

Finansdepartementet

Stockholm, 2025-06-03

Färdplan för Sverige (SOU 2025:12)

Kungl. tekniska högskolan, KTH, har anmodats att lämna remissyttrande gällande betänkandet AI-kommissionens Färdplan för Sverige (SOU 2025:12). KTH tillstyrker i huvudsak betänkandet. KTH vill anföra följande.

Generella övergripande synpunkter

KTH välkomnar AI-kommissionens betänkande och understryker behovet av skyndsamhet i att adressera, och inte endast analysera, de möjligheter och problem som utvecklingen inom AI-området för med sig.

KTH välkomnar AI-kommissionens förslag rörande högre utbildning och forskning. Utvecklingen är teknikintensiv och de mycket stora resurser som internationellt läggs på forskning inom AI-området medför att utvecklingen är mycket snabb. Hög kunskapsnivå och excellent forskning är en nödvändig förutsättning för att en liten aktör som Sverige ska vara relevant. De satsningar som AI-kommissionen föreslår avseende excellenscentra och forskarskolor med mera är därför mycket välkomna. KTH är väl rustat för att ta ansvar för att etablera och driva dessa. KTH vill framhålla vikten av att uppnå kritisk massa och betydelsen av nära samverkan med andra svenska aktörer, såväl statliga som privata, för att verksamheten ska kunna ges det djup och den bredd som krävs för excellens.

Cybercampus Sverige är navet för svensk forskning om AI och cybersäkerhet och ger KTH unika förutsättningar för verksamhet inom området. Förslaget att tilldela Cybercampus Sverige ett permanent anslag på 50 miljoner kronor per år för att bygga upp och driva en särskild forskningsavdelning för AI och säkerhet skulle, om det beslutas, säkerställa en forskningsmiljö präglad av excellens, vilket är en förutsättning för att Sverige ska vara en relevant aktör inom området.

KTH stödjer också betänkandets förslag att inrätta ett nationellt institut för AI-säkerhet i nära anslutning till KTH och Cybercampus Sverige, liksom förslaget om en forskningsfond på 50 miljoner kronor för att stimulera breda samarbeten. KTH erinrar samtidigt om att sådana breda samarbeten bör omfatta såväl inhemska som utländska samarbetsparter.

I betänkandet läggs stor vikt på positiva konsekvenser som ekonomisk tillväxt och konkurrenskraft. Samtidigt innebär AI-teknikens betydande energi- och resursbehov ett problem som varje land behöver hantera, och när betänkandet berör miljöaspekter är det oftast just i samband med energianvändning för AI-infrastruktur. KTH vill framhålla att AI-utvecklingen här har nära kopplingar till andra forskningsområden som energi- och miljöteknik, vilka bör uppmärksammas då Sverige även här ligger långt framme.

KTH efterlyser en mer djuplodande miljökonsekvensbedömning. Betänkandet berör endast i mindre utsträckning AI:s potential att bidra till eller försvåra klimatomställningen, liksom frågor som välbefinnande, rättvis fördelning av teknikens fördelar eller hur AI-tekniken kan bidra till att lösa klimat- och miljöutmaningar.

Kommentarer och ändringsförslag gällande spetsforskning och högre utbildning

Post-doktorer, gästforskare och forskarskolor

Förslagen avseende post-doktorer, gästforskare och forskarskolor är lovvärda. KTH noterar likheterna med WASP som dock har haft ett uttalat fokus mot natur- och ingenjörsvetenskaperna. AI-kommissionen skriver:

Enligt AI-kommissionen ska målet vara att ha excellent forskning på världsfrenten inom AI i sig, samtidigt som vi satsar på ämnesspecifik AI-kompetens inom breda vetenskapsområden, som natur- och teknikvetenskap, medicin och hälsa, samt humaniora och samhällsvetenskap.¹

KTH påtalar vikten av att säkerställa att satsningen görs för breda vetenskapsområden, inklusive humaniora och samhällsvetenskap, och inte enbart för natur- och ingenjörsvetenskaperna.

Excellenscenter

KTH välkomnar förslaget att ge framstående forskare långsiktig finansiering och frihet att utforska nya idéer utan kortsiktiga krav. I betänkandet föreslås att detta skulle kunna ske via inrättandet av tre excellenscenter inom AI vilka skulle utgöra en nationell forskningsinfrastruktur, i likhet med SciLifeLab inom livsvetenskaperna. KTH anser att det i så fall är viktigt att tydliggöra hur huvudmannaskapet för dessa excellenscenter ska organiseras, något som AI-kommissionen delegerat till Vetenskapsrådet att göra.

KTH anser vidare att excellenscentra kan vara ett effektivt nyttjande av forskningsmedel under förutsättning att de organiseras så att de har förmåga att anpassa sin verksamhet till och utifrån de snabba förändringar som forskningsfältet kring AI präglas av.

¹ Bilaga 2, s. 53.

Vidare bör möjligheterna med smal AI tas i beaktande i utformningen av dessa excellenscentra. Smal AI utgör ett viktigt område som bör prioriteras även om det sällan figurerar i forskningspolitiska diskussioner. I kontrast till AGI, artificiell generell intelligens, har svensk akademisk forskning och industri helt andra förutsättningar att konkurrera på högsta nivå inom smal AI. Fokus inom smal AI är mer på integration av domänspecifik expertis snarare än att skala upp generiska modeller, språkmodeller, och träna dessa mot stora datamängder, något som varit den dominerande strategin för utvecklingen av AGI. KTH anser därför att inrättandet av excellenscentra bör göras med hänsyn tagen till möjligheterna med smal AI.

