SDU RIO Legal Services
19. September 2019

Notat vedr. behandling af personoplysninger og anonymisering til forskning

1. Hvad er personoplysninger?

Personoplysninger defineres som oplysninger, der direkte eller indirekte kan identificere en fysisk person. Det kan være et navn, registreringsnummer på en bil eller oplysninger om et sygdomsforløb. Det kan også være erfaringer, oplevelser eller andre ting, som kan henføres til en bestemt person. Det betyder, at flere oplysninger, som i sammenhæng kan knyttes til en person, vil også være personoplysninger, selvom de enkelte oplysninger i sig selv, ikke ville være personoplysninger.
Man skelner således mellem direkte identifikation og indirekte identifikation. En direkte identifikation er f.eks. et navn eller personnummer. Indirekte identifikation er alder, køn, uddannelse og andre oplysninger, som ikke direkte giver mulighed for at identificere en person – men som kan anvendes, i sammenhæng med andre oplysninger, til identifikation.

Der er ingen grænser for, hvilke oplysninger, der kan være personhenførbare og det vil derfor være en konkret vurdering i hvert tilfælde. Skal du f.eks. lave en spørgeskemaundersøgelse, kan hvert svar godt være anonymt, men koblingen af svarerne vil i nogle tilfælde give mulighed for at identificere en person.
Det kan være der bliver stillet spørgsmål omkring klassetrin, skole, køn, fritidsinteresser og om forældrene er skilt. De enkelte svar vil ikke være personoplysninger, men når resultaterne af det enkelte skema samles i en helhed, vil man kunne identificere deltageren – i nogle tilfælde.

Selvom der muligvis kan være flere elever, der har svaret det samme og der dermed f.eks. vil være to, som kan identificeres, vil det fortsat være personoplysninger, da det er så snæver en gruppe, at man ikke er i tvivl. Man skal derfor se oplysningerne i en større sammenhæng og ikke se på hvert spørgsmål isoleret.
Direkte identifikatorer er dem, som umiddelbart kan identificere personen uden, du har behov for yderligere informationer. Det er f.eks. det fulde navn eller CPR.nr. på en person – hvor den ene oplysning alene, kan identificere personen.

Stærke indirekte idenfikatorer er f.eks. email-adresse eller telefonnr., som ikke umiddelbart giver dig mulighed for at identificere personen, men ved at undersøge yderligere, kan knyttes til en person. Det kræver ikke flere oplysninger end den ene, at finde personen, hvis blot man søger efter indehaveren af oplysningerne.
Indirekte identifikatorer er dem, som ikke kan stå alene, hvis de skal knyttes til en person. Postnr. er f.eks. ikke en personoplysning i sig selv, men kan være det, hvis det knyttes til andre oplysninger og dermed, i sammenhæng, gør det muligt at identificere personen.

2. Hvad er behandling af personoplysninger?

Al brug af personoplysninger er behandling i juridisk forstand. Det er således enhver aktivitet eller række af aktiviteter som personoplysninger gøres til genstand for.
Det kan være indsamling, registrering, organisering, systematisering, opbevaring, tilpasning eller ændring, genfinding, søgning, brug, videregivelse, formidling eller enhver anden form for overladelse, sammenstilling eller samkøring, begrænsning, sletning eller tilintetgørelse. Også en kiggeadgang er omfattet af begrebet behandling.
Det er således enhver håndtering af data – i bred forstand.

2.1 Databehandler vs. dataansvarlig

Det er afgørende at få klarlagt dataansvaret, når der sker behandling af personoplysninger. Den dataansvarlige er den organisation, person eller

virksomhed, der har det overordnede ansvar for behandlingen af data. Det er således den dataansvarlige, der tager initiativ til behandlingen og beslutter hvilket formål og med hvilke hjælpemidler, der skal ske behandling. Den enkelte forsker er ikke dataansvarlig i ovennævnte forstand, det er derimod SDU.

Databehandleren er en ekstern (i forhold til den dataansvarlige), som behandler personoplysninger for den dataansvarlige. Det er således efter instruks og normalt også på opfordring af den dataansvarlige. Databehandleren har ansvaret for at opbevare og behandle data i overensstemmelse med den aftale, der sætter rammer for behandlingen. Aftalen kaldes en databehandleraftale og skal indgås mellem den dataansvarlige og databehandleren. Det er vigtigt at have rollerne på plads forud for opstarten af behandlingerne.

