

Ansökan om forskningsmedel

Datum
210201

Observera att ansökan och bilagor endast ska skickas elektroniskt till konkurrensverket@kkv.se.

1 Sökande (huvudansvarig för projektet)

Namn *	Henrik Artman
Universitet/högskola eller motsvarande	Institution eller motsvarande
KTH	EECS
Postadress	Postnr och ortsnamn
Lindstedts 3	100 44 STHLM
Telefonnummer	E-post
[REDACTED]	[REDACTED]

* Huvudansvarig för projektet är ansvarig för att uppgifterna som lämnas i ansökningsblanketten är korrekta.

2 Anslagsförvaltare

Universitet/högskola eller motsvarande	Institution
KTH	EECS
Postadress	Postnr och ortsnamn
Brinellvägen 8	100 44 STHLM
Telefonnummer	
087906000	

3 Projektbeskrivning

Projekttitel	
UppFörA - Upphandling och förklarbarhet i algortimer	
Projektets relevans för Konkurrensverket (högst 240 tecken)	
Behovet att myndigheter upphandlar automatiserade beslutssystem ansvarsfullt ökar, särskilt förmågan att ställa krav på förklarbarhet i algoritmer. Projektet bidrar med kartläggning och vägledning av fenomenet.	
Projektet avses starta/startade, datum	Projektet beräknas vara slutfört, datum
2021-08-30	2023-08-31
Sammanfattning på svenska av projektets syfte, betydelse och genomförande (högst 1 000 tecken).	
I upphandling av algoritmiska beslutssystem, eller beslutsstödssystem är icke-funktionella krav en allt viktigare aspekt av en välgenomförd upphandling, och leveransen av en algoritm som är såväl ickediskriminerande. Projektet syftar till att kartlägga hur och om myndigheter ställer sådana krav idag, och att ta fram vägledning för hur icke-funktionell kravställning kan förbättras. Genom att såväl studera dokument (genom att nyttja offentlighetsprincipen) som att intervjua beställare hoppas projektet bland annat kunna kartlägga hur praktiken ser ut idag. Ett välgenomfört projekt skulle kunna stärka myndigheternas upphandlingskompetens, klargöra såväl etiska som juridiska krav för båda parter samt stärka medborgarnas tillit till myndigheters beslutsfattande. Frågeställningarna om förklarbarhet och transparens väcker stor uppmärksamhet såväl i akademien som i den offentliga debatten. De sökande har också god vana vid populärvetenskaplig presentation.	

Bifoga även en utförligare projektbeskrivning (svenska eller engelska, max 10 A4-sidor) som bilaga till denna ansökan.

4 Kostnadsredovisning

Observera att den högsta tillåtna arbetstiden för disputerad forskare, docent och professor är 75 procent av heltid. För doktorand, forskningsassistent eller liknande är den högsta tillåtna arbetstiden 85 procent av heltid.

Projektår 1					
Personalnamn och akademisk titel (bifoga CV)		Månadslön (brutto)	Anställningstid i projektet, månader	Arbetstid i procent av heltid	Lönekostnad inkl. sociala avgifter
Namn	Akademisk titel				
██████████	Professor	77 700	12	7	105 276
██████████	Doktorand	33 825	12	80	470 000
Summa lönekostnader inkl. sociala avgifter					575 276
Summa övriga kostnader (hämtas från tabell 4a)					24 000
Summa förvaltningskostnader inklusive lokalhyra					635 057
Total kostnad inklusive sociala avgifter och förvaltningsavgifter					1 234 328

Projektår 2					
Personalnamn och akademisk titel (bifoga CV)		Månadslön (brutto)	Anställningstid i projektet, månader	Arbetstid i procent av heltid	Lönekostnad inkl. sociala avgifter
Namn	Akademisk titel				
██████████	Professo	77 700	12	7	108 434
██████████	Doktorand	34 671	12	80	481 750
Summa lönekostnader inkl. sociala avgifter					590 184
Summa övriga kostnader (hämtas från tabell 4a)					23 000
Summa förvaltningskostnader inklusive lokalhyra					651 753
Total kostnad inklusive sociala avgifter och förvaltningsavgifter					1 264 937

Projektår 3					
Personalnamn och akademisk titel (bifoga CV)		Månadslön (brutto)	Anställningstid i projektet, månader	Arbetstid i procent av heltid	Lönekostnad inkl. sociala avgifter
Namn	Akademisk titel				
Summa lönekostnader inkl. sociala avgifter					
Summa övriga kostnader (hämtas från tabell 4a)					
Summa förvaltningskostnader inklusive lokalhyra					
Total kostnad inklusive sociala avgifter och förvaltningsavgifter					

4a Redovisning övriga kostnader

Maxbelopp för övriga kostnader per år är 25 000 kronor. Ifall detta belopp överskrids ska detta motiveras särskilt i ansökan.

	År 1	År 2	År 3
Material och utrustning			
Resor	24 000	23 000	
Andra kostnader			
Summa	24 000	23 000	

5 Kostnadssammanfattning (anges i kronor) för nu sökt anslag

Total projektkostnad 2 499 265

Därav söks från		Tidigare erhållna anslag från	
Konkurrensverket	Annan anslagsgivare *	Konkurrensverket	Annan anslagsgivare **
2 499 265	0	0	0

* Anslagsgivarens namn	Ansökan inlämnad, datum	Sökt belopp
** Anslagsgivarens namn	Ansökan beviljad, datum	Beviljat belopp

6 Övriga projekt som samtidigt kommer att ledas av huvudansvarig

Projekttitel 4003 FOI LVC (uppdragsforskning om 30% persontid) 4023 CDIS HA (Försvarsmaktsprojekt - endast doktorand, ej egen tid)
--

OBS! Namn och institution på personer som beviljas forskningsbidrag kommer att publiceras på Konkurrensverkets webbplats. Om en ansökan om forskningsbidrag skickas in till Konkurrensverket innebär det ett medgivande till att dina personuppgifter registreras och behandlas av Konkurrensverket samt att uppgifter om namn och institution för beviljade bidrag publiceras på webbplatsen.

