

Finansdepartementet

Dnr Fi2025/00457

fi.remissvar@regeringskansliet.se

Remissvar på AI-kommissionens rapport *Färdplan för Sverige* (SOU 2025:12)

Sveriges Författarförbund (SFF) önskar inkomma med kommentar på ovan angivna SOU. Sveriges Författarförbund är en facklig organisation för verksamma författare och litterära översättare. Vi har funnits sedan 1893 och ser som vår uppgift att tillvarata och bevaka medlemmarnas juridiska, ekonomiska och ideella intressen.

Vi välkomnar AI-kommissionens rapport *Färdplan för Sverige* och delar ambitionen att Sverige ska ta en ledande roll i utvecklingen och användningen av artificiell intelligens (AI).

Samtidigt vill vi framhålla ett antal avgörande aspekter som måste beaktas för att säkra ett hållbart och rättssäkert ekosystem för AI-utveckling, särskilt gäller detta användningen av upphovsrättsligt skyddat material och utvecklingen av svenska språkmodeller. Hittills har mycket av AI-träningen inneburit att globala techjättar annekterat rättigheter från svenska rättighetshavare, utan att rättighetshavarna kompenserats för användningen. Följderna kan bli katastrofala för det upphovsrättsliga ekosystem vi byggt upp i Sverige och EU.

1. Introduktion

Vi ser att AI kommer att bli avgörande för litteraturens framtida konkurrenskraft, på samma sätt som internet en gång omdanade litteraturmarknaden. Länder som ligger först i AI-utvecklingen får tillgång till snabbare innovation, högre produktivitet och ökad samhällsnytta. AI har potential att förstärka människors lärande, bredda kunskapstillgången och möjliggöra nya forskningsgenombrott. Att investera i AI är därför att investera i framtidens humankapital och nationella infrastruktur. Teknologin bör betraktas som en bärande del av morgondagens samhälle.

Precis som bredband och kollektivtrafik kräver AI långsiktiga och samordnade investeringar – i infrastruktur, regelverk, datatillgång, kompetensförsörjning och öppenhet. Egen AI-infrastruktur gör det möjligt att bygga lösningar på vårt språk, med vår kultur och i vår verklighet. Utan en sådan infrastruktur riskerar vi att bli ett land som är beroende av andras system och att därigenom förlora digital självständighet och

suveränitet. AI kommer att bli en underliggande kunskapsmotor som är avgörande för människors lärande och spridning av kunskap. Men för att den kunskapen ska vara korrekt, trovärdig och användbar krävs innehåll av hög kvalitet. AI-modeller lär sig från text och journalistik såväl som litteratur är de mest kunskapsrika och strukturerade källorna.

Språkmodeller tränas på enorma mängder text, men journalistiskt innehåll är unikt i sin uppbyggnad, faktagranskning och samhällsförankring. Utan kvalitetsjournalistik riskerar AI att basera sina svar på sämre, ytligare och mindre pålitligt material. På samma sätt är svensk litteratur avgörande för att modellerna ska kunna behålla språkliga nyanser och förstå svensk kulturhistorisk kontext. Medier ger AI-modeller välskrivna, aktuell och tillförlitlig information. Norska Mimirprojektet har analyserat effekten av upphovsrättsskyddat material på språkmodeller genom att träna modeller med både upphovsrättsskyddat och fritt material. Resultaten visade att modeller som tränats på nyhetsmedier och faktaböcker presterade bäst i faktakvalitet. Tilläggsstudier visar också att skönlitterär text tillför språklig finess, kulturell förståelse och historiska referensramar.