Utbildningsvetenskapliga perspektiv

Vikten av grundläggande högskoleutbildning, till exempel inom matematik, betonas i AI-kommissionens uppdrag. Detta för att förbättra grundförutsättningarna för att möta morgondagens behov inom AI-användning och AI-utveckling. KTH välkomnar förslaget på kompetensutveckling inom AI och att framtidssäkra innehållet i utbildningarna. KTH noterar att om AI ska integreras i utbildningarna kräver detta troligtvis att man etablerar särskilda beräkningsresurser som är reserverade för utbildning. Detta bör dock kunna täckas inom de 750 miljoner kronor som budgeterats för detta ändamål.

KTH anser vidare att det hade varit önskvärt att AI-kommissionens uppdrag varit bredare och inkluderat en djupare analys gällande AI i skola och utbildning. Betänkandet visar ibland brist på förankring i pedagogisk forskning.

I interaktionen med AI-läraren behöver eleven inte heller känna sig dum om den inte förstår och kan fråga hur många gånger som helst.²

Om lärare i svenska skolan får elever att känna sig dumma är det ett grundläggande pedagogiskt problem som inte bör lösas med AI-system.

Kommentarer och ändringsförslag på övriga delar i kommissionens förslag

Demokratisk kunskapsspridning och folkbildning

KTH anser att det är positivt att färdplanen betonar vikten av folkbildning och vuxenutbildning, med hem-PC-reformen som förebild. Sveriges historiska framgångar inom digitalisering har byggt på breda satsningar på folkbildning, från 70-talets kurs i datoranvändning som lockade 100 000 att läsa datakunskap, till 80-talets satsning på bredddatautbildning då löntagarfonder skulle användas för att utbilda hela Sveriges befolkning. Dessa satsningar gav hela befolkningen grundläggande datakunskap och bidrog till att Sverige var i digitaliseringens framkant. Dock bör folkbildning och kunskapslyft inte primärt syfta till att öka användningen av AI-verktyg, utan både till utvecklandet av nya AI-agenter och till demokratisk dialog om *när*, *hur* och *om* tekniken ska

² Bilaga 2, s. 65.

användas. Syftet med kunskapsspridning bör vara att ge människor möjlighet att både dra nytta av tekniken och *påverka* dess utveckling i önskad riktning. Samtidigt föreslås mer pengar för läraryft inom högre utbildning än för hela befolkningens kunskapslyft. Detta sker dessutom i en kontext där regeringen minskat studieförbundens statsanslag med 250 miljoner kronor för 2024, med ytterligare nedskärningar om sammanlagt 500 miljoner kronor fram till 2026. Medan PC-reformen kostade fyra miljarder kronor föreslås 500 miljoner kronor till AI-reformens motsvarighet. Ett grundläggande perspektiv vid tidigare reformer har varit att ge svenskarna kontroll över digitala system som påverkar dem själva. Detta perspektiv saknas i betänkandet.

Vidare saknas även diskussion om hur tillgång till dessa nya kraftfulla AI-verktyg kan påverka vårt demokratiska beslutsfattande. Om till exempel ett AI-system ansvarar för förslag om skolplacering för grundskolelever i en kommun. Har medborgarna då rätt att få insyn i hur denna AI har tränats och vad som ligger bakom förslaget? Hur ska offentliga tjänstemän som vänjer sig vid avancerat beslutsstöd i form av AI klara av att avvika från det AI:n föreslår? Likt flygplanspiloter som ibland måste ta över spakarna från autopiloten för att landa.

Datadelning och integritet

AI-kommissionen föreslår ett paradigmskifte för offentlighets- och sekretesslagstiftningen:

Utgångspunkten bör vara offentlighet mellan myndigheter i stället för sekretess när det gäller skydd för enskildas personliga och ekonomiska förhållanden.³

KTH instämmer i att lagar kring registersystem försvårar möjligheten att dela data och det finns ett tydligt behov av att modernisera lagstiftningen. Denna svårighet finns dock där av en anledning och i de konkreta förslag som färdplanen lägger fram saknas målsättningen att fortsatt bevara god integritet för personuppgifter. I färdplanen står visserligen att

AI-kommissionen anser att Sverige bör [...] ta en ledande roll inom så kallade privacy enhancing technologies (PET).⁴

Dock framstår det endast som en önskan om forskning. KTH anser att det borde vara ett krav vid framtagande av nya lagar kring och implementeringen av nya registersystem.

AI-kommissionens förslag riskerar att undergräva viktiga principer. Offentlighets- och sekretesslagen (OSL) fyller flera viktiga funktioner: Den stärker rättssäkerheten, förhindrar skapandet av ett totaliserande kontrollsystem, och minskar risken för felaktiga beslut baserade på ofullständig

³ Bilaga 2, s. 39.

⁴ Bilaga 2 s. 42.

information från andra myndigheter. OSL är konstruerad så att data *får* delas mellan myndigheter om den enskilde *samtycker*. En bättre väg framåt vore att stärka individens samtycke snarare än att försvaga lagen. Utifrån detta är det KTH:s bedömning att olika juridiska hinder kommer att vara betydande vid implementering av AI inom offentlig sektor och att betänkandet borde ha innehållit förslag om att utreda ny lagstiftning med bibehållande av god personlig integritet.

