

Datasikkerhedsregler under mikrodataordningerne herunder regler for hjemsendelse af analyseresul- tater og sanktioner ved databrud

Indhold

| | |
|---|-----------|
| Indledning..... | 2 |
| Adgang til forskermaskinerne og arbejdet med data | 3 |
| Hvor må man arbejde med adgang til mikrodata? | 3 |
| Arbejdet med mikrodata – hvad er tilladt?..... | 4 |
| Hvad må hjemsendes fra forskermaskinerne? | 4 |
| Eksempel på dominansproblem med erhvervsdata..... | 5 |
| Max, median og min | 6 |
| Figurer..... | 6 |
| Særligt omkring uddannelsesinstitutioner | 7 |
| Eksempler: | 7 |
| Hvad må IKKE hjemsendes?..... | 9 |
| Særligt omkring Institutionsregistret | 10 |
| Eksempler på mikrodata | 10 |
| Procedurer og sanktioner vedbrud på datafortroligheden | 13 |
| Sanktioner rammer både den enkelte bruger og den autoriserede institution, der ejer et projekt | 14 |
| Hjemsendelse af detaljerede tabeller | 14 |
| Hjemsendelse af mikrodata | 15 |
| Orienter os snarest hvis der sker brud på datafortroligheden | 17 |
| Gode råd..... | 17 |
| For brugeren, der har adgang til mikrodata: | 17 |
| For den ansvarlige for autorisationen:..... | 17 |
| Eksempler på overtrædelser | 19 |
| Bilag 1 – Brugerftale | 22 |
| Bilag 2 – Tilknytningsaftale | 24 |

Indledning

Danmarks Statistik stiller vores mange værdifulde data til rådighed for brugere, som har et velbegrunder ønske om at analysere data til gavn for det danske samfund. Det er vigtigt, at dette sker på en måde, hvor både datasikkerhed og datafortrolighed er i top. Den enkelte borger og virksomhed har krav på, at fortrolige data behandles med den største omhyggelighed. De generelle regler for, hvordan fortrolige data skal behandles er fastlagt i EU-forordningen om beskyttelse af personoplysninger (Persondataforordningen)

Datafortrolighed er en helt central forudsætning for, at mikrodataordningerne kan eksistere. Alle datasæt, der gives adgang til via forskerordningen, er fortrolige. Derfor forpligter I jer som brugere – gennem aftaler med Danmarks Statistik - til kun at hente generelle resultater og *ikke* personhenførbare eller individdata ud fra forskermaskinerne på Danmarks Statistik.

Nedenfor gives en beskrivelse af hovedreglerne for, hvordan man skal arbejde med data på forskermaskinerne/de hostede maskiner, hjemsendelse af filer, herunder, hvad der menes med mikrodata og detaljerede tabeller, proceduren ved et brud på datafortroligheden, samt nogle gode råd til, hvordan man undgår brud. Sidst i vejledningen findes en oversigt over nogle generelle eksempler på typer af databrud.

De regler som er beskrevet i denne vejledning er en forenklet version af de regler der findes i Danmarks Statistiks datafortrolighedspolitik. Ved hjemsendelse af resultater under mikrodataordningerne er det reglerne som beskrevet i denne vejlednings som er gældende.

Adgang til forskermaskinerne og arbejdet med data

Før man kan få adgang til data under mikrodataordningerne, er der visse grundlæggende betingelser, som skal være opfyldt. Vi kræver, at den institution, der søger om adgang til mikrodata, skal være autoriseret, dvs. at institutionen skal være godkendt af Danmarks Statistik til at kunne få adgang til mikrodata til konkrete projekter. Derudover skal alle brugere godkende en brugeraftale. Når en bruger godkender brugeraftalen, bekræfter brugeren at være bekendt med reglerne for datasikkerhed, dvs. at brugeren har sat sig godt ind i, hvad man må, og hvad man ikke må, når man har adgang til fortrolige data. Endelig skal den ansvarlige for institutionens autorisation eller dennes stedfortræder ligeledes godkende en tilknytningsaftale. Godkendelse af tilknytningsaftalen bekræfter, at institutionen tager det datasikkerhedsmæssige ansvar for en brugers adgang til pseudonymiserede data på forskermaskinen. Disse aftaler fremgår af bilag 1 og bilag 2.

