (se s. 27).

Bilaga A : Extramaterial

A.1 Matematik för kvantdatorer

Den mest fundamentala dataenheten i en klassisk digital dator är *bit* (binary digit), som kan anta värdena *ett* (1) eller *noll* (0). Logiskt sett består digital data enbart av ettor och nollor, exempelvis en 8-bitars byte 01101111, som sedan omtolkas till instruktioner, tal eller tecken beroende på kontext. För att använda digital data i praktiken (dvs. för att lagra, förmedla och processa den digitala datan) behöver den manifesteras fysiskt. Detta sker normalt antingen med hjälp av en elektrisk spänning (hög eller låg beroende på värdet) i primärminnet eller magnetisk polarisation i sekundärminnet (exempelvis hårddisk).

Kvantdatorn däremot är inte baserad på digitala bitar, utan *kvantbitar* (eng. *qubit*) [126]. En kvantbit kan representera en digital nolla och etta, uttryckt med vektorerna $|0\rangle = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$ respektive $|1\rangle = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$. Utöver dessa två grundtillstånd kan kvantbitar även representera *överlagring* (eng. *superposition*) och *sammanflätning* (eng. *entanglement*) vilket är fysiska egenskaper hos kvantfenomen.¹ En överlagrad kvantbit, $|\psi\rangle$, har ett tillstånd som är en blandning mellan $|0\rangle$ och $|1\rangle$, inom kvantfysiken uttryckt som

$$|\psi\rangle = \alpha|0\rangle + \beta|1\rangle = \begin{bmatrix} \alpha \\ \beta \end{bmatrix}, \quad (\text{A.1})$$

där koefficienterna² α och β är komplexa tal (dvs. $\alpha, \beta \in \mathbb{C}$) och $|\alpha|^2 + |\beta|^2 = 1$. $|\alpha|^2$ och $|\beta|^2$ tolkas som sannolikheten för $|0\rangle$ respektive $|1\rangle$. Beroendet mellan α och β kan hanteras genom att parameterisera dem med reella parametrarna θ och ϕ på följande vis:

$$\alpha = \alpha(\theta, \phi) = \cos \frac{\theta}{2}, \quad \beta = \beta(\theta, \phi) = e^{i\phi} \sin \frac{\theta}{2}. \quad (\text{A.2})$$

För alla värden på θ, ϕ kommer beroendet mellan $\alpha(\theta, \phi)$ och $\beta(\theta, \phi)$ att hålla, dvs. $|\alpha(\theta, \phi)|^2 + |\beta(\theta, \phi)|^2 = 1 \forall \theta, \phi \in [0, 2\pi)$. Geometriskt kan därför en (överlagrad) kvantbit representeras geometriskt som en punkt på ytan av en sfär (figur A.1), istället för en punkt på en linje mellan $|0\rangle$ och $|1\rangle$ (vilket vore fallet om koefficienterna var reella).

Överlagringen, vilken möjliggör att kvantbiten kan representera både 0 och 1 på samma gång, är den grundläggande fördelen med kvantbiten. Visserligen sammanfaller (eng. *collapse*) kvantbiten vid avläsning och antar ett av de två grundtillstånden (med sannolikheten $|\alpha|^2$ för $|0\rangle$ och $|\beta|^2$ för $|1\rangle$), men de operationer som utförs på kvanttillståndet bibehåller överlagringen (åtminstone i teorin), vilket innebär att båda grundtillstånden genomgår operationen samtidigt, en egenskap som kallas *kvantparallellism*.

För de flesta intressanta tillämpningar behöver man dock kunna representera fler tillstånd än bara en bit. Naturligtvis kan man utvidga kvanttillståndet genom att lägga till fler kvantbitar. För att kunna dra nytta av kvantberäkningar måste de använda kvantbitarna dock sammanflätas, dvs. tillstånden måste göras beroende av varandra. Detta innebär att kvantbitarna manipuleras tillsammans och när en kvantbit mäts fastställs värdet på samtliga (sammanflätade) kvantbitar. Matematiskt beskrivs sammanflätningen mellan två kvantbitar

¹Läsaren kan läsa mer om kvantfenomen i specialiserad litteratur.

²Kallas också amplituder

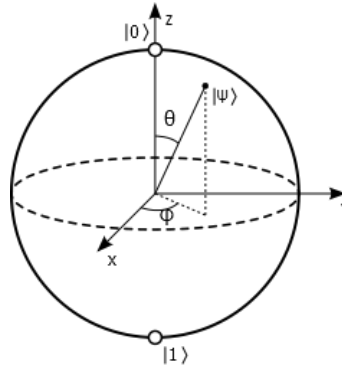


Figure A.1. En överlagrad kvantbit, $|\psi\rangle = \alpha|0\rangle + \beta|1\rangle$, kan geometriskt representeras som en punkt på ytan på en sfär (s.k. Bloch-sfär). Smite-Meister, CC BY-SA 3.0 <https://creativecommons.org/licenses/by-sa/3.0>, via Wikimedia Commons.

