

# Kungliga bibliotekets remissvar på Långsiktig reglering av forskningsdatabaser PM 2022-12-23, U2022/04089

## Sammanfattning

Kungl. biblioteket (KB) är en myndighet med en central roll i forskningen. Enligt 1 § förordning (2008:1421) med instruktion för Kungl. biblioteket, är ”Kungl. biblioteket Sveriges nationalbibliotek och en nationell forskningsinfrastruktur. Myndigheten ska samla in, beskriva, bevara och tillhandahålla fysiskt och digitalt material som behövs för att fullgöra denna uppgift. Myndigheten ska genom verksamheten bidra till den svenska forskningens kvalitet och det demokratiska samhällets utveckling.”

Som framgår av KB:s instruktion görs ingen åtskillnad mellan olika slag av källmaterial. Uppdraget omfattar således såväl fysiska som digitala källor, d.v.s. dokument som samlats in på grundval av Lag (1993:1392) om pliktexemplar av dokument, Lag (2012:492) om pliktexemplar av elektroniskt material (E-plikt), hemsidor som samlas in enligt Förordning (2002:287) om behandling av personuppgifter i Kungl. bibliotekets digitala kulturarvsprojekt (Kultarw3), samt donationer. Instruktionen gör heller inte någon åtskillnad mellan fysiska dokument som av olika skäl har digitaliserats och därefter tillhandahålls i elektroniskt format eller dokument som till sitt ursprung är elektroniska.

Som framgår av 1 § KB:s instruktion är myndigheten en nationell forskningsinfrastruktur och ska tillhandahålla källmaterial. I detta åligger det KB att ”...samla in, beskriva, bevara och tillhandahålla...digitalt material som behövs” för samhällets möjlighet att skapa ny kunskap. Det är därmed KB:s uppfattning att det inte går att isolera KB:s uppdrag från skapandet av ny kunskap och forskningsdata och forskningsdatabaser, KB:s och forskningsinstitutionernas respektive uppdrag överlappar här varandra.

Bland de forskningsförfrågningar om användning av KB:s digitala resurser som inkommit kan t.ex. nämnas en begäran om att löpande få tillgång till textdata som produceras vid digitaliseringen av den svenska dagspressen. Syftet är att över långa tidsserier med statistiska beräkningsmodeller och språkligt referensmaterial möjliggöra analys av trender och förändringar i det samtida svenska språkbruket.

KB:s resurser kan också utgöra grunddata för mer komplexa analyser som t.ex. en studie av den offentliga debatten kring frågan om migration och integration. En sådan studie fokuserar på diskursen inom tre sociala sfärer – sociala medier, dagspress och riksdagen – och möjliggör både jämförelser av dessa sfärer sinsemellan och förändring över tid. I detta projekt utgör KB:s samlingar en central delmängd och analysen bygger på samkörning

mellan delar ur den digitaliserade dagspressen och nyhetssändningar ur KB:s samlingar av radio- och tv-sändningar.

Båda dessa exempel beskriver forskningsansatser där KB:s digitala resurser fyller en avgörande funktion och där bearbetning och analys oundvikligen involverar informationsinnehåll som omfattas av dataskyddslagstiftning. I det större perspektivet utgör dock KB:s data en delmängd. Utvecklingen av nya forskningsmetoder baserade på digitala resurser, och hur fler och fler forskningsfält utnyttjar dessa, har inneburit förändrade förutsättningar. Det är KB:s förståelse att forskningsprojekt som utnyttjar KB:s resurser kommer ha behov av infrastruktur som har förutsättningar för att samla in data till forskningsdatabaser på KB.

Det har således blivit alltmer tydligt att KB:s uppdrag som nationell forskningsinfrastruktur överlappar forskningsinstitutionernas uppdrag. Att utredaren fått i uppdrag att biträda Utbildningsdepartementet med att lämna ett förslag på en långsiktig reglering av forskningsdatabaser (U2020/01944/F) ser KB därför som något positivt. Det förslag som utredaren har lämnat till lag om vissa forskningsdatabaser (promemorian sid 23) omfattar emellertid inte KB. Även om KB inte omfattas av förslaget har KB, givet vad som anförts inledningsvis, yttrat sig enligt vad som framgår nedan.