UppFörA - Upphandling och förklarbarhet i algoritmer

Bakgrund och frågeställningar

Med ökad digitalisering blir algoritmerna allt viktigare. De hjälper oss att effektivisera arbetsuppgifter, sälla, sortera och kategorisera information och i allt större utsträckning är de deltagande eller till och med utförare av mer avancerade beslutsprocesser. I takt med att mer avancerade beslut tas av algoritmer blir också konsekvenserna av sådan behandling viktigare att ta i beaktande. Vi vet att den data som vi hittills genererat om vår värld är bristfällig, och innehåller spår av förutfattade meningar och diskriminering även när vi inte avser det. För att i så stor utsträckning som möjligt undvika negativa konsekvenser krävs det mer än tidigare vad gäller konsekvenstänk, etiska krav och transparens för inköp av sådan teknik. Hur kan man då säkerställa att utvecklingen av de digitala och automatiserade systemen beaktar förklarbarhet för algoritmer när man upphandlar AI-system?

Så kallat icke-funktionella krav vid upphandling, dvs krav som avser egenskaper som systemet bör uppfylla såsom tillförlitlighet, transparens, användbarhet, effektivitet, kapacitet och säljbarhet, i upphandling är problematiska eftersom de inte kan preciseras tidigt och oftast beror på interaktion mellan ett system och dess användare. En beräkning, databasstruktur eller funktioner mellan olika komponenter och element kan preciseras tidigt och bygger vanligen på att man har en tydlig målbild av hur funktionerna ska uppnå ett givet resultat (Artman, 2002; Markensten & Artman, 2004; Zällh & Artman, 2005). Så fort man även behöver inkludera användarens interaktion med systemet finns det ett flertal faktorer som påverkar förloppet (kompetens, erfarenhet, vana, design av systemet etc.). Detta har varit beforskat för just området användbarhet under 2000-2010-talet och en del av resultatet menar att man å ena sidan bör precisera interaktionsbehovet med systemet innan man påbörjar design av funktioner, och att man kontinuerligt bör utvärdera hur pass väl användare kan använda systemet (för en populärvetenskaplig sammanfattningen av projektet Beställarkompetens finansierat av Vinnova se "Att beställa något användbart är inte uppenbart"¹, Artman et al 2010, 2009). Att beställa system som behandlar AI eller Machine Learning komponenter där man avser effektivisera eller även automatisera beslutsfattande aktualiserar problemet med att en tjänsteman eller kunden behöver ha en tydlig förståelse över hur IT-systemet har kommit fram till ett givet förklarbart beslut. Mellan 2000-2010 utvecklade dåvarande Verket för förvaltningsutveckling (Verva) en uppskattad vägledning med riktlinjer för hur kommuner, regioner och myndigheter kan och bör utveckla publika webb- och e-tjänster². Post- och Telestyrelsen, som övertog ansvaret verkar i dagsläget inte ha motsvarande vägledning för etjänster.

¹ <https://utvgrupp.files.wordpress.com/2011/10/att-besta-3a4lla-nc-3a5got-anvc-3a4ndbart-henrik-a1.pdf>

² Vägledningen för 24-timmars webben: <https://webbriktlinjer.se/>

Konkurrensverket har tillsammans med andra myndigheter efterlyst ”integrerad och tvärvetenskaplig forskning kring digitala och datadrivna marknader och hur de använder personuppgifter, användardata och andra data och vilka effekter som kan uppstå ur ett konsument-, integritets- och konkurrensperspektiv”.³ Detta projekt tar sig an ovanstående genom att undersöka och ge rekommendationer för krav på mer ansvarsfullt och transparent automatiserat beslutsfattande redan i upphandlingsförfarandet.

Automatisering och AI-användning har stor potential. I januari 2021 gick Lunds kommun ut och berättade att deras automatiska system hade hanterat 66 840 ärenden och därmed sparat in 7 150 mantimmar.⁴ Men utvecklingen ifrågasätts också. I januari 2020 anmälde Akademikerförbundet SSR Trelleborgs kommun till Justitieombudsmannen för att kommunen låter en algoritm snarare än en socialsekreterare fatta beslut om försörjningsstöd.⁵ Digitaliseringen effektiviserar men samtidigt finns det risk att varken beslutande tjänsteman eller medborgaren har tydlig inblick i lagrymden eller rationaliteten bakom en enskilt beslut. AI-system för offentlig verksamhet behöver kunna presentera en rimlig förklaring till beslutet. För att detta ska vara möjligt måste man tydligt ha uttryckt detta tydligt i upphandlingsdokumenten likväl som att dessa utvärderas i enlighet med kraven innan systemen tas i bruk.

I november 2020 presenterade Riksrevisionen en granskning om automatiserat beslutsfattande i myndigheter, särskilt utifrån balansen mellan effektivitet och rättssäkerhet. Slutsatserna från granskningen visar att även om myndigheterna fått högre effektivitet och vissa förbättrade rättssäkerhetsaspekter så brister myndigheterna i uppföljning av huruvida beslut blivit korrekt tagna, samt att det finns brister tydlig och läsbar kommunikation av automatiserad handläggning (RiR 2020:22)⁶. Även om granskningen inte rörde specifikt upphandlade tjänster är det brister som med största sannolikhet återkommer upphandlade processer. Riksrevisionen har också påpekat behov av stärkt beställarkompetens bland annat vid inköp av konsulttjänster (RiR 2018:20).

Exemplen illustrerar att inte minst offentlig sektor har ett stort behov av transparens för att skapa tillit till algoritmiskt beslutsfattande. Inom ramen för den nationella SOM-undersökningen 2018 undersöktes medborgarnas syn på automatiserat beslutsfattande i offentlig sektor—när algoritmer ersätter handläggare som beslutsfattare (Denk et al. 2018). Bara en minoritet, 20%, kände ens till att detta sker. Även om de flesta visserligen tror att automatiserade beslut blir mer opartiska så tvivlar de flesta på att de blir mer tillförlitliga. En majoritet tror också att automatiska beslut tar mindre hänsyn till människors situation och ger mindre insyn i beslutsfattandet. Det finns alltså all anledning att ta frågorna på största allvar för att säkerställa medborgarna fortsatt har anledning att känna tillit till myndigheterna.

