National Library of Sweden

Handläggare Jerker Rydén
Datum 2019-02-18
Dnr KB 2018-899

Utbildningsdepartementet
103 33 Stockholm

Yttrande över Rätt att forska – Långsiktig reglering av forskningsdatabaser (SOU 2018:36) dnr U2018/02557/F

1. Rätt att forska – Långsiktig reglering av forskningsdatabaser (SOU 2018:36)

Utredningen har antagit namnet Forskningsdatautredningen. Utredningen redovisade delar av sitt uppdrag i delbetänkandet Personuppgiftsbehandling för forskningsändamål (SOU 2017:50) till regeringen i maj 2017. KB har lämnat remissvar.

De uppdrag som lämnades i tilläggsdirektiv 2017:29 och 2017:87 redovisades separat den 9 november 2017 i promemorian Personuppgiftsbehandling i forskningsbibliotekens verksamhet (Komm2017/01529/U 2016:04, bilaga 5). KB har lämnat remissvar.

Utredningen har slutligen överlämnat slutbetänkandet Rätt att forska – Långsiktig reglering av forskningsdatabaser (SOU 2018:36). Utredningens uppdrag är därmed slutfört. Kungl. biblioteket (KB) har som remiss erhållit rubricerad SOU. KB:s svar anges nedan under 3.

2. Vad är KB/Uppdrag

Enligt KB:s instruktion (Förordning (2008:1421) med instruktion för Kungl. biblioteket, är (1§) KB Sveriges nationalbibliotek och ska främja den svenska forskningens kvalitet och en demokratisk samhällsutveckling genom att tillhandahålla källmaterial och en effektiv forskningsinfrastruktur. Härvid ska (2 §) myndigheten samla in, beskriva, bevara och tillhandahålla fysiskt och digitalt material som behövs för att fullgöra myndighetens uppgifter.

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

National Library of Sweden

in enlighet med Förordning (2002:287) om behandling av personuppgifter i Kungl. bibliotekets digitala kulturarvsprojekt (Kultarw3) samt donationer. Instruktionen gör heller inte någon åtskillnad mellan fysiska dokumenten av olika skäl har digitaliserats och därefter tillhandahålls i elektroniskt format eller dokument som till sitt ursprung är elektroniska.

Av instruktionen framgår vidare att KB (3 §) ska ha en nationell överblick över och främja samverkan inom det allmänna biblioteksväsendet i enlighet med 18 § bibliotekslagen (2013:801) samt att myndigheten ska tillhandahålla informationssystem och verka för utveckling och samordning av digitala tjänster inom det allmänna biblioteksväsendet.

KB ska enligt sin instruktion "tillhandahålla källmaterial och en effektiv infrastruktur" (§1). I detta åligger det KB att "samla in, beskriva, bevara och tillhandahålla fysiskt och digitalt material" (§2) för samhällets möjlighet att skapa ny kunskap. Det har däremot blivit alltmer tydligt att det inte går att isolera KB:s uppdrag från skapandet av ny kunskap och forskningsdata, KB:s och forskningsinstitutionernas respektive uppdrag överlappar här varandra.

Bland de forskningsförfrågningar om användning av KB:s digitala resurser som inkommit under 2018 kan t.ex. nämnas en begäran om att löpande få tillgång till textdata som produceras vid digitaliseringen av den svenska dagspressen. Syftet är att 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 kommer möjliggöra 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 eftersom KB är den enda institution som förvaltar de efterfrågade samlingarna - och där bearbetning och analys oundvikligen kommer involvera

informationsinnehåll som omfattas av både dataskydds lagstiftning. Men i det större perspektivet utgör KBs data en delmängd och KB ser framför sig att för olika forskningsprojekt kommer KB utöver att skapa förutsättningarna genom en infrastruktur tillika samla in data till en forskningsdatabas på KB.

Regeringen måste därför ta i beaktande att KB:s nu gällande instruktion kommer att behöva ändras för att KB ska kunna skapa forskningsdatabaser. Det blir avgörande att KB också ska kunna initiera forskningsprojekt för att fortsatt goda förutsättningar ska finnas för att tillhandahålla en effektiv infrastruktur för forskarsamhället.

3. Yttrande Rätt att forska – Långsiktig reglering av forskningsdatabaser (SOU 2018:36)

I det följande besvarar KB remissen i tillämpliga delar och utgår i sitt svar från de rubriker som anges i rubricerad promemoria.