I forbindelse med udarbejdelsen af en databehandleraftale, er det juristerne ved RIO, der står for forhandlingen. Vi har en standardaftale, som der tages udgangspunkt i. Det er ligeledes hos RIO, alle databehandleraftaler skal registreres og underskrives. Vi vil derfor bede dig om at sende en mail til sdu.persondata@sdu.dk, hvis du har brug for hjælp med en databehandleraftale.

3. Anonymisering vs. pseudonymisering

Det er væsentligt at gøre sig klart, om der er tale om anonyme data eller ej. Hvis data er tilstrækkeligt anonymiserede, finder reglerne i databeskyttelsesforordningen ikke anvendelse, og derfor kan man behandle data inden for en langt friere ramme. Det er imidlertid meget svært at anonymisere oplysningerne således, at de ikke længere er identificerbare. Anonymiseringen skal i øvrigt være uigenkaldelig.

3.1 Pseudonymisering

Pseudonyme data er også personoplysninger. Pseudonymisering er den handling, som ikke gør oplysningerne direkte personhenførbare, men som heller ikke sikrer total anonymitet. Den typiske situation er når du udskifter navnet med et andet navn eller CPR.nr. med en kode. Så længe, der findes en nøgle, som kan tilbageføre data til direkte personhenførbare oplysninger, er dine oplysninger ikke anonyme – men pseudonyme. Når nøglefilen slettes, bliver de pseudonyme data anonyme.
Det afgørende for, om der er tale om pseudonyme data er, om der findes en nøgle til at decode oplysninger. Så længe der er mulighed for at anvende nøglen og data sammen, og på den måde tilbageføre data til personhenførbar form. Bemærk dog, at et datasæt ikke nødvendigvis er pseudonymiseret selv om du udskifter navn, adresse og CPR-nr. med en kode. Hvis f.eks. der er tale om et spørgeskema og stillingsbetegnelse er et fritekstfelt, så vil stillingsbetegnelsen "Statsminister" eller "Højesteretspræsident" betyde, at datasættet ikke er pseudonymiseret.

Pseudonymisering er således betegnelsen for udskiftningen af den direkte identifikation med nogle pseudonymer eller koder. Koderne skal opbevares separat og skal ligeledes beskyttet teknisk, således at det ikke er muligt for en udefrakommende at sammenstille koder og data og derved identificere personerne i datasættet. Så længe nøglen til pseudonymiseringen findes, er der tale om pseudonyme data. Det er først, når nøglefilen slettes, at datasættet bliver anonymt.

Hvor det er muligt at pseudonymisere personoplysningerne bør man gøre det så hurtigt som muligt efter at oplysningerne er indsamlet. Som ovenfor nævnt indebærer pseudonymisering ikke at databeskyttelsesreglerne ikke finder anvendelse, men det er en meget effektiv sikkerhedsforanstaltning, og skulle oplysningerne komme til uvedkommendes kendskab er skaden ikke så stor som tilfældet vil være, hvis direkte personhenførbare oplysninger kommer til uvedkommendes kendskab.

Afidentificerede data

Danmarks Statistik anvender yderligere et begreb: afidentificerede data. De definerer det som *data, hvor alle formelle identifikationsoplysninger som fx navn, cpr-nr., cvr-nr. og adresse er fjernet.* Dette er ikke at forveksle med pseudonyme data, da pseudonymitet forudsætter, at datasættet er anonymt – hvis man ikke har koden. De afidentificerede data er derimod det vi kalder indirekte identificerbare data – det er altså ikke muligt at identificere deltagerne ud fra datasættet umiddelbart.
De afidentificerede data er altså fortsat personhenførbare, selvom man ikke kan finde frem til deltagerne uden at skulle koble andre oplysninger

på eller undersøge det nærmere. Så længe, det er muligt at identificere deltageren, ved hjælp af rimelige hjælpemidler, er der tale om personoplysninger. De afidentificerede data er således personoplysninger i almindelig forstand og er hverken pseudonyme eller anonyme.