På samma sätt beskrivs EU:s regelverk som problematiska och hämmande för innovation:

De regulatoriska hindren och osäkerheterna kring AI-regelverken gör också att europeiska aktörer riskerar att få tillgång till de senaste och mest avancerade AI-verktygen senare än aktörer i andra delar av världen.⁵

Denna inställning kan ifrågasättas. Om innovationer hindras av regleringar som skyddar människor, djur och natur från att komma till skada kan detta vara motiverat. Teknisk innovation och spridning är inte ett självändamål utan slutmålet bör vara AI-lösningar som skapar större välfärd och kommer till nytta för så många som möjligt på ett ansvarsfullt sätt.

Vidare betonas i betänkandet att teknisk utveckling inte är ödesbestämd utan formas av våra val. Samtidigt framhåller kommissionen ibland ökad utveckling och användning av AI som självändamål, utan tydligt beskriven samhällsnytta.

Begränsad diskussion om algoritmbias

Rapporten nämner kort problemet med bias i träningsdata men gör en begränsad analys av hur detta kan leda till diskriminering och orättvisa utfall för vissa grupper.⁶ Med tanke på hur centralt detta problem är i den internationella AI-diskussionen är detta förvånande.

Avsaknad av etisk analys kring övervakning

Vidare diskuteras hur AI kan användas för övervakning och kontroll i brottsbekämpande syfte, samtidigt som det ges begränsad uppmärksamhet åt integritetsfrågor och risker för övervakning. Detta är särskilt anmärkningsvärt i diskussionen om "biometrisk identifiering", där tekniska möjligheter beskrivs utan etisk analys.⁷

Militär användning av AI

I betänkandet står att det "[...] finns även risker i framtiden med militär tillämpning av AI. "⁸ Denna formulering anser KTH är anmärkningsvärd då vi redan de senaste åren sett militär användning av AI i stor omfattning. I kriget

⁵ Bilaga 2, s. 59.

⁶ Bilaga 2, s. 36–37.

⁷ Bilaga 2, s. 48.

⁸ Bilaga 2, s. 47.

mellan Ryssland och Ukraina finns många rapporter om målidentifikation och automatisk eldgivning med AI i drönare. I konflikten mellan Israel och Hamas finns alarmerande berättelser om hur målidentifikation har gjorts med AI-system, där människan som är satt att göra själva beslutet inte kunnat ifrågasätta AI-systemets bedömning. Militariserad AI borde rimligen diskuteras som en viktig etisk risk.

Arbetsmarknadsperspektiv

Kommissionen slår fast att människors arbetsuppgifter förvisso kommer att förändras, men oftast till det bättre, med mer intressant och givande innehåll och färre repetitiva och monotona inslag. KTH vill framhålla att detta i många fall kan komma att ske, men att negativa konsekvenser och övergångsproblem skulle kunna ges en mer utförlig beskrivning. Potentiellt omfattande strukturomvandlingar kan komma att medföra arbetslöshet inom vissa sektorer. Även om nya jobb skapas kan stora behov av omskolning tillkomma. Tidsdiskrepansen mellan när nya jobb skapas jämfört med när de försvinner, bör heller inte underskattas. KTH anser att dessa effekter bör studeras och följas upp.

AI i offentlig sektor

AI-kommissionen betonar vikten av samverkan mellan offentlig och privat sektor, men har otillräcklig diskussion om risker när privata vinstdrivande aktörer får inflytande över gemensamma offentliga resurser, särskilt kring data och kritisk infrastruktur. Den föreslagna AI-verkstaden för offentlig sektor ska exempelvis göra det möjligt för privata aktörer att utveckla lösningar för offentliga utmaningar, men diskussionen om hur maktbalansen ska upprätthållas är begränsad. Detta jämför AI-implementering med arbete mot diskriminering, hållbarhet och pandemihantering, vilket kan ifrågasättas. KTH anser att användning av AI bör drivas av tydlig samhällsnytta, inte av implementering för implementeringens skull.

Vidare framhålls AI som en lösning på den demografiska utmaningen för den offentliga sektorn.

I takt med att samhället i övrigt utvecklas ökar också kraven på offentliga tjänster. Inom ett par år bedöms sektorn behöva leverera 125 procent av dagens välfärd, men med 75 procent av dagens bemanning.⁹

KTH menar att detta riskerar att framställa AI som en enkel teknisk lösning på komplexa samhällsutmaningar, utan tillräcklig hänsyn till att vård och omsorg i grunden handlar om mänskliga relationer och beslutsfattning utifrån en helhetssyn.

⁹ Bilaga 2, s. 16.

Ekonomiska aspekter

AI-kommissionen föreslår omfattande satsningar (totalt 16,7 miljarder kronor över 10 år) vilka skulle motiveras av att den ekonomiska nyttan av ett ”fullständigt införande av AI i svensk offentlig förvaltning” skulle uppgå till 140 miljarder kronor per år. Här lutar sig AI-kommissionen bland annat på beräkningar av Myndigheten för digital förvaltning. En redovisning av hur dessa beräkningar gjorts skulle underlätta förståelsen för Kommissionens förslag.

AI-verkstad för offentlig sektor

KTH anser att idén om att utveckla ett ramverk för AI anpassat för offentlig sektor är en bra ambition. KTH anser dock att förslaget att ge Försäkringskassan och Skatteverket i uppdrag att tillsammans etablera och förvalta en AI-verkstad för offentlig sektor är orealistiskt på flera plan.