1.1 Förslag till lag om forskningsdatabaser

3 § Personuppgifter som behandlas för de ändamål som anges i 2 kap. 2 § får också lämnas ut till en utredning av oredlighet i forskning eller för att ett yttrande ska kunna lämnas i samband med en sådan utredning.

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 inlåsningseffekter, vilka är synnerligen bekymmersamma. KBs uppgift är inte enbart att möjliggöra forskningen utan även att möjliggöra en kritisk granskning av forskningen. I denna del är utredningen tyst och det ser KB som en betydande svaghet i utredningen. Ur ett demokrati perspektiv är det av väsentlig betydelse att i vart fall journalister kan granska forskningen.

2. Utredningens uppdrag/arbete

Det är enligt utredningens bedömning en integrerad del av forskningsprocessen att skapa forskningsinfrastrukturer som har bredare formulerade forskningsändamål än de som förekommer i enskilda forskningsprojekt. KB instämmer i detta. Det är enligt utredningens bedömning rimligt att uppbyggnaden av en forskningsdatabas kan

bedömas som personuppgiftsbehandling för forskningsändamål och därmed utgöra en del av själva forskningen.

KB ser sin verksamhet som en del av forskningsprocessen och att den infrastruktur för forskning som KB har byggt upp och nyligen etablerat med ett nationellt datalabb utgör ett sådant exempel som utredningen åsyftar.

Utredningen pekar på sidan 75 av betydelse för registerbaserad forskning är att förordningen i skäl 157 lyfter fram att forskare kan erhålla ny kunskap av stort värde genom att koppla samman information från olika register. För att underlätta vetenskaplig forskning får personuppgifter behandlas för vetenskapliga forskningsändamål, med förbehåll för lämpliga villkor och skyddsåtgärder i unionsrätten eller i medlemsstaternas nationella rätt.

Det nationella datalabb som nyligen etablerats på KB utgör ett sådant exempel på behovet av att koppla samman information från olika aktörer. Att bedriva TDM i vart och en av dessa databaser tjänar inte syftet. Ur ett vetenskapligt perspektiv är det många gånger en förutsättning att koppla samman information från olika aktörer. Om det inte är möjligt blir resultatet inte vetenskapligt eftersom relevanta källor saknas.

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 forskningen genom TDM har utförts.

Källorna 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 är känd och dennes personuppgifter under det att i KBs databas med digitala samlingar av biblioteks dokument personen inte är i allmänhet känd eller registrerad.

En ytterligare skillnad är att KBs samlingar av dokument jämte metadata (data om data i dokument i KBs samlingar) inte är allmänna handlingar och ej heller de data som raffineras ur KBs samlingar –

dessa är att likställa med doktorsavhandlingar som tillförs KB. samlingar, dvs resultatet av forskningen på källmaterial i KB:s samlingar.

Eftersom det utretts särskilt om KB och dataskydd får här påtalas att den utredningen tog sikte på KB:s behandling av personuppgifter i sin verksamhet när personuppgifter uteslutande i KB:s samlingar behandlas men inte för det fall KB:s samlingar av dokument behandlas jämte dokument från andra aktörer i en forskningsdatabas som KB har som huvudman.

KB har hemställt att få sin instruktion ändrad till att KB får bedriva forskning. Detta för att möjliggöra sin nya verksamhet under uppbyggnad med ett nationellt datalabb.

Utredningen definierar så forskningsdatabas som en databas (uppgiftssamling) med personuppgifter som samlas in från enskilda individer eller från andra källor och som är en nationell resurs för att tillhandahålla uppgifter till flera olika forskningsprojekt (avsnitt 8.7.1). Utanför lagens definition faller således forskningsregister som har skapats för ett visst projekt, en del av ett projekt eller på ett på liknande sätt bestämd forskning.

Syftet att på KB samla in källor från olika aktörer är att skapa förutsättningarna för relevant forskning – inkludera adekvat mängd data ur ett vetenskapligt perspektiv.

Utredningen skiljer på de uppgiftssamlingar som dess förslag syftar till att möjliggöra (forskningsdatabaser) och de tjänster som tillgängliggör uppgifterna från dessa uppgiftssamlingar. De senare kallar utredningen med ett gemensamt begrepp databasinfrastruktur för forskning. Sådan databasinfrastruktur för forskning är en form av forskningsinfrastruktur, men eftersom det senare begreppet även omfattar basutrustning, forskningsmaterial, IKT-infrastrukturer m.m. är det för brett för att användas i denna framställning.