Det behov som föreligger att skapa förutsättningar för en långsiktig reglering av forskningsdatabaser instämmer KB i. Att utredningen inte lämnar förslag som omfattar KBs verksamhet är beklagligt. KB vill härvid understryka vikten av att regeringen beaktar de synpunkter KB lämnar i detta yttrande och behovet av ett reformerat regelverk för att KB ska kunna skapa och utveckla forskningsdatabaser. Detta blir avgörande för att KB via Forskning och utveckling (FoU) i enlighet med exempelvis Riksrevisionens rapport om statliga myndigheters FoU-verksamhet (rskr 2021/22:7) fortsatt ska ha goda förutsättningar för att tillhandahålla en nationell forskningsinfrastruktur som bidrar till svensk forsknings kvalitet. Eftersom frågan om rättslig grund för KB:s personuppgiftsbehandling alltjämt bereds på Utbildningsdepartementet (tilläggsdirektiv till Forskningsdatautredningen (U 2016:04) delredovisning enligt dir. 2017:29 och 2017:87 i en promemorian *Personuppgiftsbehandling i forskningsbibliotekens verksamhet*, Komm2017/01529/U 2016:04) välkomnar KB en lagstiftning som omfattar all nationell forskningsinfrastruktur av det slag som utredningen pekar på, d.v.s. att lagen inkluderar även KB och KBs specifika uppdrag. Försåvitt regeringen inte bedömer att förslaget till lag bör omfatta KB m.h.t. myndighetens särställning som nationalbibliotek och nationell forskningsinfrastruktur, välkomnar KB att frågan om tydligare rättslig grund för KBs personuppgiftsbehandling bereds färdigt och att det sker skyndsamt.

## KB:s ställningstagande och skälen för dessa

### Promemorians förord sid 2: Långsiktiga förutsättningar för forskningsdatabaser av infrastrukturkaraktär

”Det finns således ett behov av att skapa långsiktiga förutsättningar för forskningsdatabaser av **infrastrukturkaraktär** som innehåller personuppgifter som kan utgöra underlag för flera olika framtida forskningsprojekt inom ett eller flera forskningsområden.”

KB:s instruktion; Uppgifter

1 § Kungl. biblioteket är Sveriges nationalbibliotek och **en nationell forskningsinfrastruktur**. Myndigheten ska samla in, beskriva, bevara och tillhandahålla fysiskt och digitalt material som behövs för att fullgöra denna uppgift. Myndigheten ska genom verksamheten **bidra till den svenska forskningens kvalitet** och det demokratiska samhällets utveckling.

### Skäl för KB:s ställningstagande

KB är således en nationell forskningsinfrastruktur som ska bidra till den svenska forskningens kvalitet. KB instämmer med utredningen att det finns ett behov av att skapa långsiktiga förutsättningar för forskningsdatabaser av infrastrukturkaraktär som innehåller personuppgifter som kan utgöra underlag för flera olika framtida forskningsprojekt inom ett eller flera forskningsområden. I KB:s samlingar finns forskningsdatabaser som ingår i den nationella forskningsinfrastruktur som KB utgör. Utgångspunkten för utredningen är således i allra högsta grad relevant för KB:s uppdrag och den verksamhet KB bedriver.

En integrerad del av forskningsprocessen är att utnyttja forskningsinfrastrukturer som har bredare formulerade forskningsändamål än de som förekommer i enskilda forskningsprojekt. För att KB ska kunna bidra till den svenska forskningens kvalitet är därför utveckling av forskningsinfrastrukturer på KB avgörande för att uppnå instruktionens skrivningar. För att ligga i framkant med den akademiska forskningsfronten sker detta i samverkan med forskare på lärosäten. Den infrastruktur för forskning som KB har byggt upp och nyligen etablerat med ett nationellt datalabb (KB-labb) utgör därmed ett sådant exempel som utredningen åsyftar. KB-labb utgör ett exempel på behovet av att koppla samman information från olika aktörer. Att bedriva analyser genom det som benämns text- och datautvinning (Text and Data Mining, TDM) i vart och en av relevanta databaser tjänar inte syftet. Ur ett vetenskapligt perspektiv är det många gånger en förutsättning att koppla samman information från olika aktörer.

