



Utbildningsdepartementet
u.remissvar@regeringskansliet.se
Er ref: U2022/04089

Yttrande rörande ”Långsiktig reglering av forskningsdatabaser” (U2022/04089)

Institutet för framtidsstudier (IFFS) är en tvärvetenskaplig forskningsinstitution med stor erfarenhet av databaserad forskning om frågor som rör samhällets utmaningar. Institutet har synpunkter på förslagen som ges angående långsiktig reglering av forskningsdatabaser.

Institutet för framtidsstudier *ställer sig bakom* ett flertal av förslagen i promemorian:

- Forskningsdatabaser behöver inte etikprövas i sig, utan forskningsprojekt som använder dem är de som ska etikprövas
- Forskningsdatabaser bör hållas av lärosäten
- Forskningsdatabaser ska kunna kompletteras med data från myndighetsregister
- Forskningsdatabaser ska som regel inte gallras

Institutet för framtidsstudier har följande sex *invändningar* över förslag i promemorian. Invändningarna utvecklas nedan.

1. Forskningsdatabaser som består av bearbetade uppgifter från myndighetsregister bör också definieras som forskningsdatabaser och regleras i lag
2. Regeringen bör inte ges avgörande inflytande över beslut om forskningsdatabaser
3. En systematisk och regelbunden kontroll över inloggningar till en forskningsdatabas som är pseudonymiserad är både orimlig, eftersom inget sådant system finns i bruk, och poänglös, eftersom sådan kontroll knappast kan – och enligt utredningen inte heller förväntas göras – av forskningsprojekt som utnyttjar data
4. Vi motsätter oss starkt förslaget att forskare inte ska kunna inhämta uppgifter om en respondents biologiska släktingar utan deras ”informerade och uttryckliga samtycke”



5. Uppgiftslämnare bör ha rätt att få sina identifierade uppgifter i kodnyckeln raderade, men inte övriga uppgifter, som utredningen föreslår
6. Tillfällig tillgång till uppgifter om individer i urvalsramen som ej givit samtycke till uppgiftslämnande bör utsträckas till att innefatta rimlig tid för konstruktion av vikter och bortfallsanalys

1 Begränsningen till databaser baserade på samtycke

Promemorian sägs handla om ”forskningsdatabaser” men berör huvudsakligen det som inom forskningen kallar ”surveydatabaser”, som dominerar databaser baserade på samtycke. Det promemorian kallar ”forskningsdatabaser” exkluderar således det som inom forskningen oftast kallas ”registerdatabaser”, alltså bearbetade sammansättningar av myndighetsregister utan samtycke från de ingående personerna. Forskares tillgång till registerdatabaser är något som gör svensk forskning internationellt konkurrenskraftig, och att få processen runt dem lättare att hantera kan göra forskningen betydligt billigare och mer effektiv. Att slutligen få en lagstiftning som reglerar registerdatabaser måste därför ha hög prioritet och det hade varit önskvärt att detta skulle ingått i lagförslaget.

2 Regeringens inflytande

Regeringen bör hållas på armlängds avstånd från att besluta om praktiska detaljer i frågor om forskningsdatabaser. Regeringen ska enligt lagförslaget utse, skriva förordning för och ge föreskrifter för enskilda forskningsdatabaser, bestämma om arkivering, etc (sid 223), vilket bland annat innebär att bestämma vilka forskningsområden som en enskild databas ska skapa underlag till och lämna ut uppgifter till (sid 158). Dessa åtgärder kräver stor kännedom om forskning och forskningsverksamhet och bör vara forskarstyrda, eller skötas av en myndighet som har expertkunskap om forskning. Viktigt är att beslutet bör fattas av en myndighet för att beslutet ska kunna överklagas. Pondera att det ryktas om att genetiska data i någon forskningsdatabas har läckt och det plötsligt reses stora krav från sociala och andra media på regeringen att ingripa. I en sådan situation bör det finnas ett skalskydd för forskningsdatabaser, med krav på en objektiv utredning och överklagandemöjligheter för eventuella sanktioner. Möjliga alternativa instanser till regeringen är Vetenskapsrådet och Etikprövningsmyndigheten (EPM).

3 Användandekontroll

Den mer praktiska frågan handlar om användande, dvs forskningsbarheten. Detta diskuteras nästan inte alls i promemorian, men ett av förslagen gör att frågan aktualiseras. Det gäller kravet på att logga varje användande, för att kontrollera att ingen obehörig använder data (sid. 191–192), något som ska ske inte bara efter



misstanke utan systematiskt som en permanent övervakning. Man kunde tänka sig att de stora universiteten har kapacitet att implementera ett sådant system längre fram i tiden, men knappast alla lärosäten. Systemet är också opraktiskt, eftersom alla användare blir beroende av en fungerande teknisk infrastruktur som kräver ständigt underhåll, antagligen inte alltid är åtkomlig för distansarbete, och kanske inte ens på laglig väg kan användas utomlands. Dessutom är det sannolikt inte nödvändigt för databaser som är pseudonymiserade (vilket torde gälla nästan alla forskningsdatabaser).