KB:s datalabb kommer som anförts förutsätta en databas (uppgiftssamling) med personuppgifter som samlas in från enskilda individer eller från andra källor och som är en nationell resurs för att tillhandahålla uppgifter till flera olika forskningsprojekt. Syftet är att möjliggöra fjärråtkomst till datalabbet.

National Library of Sweden

7.4.2 Forskningsdatabaser som vi avser att reglera

Utredningen finner också att forskningens frihet innebär att initiativet till att skapa en forskningsdatabas bör komma från forskningshuvudmännen och inte från regeringen. Vi konstaterar att det är i forskningsverksamheten som behoven av uppgiftssamlingar bäst kan identifieras och formuleras.

KB instämmer i detta men vill påtala att utredningen har till synes inte 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. Det är bland i detta syfte som KB har beslutat om att inrätta ett nationellt datalabb som är en resurs för samtliga forskningsdiscipliner.

KB har alltså identifierat behov 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, dvs. longitudinellt.

Vidare bör noteras att 12 kap. 11 § 3 TF anges följande. Som allmän handling anses ej tryckt skrift, ljud- eller bildupptagning eller annan handling som ingår i bibliotek eller som från enskild har tillförts allmänt arkiv uteslutande för förvaring och vård eller forsknings- och studieändamål eller privata brev, skrifter eller upptagningar som eljest ha överlämnats till myndighet uteslutande för ändamål som nu angivits. Upptagning av innehållet i sådan handling ska enligt 2 kap. 1 § 4 TF inte heller anses som allmän handling, om upptagningen förvaras hos myndighet där den ursprungliga handlingen ej skulle vara att anse som allmän. Den s.k. biblioteksregeln i 2 kap. 11 § 3 TF gäller såväl biblioteksmyndigheter som andra myndigheter som har bibliotek för sin egentliga verksamhet. För att en handling ska omfattas av bestämmelsen krävs att den ingår som en beståndsdel i en samling av handlingar organiserad på sådant sätt att den enligt vanligt språkbruk är att betrakta som bibliotek (prop. 1975/76:160 s. 179 f).

8. Förslag till reglering av forskningsdatabaser

8.8.4 Grundläggande villkor för skapandet av en forskningsdatabas

Beslut om skapandet av forskningsdatabasen och om vilka uppgifter som ska behandlas och för vilka särskilda ändamål

Utredningens förslag är att en forskningshuvudman ska i ett beslut om att skapa en forskningsdatabas fastställa vilka uppgifter som ska behandlas i forskningsdatabasen och för vilka särskilda ändamål dessa uppgifter är nödvändiga att behandla.

Mot bakgrund av att KB kommer samla in data - med personuppgifter - av annan karaktär än många andra forskningsutövare dvs vilket slag av personuppgifter som aktualiseras kan inte med säkerhet fastställas endast att de i mer eller mindre omfattning kan förekomma i insamlad data och i förekommande fall vilket slags personuppgifter.

Enligt utredningen innebär detta att innan forskningshuvudmannen börjar samla in uppgifter till en forskningsdatabas eller omvandlar en existerande uppgiftssamling till en forskningsdatabas, måste huvudmannen således kunna precisera ändamålet och vilka uppgifter som ska samlas in något mer än vad som framgår av det övergripande ändamålet för forskningsdatabaser; att skapa underlag för olika forskningsprojekt. Det kan t.ex. handla om att precisera vilka typer av forskningsprojekt som forskningsdatabasen ska kunna utgöra underlag till, vilka forskningsfrågor som ska besvaras, inom vilka forskningsområden projekten ska genomföras och vilken typ av uppgifter som ska samlas in till en viss närmare angiven typ av forskning.

Även om ändamålet för en forskningsdatabas måste vara preciserat och alltså inte utgöras av ett generellt ändamål såsom "stöd till framtida forskning" bör det enligt utredningen ändå kunna vara relativt brett. Det bör kunna utformas på motsvarande sätt som i LifeGene-lagen: "...att skapa underlag för olika forskningsprojekt om vad arv och miljö betyder för uppkomsten och utvecklingen av olika typer av sjukdomar och för människors hälsa i övrigt.". Den demografiska databasen (DDB) vid Umeå universitet har en bred uppgift. Enligt dess fortfarande gällande förordning är den "att registrera och bearbeta huvudsakligen

demografiska och historiska data för forsknings-, utbildnings- och arkivändamål samt att göra dessa data tillgängliga för forskare." DDB ska dessutom enligt förordningen verka för vetenskapligt samarbete och metod- utveckling.