KTH menar att det är oklart vad som ska byggas och i synnerhet vad som menas med en minsta livskraftig produkt (MVP) i detta sammanhang. Att bygga ett ramverk för de olika typerna av verksamheter inom offentlig sektor utan en tydlig målbild brukar sällan lyckas. AI-kommissionen kan eventuellt anse att specifikationen av en sådan målbild ingår i uppdraget till Försäkringskassan och Skatteverket, men detta är oklart. En bättre strategi kan vara att inleda med att utveckla några AI-baserade lösningar (smal AI) för specifika verksamheter inom offentlig sektor och sedan successivt bygga vidare på dessa.

KTH anser vidare att beloppen som reserveras för en AI-verkstad och tillvägagångssättet är orealistiska. Gissningsvis baseras den tekniska plattformen för AI-verkstaden på något av de stora ramverken för AI. Det är ett mycket kvalificerat arbete att på denna grund bygga datastrukturer och beräkningskapacitet som är anpassade för AI inom offentlig sektor och som det går att bygga vidare på. Det går att göra men skulle kräva ett team i världsklass, i nivå med det team på Meta som byggde pyTorch eller det på Google Brain som byggde TensorFlow. Att attrahera nödvändig expertis till ett sådant team kommer kräva lönenivåer och tillgång till en FoU-miljö som varken Försäkringskassan eller Skatteverket kan erbjuda. Kostnaden riskerar dessutom att bli betydligt högre än de 500 miljoner kronor som AI-kommissionen föreslår. Jämförelsevis anger AI-kommissionen att den totala kostnaden för att utveckla en svensk hälsomodell under en femårsperiod är 5 miljarder kronor.

Svensk språkmodell

KTH håller med om vikten av att ha en språkmodell för svenska språket. En sådan skulle kunna utgöra ett stort stöd för bland annat offentlig förvaltning. KTH är dock mycket tveksam till det av AI-kommissionen föreslagna arbetssättet.

Att finansiera och driva arbetet via en nationell samordnare som leder en kommitté med flera olika aktörer riskerar att bli tungrovt. Risken är stor att det som utvecklas och släpps snabbt blir inaktuellt. Detta var tydligt redan med AI

Swedens satsning på GPT-SW3, vilket är en språkmodell baserat på GPT 3.5 som anpassats för de nordiska språken genom att den tränats mot utvalda svenska texter (1.2 terabyte träningsdata). Modellen byggde på AI-teknik som var föråldrad redan då modellen släpptes. En bättre användning av medlen skulle vara att ta fram högkvalitativa textdata som kan användas för att utveckla språkmodeller för svenska språket. Själva utvecklingen av språkmodellerna kan göras av olika forskargrupper eller företag.

Vidare menar KTH att dessa forskargrupper eller företag borde få tillgång till den rikedom av data som det offentliga besitter, särskilt Kungliga biblioteket. KTH föreslår ett upplägg liknande hur licenser för 3G delades ut. Data delas ut kostnadsfritt till ett fåtal aktörer. Dessa aktörer kan då träna sina språkmodeller på denna data och sedan ge myndigheter, kommuner och regioner gratis tillgång till de tränade modellerna.

Digital infrastruktur

KTH välkomnar att kommissionen uppmärksammat vikten av digital infrastruktur. Det finns flera pågående initiativ, både nationellt (NAISS) och inom EU (EuroHPC), för att bygga upp en beräkningsinfrastruktur för AI inom akademisk forskning. Vidare håller AI Factories på att etableras i Linköping vilket kommer erbjuda liknande beräkningsinfrastruktur för start-ups samt små- och medelstora företag.

KTH noterar dock att det saknas en motsvarande infrastruktur för lagring och delning av tränings- och testdata. Denna brist är ett återkommande problem även i de fall data hämtats från en extern part som hanterar lagring. Det beror på att den rådata som externa parter lagrar ofta behöver förbehandlas. Detta innebär att det behövs säker lagring, i europeisk ägo, som uppfyller europeiska lagkrav och integritetsaspekter. En sådan viktig infrastruktur kräver finansiering, något som inte beaktats i betänkandet.

KTH anser vidare att en än mer akut problematik är svårigheten att långsiktigt underhålla högkvalitativ mjukvara som utvecklats inom akademiska forskningsprojekt. Sådan mjukvara utgör en del av forskningsinfrastrukturen, speciellt inom AI-forskning. KTH ser gärna att avancerat stöd och utveckling av mjukvara ingår i begreppet digital infrastruktur.



BESLUT

Datum för beslut:
2025-06-03

Diarienummer:
HS-2025-0570

Beslut om remissvar på betänkandet AI-kommissionens Färdplan för Sverige (SOU 2025:12)

Beslutet

Rektor beslutar att Kungl. Tekniska högskolan lämnar remissvar enligt bilaga 1 på betänkandet AI-kommissionens Färdplan för Sverige (SOU 2025:12).

Ärendet

Kungl. Tekniska högskolan har av Finansdepartementet anmodats att lämna synpunkter på betänkandet AI-kommissionens Färdplan för Sverige (SOU 2025:12).