3

<https://www.konkurrensverket.se/contentassets/a286ebcb7ccf442dba171f119e7a68bf/konsumentverks-konkurrensverkets-och-datainspektionens-inspel-till-forskningspolitiska-propositionen.pdf>

⁴ <https://computersweden.idg.se/2.2683/1.745387/lund-kommunala-bottar>, läst 2021-01-13

⁵ <https://www.svd.se/obehorig-algoritm-tar-beslut-i-socialtjansten>, läst 2020-03-11

⁶ Såväl fallet Trelleborgs kommun som Riksrevisionens granskning behandlas i artikel i Dagens Nyheter, 2021-01-31:

<https://www.dn.se/ekonomi/svart-fa-insyn-i-algoritmerna-som-styr-robotarna-i-jobbet/>

För att den offentliga sektorn ska kunna dra nytta av fördelarna med automatisering utan att drabbas av dess nackdelar fick Lantmäteriet och DIGG i december 2019 ett regeringsuppdrag om att testa ny teknik vid automatisering inom offentlig förvaltning.⁷ I den slutrapport som presenterades ett år senare föreslår myndigheterna en förtroendemodell som syftar till att skapa tillit till myndighetsbeslut – även när dessa går emot ens förhoppningar.⁸ Relaterat till det uppdraget fick DIGG så sent som den 28 januari i år ett regeringsuppdrag om att bistå med rättsligt stöd till förvaltningsgemensam digitalisering.⁹

Ytterligare en komplicerande faktor är att den stora majoriteten av företag, myndigheter och organisationer inte själva bygger sina algoritmer (inkl. AI) vare sig nu eller i framtiden. Istället *upphandlar* man den från andra – underleverantörer som är experter på mjukvara för exempelvis bildigenkänning, texthantering eller riskbedömning. Å ena sidan är detta helt rimligt: en leverantör som specialiserar sig på att göra en enskild produkt så bra som möjligt skapar mycket högre kvalitet än om alla som har nytta av en sådan produkt försöker uppfinna hjulet själva. Å andra sidan finns det en risk för att såväl slutanvändaren som beställaren hamnar i en onödig beroendeställning till leverantören och att produkten ifråga behandlas som en svart låda. Det kan i så fall leda till ett obehagligt uppvaknande om produkten ifråga visar sig fatta beslut som diskriminerar vissa grupper, exempelvis minoriteter eller människor med knappa ekonomiska resurser, vilket har skett upprepade gånger mot olika slags algoritmer (Nature, 2016). Även om den tekniska lösningen är inköpt utifrån så är det svårt att svära sig fri från ansvaret för hur den används i den egna verksamheten.

Det är mot denna bakgrund som det föreslagna projektet ställer följande forskningsfrågor:

- Vilka krav på transparent AI och algoritmer ställer företag, myndigheter och organisationer idag?
- Genom vilka begrepp och mätmetoder kan företag, myndigheter och organisationer utvecklas och ställa mer ändamålsenliga krav på transparent AI? Om sådana krav ställs, hur utformas de mer konkret?
- Finns det etablerad best practice inom myndigheter som andra kan lära sig av?
- Har mindre myndigheter möjlighet att ställa verksamhetsanpassade och tillika förklarbara krav för algoritmiska lösningar?

7

<https://www.regeringen.se/regeringsuppdrag/2019/12/uppdrag-om-att-testa-ny-teknik-vid-automatisering-inom-offentlig-forvaltning/>

8

<https://www.digg.se/om-oss/nyheter/2020/en-fortroendemodell-skapar-tydlighet-och-tillit-i-myndigheters-automatisering-med-ny-teknik>

9

<https://www.regeringen.se/pressmeddelanden/2021/01/digg-ska-ge-rattsligt-stod-i-digitaliseringsfragor/>

Befintlig forskning

Det finns mycket tekniskt inriktad forskning om förklarbar AI; det som på engelska kallas för *explainable AI* (förkortat XAI). Bra litteraturöversikter ges exempelvis av Guidotti et al. (2018) och Du et al. (2019). I korthet kan man säga att den tekniska litteraturen framförallt handlar om hur man kan omvandla mindre förklarbara AI-modeller som (djupa) artificiella neuronnät eller supportvektormaskiner till mer förklarbara modeller i form av exempelvis beslutsträd (Andrews et al., 1995; Barakat & Bradley, 2010) eller s.k. *heat* respektive *saliency maps* (Samek et al., 2016; Adebayo et al., 2018) som grafiskt illustrerar vad som är viktigt för algoritmernas utfall. En viktig insikt är dock att ökad förklarbarhet inte är gratis – förklarbara modeller brukar bara förklara den fullständiga modellen lokalt, alternativt efterlikna den men med sämre prestanda (exempelvis större typ I- och typ II-fel). Det finns alltså en genuin avvägning att göra: mer förklarbarhet sker på bekostnad av andra önskvärda egenskaper.

Det finns också en växande litteratur inriktad på etiska aspekter av förklarbar AI. Förklarbarhet och transparens betraktas nästan alltid som något *prima facie* gott. Inte minst möjliggör förståelse för hur ett beslutsstödssystem fungerar att dess användare kan agera mer ansvarsfullt (Fleischmann & Wallace, 2005), vilket är ett starkt skäl att eftersträva förklarbarhet. Samtidigt kan förklarbarhet vara något instrumentellt gott snarare än något som är gott i sig: transparens är en *pro-etisk omständighet*, skriver Turilli & Floridi (2009), som i sin tur möjliggör etiskt handlande. Däremot är det få som förespråkar förklarbarhet till varje pris. de Laat (2018) argumenterar för att alltför omfattande transparens kan få negativa effekter, exempelvis genom att möjliggöra oönskat beteende (transparens i polisarbete eller försäkringssystem kan exempelvis underlätta för tjuvar eller försäkringsbedragare). Zerilli et al. (2019) påpekar också att eftersom vi människor ofta överskattar hur transparenta vi själva är kanske vi också kräver för mycket transparens av AI.