För KBs del skulle det visserligen vara möjligt att ange en bred uppgift för en forskningsdatabas men att det samtidigt skulle undergräva syftet att på förhand utesluta data eftersom data de facto kan utgöra en alltjämt okänd delmängd av den totala datamängden.

Syftet med att tillhandahålla data är för forskningen är att raffinera nämna data. Detta är en etablerad vetenskaplig metodik som möjliggör ny banbrytande forskning eftersom människan inte har möjlighet att bearbeta och systematisera data i så stora mängder som det är fråga om.

Det slag av forskningsdatabas som KB kommer att etablera kommer själva forskningen att utgöra ett delmoment i arbetet med att etablera forskningsdatabasen. Dvs ett beslut av en etisk forskningsnämnd bör godta att sk TDM/AI används för att bestämma vad som ska utgöra databasen baserat på de kriterier som används för att utforma de algoritmer som används inom TDM/AI. Eftersom tekniken ständigt utvecklas kommer det finnas behov att utveckla detta slag av forskningsdatabaser.

Det kommer därför vara svårt för att inte säga omöjligt att på förhand ange vilken typ av uppgifter som ska samlas in till en viss närmare angiven typ av forskning. Därför kommer KB vara nödgad att ange att i princip alla möjliga uppgifter kommer samlas in.

8.8.5 En forskningsdatabas ska etikprövas

KB ställer sig positiv till det synsätt som präglar utredningens förslag.

8.8.6 Etikprövningen av en forskningsdatabas

Utredningen tar i sina resonemang sin utgångspunkt i etablerad etikprövning och hänvisar till att de deltagande försökspersonerna behöver skyddas från risker att skadas fysiskt, psykiskt eller integritetsmässigt i samband med forskning. Härvid hänvisar utredningen till 7–11§§ nu gällande etikprövningslagen som anger de allmänna utgångspunkter som gäller vid etikprövningen. Bestämmelserna har sin bakgrund i artikel 16 i Europarådets konvention om mänskliga rättigheter och biomedicin. Syftet med

bestämmelserna är att skydda både människovärdet i stort och den enskilde forskningspersonens välbefinnande.

Det slags forskning som en forskningsdatabas på KB skulle tjäna kan inte avgränsas så enkelt som forskning inom biomedicin. Urvalet av personer som berörs är till övervägande delen inte kända och en anonymisering är inte alltid till gagn för forskningen. Att avgränsa urvalet av data kan i många fall motverka syftet med en forskningsdatabas. Jämför med ett forskningsbibliotek skulle utesluta forskares från tillgång till samtligt källmaterial i bibliotekets samlingar.

KB är införstådd med innebörden av förordningen, skyddsbehovet och den föreslagna lagstiftningen men menar att om datadriven forskning ska kunna bedrivas med utgångspunkt från det slags av data som KB förvaltar genom sitt uppdrag och det slags data som är önskvärt att samla i en forskningsdatabas tillsammans med KBs egna samlingar. Rimligen måste etikprövningen ta sin utgångspunkt i hur de algoritmer som är tänkta att användas för att utvinna relevant data och analysera samband mellan data utgöra utgångspunkten för prövningen. I annat fall kommer forskningen inte kunna bedrivas och etikprövningen kommer ge upphov till betydande inlåsningseffekter.

8.10.7 Extern uppföljning och utvärdering

KB välkomnar utredningens förslag att Vetenskapsrådet ska regelbundet följa upp och utvärdera verksamheten i en forskningsdatabas. Vetenskapsrådet ska enligt utredningens förslag i sin utvärdering granska forskningsdatabasen avseende säkerhet och ändamålsenlighet.

Det slags forskningsdatabas som KB ser framför sig i sitt nationella datalabb är en helt ny verksamhet. Att Vetenskapsrådet föreslås få som uppgift att regelbundet följa upp och utvärdera verksamheten i en forskningsdatabas välkomnar KB. Att utredningen föreslår att VR ska initiera och stödja forskningshuvudmännen i utarbetandet av gemensamma uppförandekoder enligt art.40 Europaparlamentets och rådets förordning (EU) 2016/679 av den 27 april 2016 om skydd för fysiska personer med avseende på behandling av personuppgifter och om det fria flödet av sådana uppgifter (allmän data- skyddsförordning) välkomnas likaledes.