Remissvaret har handlagts inom ledningskansliet och beretts med KTH:s forskningsberedning samt en expertgrupp. Expertgruppen bestod av professor Joakim Gustafsson, institutionen för intelligenta system, universitetslektor Arvind Kumar, institutionen för datavetenskap, professor Lukas Käll, institutionen för kemi, bioteknolog och hälsa, professor Svante Linusson, institutionen för matematik, professor Sebastiaan Meijer, institutionen för medicinteknik och hälsosystem, universitetslektor Lina Rahm, institutionen för filosofi och historia och professor Ozan Öktem, institutionen för matematik,

Detta beslut har fattats av rektor Anders Söderholm efter föredragning av utbildningsledare Johanna Andersson Raeder. Närvarande vid beslutet var universitetsdirektör Kerstin Jacobsson, biträdande universitetsdirektör Fredrik Oldsjö, ordförande för Tekniska högskolans studentkår Gustav Heldt och mötets sekreterare Annette Grahn.

Denna handling har hanterats digitalt och är därför inte undertecknad.

Bilaga 1: Kungl. Tekniska högskolans remissvar på betänkandet AI-kommissionens Färdplan för Sverige (SOU 2025:12).

Sändlista

Finansdepartementet fi.remissvar@regeringskansliet.se och fi.ofa.dof.remisser@regeringskansliet.se

Expeditionsdatum:

2025-06-03

Bilaga 1: Kungl. Tekniska högskolans remissvar på betänkandet AI-kommissionens Färdplan för Sverige (SOU 2025:12)

Kungl. tekniska högskolan, KTH, har anmodats att lämna remissyttrande gällande betänkandet AI-kommissionens Färdplan för Sverige (SOU 2025:12). KTH tillstyrker i huvudsak betänkandet. KTH vill anföra följande.

Generella övergripande synpunkter

KTH välkomnar AI-kommissionens betänkande och understryker behovet av skyndsamhet i att adressera, och inte endast analysera, de möjligheter och problem som utvecklingen inom AI-området för med sig.

KTH välkomnar AI-kommissionens förslag rörande högre utbildning och forskning. Utvecklingen är teknikintensiv och de mycket stora resurser som internationellt läggs på forskning inom AI-området medför att utvecklingen är mycket snabb. Hög kunskapsnivå och excellent forskning är en nödvändig förutsättning för att en liten aktör som Sverige ska vara relevant. De satsningar som AI-kommissionen föreslår avseende excellenscentra och forskarskolor med mera är därför mycket välkomna. KTH är väl rustat för att ta ansvar för att etablera och driva dessa. KTH vill framhålla vikten av att uppnå kritisk massa och betydelsen av nära samverkan med andra svenska aktörer, såväl statliga som privata, för att verksamheten ska kunna ges det djup och den bredd som krävs för excellens.

Cybercampus Sverige är navet för svensk forskning om AI och cybersäkerhet och ger KTH unika förutsättningar för verksamhet inom området. Förslaget att tilldela Cybercampus Sverige ett permanent anslag på 50 miljoner kronor per år för att bygga upp och driva en särskild forskningsavdelning för AI och säkerhet skulle, om det beslutas, säkerställa en forskningsmiljö präglad av excellens, vilket är en förutsättning för att Sverige ska vara en relevant aktör inom området.

KTH stödjer också betänkandets förslag att inrätta ett nationellt institut för AI-säkerhet i nära anslutning till KTH och Cybercampus Sverige, liksom förslaget om en forskningsfond på 50 miljoner kronor för att stimulera breda samarbeten. KTH erinrar samtidigt om att sådana breda samarbeten bör omfatta såväl inhemska som utländska samarbetsparter.

I betänkandet läggs stor vikt på positiva konsekvenser som ekonomisk tillväxt och konkurrenskraft. Samtidigt innebär AI-teknikens betydande energi- och resursbehov ett problem som varje land behöver hantera, och när betänkandet berör miljöaspekter är det oftast just i samband med energianvändning för AI-infrastruktur. KTH vill framhålla att AI-utvecklingen här har nära kopplingar till andra forskningsområden som energi- och miljöteknik, vilka bör uppmärksammas då Sverige även här ligger långt framme.

KTH efterlyser en mer djuplodande miljökonsekvensbedömning. Betänkandet berör endast i mindre utsträckning AI:s potential att bidra till eller försvåra klimatomställningen, liksom frågor som välbefinnande, rättvis fördelning av teknikens fördelar eller hur AI-tekniken kan bidra till att lösa klimat- och miljöutmaningar.

Kommentarer och ändringsförslag gällande spetsforskning och högre utbildning

Post-doktorer, gästforskare och forskarskolor

Förslagen avseende post-doktorer, gästforskare och forskarskolor är lovvärda. KTH noterar likheterna med WASP som dock har haft ett uttalat fokus mot natur- och ingenjörsvetenskaperna. AI-kommissionen skriver:

Enligt AI-kommissionen ska målet vara att ha excellent forskning på världsfrenten inom AI i sig, samtidigt som vi satsar på ämnesspecifik AI-kompetens inom breda vetenskapsområden, som natur- och teknikvetenskap, medicin och hälsa, samt humaniora och samhällsvetenskap.¹

KTH påtalar vikten av att säkerställa att satsningen görs för breda vetenskapsområden, inklusive humaniora och samhällsvetenskap, och inte enbart för natur- och ingenjörsvetenskaperna.

Excellenscenter

KTH välkomnar förslaget att ge framstående forskare långsiktig finansiering och frihet att utforska nya idéer utan kortsiktiga krav. I betänkandet föreslås att detta skulle kunna ske via inrättandet av tre excellenscenter inom AI vilka skulle utgöra en nationell forskningsinfrastruktur, i likhet med SciLifeLab inom livsvetenskaperna. KTH anser att det i så fall är viktigt att tydliggöra hur huvudmannaskapet för dessa excellenscenter ska organiseras, något som AI kommissionen delegerat till Vetenskapsrådet att göra.