Det är därför avgörande att det tydliggörs att:

- Skapandet och användningen av språkmodeller förutsätter tillgång till böcker och nyhetsmaterial på svenska.
- Träning och användning av språkmodeller innebär ett intrång i upphovsrätten, vilket i sin tur kräver både samtycke och ekonomisk ersättning.
- Upphovsrättsinnehavare får skälig ersättning – inte bara för att det är juridiskt nödvändigt, utan för att modellerna kan användas för att skapa tjänster som i praktiken riskerar att konkurrera ut de ursprungliga innehållsskaparna.
- Vi ser att Sveriges författare skulle kunna bidra aktivt till att skapa stora svenska språkmodeller av hög kvalitet – tillgängliga för mediebranschen, näringslivet, forskningen, vården, skolan och myndigheterna – under förutsättning att både träning och användning av dessa modeller sker med respekt för upphovsrätten och genom en fungerande licensmodell.

2. Nyttan av språkmodeller på svenska

Varför behöver vi då en språkmodell på svenska? Under Jensen Huangs, grundare och vd Nvidia, besök i Sverige i maj 2025, där han utnämndes till hedersdoktor vid Linköpings universitet, uppmanade Huang Sverige att utveckla en "nationell AI" för att inte vara beroende av USA och Kina. Han betonade att AI kodifierar en nations kunskap, kultur och värderingar, och att det är viktigt att denna intelligens inte outsourcas. Han uttryckte att varje land bör "äga produktionen av sin egen intelligens", vilket innebär att länder bör ha kontroll över sina data och den intelligens som genereras från dem. Han har många gånger fört fram att detta är avgörande för att bevara nationell kultur, samhällsintelligens och historia. Huang uppmanade världsledare att inte vara

överväldigade av AI, utan att omfamna teknologin och integrera den med lokala språk och expertis. Det finns fyra huvudsakliga argument att föra fram:

a) Identitet. Enligt studier, bland annat vid Stanford University finns det betydande utmaningar i dagens stora språkmodeller (LLM). De har svårt att fånga språkliga och kulturella nyanser samt de unika uttryck som speglar normer, attityder, traditioner och värderingar i icke-engelska språk. När en språkmodell tränas på ett annat språk – vanligtvis amerikansk engelska – och sedan anpassas till svenska, går viktiga nyanser, lokal förankring och historiska referenser förlorade. Resultatet blir ofta översättningar med en språklig logik som känns amerikaniserad. En annan studie från Columbia University visar att GPT-4o:s förmåga att spegla ett lands samhällsvärderingar, mätt genom World Values Survey, korrelerar med tillgången på digitala resurser och träningsdatapå det språket. Felprocenten var mer än fem gånger högre för språk med lägst andel träningsdata och resurser jämfört med språk med högst resurser.

Vi vill säkerställa att de värderingar, normer, den historia och kultur som är en integrerad del av det svenska språket också återspeglas i de språkmodeller som genererar ny information och nya berättelser. Språk är inte bara ett verktyg för kommunikation, utan en representation av världens underliggande komplexitet – en modell för hur vi förstår verkligheten. En språkmodell som i huvudsak tränats på en annan kulturs perspektiv, normer och världssyn riskerar att påverka varje nations unika identitet på ett osynligt men genomgripande sätt.

b) Nationell konkurrenskraft. Tillgång till en avancerad språkmodell på svenska är att betrakta som grundläggande digital infrastruktur. Att utveckla landsspecifika språkmodeller kan bidra till teknisk innovation, ekonomisk tillväxt och stärkt självständighet. Genom att investera i AI-forskning och -utveckling kan Sverige främja innovation, skapa arbetstillfällen och stärka sin globala position inom AI. Genom att anpassa språkmodeller till svenska behov kan tekniken optimeras inom sektorer som utbildning, vård, media, myndigheter och näringsliv.

c) Nationell säkerhet. Ur ett säkerhetsperspektiv är det särskilt viktigt att behålla möjligheten att finjustera modeller med hjälp av konfidentiell och känslig nationell data. I vissa situationer kan det vara avgörande att denna data inte lämnar landets gränser. Ett visst inflytande över språkmodellerna är därför nödvändigt för att värna både vår informationssuveränitet, integritet och motståndskraft.

d) Resiliens och förtroende. För att skapa samhälleligt förtroende krävs att människor vet att de språkmodeller som används tränats på transparenta och trovärdiga källor. I dagsläget saknas ett effektivt system för att reglera, granska eller korrigera de narrativ som genereras av AI-system. Som enskild nation är det svårt att säkerställa att

logiken i stora språkmodeller främjar det gemensamma bästa – snarare än att driva specifika kommersiella eller geopolitiska intressen.