Forskning om att AI har använts för att välja och värdera offerter vid upphandling (ex. Alhazmi & McCaffer, 2000; Cui, Li, Zhan, 2020). Det har dock inte fått genomslag inom svensk förvaltning och skulle potentiellt även kräva explicita kriterier för urval och slutgiltigt val. Så även om det potentiellt skulle finnas en nytta med att använda AI inom upphandlingen så återstår problemet med hur algoritmer för urval och kriterier för urvalet görs förklarbart för såväl tjänsteperson vid myndigheten som för offertgivare. Detta forskningsprojekt avser alltså att fokusera på hur en myndighet kan upphandla design av automatiserade urvals och beslutsalgoritmer som tydligt kan förklara för samtliga involverade parter beslutskriterier och hur de tillämpas. Eftersom myndigheters automatiserade beslutsprocesser inom exempelvis deklarationer, skuldsanering, bidragsbeslut, körkortstillstånd ökar och kan förväntas öka inom de närmaste åren med allt mer sofistikerade system och algoritmer är det av essentiell vikt att systemen i högre grad även kan stödja förklarbarhet. Utifrån vår sökning av forskningsdatabaser så finner vi en avsaknad av forskning kring hur AI upphandlas och kravställs. Detta understryker vikten av projektets forskningsfrågor ur det vetenskapliga perspektivet – området ligger öppet för nydanande forskning.

En tidigare artikel av de sökande har gått igenom nationella AI-strategier för att undersöka vilken etisk vägledning som ges för AI-implementationer, och visade där att riktlinjerna i de

behandlade dokumenten är nästintill obefintlig. Det finns således en lucka mellan önskvärd praktik och sätt att nå den praktiken som behöver undersökas vidare (Dexe et al. 2020b).

Syfte, metod och data

Syftet med projektet är att undersöka hur offentliga verksamheter (myndigheter, regioner, kommuner) upphandlar automatiserade beslutssystem med särskilt fokus på hur man uttrycker krav på förklarbarhet för involverade parter. Målet med projektet är formulera såväl processmodell som begreppskarta för vad som kan vara lyckade upphandlingsunderlag inom ramen för förklarbarhet.

Ett skäl till bristen på forskning kring upphandling och kravställning av AI är svårigheten att få tag på upphandlingsunderlag. Här erbjuder emellertid den svenska offentlighetsprincipen, tillsammans med GDPR, en ur ett internationellt perspektiv unik möjlighet till datainsamling.

Projektet struktureras enligt följande två-stegs datainsamlings och analysmodell.

1. Offentliga upphandlingsunderlag av automatiserade beslutssystem
 - a. **Förundersökning:** En bred genomlysning av automatiserade beslutsprocesser vid offentliga verksamheter
 - b. **Urval och datainhämtning:** Baserat på urvalskriterier, såsom att exempelvis beslut ska nå medborgare eller annan aktör bortom tjänsteman på myndigheten, begära ut upphandlingsunderlag eller kravspecifikationer vid internutveckling av systemet. Gissningsvis kommer i första hand större statliga myndigheter såsom Skatteverket, Försäkringskassan, Kronofogden vara särskilt intressanta eftersom de i hög grad använder etjänster med behov av förklarbarhet.
 - c. **Analys:** begrepp och effektmodeller relativt förklarbarhet: Finns det mätetal etc. i krav? Användarcentrerat? Hur omnämns slutanvändare? Ställdes det krav på utvärdering med slutanvändare (tjänsteperson/medborgare) för att säkerställa förklarbarhet? Metoden för analysen kommer vara baserad interbedömarreliabilitet genom att minst två personer kommer bedöma klassificering av upphandlingsunderlagen och dess begreppsliga klassificering (se Dexe et al, 2020b för liknande klassificeringsmetod).
 - d. **Utkomst:** Process- och begreppsmodell. Eventuellt kan det skapas olika relationsmodeller beroende på komplexiteten i det upphandlade systemet och eller den offentliga verksamhetens behov. Det finns förstås viss risk för att upphandlingsunderlagen inte innehåller begrepp om förklarbarhet eller liknande.
2. Intervjuer med syfte att undersöka/skapa "best practice" modell
 - a. **Intervjuer:** Semi-strukturerad intervjumetodik med upphandlingsansvarig på myndigheten. Process- och begreppsmodellen i steg 1 används som bas för intervjuerna för att kunna förankra intervjuerna i upphandlingsunderlag och förklarbarhet.
 - b. **Intervjuer:** Semi-strukturerad intervjumetodik med leverantörsansvarig på utvecklarorganisationen med process- och begreppsmodellen enligt steg 1 som bas

- c. **Analys:** Intervjuernas analyseras utifrån frågor såsom; Vilka begrepp är lämpliga? Vilka begrepp och vilka begreppsrelationer anser respondenterna är mest effektiva? Hur ser man att begreppen förhåller sig till varandra?
- d. **Utkomst:** En utvecklad och enhetlig process och begreppsmodell

En tänkbar uppföljning av intervjuerna, och i viss mån för spridning av kunskap, är att anordna workshops eller fokusgruppsintervjuer för att ytterligare och under interaktiva former belysa analysen. Potentiellt skulle projektet kunna koppla an mot metoden Value Sensitive Design (Davis & Nathan, 2015) för att konstruktivt kunna såväl klassificera som att med myndigheterna arbeta fram en mer ansvarsfull etisk kravställning av förklarbarhet.