Det är på denna punkt s.k. digital humaniora eller mer adekvat uttryckt beforskandet av stora datamängder tvärsektoriellt/tvärvetenskapligt som behovet av att sammankoppla stora datamängder gör sig gällande på ett annat sätt än för definierade datamängder såsom biobanksregister för att ta ett exempel. Vetskapen finns att data som existerar i olika

register/databaser är vetenskapligt relevanta, men det finns ingen explicit kunskap om vilka specifika data som är relevanta innan forskning genom TDM har utförts.

KB:s samlingar är mycket omfattande till sitt innehåll och finns dokumenterade på KB på dokumentbasis, inte på basis av person. Detta att jämföra med myndigheter som behandlar personuppgifter i vilket fall personen och dennes personuppgifter är kända. I KB:s databas med digitala samlingar av biblioteks dokument är personen i allmänhet inte känd eller registrerad.

KB instämmer i det som står i promemorian men vill påtala att utredningen till synes inte har ett fullt ut dynamiskt synsätt på forskningsdatabaser. Insamling kan ske från fler än de som utredningen pekar på – enskilda, myndigheter, företag och andra civila aktörer, samt en kombination av dessa. Inom datadriven forskning, och i synnerhet om denna är tvärvetenskaplig, aktualiseras insamling av data från långt fler aktörer. Syftet med inrättandet av ett nationellt datalabb på KB har varit att möta forskningens behov och vara en resurs för samtliga forskningsdiscipliner. KB har identifierat att det finns ett stort forskarbehov av en uppgiftssamling med breda ändamål som servar flera forskningsprojekt eller program. Tillika finns det behov av att forska med hjälp av insamlade uppgifter över tid, d.v.s. longitudinellt.

Eftersom det utretts särskilt om KB och dataskydd (Tilläggsdirektiv till Forskningsdatautredningen (U 2016:04) delredovisning enligt dir. 2017:29 och 2017:87 i en promemorian *Personuppgiftsbehandling i forskningsbibliotekens verksamhet*, Komm2017/01529/U 2016:04) får här påtalas att den utredningen tog sikte på KB:s behandling av personuppgifter i sin verksamhet, men inte för det fall KB:s samlingar av dokument behandlas jämte dokument från andra aktörer i en forskningsdatabas som KB har ansvaret för.

## **Promemorians sammanfattning sid 14: Utgångspunkter för en långsiktig reglering**

”Även i denna promemoria används begreppet forskningsdatabas för beskrivning av register eller annan samling av personuppgifter som syftar till att skapa underlag för framtida ännu inte definierade forskningsprojekt. Det är värdefullt för forskningen om sådana forskningsdatabaser kan inrättas och utvecklas.”

### **Skäl för KB:s ställningstagande**

KB anser att dess samlingar uppfyller det syfte som beskrivs ovan. Att en databas på KB (samling) ska kunna utvecklas framgick av KB:s tidigare instruktion som föreskrev att KB skulle ”förkovra” (jämför ”utvecklas” ovan i promemorian) sina samlingar. Någon ändring i sak har inte skett och ej heller avsetts. KB:s insamling av den svenska delen av internet, Kulturarw3, fyller samma funktion och pekas ut i förarbetena till Lag (2012:492) om pliktexemplar av elektroniskt material som ett komplement – aktiv insamling.