Många länder, såsom Japan, Taiwan, Island, Danmark, Norge, Spanien och Nederländerna, har redan påbörjat initiativ för språkmodeller tränade på det egna språket – ofta i konsortier där forskning, näringsliv och myndigheter samverkar, med delvis statlig finansiering. Det är i sådana konsortiebildningar viktigt att säkerställa att även mindre kommersiella aktörer uppmuntras att delta, för att skapa ett levande och konkurrensutsatt ekosystem.

Vi vill bidra till utvecklingen av stora svenska språkmodeller genom att tillgängliggöra mediedata av hög kvalitet, förutsatt att det sker inom ramen för en långsiktigt hållbar licensmodell, samt genom att delta aktivt i kravställningen och användningen av dessa modeller.

3. Etablering av en ny struktur för utveckling av svenska språkmodeller

Färdplanen föreslår att staten bör ta ett övergripande ansvar för utvecklingen av svenska språkmodeller (kap. 6.2.5). Vistödjer detta, men vill poängtera behovet av att skapa ett nytt initiativ som organiseras på ett sätt som säkerställer att utvecklingen sker i nära förbindelse med samtliga berörda parter, dvs både upphovsrättsinnehavare till den data som är nödvändig för att träna språkmodellerna, men också framtida användare av språkmodellerna såsom oberoende ledande forskare samt representanter för behovsägare, dataägare, forskare och infrastrukturägare. Detta organ bör ges ett tydligt mandat och finansiering för att bygga svenska språkmodeller med hög kvalitet och öppenhet - här kan sannolikt inspiration hämtas från Kanadas Mitacs-program som *Färdplan för Sverige* lyfter fram (kap. 6.2.1) - samt ges ett tydligt mandat och ansvar för att utveckla en långsiktigt hållbar licensmodell (se punkt 3).

Vid ett beslut om att initiera ett projekt för en nationell språkmodell krävs starka partnerskap för teknisk infrastruktur. Det handlar om att samla beräkningskraft, datalagring, säkerhet och mjukvarustöd, vilket förutsätter samverkan mellan stat, akademi, näringsliv och tekniska leverantörer för att säkerställa skalbarhet, säkerhet och prestanda. Utan sådana partnerskap riskerar satsningen att bli för långsam, dyr eller tekniskt otillräcklig. Utgångspunkten bör vara att använda en redan existerande öppen LLM (Large Language Model) som basmodell (t.ex. Mistral, LLaMA, Falcon eller GPT-NeoX). Detta sparar tid, resurser och miljöpåverkan jämfört med att träna en modell från grunden. Med svensk data på plats kan man därefter finjustera modellen på svensk text för bättre språkförståelse och instruktionsträna modellen med svenska värderingar och förutsättningar. För detta krävs ett antal avgörande komponenter:

- a) Beräkningskapacitet (compute): Finjustering eller träning av stora språkmodeller kräver tillgång till avancerad hårdvara, såsom GPU-kuster eller specialiserade AI-processorer. Detta kan ske genom partnerskap med nationella superdatorcenter, molnleverantörer eller industriaktörer.
- b) Datainfrastruktur och lagring: Säker, redundant och snabb lagring för enorma datamängder krävs både för träning, utvärdering och löpande drift. Samtidigt måste tillgången till data regleras enligt upphovsrätt, GDPR och andra relevanta lagar.
- c) Energiförsörjning och hållbarhet: AI-träning är energikrävande. Därför krävs planering för effektiv energianvändning och gärna lokalisering i regioner med tillgång till fossilfri el.
- d) IT-säkerhet och nationell datasuveränitet: Språkmodeller som används i offentliga och känsliga sammanhang måste drivas inom säkra miljöer, där dataintegritet, spårbarhet och skydd mot intrång garanteras – något som kräver särskild kompetens och teknik.
- e) Mjukvaruramverk och driftplattformar: För att finjustera eller träna, och använda, språkmodellen behövs stabila ramverk, orkestreringsverktyg, utvärderingsverktyg och API-strukturer. Dessa kräver kontinuerlig uppdatering och expertkompetens.