Når man får adgang til data under mikrodataordningerne, får man adgang til data under et projektnummer. Det kan være et konkret projekt, en projektdatabase eller en myndighedsordning.

Fælles for disse projekter er, at de data, som man får adgang til, indeholder fortrolige oplysninger. Adgangen til data er derfor beskyttet med personlige passwords.

De passwords, som man får udleveret for at få adgang til data på forskermaskinerne, er personlige og må ikke udleveres til andre

Man må aldrig overlevere disse passwords til andre. Det betyder også at man ikke må forlade eller overlade sin computer til andre, når man er logget ind på sit projekt på forskermaskinerne. Heller ikke, hvis de selv har godkendt adgang til det samme projekt. Ved en godkendt adgang forstås, at en bruger har godkendt en brugeraftale, er blevet tilknyttet institutionen via en godkendt tilknytningsaftale og har fået adgang til det konkrete projekts data.

Skal man forlade skærmen, både i kortere eller længere tid, så skal adgangen til forskermaskinen afbrydes (man logger ud), så det sikres, at andre ikke kan få adgang. Det er altså ikke nok at sætte en screensaver på - adgangen til forskermaskinen skal helt afbrydes. Hvis man har mistanke om, at andre kender ens password, så skal man hurtigst muligt ændre sit password.

Hvor må man arbejde med adgang til mikrodata?

Det er kun tilladt at arbejde med mikrodata under mikrodataordningerne, nedenstående steder:

- På arbejdspladsen, dvs. den autoriserede institution
- Via en hjemmearbejdsplads

Ovenstående regler betyder også, at man ikke må arbejde med data offentlige steder, fx på biblioteker, cafeer eller i toget.

Arbejdet med mikrodata – hvad er tilladt?

Det er tilladt både at se og arbejde med mikrodata på skærmen. Det er fx relevant at se på en lille stikprøve af mikrodata, hvis man, som en kontrol, skal se om dannelsen af en ny variabel er gået godt. Men man må aldrig forsøge at identificere enkeltpersoner - heller ikke sig selv.

Det kan også ske, at man har brug for hjælp fx til programmering. Det er tilladt at få hjælp af en person, der også har en godkendt adgang til projektets data – forudsat at man IKKE forlader skærmen, så længe hjælpen eller vejledningen pågår. Dvs. hvis der er behov for, at en person skal hjælpe, så skal det først sikres, at denne person har adgang til projektet. Adgangen kan jeres institutions FSE-BOA-administratorer give via. FSEBOA.

Dvs. at det er tilladt at:

- se og arbejde med mikrodata på skærmen fx for at kontrollere at beregninger er gået godt
- få hjælp ved skærmen af en person, der også har en godkendt adgang til projektets data

Hvad må hjemsendes fra forskermaskinerne?

Al hjemsendelse af resultater skal foregå via hjemsendelsesværktøjet. Kun ved at bruge hjemsendelsesværktøjet sikres det at de hjemsendte filer bliver skannet for mikrodata og man derfor kan få en advarsel herom før materialet hjemsendes og derved undgå et muligt databrud. En vejledning i hjemsendelse af resultater findes under linket: <https://www.dst.dk/da/TilSalg/Forskningsservice/hjemtagelse-af-analyseresultater> - hjemsendelse af filer.

Som hovedregel må der kun hjemsendes analyseresultater, aggregerede tabeller eller figurer, hvor det ikke er muligt at identificere enkelte enheder, dvs. hverken personer, husstande, familier, virksomheder og andre enheder med afidentificerede løbenummer. Dette gælder også, selvom det afidentificerede løbenummer er fjernet. **Det må med andre ord ikke være muligt at genkende nogen enkeltpersoner eller -virksomheder i det hjemsendte materiale.** Generelt gælder nedenstående tommelfingerregler for aggregerede data.

Hjemsendte resultater bør som udgangspunkt have en sådan form, at de umiddelbart kan anvendes i en publikation

Men det er tilladt at arbejde videre med det hjemsendte *aggregerede* materiale, fx til dannelse af figurer eller videre analyser i statistikprogrammer, forudsat det hjemsendte materiale opfylder ovenstående betingelser.