KTH anser vidare att excellenscentra kan vara ett effektivt nyttjande av forskningsmedel under förutsättning att de organiseras så att de har förmåga att anpassa sin verksamhet till och utifrån de snabba förändringar som forskningsfältet kring AI präglas av.

Vidare bör möjligheterna med smal AI tas i beaktande i utformningen av dessa excellenscentra. Smal AI utgör ett viktigt område som bör prioriteras även om det sällan figurerar i forskningspolitiska diskussioner. I kontrast till AGI, artificiell generell intelligens, har svensk akademisk forskning och industri helt andra förutsättningar att konkurrera på högsta nivå inom smal AI. Fokus inom smal AI är mer på integration av domänspecifik expertis snarare än att skala upp generiska modeller, språkmodeller, och träna dessa mot stora datamängder, något som varit den dominerande strategin för utvecklingen av AGI. KTH anser därför att inrättandet av excellenscentra bör göras med hänsyn tagen till möjligheterna med smal AI.

Utbildningsvetenskapliga perspektiv

Vikten av grundläggande högskoleutbildning, till exempel inom matematik, betonas i AI-kommissionens uppdrag. Detta för att förbättra grundförutsättningarna för att möta morgondagens behov inom AI-användning och AI-utveckling. KTH välkomnar förslaget på kompetensutveckling inom AI och att framtidssäkra innehållet i utbildningarna. KTH noterar att om AI ska integreras i utbildningarna kräver detta troligtvis att man etablerar särskilda beräkningsresurser som är reserverade för utbildning. Detta bör dock kunna täckas inom de 750 miljoner kronor som budgeterats för detta ändamål.

¹ Bilaga 2, s. 53.

KTH anser vidare att det hade varit önskvärt att AI-kommissionens uppdrag varit bredare och inkluderat en djupare analys gällande AI i skola och utbildning. Betänkandet visar ibland brist på förankring i pedagogisk forskning.

I interaktionen med AI-läraren behöver eleven inte heller känna sig dum om den inte förstår och kan fråga hur många gånger som helst.²

Om lärare i svenska skolan får elever att känna sig dumma är det ett grundläggande pedagogiskt problem som inte bör lösas med AI-system.

Kommentarer och ändringsförslag på övriga delar i kommissionens förslag

Demokratisk kunskapsspridning och folkbildning

KTH anser att det är positivt att färdplanen betonar vikten av folkbildning och vuxenutbildning, med hem-PC-reformen som förebild. Sveriges historiska framgångar inom digitalisering har byggt på breda satsningar på folkbildning, från 70-talets kurs i datoranvändning som lockade 100 000 att läsa datakunskap, till 80-talets satsning på bredddatautbildning då löntagarfonder skulle användas för att utbilda hela Sveriges befolkning. Dessa satsningar gav hela befolkningen grundläggande datakunskap och bidrog till att Sverige var i digitaliseringens framkant. Dock bör folkbildning och kunskapslyft inte primärt syfta till att öka användningen av AI-verktyg, utan både till utvecklandet av nya AI-agenter och till demokratisk dialog om *när*, *hur* och *om* tekniken ska användas. Syftet med kunskapsspridning bör vara att ge människor möjlighet att både dra nytta av tekniken och *påverka* dess utveckling i önskad riktning. Samtidigt föreslås mer pengar för lärarlyft inom högre utbildning än för hela befolkningens kunskapslyft. Detta sker dessutom i en kontext där regeringen minskat studieförbundens statsanslag med 250 miljoner kronor för 2024, med ytterligare nedskärningar om sammanlagt 500 miljoner kronor fram till 2026. Medan PC-reformen kostade fyra miljarder kronor föreslås 500 miljoner kronor till AI-reformens motsvarighet. Ett grundläggande perspektiv vid tidigare reformer har varit att ge svenskarna kontroll över digitala system som påverkar dem själva. Detta perspektiv saknas i betänkandet.

Vidare saknas även diskussion om hur tillgång till dessa nya kraftfulla AI-verktyg kan påverka vårt demokratiska beslutsfattande. Om till exempel ett AI-system ansvarar för förslag om skolplacering för grundskolelever i en kommun. Har medborgarna då rätt att få insyn i hur denna AI har tränats och vad som ligger bakom förslaget? Hur ska offentliga tjänstemän som vänjer sig vid avancerat beslutsstöd i form av AI klara av att avvika från det AI:n föreslår? Likt flygplanspiloter som ibland måste ta över spakarna från autopiloten för att landa.

Datadelning och integritet

AI-kommissionen föreslår ett paradigmskifte för offentlighets- och sekretesslagstiftningen:

Utgångspunkten bör vara offentlighet mellan myndigheter i stället för sekretess när det gäller skydd för enskildas personliga och ekonomiska förhållanden.³

KTH instämmer i att lagar kring registersystem försvårar möjligheten att dela data och det finns ett tydligt behov av att modernisera lagstiftningen. Denna svårighet finns dock där av en anledning och i de konkreta förslag som färdplanen lägger fram saknas målsättningen att fortsatt bevara god integritet för personuppgifter. I färdplanen står visserligen att

AI-kommissionen anser att Sverige bör [...] ta en ledande roll inom så kallade privacy

² Bilaga 2, s. 65.