Det är viktigt att poängtera att projektet i stora delar är explorativt eftersom det inte finns relaterad forskning, såväl internationellt som nationellt, som kan axla projektet. Vi är fullt beredda på att anpassa planerad metod, datakällor och arbeta direkt med intressenter i konkreta projekt (exempelvis med specifik myndighet eller SKR).

Tidsplan

Projektet bemannas huvudsakligen med en doktorand, Jacob Dexe, i två år. Jacob har nyligen genomfört 50%-seminarium (motsvarande licentiatseminarium) och ger sig i och med det föreslagna projektet in på en delvis ny forskningsinriktning. Professor Henrik Artman deltar i sin egenskap av huvudhandledare.

Eftersom doktoranden redan finns på plats kan projektet påbörjas omedelbart efter beslut och beräknas sedan pågå i två år fram till disputation. En tentativ tidsplan följer:

- Q1-Q2: Offentliga handlingar begärs ut och intial innehållsanalys genomförs.
- Q2-Q3: Intervjuer bokas och genomförs.
- Q4-Q5: Första publikation.
- Q5-Q6: Kompletterande datainsamling genomförs i form av dokumentstudier och eventuell GDPR-begäran.
- Q7-Q8: Andra publikation.
- Q8: Disputation.

Förväntat resultat

Leverabler till Konkurrensverket:

- Ensade termer och begrepp kopplat till förklarbarhet, transparens och dylikt.
- Omvärldsbevakning -- hur ser det ut med upphandlingen just nu? (Deskriptivt)
- Processmodell för icke-funktionella krav inom AI (Preskriptivt)
- Skillnader mellan medborgares respektive tjänstepersoners uppfattning av och förståelse för automatiskt fattade beslut.
- Grund för framtida vägledning för beställning av AI-tjänster med krav på förklarbarhet

Spridning

Frågeställningarna om förklarbarhet och transparens väcker stor uppmärksamhet såväl i akademien som i den offentliga debatten. Det bör därför finnas goda möjligheter att nå ut med forskningsresultat i många sammanhang. De sökande har också god vana vid populärvetenskaplig presentation.

Utbildningsprogrammen på KTH och även andra universitet har i hög grad behov att formulera kurser som lyfter etiska frågor inom algoritmiskt beslutsfattande, AI eller Machine Learning aspekter. Vi ser att framtida ingenjörer inom såväl traditionell datalogisk utbildning som inom specialiseringar inom AI och maskininlärning har ett behov av att å ena sidan kunna kravställa systemupphandlingar eller ta ansvar som systemleverantör.

Projektet berör ett av konkurrensverkets huvudområden, upphandling. Projektet kan ha framtida betydelse för hur man upphandlar AI-system, automatiserat beslutsfattande och andra automatiserade processer som berör såväl myndigheters som allmänhetens behov att förstå bevekelsegrunderna för enskilda beslut. I förlängningen kan detta även påverka hur IT-marknaden generellt beaktar såväl människans roll i systemutveckling (såsom genom utvärderingar och processer för att beakta förklarbarhet) som att "mjuka" (icke-funktionella) krav kan utgöra en viktig konkurrensfördel som blir avgörande för val mellan olika tekniska lösningar.

Referenser

- Adebayo J, Gilmer J, Muelly M, Goodfellow I, Hardt M, Kim B (2018) Sanity checks for saliency maps. In: *Advances in Neural Information Processing Systems*, pp 9505–9515
- Alhazmi, T. McCaffer, R. (2000) Project Procurement System Selection Model *Journal of Construction Engineering and Management* Vol. 126, Issue 3 (May 2000), [https://doi.org/10.1061/\(ASCE\)0733-9364\(2000\)126:3\(176\)](https://doi.org/10.1061/(ASCE)0733-9364(2000)126:3(176))
- Andrews R, Diederich J, Tickle AB (1995) Survey and critique of techniques for extracting rules from trained artificial neural networks. *Knowledge-based systems* 8(6):373–389
- Artman, H. "Procurer usability requirements : Negotiations in contract development," in *NordiCHI '02 Proceedings of the second Nordic conference on Human-computer interaction*, 2002, pp. 61-70.
- Artman, H. and S. Zällh, "Finding a way to Usability : Procurement of a taxi dispatch system," *Cognition, Technology & Work*, vol. 7, no. 3, pp. 141-155, 2005.
- Artman, H. et al., *Acquisition of usable IT : Acquisition projects to reflect on*. Stockholm : KTH Royal Institute of Technology, 2009.
- Artman, H. et al., *Att beställa något användbart är inte uppenbart : En motiverande bok om att beställa användbarhet*. US-AB, 2010.
- Barakat N, Bradley AP (2010) Rule extraction from support vector machines: a review. *Neurocomputing* 74(1-3):178–190