3.2 Anonymisering

Anonyme oplysninger er ikke personhenførbare og at det ikke er muligt at identificere en person ud fra oplysninger i datasættet. Der findes nogle forskellige værktøjer til at anonymisere og sikre, at der reelt bliver tale om anonyme data. Hvis oplysningerne, med en rimelig indsats, kan anvendes til at identificere en fysisk person, er der ikke tale om anonyme oplysninger. Hvad der nærmere ligger i begrebet *en rimelig indsats*, er svært at klarlægge. Det er en vurdering af de tekniske muligheder, der er, for at identificere personen. Man skal tage højde for alle de hjælpemidler, der, med rimelighed, kan tænkes anvendt i forsøget på at identificere en person.
Når det handler om andre muligheder for at identificere information i åbne spørgsmål og risikoen for at oplysningerne bliver tilgængelig, skal denne risiko vurderes i hvert tilfælde ud fra emnet for forskningsprojektet og de bagvedlæggende omstændigheder.
Det vil formentligt være svært at lave en fuldstændig anonymisering, men data anses for at være anonyme, hvis det ikke er muligt, med en rimelig indsats, at identificere en given person. Når det skal vurderes om det er muligt at identificere en person, skal man tage alle hjælpemidler, der med rimelighed kan tænkes bragt i anvendelse for at identificere den pågældende person, i betragtning. Man skal også have for øje, om andre er i besiddelse af oplysninger, der gør det muligt at identificere den enkelte.

4. Valg af behandlingshjemmel

Når du skal behandle personoplysninger til forskningsbrug, er der flere valgmuligheder i forhold til lovhjemmel. For det første, skal du gøre dig klart om der er tale om personoplysninger. Herefter skal du vurdere hvilken hjemmel, der passer bedst til dit projekt og giver deltagerne de bedste rettigheder. Der er forskel på rettighederne og kravene i forhold til de forskellige lovhjemler.
Uanset hvilken hjemmel du vælger, skal oplysningspligten opfyldes, så deltageren ved, hvad vedkommendes personoplysninger anvendes til. Hvis du bruger SDU's skabeloner, er du dækket ind i forhold til kravene til oplysningspligt.

Generelt skal du tænke på, at du ikke må behandle flere oplysninger, end det er relevant for dit projekt. Du skal derfor minimere dine behandlinger mest muligt, så du alene beder om de oplysninger, der er nødvendige. Dernæst kan du minimere behandlingerne ved at slette/anonymisere unødvendige oplysninger i takt med, du ikke længere har behov for dem. Dette er naturligvis alene en mulighed i forhold til de oplysninger, du ikke har brug for i forhold til verificering af forskningsresultater.

Hvis du skal indsamle, bruge osv. personoplysninger til et forskningsprojekt og du ved, at oplysningerne ikke senere skal bruges til andet end forskning, vil det ofte være en mulighed at bruge databeskyttelseslovens § 10 og/eller databeskyttelsesforordningens artikel 6, stk. 1. e) som hjemmelsgrundlag. Derved undgår du de risici der kan være forbundet med at benytte samtykke som hjemmelsgrundlag, se nedenfor.
Den næste hjemmel er samtykket i sin klassiske forstand – et samtykke til at deltage og til at SDU må behandle deltagerens personoplysninger i overensstemmelse med reglerne i databeskyttelsesforordningen.

Hvis du vælger at anvende samtykket som behandlingshjemmel, har deltagerne ret til tilbagekalde/fortryde deres samtykke og forlange sletning af deres personoplysninger. Det kan jo have betydning for dit projekt, hvis en eller flere deltagere trækker deres samtykke tilbage. Hvis du kan anonymisere oplysninger, vil de ikke længere være omfattet af reglerne om databeskyttelse og derfor har deltageren ikke ret til at få slettet de anonymiserede oplysninger (det er jo heller ikke muligt at finde den pågældende deltagers oplysninger, da de er anonymiserede).
Vælger du § 10/art. 6, er der ikke noget samtykke at tilbagekalde, og du risikerer derfor ikke at skulle slette oplysninger af den grund.

Vi har lavet en skabelon, som giver deltagerne mulighed for at 'samtykke' til at deltage – men samtykket dækker ikke selve behandlingen af oplysninger, da den er hjemlet i § 10/art. 6. Deltagerne skal derfor 'samtykke' til at deltage og kan trække dette samtykke tilbage, hvilket vil medføre at der ikke indsamles nye oplysninger, men forskeren kan stadig behandle de oplysninger, der er indsamlet inden 'samtykket' blev trukket tilbage. Der er således mulighed for at fortsætte behandlingen af de eksisterende oplysninger, idet indsamling, brug, analyse, videregivelse osv. sker med hjemmel i § 10/art. 6.

Det vigtigt, du vælger den korrekte hjemmel før du starter indsamlingen af personoplysninger.