Dessutom verkar hela frågan om loggning vara poänglös, eftersom inga sådana krav ställs – eller kan ställas – på externa användare (forskningsprojekt), vilka inkluderar icke-offentliga och utländska användare. Den enda ledtråd som ges i lagförslaget är att den som ansöker om att få ta del av data ”måste kunna visa vid sin ansökan till forskningsdatabasen att utlämnade personuppgifter får ett fortsatt fullgott integritetsskydd” (sid 197) – det är dock en helt annan sak än de starka kraven på systematisk kontroll av obehöriga användare genom loggning, som man menar bör läggas på personuppgiftsansvarig för forskningsdatabasen.

4 Biologisk släkting

Lagförslaget (sid 186) innebär att forskare inte ska kunna inhämta uppgifter om en respondents biologiska släktingar ”utan dennes informerade och uttryckliga samtycke”, som en tidigare regering också uttalat i fallet med LifeGene (Prop 2012/13: 163, sid 42). Detta skulle i stort sett omöjliggöra betydande forskningsområden som rör intergenerationella processer och relationer mellan individer inom familjer. I datainsamlingar där ett syfte är att studera social rörlighet, vanliga inom sociologi och nationalekonomi som indikatorer på ojämlikhet, frågas t.ex. om mammans och pappans utbildning och yrke under respondentens uppväxt. Sådan information används också för att studera bakgrunden till människors politiska preferenser och röstbeteende i statsvetenskapliga studier, liksom uppgifter om föräldrar är centrala inom psykologi. I demografiska studier finns det förstås starka skäl att samla in uppgifter om föräldrar, barn och ibland syskon. Ett förbud mot inhämtande av sådana uppgifter skulle innebära en utomordentligt allvarlig begränsning av angelägen forskning.

Här borde man beakta vilken typ av uppgifter om biologiska släktingar det handlar om. Om de är av den karaktären att avsikten inte är att särskilt studera tredje person, utan att uppgifterna används som en indikator på respondentens förhållanden (t.ex. uppväxtmiljö), bör information kunna tas in via en respondent som givit samtycke. Alternativet, att söka informerat samtycke från ofta vuxna människors föräldrar och fråga dem om utbildning och yrke under respondentens uppväxt, torde vara praktiskt ogenomförbart i de flesta fall, och går emot principen om uppgiftsminimering. Grundprincipen bör alltså vara att forskare kan ta in



information om biologiska släktingar. Frågan om det kan finnas undantag för viss typ av uppgifter om biologiska släktingar som inte bör tillåtas (t.ex. känsliga uppgifter som hälsa, brott, eller politiska preferenser) bör utredas.

5 Användandet av uppgifter från myndighetsregister under kortare tid

Lagförslaget ger forskare möjlighet att använda registerdata även för personer som inte ger samtycke för användande, vid urvalsdragning, vilket är insiktsfullt. Restriktionen är att ”behandlingen inte får pågå under längre tid än vad som är nödvändigt” (sid 167). Men nödvändigt för vad? Det som nämns är ”tiden innan varken information eller samtycke kan lämnas”, liksom ’urval’. Det bör tydligt framgå att användandet av sådana uppgifter inte enbart gäller urval. Det är helt centralt för upprättandet av forskningsdatabaser att dessa uppgifter också kan användas för konstruktion av vikter (poststratifiering) liksom för bortfallsanalys. Riskerna för integritetsintrång med sådant användande är i stort sett obefintliga, men vinsterna för forskningsdatabasers kvalitet påtagliga.

6 Raderande av uppgifter enligt uppgiftslämnarens önskemål

Uppgiftslämnarens rätt att få sina uppgifter raderade tycks bryta mot lagen (2019:504) om ansvar för god forskningssed och prövning av oredlighet i forskning. Om man tillåter raderande av data kan varken replikationer eller prövning av oredlighet ske med någon precision (det går att föreställa sig situationer där många uppgiftslämnare ser en orsak att lämna en databas). Det finns flera andra skäl att vara tveksam till idén att raderandet av hela poster är rätta vägen att gå för att tillfredsställa uppgiftslämnarens integritetsskydd. Ett nämns av utredningen, nämligen att hela processen att aktivera kodnyckel och söka efter personuppgifter innebär en viss risk i sig. Många tidskrifter kräver nuförtiden också att analyser i publikationer ska kunna replikeras. För att garantera uppgiftslämnarens trygghet bör det vara tillfredsställande att anonymisera posten, dvs. ta bort kopplingen i kodnyckeln, samt i förekommande fall också radera eller t.ex. trunkera variabler som annars i kombination kunde vara identifierande, men låta uppgifterna finnas kvar i databasen. Detta har varit rutin i ett flertal dataset som lämnats till SND och bör relativt enkelt kunna genomföras av forskningsdatabasens forskare och huvudmän.

Detta remissvar har beslutats av professor Gustaf Arrhenius, vd för Institutet för framtidsstudier. Underlaget har tagits fram av professor Jan Jonsson. Även professor Carina Mood, professor Jerzy Sarnecki, och docent Anna Tyllström har lämnat synpunkter.

Gustaf Arrhenius
VD, Institutet för Framtidsstudier