Uppbyggnad av det slags infrastruktur och forskningsdatabaser som KB har inlett är i sin utvecklingsfas och det är därför svårt att med säkerhet bedöma vad för slags reglering som är ändamålsenlig för forskningen givet dataskyddslagen och förordningen. Det skulle därför vara till stor fördel om det den föreslagna lagstiftningen gav VR utrymme att ge ett adekvat stöd givet att kunskap och erfarenhet av det slags av forsningsdatabas och nationellt datalabb inte existerar. Det inryms så pass stor dynamik i detta slag av forskning givet tekniken ständigt utvecklas och frågor om vetenskaplig metodik att det vore till gagn för forskningen och samhället i stort om VR kunde ges lämpligt utrymme att stödja exempelvis KB som en forskningshuvudman med utarbetande av uppförandekoder.

8.12 Avvecklingen av en forskningsdatabas

8.12.1 Behov av reglering i detta steg av livscykeln

En forskningsdatabas kan av olika anledningar behöva avvecklas. Behovet kan vara täckt av andra databaser, forskningsproblemet kan ha upphört att vara intressant, eller ny teknik kan ha utvecklats så att uppgifterna kan nås på annat sätt. Till detta steg i en forskningsdatabas livscykel, avvecklingen, hör frågan om uppgifterna ska gallras eller bevaras genom arkivering.
I detta sammanhang får erinras om att KB:s samlingar inte är allmänna handlingar och de dokument som tillförs en forskningsdatabas som KB har som huvudman och vilken databas utgör en del av biblioteket enligt biblioteksregeln i 12 kap. 11 § 3 TF inte utgör en allmän handling.

Antingen gallras data eller återförs till den från vilken data har insamlats alternativt inkorporeras data tillsvidare i KBs samling. Oavsett vilket tillämpas inte arkivlag eller OSL.

Sett ur utredningens syn på en forskningsdatabas cykler kan KB redan nu skönja en framtid där en forskningsdatabas antingen avslutas och resultatet endast inordnas i KBs samling eller att den till del eller sin helhet införs i en ny forskningsdatabas. Processen skulle kunna beskrivas som en till sin helhet eller del iterativ.

National Library
of Sweden

Beslut i detta ärende har fattats av biträdande riksbibliotekarien Lars Ilshammar efter föredragning av verksjuristen Jerker Rydén.

Lars Ilshammar
Biträdande riksbibliotekarie

Jerker Rydén
Verksjurist

informationsinnehåll som omfattas av både dataskydds lagstiftning. Men i det större perspektivet utgör KBs data en delmängd och KB ser framför sig att för olika forskningsprojekt kommer KB utöver att skapa förutsättningarna genom en infrastruktur tillika samla in data till en forskningsdatabas på KB.

Regeringen måste därför ta i beaktande att KB:s nu gällande instruktion kommer att behöva ändras för att KB ska kunna skapa forskningsdatabaser. Det blir avgörande att KB också ska kunna initiera forskningsprojekt för att fortsatt goda förutsättningar ska finnas för att tillhandahålla en effektiv infrastruktur för forskarsamhället.

3. Yttrande Rätt att forska – Långsiktig reglering av forskningsdatabaser (SOU 2018:36)

I det följande besvarar KB remissen i tillämpliga delar och utgår i sitt svar från de rubriker som anges i rubricerad promemoria.

1.1 Förslag till lag om forskningsdatabaser

3 § Personuppgifter som behandlas för de ändamål som anges i 2 kap. 2 § får också lämnas ut till en utredning av oredlighet i forskning eller för att ett yttrande ska kunna lämnas i samband med en sådan utredning.

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 inlåsningseffekter, vilka är synnerligen bekymmersamma. KBs uppgift är inte enbart att möjliggöra forskningen utan även att möjliggöra en kritisk granskning av forskningen. I denna del är utredningen tyst och det ser KB som en betydande svaghet i utredningen. Ur ett demokrati perspektiv är det av väsentlig betydelse att i vart fall journalister kan granska forskningen.

2. Utredningens uppdrag/arbete

Det är enligt utredningens bedömning en integrerad del av forskningsprocessen att skapa forskningsinfrastrukturer som har bredare formulerade forskningsändamål än de som förekommer i enskilda forskningsprojekt. KB instämmer i detta. Det är enligt utredningens bedömning rimligt att uppbyggnaden av en forskningsdatabas kan