## **Promemorians sammanfattning sid 14 f: Tydligt rättsligt stöd för personuppgiftsbehandling**

”Dessa, liksom en del redan befintliga centrala uppgiftssamlingar behöver dock ett tydligt rättsligt stöd för sin personuppgiftsbehandling. Det föreslås därför en långsiktig generell reglering som syftar till att ge ett sådant rättsligt stöd samtidigt som den tillvaratar intresset av ett gott integritetsskydd i verksamhet med forskningsdatabaser.”

### **Skäl för KB:s ställningstagande**

Det ovan redovisade utdraget i promemorian instämmer KB med. KB vill ha ett tydligare rättsligt stöd och välkomnar ett förslag som omfattar KB:s verksamhet.

## **Promemorians sammanfattning sid 15: En långsiktig reglering bör omfatta forskningsdatabaser som baserar sig på de registrerades frivilliga medverkan**

”En långsiktig reglering bör...omfatta forskningsdatabaser som baserar sig på de registrerades frivilliga medverkan och som, i huvudsak, består av personuppgifter som samlas in från dem själva eller med deras aktiva medverkan.”

### **Skäl för KB:s ställningstagande**

Att promemorian tar som utgångspunkt forskningsdatabaser som baserar sig på de registrerades frivilliga medverkan och som, i huvudsak, består av personuppgifter som samlas in från dem själva eller med deras aktiva medverkan är ur ett dataskyddsperspektiv en rimlig lösning, men inte ur ett forskningsperspektiv. Att ta som utgångspunkt personuppgifter som samlas in genom samtycke snävar in omfånget av möjliga forskningsdata. Formuleringen utesluter KB som en nationell forskningsinfrastruktur och förhindrar KB att bidra till den svenska forskningens kvalitet.

## **Promemorians sammanfattning sid 15: Det är universitet och högskolor vid vilka forskning bedrivs som bedöms ha den kapacitet som krävs för att utveckla och förvalta forskningsdatabaser**

”Vidare bör en reglering vara inriktad på sådana forskningsdatabaser som är centrala i den meningen att de är av särskilt vetenskapligt värde för forskningen i ett långsiktigt perspektiv, och då inte bara för forskningsprojekt som forskningshuvudmannen ansvarar för utan även för forskningsprojekt som bedrivs av andra inom ramen för det eller de forskningsområden som varje forskningsdatabas skapar underlag för (se vidare nedan). Det är universitet och högskolor vid vilka forskning bedrivs som bedöms ha den kapacitet som krävs för att utveckla och förvalta forskningsdatabaser av det beskrivna slaget och som därför bör ges möjlighet att göra detta genom en ny reglering.”

## **Skäl för KB:s ställningstagande**

Härvid gör utredningen ännu en avgränsning med avseende på vilken aktör som ska få en rättslig grund för forskningsdatabaser av detta slag. Utredningen anför att det är universitet och högskolor vid vilka forskning bedrivs som bedöms ha den kapacitet som krävs för att utveckla och förvalta forskningsdatabaser av det beskrivna slaget och som därför bör ges möjlighet att göra detta genom en ny reglering. Genom EU-direktivet om upphovsrätt på den digitala inre marknaden (DSM-direktivet) som genomfördes i lag (1960:729) om upphovsrätt till litterära och konstnärliga verk den 1 januari 2023 ges i lagens 15 b § utöver forskningsinstitutioner (lärosäten) även kulturarvsinstitutioner såsom KB, rätt att enligt en inskränkning bedriva forskning med stöd av AI (TDM). Enligt KB är promemorians slutsatser i otakt med den utveckling som sker inom forskningen och KB:s uppdrag, se ovan, dvs att vid användning av AI (TDM) har KB en tydlig rättslig grund för vetenskaplig forskning men inte motsvarande för behandling av personuppgifter.

## **Promemorians sammanfattning sid 16: Primära ändamålen skapa underlag för olika framtida forskningsprojekt inom för forskningsdatabasen angivna forskningsområden samt att lämna ut uppgifter till sådana forskningsprojekt**

Att KB som nationell forskningsinfrastruktur med KB-labb och dess forskningsdatabaser är av betydelse framgår av vad som står i PM, se nedan.