- Cui, Ruomeng and Li, Meng and Zhang, Shichen, *AI and Procurement* (April 7, 2020). *Manufacturing & Service Operations Management*, Available at SSRN: <http://dx.doi.org/10.2139/ssrn.3570967>
- Davis, Janet & Nathan, Lisa. (2015). *Value Sensitive Design: Applications, Adaptations, and Critiques. Handbook of Ethics, Values, and Technological Design: Sources, Theory, Values and Application Domains.* 11-40. https://doi.org/10.1007/978-94-007-6970-0_3.
- de Laat PB (2018) *Algorithmic decision-making based on machine learning from big data: Can transparency restore accountability? Philosophy & Technology* 31(4):525–541.
- Denk, T., Hedström, K., & Karlsson, F. (2019). *Medborgarna och automatiserat beslutsfattande. I Storm och Stiltje : SOM-institutets 74:e forskarantologi (1:a uppl., s. 183–196). Göteborg: SOM-institutet. Hämtad från <http://urn.kb.se/resolve?urn=urn:nbn:se:oru:diva-74823>*
- Dexe, J., Franke, U., & Rad, A. (2020b, december). *Nordic lights? National AI policies for doing well by doing good. In Journal of Cyber Policy. Vol 5, issue 3. Routledge.*
- Dexe, J., Ledendal, J., & Franke, U. (2020a, september). *An Empirical Investigation of the Right to Explanation Under GDPR in Insurance. In International Conference on Trust and Privacy in Digital Business (pp. 125-139). Springer, Cham. https://doi.org/10.1007/978-3-030-58986-8_9*
- Du M, Liu N, Hu X (2019) *Techniques for interpretable machine learning. Communications of the ACM* 63(1):68–77, <https://doi.org/10.1145/3359786>
- Fleischmann KR, Wallace WA (2005) *A covenant with transparency: Opening the black box of models. Communications of the ACM* 48(5):93–97.
- Guidotti R, Monreale A, Ruggieri S, Turini F, Giannotti F, Pedreschi D (2018) *A survey of methods for explaining black box models. ACM computing surveys (CSUR)* 51(5):1–42, <https://doi.org/10.1145/3236009>
- RiR 2018:20: *Staten som inköpare av konsulttjänster – tänk först och handla sen. Rapport från Riksrevisionen.*
- RiR 2020:22: *Automatiserat beslutsfattande i statsförvaltningen – effektivt, men kontroll och uppföljning brister. Rapport från Riksrevisionen.*
- Markensten, E. and H. Artman, "Procuring a usable system using unemployed personas," in *ACM International Conference Proceeding Series*, 2004, pp. 13-22.
- Nature (2016) *More accountability for big-data algorithms. Nature* 537(7621):449, <https://doi.org/10.1038/537449a>
- Samek W, Binder A, Montavon G, Lapuschkin S, Müller KR (2016) *Evaluating the visualization of what a deep neural network has learned. IEEE transactions on neural networks and learning systems* 28(11):2660–2673
- Turilli M, Floridi L (2009) *The ethics of information transparency. Ethics and Information Technology* 11(2):105–112
- Zerilli J, Knott A, Maclaurin J, Gavaghan C (2019) *Transparency in algorithmic and human decision- making: Is there a double standard? Philosophy & Technology* 32(4):661–683

1. CV Henrik Artman

1.1 Personal Data

Name: Henrik Artman

[REDACTED]

[REDACTED]

[REDACTED]

Language: Swedish (native), English (Fluent)

Current employment

Professor in Human-Computer Interaction at KTH CSC. Date of employment at KTH: 2001-01-01,
Professor since: 2011-01-01

1.2 Diploma

1992 Bachelor of Philosophy at the department of psychology, Stockholm university PLUS
50 points in courses of Computer Science, Stockholm University

1999 Ph.D. in Communication Studies at the department Tema at Linköpings
University,

2005 Appointed associate professor (docent) at Royal Institute of Technology (KTH).

1.3. Employment's

2011- Professor at KTH

2010- FOI, Research Director, ("Forskningschef"), on commissioned research

2001-11 KTH, Senior Lecturer ("Universitetslektor"), on leave ("tjänstledig") 80%

2009-10 FOI, "Laborator" (50% may-july, 80% 090901-091231)

2007-8 Verva (Swedish Administrative Development Agency) , 80%

2001- KTH: Since 2001 Artman is employed as a full time university lecturer at KTH NADA.

1.4. Research Interests

User Centered Design in Procurement Processes

Already in 2001 I started a research project how procurer organizations could work with user centered design and usability. I have since then published several articles on how to organize the acquisition process with usability in mind, as well as how to teach engineers to think in UCD-terms.

Service design

Currently I work together with ABB to organize their service design processes of maintaining mining robots. I also work with service design in customer-advisory meetings at banks and governmental energy advisory (energimyndigheten). Both projects focus of designing services and relation to information technology.

E-government and information technology in citizen communication

Webb-services and the design of citizen communication systems is one of my areas. Lately my work has focused on explainable AI as well as Cyber Situational awareness from a human-computer interaction perspective.

2 List of selected publications

2.1. Refereed journal articles

1. Aronsson, S., Artman, H., Brynielsson, J., Lindquist, S., Ramberg, R. (2019). The Pedagogics of Simulator Training: An Activity Theory Based Analysis and Comparison of Complex Decision Making Training Simulator Facilities. "Cognition, Technology and Work".
2. Aronsson, S., Artman, H., Lindquist, S., Mitchell, M., Persson, T., Ramberg, R., Romero, M., ter Vehn, P. (2019). Supporting After Action Review in Simulator Mission Training: Co-creating visualization concepts for training of fast-jet fighter pilots. *Journal of Defense Modeling and Simulation, JDMS*. SAGE. DOI: 10.1177/1548512918823296
3. Hulten, M., House, D., Artman, H. (2016) A Model to analyse students cooperative Ideageneration in conceptual design. *International journal of technology and design education*, ISSN 0957-7572, E-ISSN 1573-1804
4. Ramberg, R., Artman, H., Karlberg, K. (2013). Designing Learning Opportunities in Interaction Design: Interactionaries as a means to study and teach student design processes. *Int. Journal Designs for learning*, vol 6.
5. Arvola, M., Artman, H. (2008). Studio Life: The construction of interaction Design. *Digital Kompetanse /Nordic Journal of Digital Literacy*. Vol. 3., pp.78-96.
6. Artman, H., Arvola, M. (2006) Enactments in Interaction Design: How Designers Make Sketches Behave. *Artifact*, 99999 (1), 1749-3463.
7. Artman, H., Zäll, S. (2005) Finding a way to Usability. *Cognition, Technology and Work*, 7, pp. 141-155
8. Johansson B, Artman H & Waern Y (2001) Technology in Crisis Management Systems - ideas and effects. *Document Design Journal of research and problem solving in organizational communication special issue Pragmatics in Crisis Vol 2, Issue 3*, pp. 247-258.
9. Artman, H. (2000) Team situation assessment and information distribution. *Ergonomics*, vol. 43, no 8, 1111-1128
10. Artman, H., Waern, Y. (1999) Distributed cognition in an emergency Co-ordination Center. *Cognition, Technology & Work*, 1: 237-246.
11. Artman, H. (1999). Situation awareness and co-operation within and between Hierarchical Units in Dynamic Decision Making. *Ergonomics*, vol. 42, no.11, 1404-1417.