For tabeller gælder nedenstående:

Tabeller skal indeholde mindst 3 observationer pr. celle

For tabeller er tommelfingerreglen, at tabeller skal indeholde mindst 3 observationer, men aggregeringsniveauet skal altid afhænge af en konkret vurdering

af, hvad tabellen indeholder. Hvis man er i tvivl om 3 observationer er nok til at sikre anonymiteten, så skal der aggregeres yderligere. Med aggregering menes, at celler/tal slås sammen, indtil de overholder antallet af observationer og sikrer anonymitet.

Kravet gælder det laveste aggregeringsniveau. Dvs. hvis man kombinerer flere variabler i en tabel, så skal der mindst være tre observationer i alle celler i de variabler, man kombinerer. Hvis minimumskrevet ikke vurderes af være tilstrækkeligt, så skal der aggregeres yderligere. Det helt centrale er at aggregering sikrer anonymiteten, og man skal desuden være opmærksom på, at kolonner med totaler ikke afslører et evt. sløret antal observationer.

Hvis man arbejder med erhvervsdata, så skal man, ud over kravet om mindst tre observationer i hver celle, ligeledes overholde dominanskriteriet.

Ved erhvervsdata skal de to største statistiske enheder (virksomheder) samlet ikke udgøre mere end 85 pct. af omsætningen i en given tabelcelle

Bemærk, at dominanskriteriet altid beregnes på baggrund af omsætningen, uanset om der oplyses fuldtidsansatte og/eller en anden økonomisk variabel, fx. værditilvækst i tabellen. Nedenfor vises et eksempel på en tabel hvor dominanskriteriet ikke er opfyldt.

Hvis kriterierne ikke er opfyldt, så er det nødvendigt at "sløre" data, så enkelt enheders oplysninger ikke kan identificeres.

Dette kan gøres ved følgende:

1. Sammenlægning af undergrupper (aggregering).
2. Sletning af dominerende værdier/out liere.

Ved sammenlægning af grupper tjekkes det efterfølgende, om den nye gruppering af enheder opfylder begge diskretioneringskriterier, dvs. antalskriteriet (minimum 3 observationer) og dominanskriteriet. Ved sletning af oplysninger skal man være opmærksom på evt. behov for sekundær diskretionering (også kaldet følgediskretionering), således at de slettede oplysninger ikke kan afledes ud fra totaler, subtotaler eller lignende.

Eksempel på dominansproblem med erhvervsdata

Tabellen nedenfor viser et eksempel på et diskretioneringsproblem som følge af at dominanskriteriet ikke bliver overholdt.

Eksempel på dominansproblem

| Diskretionering | Branche | Antal virksomheder | Omsætning | Omsætning af to største virksomheder – tjek af dominans |
|-----------------|--------------|--------------------|------------------------|--|
| Nej | Hovedbranche | 25 | 100 mio. kr. | 70 mio. kr. |
| Ja, sekundær | Sektor 1 | 5 | 30 mio. kr. | 20 mio. kr. |
| Ja, primær | Sektor 2 | 20 | 70 mio. kr. | 65 mio. kr. |

Tabellen viser, at der er mere end 3 virksomheder i hver gruppe, dvs. at tabellen opfylder minimumskravet for aggregering. Da der indgår virksomhedsdata i tabellen, så skal der også kontrolleres for dominerende virksomheder. Det gøres ved at se på omsætningen af de to største virksomheder i de tre grupper, se

kolonne 5 i tabellen. For sektor 1 udgør de to største CVR-numre 67 pct. (20/30) af omsætningen, så her er der ikke et dominansproblem. Men for sektor 2 udgør de to største virksomheder 93 pct. (65/70) af omsætningen, og dermed har vi et dominansproblem. Derfor bortdiskretioneres først omsætningen for sektor 2 (primær diskretionering), og herefter bortdiskretioneres også omsætningen for sektor 1 (følgediskretionering), da det ellers vil være muligt at udlede omsætningen for sektor 2.

Hvis tabellen også indeholdt andre oplysninger – fx værditilvækst eller antal ansatte i tabellen ovenfor, så skal disse oplysninger på lige fod med omsætningen også diskretioneres i sektor 1 og sektor 2, da der er tale om økonomisk variable og disse oplysninger derfor også skal opfylde dominanskriteriet.