5. Anonymiseringsmetoder

Der findes nogle forskellige teknikker til at anonymisere**.** For det første kan du gøre dig nogle tanker forud for indsamlingen, så du tilpasser din indsamling af data, på en måde, så det alene er de relevante oplysninger, der indsamles og i en form, som er så lidt personhenførbar som muligt. Det kan f.eks. være ved at lave lukkede spørgsmål, hvor deltageren ikke har mulighed for at tilføje egne oplysninger. På den måde får du ikke flere detaljer end dem, du har behov for. Du skal derfor være påpasselig med at indsætte spørgsmål, hvor det er muligt at skrive fritekst, men afgrænse til afkrydsning af forudbestemte muligheder.
Det er også muligt at ændre oprindelseskommune til at opdele i kategorier som urban, semi-urban, landlig osv. Dette vil minimere risikoen for identifikation – men kan stå ikke alene. Det vil være en del af en større anonymiseringsproces.

Man kan også tilføje *noise* eller støj, så oplysningerne tilføjes nogle ekstra oplysninger eller detaljer. Det kan f.eks. være ved at ændre alderen på deltageren med +- 2 år. Så vil data, til en vis grad være nøjagtige men ikke umiddelbare tilgængelig for uvedkommende – hvis støjen er en del af den fulde anonymisering.
Det kan også være en teknik at fjerne eller ændre oplysninger, så f.eks. personer med AIDS bliver kategoriseret som personer med svær, langvarig sygdom og andre personer er blot kategoriserede som rask eller syge. På den måde skelnes der ikke mellem de forskellige sygdomme men deltagerne deles op i grupper, som fortsat er relevante for forskningsprojektet.

Man kan også kategorisere de øvrige oplysninger i grupper, så der ikke fremgår så mange personhenførbare detaljer. Ved at dele deltagerne op i kategorier f.eks. i forhold til alder, uddannelsessted, arbejdsplads, bosted og husholdning. Det kan kategoriseres som 41-45 år, længevarende uddannelse, butik, by i Jylland og kone og to børn. På den måde er oplysningerne så generelle, at det er svært at identificere personen, da beskrivelsen kan passe på flere forskellige personer.
Nedenfor kan du se en overordnet oversigt over typen af oplysninger og hvad der kan gøres for at anonymisere dem. De **følsomme** oplysninger er markeret med fed.

Konkrete eksempler

I marts 2019 udtalte Datatilsynet alvorlig kritik vedr. et taxaselskabs behandling er personoplysninger. Selskabet havde 'anonymiseret' kundernes oplysninger ved at slette kundernes navne men ikke deres telefonnr. Dermed er der fortsat mulighed for at identificere den enkelte kunde – ud fra oplysninger om dato for kørslen, kørslens begyndelses- og sluttidspunkt, antal kørte kilometer, betalingen, GPS-koordinater. Dermed medfører at oplysningerne alene er pseudonyme og derfor stadig personhenførbare.

6. Anmeldelse af forskningsprojekt

Forskningsprojekter skal anmeldes til SDU's fortegnelse, hvis der behandles personoplysninger i projektet. Anmeldelsen skal ske via universitetets fortegnelse, som SDU RIO forvalter. SDU RIO skal, på forespørgsel fra Datatilsynet, kunne fremvise en fortegnelse over igangværende projekter på universitetet, hvor der behandles personoplysninger, til brug for Datatilsynets inspektion.
For at anmelde projektet skal der udfyldes et anmeldelsesskema i den elektroniske fortegnelse. Det gør du ved at tilgå SDU's fortegnelsesside. Herefter kommer en række spørgsmål, som alle relaterer sig til projektet – herunder om du behandler almindelige eller følsomme oplysninger, hvor længe du forventer, at dit projekt vil vare osv.
Når du har udfyldt skemaet, trykker du indsend. SDU RIO gennemgår herefter det anmeldte materiale med henblik på at sikre, at de databeskyttelsesretlige krav er opfyldt. En udtalelse fremsendes herefter til anmelderen.

7. Godkendelse fra andre myndigheder

Nogle projekter kræver yderligere godkendelse, end den, man får hos RIO. Sundhedsvidenskabelige projekter, hvor der indgår biologisk materiale fra mennesker, skal godkendes af Den Nationale Videnskabsetiske Komité. Ønsker du at få udtræk af oplysninger i patientjournaler, uden samtykke fra patienterne, kræver det en godkendelse fra Styrelsen for Patientsikkerhed.