³ Bilaga 2, s. 39.

enhancing technologies (PET).⁴

Dock framstår det endast som en önskan om forskning. KTH anser att det borde vara ett krav vid framtagande av nya lagar kring och implementeringen av nya registersystem.

AI-kommissionens förslag riskerar att undergräva viktiga principer. Offentlighets- och sekretesslagen (OSL) fyller flera viktiga funktioner: Den stärker rättssäkerheten, förhindrar skapandet av ett totaliserande kontrollsystem, och minskar risken för felaktiga beslut baserade på ofullständig information från andra myndigheter. OSL är konstruerad så att data *får* delas mellan myndigheter om den enskilde *samtycker*. En bättre väg framåt vore att stärka individens samtycke snarare än att försvaga lagen. Utifrån detta är det KTH:s bedömning att olika juridiska hinder kommer att vara betydande vid implementering av AI inom offentlig sektor och att betänkandet borde ha innehållit förslag om att utreda ny lagstiftning med bibehållande av god personlig integritet.

På samma sätt beskrivs EU:s regelverk som problematiska och hämmande för innovation:

De regulatoriska hindren och osäkerheterna kring AI-regelverken gör också att europeiska aktörer riskerar att få tillgång till de senaste och mest avancerade AI-verktygen senare än aktörer i andra delar av världen.⁵

Denna inställning kan ifrågasättas. Om innovationer hindras av regleringar som skyddar människor, djur och natur från att komma till skada kan detta vara motiverat. Teknisk innovation och spridning är inte ett självändamål utan slutmålet bör vara AI-lösningar som skapar större välfärd och kommer till nytta för så många som möjligt på ett ansvarsfullt sätt.

Vidare betonas i betänkandet att teknisk utveckling inte är ödesbestämd utan formas av våra val. Samtidigt framhåller kommissionen ibland ökad utveckling och användning av AI som självändamål, utan tydligt beskriven samhällsnytta.

Begränsad diskussion om algoritmbias

Rapporten nämner kort problemet med bias i träningsdata men gör en begränsad analys av hur detta kan leda till diskriminering och orättvisa utfall för vissa grupper.⁶ Med tanke på hur centralt detta problem är i den internationella AI-diskussionen är detta förvånande.

Avsaknad av etisk analys kring övervakning

Vidare diskuteras hur AI kan användas för övervakning och kontroll i brottsbekämpande syfte, samtidigt som det ges begränsad uppmärksamhet åt integritetsfrågor och risker för övervakning. Detta är särskilt anmärkningsvärt i diskussionen om "biometrisk identifiering", där tekniska möjligheter beskrivs utan etisk analys.⁷

Militär användning av AI

I betänkandet står att det "[...] finns även risker i framtiden med militär tillämpning av AI."⁸ Denna formulering anser KTH är anmärkningsvärd då vi redan de senaste åren sett militär användning av AI i stor omfattning. I kriget mellan Ryssland och Ukraina finns många rapporter om målidentifikation och automatisk eldgivning med AI i drönare. I konflikten mellan Israel och Hamas finns alarmerande berättelser om hur målidentifikation har gjorts med AI-

⁴ Bilaga 2 s. 42.

⁵ Bilaga 2, s. 59.

⁶ Bilaga 2, s. 36–37.

⁷ Bilaga 2, s. 48.

⁸ Bilaga 2, s. 47.

system, där människan som är satt att göra själva beslutet inte kunnat ifrågasätta AI-systemets bedömning. Militariserad AI borde rimligen diskuteras som en viktig etisk risk.

Arbetsmarknadsperspektiv

Kommissionen slår fast att människors arbetsuppgifter förvisso kommer att förändras, men oftast till det bättre, med mer intressant och givande innehåll och färre repetitiva och monotona inslag. KTH vill framhålla att detta i många fall kan komma att ske, men att negativa konsekvenser och övergångsproblem skulle kunna ges en mer utförlig beskrivning. Potentiellt omfattande strukturomvandlingar kan komma att medföra arbetslöshet inom vissa sektorer. Även om nya jobb skapas kan stora behov av omskolning tillkomma. Tidsdiskrepansen mellan när nya jobb skapas jämfört med när de försvinner, bör heller inte underskattas. KTH anser att dessa effekter bör studeras och följas upp.

AI i offentlig sektor

AI-kommissionen betonar vikten av samverkan mellan offentlig och privat sektor, men har otillräcklig diskussion om risker när privata vinstdrivande aktörer får inflytande över gemensamma offentliga resurser, särskilt kring data och kritisk infrastruktur. Den föreslagna AI-verkstan för offentlig sektor ska exempelvis göra det möjligt för privata aktörer att utveckla lösningar för offentliga utmaningar, men diskussionen om hur maktbalansen ska upprätthållas är begränsad. Detta jämför AI-implementering med arbete mot diskriminering, hållbarhet och pandemihantering, vilket kan ifrågasättas. KTH anser att användning av AI bör drivas av tydlig samhällsnytta, inte av implementering för implementeringens skull.

Vidare framhålls AI som en lösning på den demografiska utmaningen för den offentliga sektorn.

I takt med att samhället i övrigt utvecklas ökar också kraven på offentliga tjänster. Inom ett par år bedöms sektorn behöva leverera 125 procent av dagens välfärd, men med 75 procent av dagens bemanning.⁹

KTH menar att detta riskerar att framställa AI som en enkel teknisk lösning på komplexa samhällsutmaningar, utan tillräcklig hänsyn till att vård och omsorg i grunden handlar om mänskliga relationer och beslutsfattning utifrån en helhetssyn.