4. Skönlitteraturens roll i att bygga upp en nationell språkmodell

Kunskapskälla. Journalistik och faktalitteratur är en av de mest kunskapsrika källorna. Utan kvalitetsjournalistik och faktalitteratur riskerar AI att basera sina svar på sämre, ytligare och mindre trovärdigt innehåll.

Förstå samtiden. Det är genom medier AI "förstår världen" i termer av samhälle, demokrati och mänskliga perspektiv. Journalistik och litteratur speglar och förklarar samtid, vilket AI-modeller annars saknar. Nyhetsmedier sätter händelser i ett sammanhang, analyserar politik, tolkar trender, ofta med lokal och tidsbunden relevans. Skönlitteratur är en bärare av nationell, regional och historisk identitet. Den ger kontextuella beskrivningar och åtar sig tidens värderingar, normer, konflikter och ideal. AI utan tillgång till medier förlorar i förmåga att förstå vad som händer i samhället – här och nu.

Mänskliga perspektiv. Skönlitteratur gestaltar komplexa inre liv, moraliska dilemman, känslor och mellanmänskliga relationer. Romaner speglar det mänskliga psyket på ett sätt som faktatexter sällan gör. Stora språkmodeller behöver sådant material för att bättre hantera empatiska och emotionellt nyanserade uppgifter

Språklig och kulturell mångfald. Skönlitteratur innehåller ett varierat, nyanserat och kreativt språk. Detta stärker modellens stilistiska bredd och förståelse för kontextuella skiftningar. Det exponerar modellerna för metaforer, symbolik och idiomatiska uttryck såväl som historiska och stilistiska språkexempel. Lokala medier och lokala berättelser tillför dialekter, sociolekter, olika berättarperspektiv och kulturella referenser som globala AI-modeller annars saknar. Det stärker AI-modellers förmåga att vara relevanta i olika regioner och kontexter. Utan innehåll från medier blir AI-modeller ensidiga, anglocentriska och dåligt anpassade till olika målgrupper.

Marginaliserade röster och specifika perspektiv. Medier och litteratur lyfter fram minoriteter, maktmissbruk, klimatfrågor och granskningar som inte nödvändigtvis dominerar andra informationskällor. Skönlitteratur gestaltar hur olika samhällen ser på kön, klass, etnicitet och makt. AI behöver detta innehåll för att inte reproducera maktobalanser och blinda fläckar. AI-modeller med god kvalitet kräver ett brett, representativt och undersökande innehåll – och det är precis vad medier tillför.

Faktakollade källor. Ett filter i form av kvalitetssäkrade medier är avgörande i en tid av desinformation. AI-modeller riskerar att reproducera falska påståenden om de inte tränas på pålitliga källor. Litteratur och redaktionellt innehåll har redan genomgått kvalitetsgranskning, vilket gör det ovärderligt för modellens robusthet.

5. Ersättning för användning av upphovsrättsskyddat material i AI-träning

Färdplan för Sverige lyfter vikten av att ta hänsyn till upphovsrätten vid träning av språkmodeller och föreslår en ersättningsmodell baserad på avtalslicenser (kap. 6.2.5 och kap. 9.3). Författarförbundet vill understryka att en sådan modell är avgörande för att möjliggöra ett sunt och rättvist AI-ekosystem där rättighetshavare, såsom nyhetsmedier och litteratur, får rimlig ersättning för det innehåll som används i träning av AI-modeller och då AI-modeller används.