2.2. Selected refereed conference proceedings

12. Heyman, S., Artman, H. (2015). Computer Support for Financial Advisors and Their Clients: Co-creating an Investment Plan. Accepted to CSCW 2015.
13. Tariq, A.M., Brynielsson, J., Artman, H. (2014). The Security Awareness Paradox: A case study. In *Proceedings of IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining 2014 (ASONAM 2014)*, Beijing, China.

14. Artman, H., House, D., Hulten, M. (2014) Design Learning Opportunities in Engineering Education: A case study of students solving an interaction–design task. Conference on Designs for Learning.
15. Eriksson, E., Artman, H., Swartling, A. (2013) The Secret Life of a Persona: when the personal becomes private. 31st Annual CHI Conference on Human Factors in Computing Systems: Changing Perspectives, CHI 2013; Paris; France; 27 April 2013 through 2 May 2013, 2677-2686
16. Tariq, M. A., Brynielsson, J. & Artman, H. (2012). Storytelling for Tackling Organized Cybercrime. In: Proceedings of the 26th BCS Conference on Human Computer Interaction. Paper presented at 26th BCS Conference on Human Computer Interaction, 12th - 14th September 2012, Birmingham, UK (pp. 1-4). BISL.
17. Tariq, M. A., Brynielsson, J. & Artman, H. (2012). Framing the attacker in organized cybercrime. In: Proceedings - 2012 European Intelligence and Security Informatics Conference, EISIC 2012. Paper presented at 2012 European Intelligence and Security Informatics Conference, EISIC 2012, 22 August 2012 through 24 August 2012, Odense (pp. 30-37). IEEE.
18. Artman, H., Brynielsson, J., Johansson, B. J. & Trnka, J. (2011). Dialogical Emergency Management and Strategic Awareness in Emergency Communication. In: Proceedings of the 8th International ISCRAM Conference. Paper presented at ISCRAM. Lisbon, Portugal, May 2011.
19. Artman, H., Ramberg, R., Karlgren, K. & Strååt, B. (2012). Designing Interaction in Interaction Design: Using interactionaries in order to understand student use of interaction design concepts. In: Designs for Learning 2012: Conference Proceedings: . Paper presented at Designs for Learning 2012, 3rd International Conference Exploring Learning Environments 25-27 April 2012, Copenhagen, Denmark (pp. 14-16). Copenhagen, Denmark: Aalborg University.
20. Holmlid, S., Lantz, A., Artman, H. (2008). Design management of interaction design. The fourth Art of Management and Organization Conference, The Banff centre, Alberta, Canada, 9th-12th September 2008.
21. Lantz, A., Artman, H., Ramberg, R. (2005). Interaction Design as Experienced by Practitioners. Proceedings of In The Making. Copenhagen.
22. Artman, H. Ramberg, Sundholm, H. Cerratto-Pargman. T. (2005). Action Context and Target Context Representations. International Conference on Computer Supported Collaborative Learning (CSCL'05). Taiwan.
23. Markensten, E., Artman, H. (2004). Procuring Usable Systems Using Unemployed Personas. Proceedings of NordiCHI'04, Tampere, Finland, October, 13-23
24. Sundholm, H., Ramberg, R., & Artman, H., (2004). Learning Conceptual Design: Collaborative Activities with Electronic Whiteboards. CADE04 - Computers in Art and Design Education., Copenhagen, June, 2004.
25. Sundholm, H., Artman, H., Ramberg, R. (2004) Backdoor Creativity – Technological support and Collaborative Creativity. In F.Darses, Zarate, P., Zackland, M. (Eds.) Proceedings of Conference on the Design of Cooperative Systems (COOP'04). Heyeres, France, May, 2004, 99-114
26. Artman, H. (2002) Procurer Usability Requirements: Negotiations in contract development. Proceedings of NORDICHI 02, 61-70.

27. Arnborg, S., Artman, H., Brynielsson, J., and Wallenius, K. (2000) Information Awareness in Command and Control: Precision, Quality, Utility Third International Conference on Information Fusion, ThB1/25-32, Paris, France, 2000.
28. Artman, H., Persson, M. (2000) Old Practices - new technology: Observations of how established practices meet new technology. In: R. Dieng, A. Gibson, L. Kersenty, G. De Michelis (eds.) Designing Cooperative Systems. p. 35-49. Amsterdam: Ios Press Ohmsha.
29. Artman, H., Granlund, R. (1998) Team Situation Awareness Using Graphical or Textual Databases in Dynamic Decision Making. In T. Green, L. Bannon, C. Warren, J. Buckley (Eds.) Cognition and cooperation. Proceedings of 9th Conference of Cognitive Ergonomics. (pp. 149-155). Limerick: Ireland.
30. Artman, H., Garbis, C. (1998). Situation Awareness as Distributed Cognition. In T. Green, L. Bannon, C. Warren, Buckley (Eds.) Cognition and cooperation. Proceedings of 9th Conference of Cognitive Ergonomics. (pp. 151-156). Limerick: Ireland.J.
31. Garbis, C., Artman, H. (1998). Co-ordination and Communication as Distributed Cognition. In F. Darses, Zarate, P. (Eds.) Proceedings of Conference on the Design of Cooperative Systems (COOP'98). (pp.1-12). Cannes, France.