$|\psi_1\rangle$ och $|\psi_2\rangle$ med

$$\begin{aligned} |\psi_1\rangle \otimes |\psi_2\rangle &= (\alpha_1|0\rangle + \beta_1|1\rangle) \otimes (\alpha_2|0\rangle + \beta_2|1\rangle) \\ &= \alpha_1\alpha_2|00\rangle + \alpha_1\beta_2|01\rangle + \beta_1\alpha_2|10\rangle + \beta_1\beta_2|11\rangle, \end{aligned}$$

där $\otimes$ är en operator (kallad tensorprodukt) som skapar flerbitstillstånd och exempelvis

$$|01\rangle = \begin{bmatrix} 0 \\ 1 \\ 0 \\ 0 \end{bmatrix}$$

är grundtillståndet 01 för två sammanflätade kvantbiter (värdena 0 och 1 för den andra respektive den första kvantbiten, numrerade från höger till vänster). Analogt med den enskilda kvantbiten är summan av sannolikheterna för kvantbitarna 1, dvs. $|\alpha_1\beta_1|^2 + |\alpha_1\beta_2|^2 + |\beta_1\alpha_2|^2 + |\beta_1\beta_2|^2 = 1$.³ Sammanflätningen kan (i teorin) generaliseras till godtyckligt antal (n) kvantbiter, $|\psi\rangle = \otimes_{i=1}^n |\psi_i\rangle$, och koefficienterna summerar till 1.⁴

Det finns olika metoder för att utföra beräkningar med kvantbiter, bland annat *kvantgrindar* och så kallad adiabatisk [7]. Kvantgrindar är kvantdatorns motsvarighet till logiska grindar i digitala datorer. Att använda en viss kvantgrind U på ett kvanttillstånd $|\psi_{\text{in}}\rangle$ och få resultatet $|\psi_{\text{ut}}\rangle$ uttrycks $U|\psi_{\text{in}}\rangle = |\psi_{\text{ut}}\rangle$ matematiskt och kan även uttryckas grafiskt som i figur A.2.

$$|\psi_{\text{in}}\rangle \xrightarrow{U} |\psi_{\text{ut}}\rangle$$

Figure A.2. Visualisering en kvantgrind U som överför ett kvanttillstånd, $|\psi_{\text{in}}\rangle$, till ett annat, $|\psi_{\text{ut}}\rangle$.

Ett exempel på en kvantgrind är Pauli-X (med symbolen X) som är kvantdatorns motsvarighet till en digital icke-grind (NOT gate) och beskrivs matematiskt för $|\psi\rangle$ (som i ekv. A.1):

³Notera att de nya koefficienterna inte behöver normaliseras då (eftersom för $z_j = r_j e^{i\theta_j} \in \mathbb{C} : |z_j| = r_j, |z_1 \cdot z_2| = r_1 r_2 = |z_1| |z_2|$, och per definition $|\alpha_j|^2 + |\beta_j|^2 = 1$) $|\alpha_1\alpha_2|^2 + |\alpha_1\beta_2|^2 + |\beta_1\alpha_2|^2 + |\beta_1\beta_2|^2 = |\alpha_1|^2|\alpha_2|^2 + |\alpha_1|^2|\beta_2|^2 + |\beta_1|^2|\alpha_2|^2 + |\beta_1|^2|\beta_2|^2 = |\alpha_1|^2(|\alpha_2|^2 + |\beta_2|^2) + |\beta_1|^2(|\alpha_2|^2 + |\beta_2|^2) = |\alpha_1|^2 + |\beta_1|^2 = 1$. ■

⁴Dvs. $\alpha_1 \cdots \alpha_n + \dots + \alpha_1 \cdots \alpha_j \beta_{j+1} \cdots \beta_n + \dots + \beta_1 \cdots \beta_n = 1$

$$X|\psi\rangle = \overbrace{\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}}^X \overbrace{\begin{bmatrix} \alpha \\ \beta \end{bmatrix}}^{|\psi\rangle} = \begin{bmatrix} \beta \\ \alpha \end{bmatrix} = \beta|0\rangle + \alpha|1\rangle = |\psi_{\text{ut}}\rangle$$

A.2 Mjukvaror

Nedan följer en lista av länkar till ett urval av mjukvaror för neuro-symboliskt lärande och resonerande och kvantberäkningar.

A.2.1 Neurosymboliskt lärande och resonerande

Samtliga sidor besöktes i februari 2022.

- ANFIS, <https://www.mathworks.com/help/fuzzy/anfis.html>
- DeepProbLog, <https://github.com/ML-KULEuven/deepproblog>
- KBANN, <https://github.com/bits-mayank/KBANN>
- LogicTensorNetworks, <https://github.com/logictensornetworks/logictensornetworks>
- NeuraLogic, <https://github.com/GustikS/NeuraLogic>

A.2.2 Kvantdatorberäkningar

Samtliga sidor besöktes i februari 2022.

- cirq, <https://quantumai.google/cirq>
- PennyLane, <https://pennylane.ai/>
- pyQuil, <https://github.com/rigetti/pyquil>
- QIO, <https://hackage.haskell.org/package/QIO>
- qiskit, <https://qiskit.org/>
- Quantum algorithm zoo, <https://quantumalgorithmzoo.org/>
- Quantum open source foundation, <https://qosf.org/>
- TensorFlow Quantum[27], <https://ai.googleblog.com/2020/03/announcing-tensorflow-quantum-open.html>

FOI är en huvudsakligen uppdragsfinansierad myndighet under Förvarsdepartementet. 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