Ekonomiska aspekter

AI-kommissionen föreslår omfattande satsningar (totalt 16,7 miljarder kronor över 10 år) vilka skulle motiveras av att den ekonomiska nyttan av ett ”fullständigt införande av AI i svensk offentlig förvaltning” skulle uppgå till 140 miljarder kronor per år. Här lutar sig AI-kommissionen bland annat på beräkningar av Myndigheten för digital förvaltning. En redovisning av hur dessa beräkningar gjorts skulle underlätta förståelsen för Kommissionens förslag.

AI-verkstad för offentlig sektor

KTH anser att idén om att utveckla ett ramverk för AI anpassat för offentlig sektor är en bra ambition. KTH anser dock att förslaget att ge Försäkringskassan och Skatteverket i uppdrag att tillsammans etablera och förvalta en AI-verkstad för offentlig sektor är orealistiskt på flera plan.

KTH menar att det är oklart vad som ska byggas och i synnerhet vad som menas med en minsta livskraftig produkt (MVP) i detta sammanhang. Att bygga ett ramverk för de olika typerna av verksamheter inom offentlig sektor utan en tydlig målbild brukar sällan lyckas. AI-kommissionen kan eventuellt anse att specifikationen av en sådan målbild ingår i uppdraget till Försäkringskassan och Skatteverket, men detta är oklart. En bättre strategi kan vara att inleda

⁹ Bilaga 2, s. 16.

med att utveckla några AI-baserade lösningar (smal AI) för specifika verksamheter inom offentlig sektor och sedan successivt bygga vidare på dessa.

KTH anser vidare att beloppen som reserveras för en AI-verkstad och tillvägagångssättet är orealistiska. Gissningsvis baseras den tekniska plattformen för AI-verkstaden på något av de stora ramverken för AI. Det är ett mycket kvalificerat arbete att på denna grund bygga datastrukturer och beräkningskapacitet som är anpassade för AI inom offentlig sektor och som det går att bygga vidare på. Det går att göra men skulle kräva ett team i världsklass, i nivå med det team på Meta som byggde pyTorch eller det på Google Brain som byggde TensorFlow. Att attrahera nödvändig expertis till ett sådan team kommer kräva lönenivåer och tillgång till en FoU-miljö som varken Försäkringskassan eller Skatteverket kan erbjuda. Kostnaden riskerar dessutom att bli betydligt högre än de 500 miljoner kronor som AI-kommissionen föreslår. Jämförelsevis anger AI-kommissionen att den totala kostnaden för att utveckla en svensk hälsomodell under en femårsperiod är 5 miljarder kronor.

Svensk språkmodell

KTH håller med om vikten av att ha en språkmodell för svenska språket. En sådan skulle kunna utgöra ett stort stöd för bland annat offentlig förvaltning. KTH är dock mycket tveksam till det av AI-kommissionen föreslagna arbetssättet.

Att finansiera och driva arbetet via en nationell samordnare som leder en kommitté med flera olika aktörer riskerar att bli tungrott. Risken är stor att det som utvecklas och släpps snabbt blir inaktuellt. Detta var tydligt redan med AI Swedens satsning på GPT-SW3, vilket är en språkmodell baserat på GPT 3.5 som anpassats för de nordiska språken genom att den tränats mot utvalda svenska texter (1.2 terabyte träningsdata). Modellen byggde på AI-teknik som var föråldrad redan då modellen släpptes. En bättre användning av medlen skulle vara att ta fram högkvalitativa textdata som kan användas för att utveckla språkmodeller för svenska språket. Själva utvecklingen av språkmodellerna kan göras av olika forskargrupper eller företag.

Vidare menar KTH att dessa forskargrupper eller företag borde få tillgång till den rikedom av data som det offentliga besitter, särskilt Kungliga biblioteket. KTH föreslår ett upplägg liknande hur licenser för 3G delades ut. Data delas ut kostnadsfritt till ett fåtal aktörer. Dessa aktörer kan då träna sina språkmodeller på denna data och sedan ge myndigheter, kommuner och regioner gratis tillgång till de tränade modellerna.

Digital infrastruktur

KTH välkomnar att kommissionen uppmärksammat vikten av digital infrastruktur. Det finns flera pågående initiativ, både nationellt (NAISS) och inom EU (EuroHPC), för att bygga upp en beräkningsinfrastruktur för AI inom akademisk forskning. Vidare håller AI Factories på att etableras i Linköping vilket kommer erbjuda liknande beräkningsinfrastruktur för start-ups samt små- och medelstora företag.

KTH noterar dock att det saknas en motsvarande infrastruktur för lagring och delning av tränings- och testdata. Denna brist är ett återkommande problem även i de fall data hämtats från en extern part som hanterar lagring. Det beror på att den rådata som externa parter lagrar ofta behöver förbehandlas. Detta innebär att det behövs säker lagring, i europeisk ägo, som uppfyller europeiska lagkrav och integritetsaspekter. En sådan viktig infrastruktur kräver finansiering, något som inte beaktats i betänkandet.

KTH anser vidare att en än mer akut problematik är svårigheten att långsiktigt underhålla högkvalitativ mjukvara som utvecklats inom akademiska forskningsprojekt. Sådan mjukvara utgör en del av forskningsinfrastrukturen, speciellt inom AI-forskning. KTH ser gärna att avancerat stöd och utveckling av mjukvara ingår i begreppet digital infrastruktur.