För licensavgifter för vidareanvändning av modellerna förordar vi differentierade nivåer beroende på användningsområde. Detta är viktigt också då det finns en uppenbar risk att en rättighetshavares data kan komma att ligga till grund för en konkurrerande tjänst som underminerar rättighetshavarens egen affärsverksamhet. Det behövs därför en differentierad licensmodell där forskning under vissa premisser undantas från avgifter, medan kommersiellt bruk bär full kostnad. Forskningens särställning bör kodifieras, men även i offentlig verksamhet såsom myndigheter och skola bör en lägre men tydlig ersättningsmodell gälla. Det blir här särskilt viktigt att klargöra när forskningsändamålen slutar och en annan typ av användning tar vid. Bara för att ett ursprungligt användande utgör forskning kan detta inte utsträckas till kommande användningsområden då detta lätt skulle kunna leda till missbruk av undantagsbestämmelser med mera.

Det är centralt att förstå att när en AI-modell väl tränats på data är processen oåterkallelig – det går inte att i efterhand välja bort sitt innehåll. Därför måste ersättningsmodellerna inte endast vara tydliga, rättvisa och långsiktigt hållbara, utan även vara implementerade redan från början.

6. Sverige som förebild för licensiering av kvalitetsdata

Sverige har en lång och internationellt erkänd tradition av kollektiv rättighetsförvaltning och licensiering (STIM, Copyswede, Bonus Copyright Access m.fl.). Detta ger Sverige ett bra utgångsläge då ersättningsmodeller för generativ AI ska tas fram. Genom att utgå

från de redan befintliga licensmodeller som finns för upphovsrättsskyddat innehåll kan Sverige med en svensk språkmodell stärka Sveriges konkurrenskraft med respekt för upphovsrätten och samtidigt skapa incitament för innehållsskapande.

Transparens är ett grundvillkor för att överhuvudtaget kunna tillämpa en ersättningsmodell. För att rättighetshavare ska kunna kräva ersättning krävs att det går att fastställa om deras innehåll faktiskt har använts. Om AI-aktörer inte redovisar vilka källor som ingår i träningsdata blir det omöjligt att identifiera, värdera eller reglera användningen. En rättvis modell för kompensation kräver att man vet vilka verk som används, i vilken omfattning och med vilken betydelse för modellens funktion. Det krävs för att kunna:

- Fördela ersättning mellan aktörer
- Avgöra vilka innehållstyper som har störst värde
- Sätta rimliga licenspriser

Även för att kollektiva ersättningsmodeller ska kunna fungera måste det finnas spårbar data om användning. Det är en förutsättning för att fördela ersättningen rättvist över tid, med respekt för olika källors bidrag.

Det är därför också av vikt att Sverige står upp för upphovsrätten och krav på transparens i träningsdata i bland annat det pågående implementeringsarbetet kring EU:s AI förordning och DSM direktivet.


7. Avslutande synpunkter

Författarförbundet vill avslutningsvis betona att utvecklingen av svenska språkmodeller inte kan ske på bekostnad av författares rättigheter, samt att utvecklingen bör ledas av en oberoende aktör i samarbete med en grupp där både upphovsrättsinnehavare och framtida användare är representerade.

Vi efterlyser därför:

- Ett tydligt statligt uppdrag för att skapa förutsättningarna för att en stor språkmodell fungerar väl på svenska och speglar svenska värderingar och förutsättningar
- Ett tydligt statligt uppdrag att skapa en fungerande licensstruktur som säkerställer rättvis ersättning och skydd för innehållsskapare.
- Ett nära samarbete mellan rättighetshavare, forskningsinstitutioner och myndigheter.


Författarförbundet genom Åsa Anesäter, förbundsjurist


Anja Gatu, ordförande