”I lagen regleras uttömmande för vilka ändamål personuppgifter får behandlas i en forskningsdatabas. De primära ändamålen är att just skapa underlag för olika framtida forskningsprojekt inom för forskningsdatabasen angivna forskningsområden samt att lämna ut uppgifter till sådana forskningsprojekt. Sekundära ändamål för behandling av redan insamlade ändamål är utlämnande till dels forskningshuvudmän i samband med utredning av oredlighet i forskning eller annan avvikelser från god forskningssed, dels fullgörande av uppgiftsskyldighet enligt lag eller förordning.”

## **Skäl för KB:s ställningstagande**

Vad som inte uttalas i promemorian är behovet av referensdata och verifiering av forskning och att detta kommer förutsätta att KB-labb och KB:s forskningsdatabaser kan användas till avsett syfte. Att förutsättningarna är uppfyllda i nämnda avseende är helt avgörande för denna viktiga roll som KB har som nationell forskningsinfrastruktur.

I detta sammanhang aktualiseras transparens i övrigt. Hur ska exempelvis journalister kunna granska forskningen? Och vad gäller andra forskare – fångas de upp genom 1 §? Om inte uppkommer inlåsnings effekter, vilka är bekymmersamma. KB:s uppgift är inte enbart att möjliggöra forskning utan även att möjliggöra en kritisk granskning av forskningen. I denna del är utredningen tyst och det ser KB som en brist. KB har i uppdrag att ”bidra till den svenska forskningens kvalitet och det demokratiska samhällets utveckling”. Ur ett demokratiperspektiv är det av väsentlig betydelse att forskning kan granskas.

## Övrigt

I sitt yttrande över Europeiska kommissionens förslag till förordning om harmoniserade regler för artificiell intelligens, Proposal for a Regulation laying down harmonised rules on artificial intelligence (artificial intelligence act) and amending certain union legislative acts (COM (2021) 206), I2021/01304 (KB 2021-526) har myndigheten förordat att ärendena om behandling av personuppgifter vid AI och TDM och behandling av personuppgifter i forskningsdatabaser bereds färdigt i närtid samt att en översyn av hur Infosoc-direktivet har genomförts i svensk rätt görs snarast möjligt. KB konstaterar med tillfredsställelse att en översyn av hur Infosoc-direktivet har genomförts i svensk rätt har inletts (Utredningen om en översyn av vissa frågor om upphovsrätt (Ju 2022:16)).

## Skäl för KB:s ställningstagande

I regeringskansliet bereds fortfarande ärendena om dels KB:s behandling av personuppgifter vid AI/TDM/ML, dels behandling av personuppgifter i forskningsdatabaser - en långsiktig reglering av forskningsdatabaser. Vidare bereds i regeringskansliet en översyn av inskränkningarna i upphovsrättslagen, varvid frågan nu bereds av Utredningen om en översyn av vissa frågor om upphovsrätt (Ju 2022:16). Samtliga dessa ärenden har en avgörande betydelse för KB i sin tillämpning av AI i forskningens infrastruktur.

Genom storskalig digitalisering på KB och leveranser av digitala dokument enligt lag om elektronisk pliktleverans uppstår de facto en eller flera databaser på KB som ska kunna användas inom forskningsinfrastrukturer med stöd av AI/TDM. Att tydligare rättslig grund för behandling av personuppgifter för AI/TDM och behandling av personuppgifter i forskningsdatabaser fortfarande bereds i regeringskansliet har en hämmande effekt på användning av AI-system och TDM i forskningen och begränsar i avsevärd mån KB:s möjlighet att utföra sitt uppdrag.

Beslut i detta ärende har fattats av riksbibliotekarien Karin Grönvall efter föredragning av verksjuristen Jerker Rydén och forskningsstrateg Christopher Natzén. I den slutliga handläggningen har även enhetschefen Love Börjeson deltagit.

Beslutet har signerats elektroniskt.

Karin Grönvall  
Riksbibliotekarie

Jerker Rydén  
Verksjurist

Christopher Natzén

Forskningsstrateg