Derudover kan der være tilfælde, hvor du ønsker at modtage oplysninger fra andre myndigheder, virksomheder eller registre mv., og her kræver det den afgivende parts godkendelse eller accept, før du kan modtage oplysningerne.
De fleste har et videregivelsessystem, hvor du kan søge om at modtage oplysninger. Andre ønsker at udarbejde en videregivelseserklæring, som du kan sende forbi SDU RIO til gennemlæsning. Når du modtager oplysningerne fra modparten, bliver SDU dataansvarlig for oplysningerne og det skal anmeldes til fortegnelsen.

8. Videregivelse/overladelse af forskningsdata

Juristerne i SDU RIO skal inddrages, når forskningsdata ønskes videregivet/overladt til andre med henblik på at sikre, at de databeskyttelsesretlige regler bliver overholdt.
Videregivelse er den situation, hvor du videregiver personoplysninger til eksterne parter uden for SDU. Det er altså personer, virksomheder, kommuner eller andre, som ikke er knyttet til eller ansat ved SDU.
Videregivelse kan også ske til parter uden for EU.
Overladelse er den situation, hvor du overlader data til en anden SDU-kollega og dermed forbliver data internt på SDU. Det vil typisk være en forsker, som skal bruge data til et nyt projekt, og ikke tidligere har haft adgang til data.
Du kan finde overladelses- og videregivelsesblanketter på SDU's hjemmeside.

9. Har du spørgsmål?

Spørgsmål vedr. behandling af persondata stiles til sdu.persondata@sdu.dk.

Identifikations type	Direkte identifikator	Stærk indirekte identifikator	Indirekte identifikator	Anonymiseringsmetode
CPR.nr.	X			Sletning
Fulde navn	X			Sletning/ændring
Emailadresse	X	X		Sletning
Telefonnr.		X		Sletning
Postnummer			X	Sletning/Kategorisering
Landsdel			X	Kategorisering
Kommune			X	Kategorisering
Region			X	
Lydfil (af stemme)	X			Sletning/sløring
Video (af person)	X			Sletning/sløring
Foto (af person)	X			Sletning/sløring
Fødselsår			X	Kategorisering
Alder			X	Kategorisering
Køn			X	
Ægteskabelig status			X	
Husholdnings-sammensætning			X	Kategorisering
Erhverv			X	Kategorisering
Uddannelse			X	Kategorisering
Arbejdssituation		(X)	X	
Modersprog			X	Kategorisering
Nationalitet			X	Kategorisering
Arbejdsplads		(X)	X	Kategorisering
Registreringsnummer på bil		X		Sletning
Hjemmeside		(X)	X	Sletning
Studenternr.		X		Sletning
Kontonr.		X		Sletning
Helbredsoplysninger		(X)		Kategorisering/sletning
IP-adresse		X		Sletning
Etnicitet		(X)		Kategorisering/sletning
Strafbare forhold			X	Kategorisering/sletning
Politiske forhold			X	Kategorisering
Religiøse forhold			X	Kategorisering
Sociale forhold			X	Kategorisering/sletning
Seksuel orientering			X	Sletning

Biometriske data		X		Sletning


Data til forskningsprojekt
Persondata (også pseudonyme) Pkt. 1
Anonyme data Pkt. 1
Omfattet af GDPR
Ikke omfattet af GDPR
SDU er dataansvarlig Pkt. 2.1
SDU er databehandler Pkt. 2.1
Du skal anmelde projektet til RIO Pkt. 6
Der skal laves en aftale med den dataansvarlige
Skal ikke-SDU-ansatte have adgang til data?
Hvor får du data fra?
Du må som udgangspunkt ikke bruge data til din egen forskning.
Videregivelse (Eksterns formål) Pkt. 8
Databehandling (SDU's formål) Pkt. 2.1
Indsamler selv fra deltageren Pkt. 4
Modtager fra andre end deltageren selv. F.eks. fra register, patientjournal mv. Pkt. 4
Der skal anomdes om videregivelse via hjemmesiden.
Der skal indgås en databehandleraftale. Pkt. 2.1
Samtykke + Oplysningspligt Pkt. 4
DBL § 10 (+ Oplysningspligt) Pkt. 4
DBL § 10 Pkt. 4
Godkendelse fra kompetent myndighed Pkt. 7