2.3. Selected books and other publications

32. Artman, H., Dovhammar, U., Holmlid, S., Lantz, A., Lindquist, S., Markensten, E., Swartling, A. (2010). Att beställa någon användbart är inte uppenbart. Published at KTH.
33. Artman, H. (2010). Att beställa e-tjänster. Beställarens roll i systemutveckling. In K. Lindblad-Gidlund, S. Eriksen, A. Ranerup (red.). Förvaltning och medborgarskap i förändring. Lund: Studentlitteratur.
34. Artman, H., Ramberg, R. (2009). Delade representationer och Kollaborativt Lärande av Interaktionsdesign [DEKAL]. Resultatdialog 2009. Vetenskapsrådet. WWW.VR.SE
35. Artman, H., Lindquist, S. (eds.) (2009). Acquisition of Usable IT-systems: Acquisition projects to reflect on. HCI-66. Published as free PDF at KTH CSC.
36. Markensten, E., Artman, H. (2009). Bridging the gap. In Artman, H., Lindquist, S. (eds.) (2009; Forthcoming). Acquisition of Usable IT-systems: Acquisition projects to reflect on. HCI-66. Published as free PDF at KTH CSC.
37. Swartling, A., Artman, H., Dovhammar, U. (2009). A Game of Usability: Visualising responsibility for Usability Through a Football Metaphor. In Artman, H., Lindquist, S. (eds.) (2009; Forthcoming). Acquisition of Usable IT-systems: Acquisition projects to reflect on. HCI-66. Published as free PDF at KTH CSC.
38. Følstad, H., Artman, H., Krogstie, J. (eds.) (2006) User Involvement and representation in e-Government projects. NordiCHI, Oslo, Oct. 15, 2006

3. Selected funding of research

- VINNOVA – Procurement Competence, 01-04, Main Applicant and project leader Henrik Artman, Co-applicant Stefan Holmlid, 3.3 Msek,
- Swedish Research Council – Shared Representations and Collaborative Learning and Interaction design (DeKaL), 04-06, Main applicant Robert Ramberg, Principal Investigators Henrik Artman, Ann Lantz, Tessa Ceratto, Klas Karlgren, 4.5 Msek
- FMV – Usability Requirements and Responsibility, 05-07, Main Applicant and project leader Henrik Artman, 1.5 Msek

- VINNOVA - Enabling Technology and User-Centered Design on the Procurer Organizations Terms, 2006-2008, Vinnova, Main Applicant and Project leader Henrik Artman. Co-applicant Stefan Holmlid. 3.3 Msek
- FMV – Usability in Procurement – 2007, Main Applicant and project leader Henrik Artman, 1.3 Msek.
- VINNOVA – Lineage of Conquest – 2009, Main applicant: Yashar Moradbahkti. Project leader from KTH, Henrik Artman, 2 Msek
- Naturvårdsverket - EFFICIENT ENVIRONMENTAL INSPECTION AND ENFORCEMENT – 2009-2013, Main Applicant: Adam Jacobson, SU, Work package project leader : Henrik Artman, 18 Msek.
- Länsförsäkringar, Talför -Transparenta algoritmer i
- försäkringsbranschen (TALFÖR), 2018-2020, Main Applicant: Ulrik Franke. Approx. 6 Msek
- I am for the moment working on commissioned research at FOI, 30%

Curriculum Vitae

KONKURRENSVERKET

2021-02-01

Avd AF

Dnr 27/2021

KSnr 363 Aktbil 25

Name	Jacob Dexe
██████████	██████████
Position	Researcher, RISE – Research institutes of Sweden. Industrial PhD-student – KTH.
██████████	██████████
██████████	██████████
Degrees	2015, master, political science, Lund University.

Currently an industrial PhD-student at RISE and KTH working on transparency and ethics in algorithmic decision making. Previously worked extensively on issues relating to privacy and personal data collection. Served as an expert on the opinion poll Delade Meningar for seven years, previously head of business area The Connected Individual at RISE, manager of the Engaging Privacy project, and advisor to the think tank Digital Utmaning.

Publications (journals)

- Dexe, Jacob, Franke, Ulrik, 2020. Nordic lights? National AI policies for doing well by doing good. *Journal of Cyber Policy*, 5:332–349. <https://doi.org/10.1080/23738871.2020.1856160>
- Dexe, Jacob, Franke, Ulrik, Rad, Alexander, 2021. Transparency and insurance professionals – A study of Swedish insurance practice attitudes and future development. *The Geneva Papers on Risk and Insurance – Issues and Practice*. Forthcoming. <https://doi.org/10.1057/s41288-021-00207-9>

Publications (other)

- Dexe, Jacob, Franke, Ulrik, Anneli, Avatare-Nöu, Rad, Alexander, 2020. Towards Increased Transparency with Value Sensitive Design. I: *Artificial Intelligence in HCI. HCI International 2020*, ss 3–15. Springer. https://doi.org/10.1007/978-3-030-50334-5_1
- Dexe, Jacob, Ledendal, Jonas, Franke, Ulrik, 2020. An empirical investigation of the right to explanation under GDPR in insurance. I: *Trust, Privacy and Security in Digital Business. The 17th International Conference on Trust, Privacy and Security in Digital Business – TrustBus 2020*. Springer. https://doi.org/10.1007/978-3-030-58986-8_9

Publications (non academic)

- Digital Utmaning, 2017. Företagens lilla katekes om integritet i en uppkopplad värld – Termer och begrepp och vad de betyder. (Main author).
- Lundblad, Joakim, Dexe, Jacob, 2016. Rewiring Europe: Five priorities for a Lasting Digital economy. Report from Wilfried Martens Centre.

Employment

- RISE – Research institutes of Sweden. Researcher. 2015 – present
Currently working as an Industrial PhD-student. Previous positions include project manager, head of a business area, and internal advocate for privacy projects.

- Fores. Research and communications assistant. 2013 – 2014
Worked with the Digital Society-program at the think tank Fores. The job included coordination of research projects, arranging seminars, writing and publishing opinion pieces and copyediting.
- Social Sciences Student Association at Lund University. Vice chairman. 2012 – 2013
Full time position. Responsible for educational matters, such as coordinating student representation, participating as a student representative and working for student rights.

Education

- MSc in Political Science. Lund University. 2011 – 2015
Master's degree in political science. Focused on policy processes, policy making, internet governance, and political action. The master thesis "Constraining the digital world" covered different perspectives on regulation on network neutrality. Hiatus for working at the student union and then went on an internship in the fall of 2013. Finished the thesis during the summer of 2015.