For hovedbranchen samlet udgør de to største CVR-numre 70 pct. (70/100) af omsætningen, så her skal ikke diskretioneres. Alternativt til at fjerne omsætningen i de to sektorer, kan man i stedet vælge kun at oplyse omsætningen for sektor 1 og derved bortdiskretionere omsætningen for sektor 2 og totalen for hovedbranchen. Dog skal man her være absolut sikker på at brugeren ikke har mulighed for at få omsætningen for hovedbranchen fra anden kilde, ex. offentliggjorte tal i Statistikbanken eller tidligere leveringer.

Hvis man vælger at diskretionere omsætningen i de to sektorer, så vil den diskretionerede tabel se ud som følger:

Eksempel på dominansproblem – efter diskretionering

| Diskretionering | Branche | Antal virksomheder | Omsætning |
|-----------------|--------------|--------------------|--------------|
| Nej | Hovedbranche | 25 | 100 mio. kr. |
| Ja, sekundær | Sektor 1 | 5 | - |
| Ja, primær | Sektor 2 | 20 | - |

Max, median og min

Det er under mikrodataordningerne tilladt at hjemsende værdier, der beskriver en fordeling, som max, min og median, der ofte henviser til enkeltindivider. Men det er en forudsætning for at hjemsende disse nøgletal, at der ikke er fare for identifikation af enkeltindivider. Datasikkerheden skal altid prioriteres højest, og hvis der er tvivl om identifikation ved brug af disse værdier, så må de ikke hjemsendes.

Max og min og percentiler som fx median må kun hjemsendes, hvis der ikke er fare for identifikation af enkeltindivider. Ingen personer må derfor kunne genkendes i det hjemsendte output

Man skal være særligt opmærksom på, om et materiale indeholder outliers, dvs. ekstreme værdier og evt. fjerne dem, før man hjemsender fx max-værdier. Skal max og min værdier sendes hjem, skal observationer i begge ender af fordelingen ligge tæt på hinanden, så ingen observationer skiller sig ud og dermed skaber risiko for identifikation.

Figurer

Figurer må kun hjemsendes, hvis de ikke indeholder identificerbar information. Man skal ved figurer ligeledes være særligt opmærksomme på såkaldte outliere/ekstreme værdier.

Det skal derudover sikres, at figurerne ikke indeholder indlejrede mikrodata, dvs. at observationerne der danner grundlag for figuren også gemmes i filen. Sidstnævnte kan fx, sikres ved at hjemsende figuren i pdf-format.

Generelt må figurer kun hjemsendes, hvis man er sikker på, at de ikke indeholder identificerbar information. Dvs. at man fx med Kaplan Meier kurver skal være opmærksom på små populationer og ved Scatterplots skal være særlig opmærksom på outliere. Dvs. at figurer generelt kun må sendes hjem, hvis man er helt sikker på, at indholdet ikke på nogen måde kan identificere enkeltindivider eller virksomheder. Hvis man er i tvivl, så skal man ikke sende dem hjem. I stedet bør man overveje, om man kan transformere data hensigtsmæssigt for at sikre anonymiteten fx ved at bruge gennemsnitsobservationer baseret på fx 5 observationer – hvis det giver mening.

Særligt omkring uddannelsesinstitutioner

Det er tilladt at hjemsende oplysninger på institutionsniveau. Dvs. at man godt må hjemsende oplysninger aggregeret på et enkelt institutionsnummer, forudsat hjemsendelsen opfylder ovenstående krav, dvs. at det ikke er muligt at identificere enkeltindivider. Vær opmærksom på at aggregeringskravet evt. skal skærpes, da man hjemtager oplysninger på et identificerbart område. Fx vil små populationer med kombinerede oplysninger om fx alder og køn øge risikoen for, at hjemsendte oplysninger er personhenførbare.

Eksempler:

Nedenfor præsenteres to eksempler på et sæt fiktive tabeller, hvor hjemsendelse vil være hhv. forbudt og tilladt.

Tabellerne nedenfor må ikke sendes hjem. Tabellerne opfylder ikke antalskriteriet, da to af observationerne har under tre observationer (Herning og Odense). I den nederste tabel er tilføjet en yderligere variabel indkomst og her får man yderligere information, om indkomsten for den enkelte person fra Herning. Reelt indeholder denne tabel individinformation, og det må ikke forekomme. Dvs. at disse tabeller er for små til at hjemsende.

| Bopæl | Antal personer | |
|------------|----------------|-----------------|
| Bornholm | 3 | |
| Herning | 1 | |
| København | 4 | |
| Odense | 2 | |
| Aarhus | 3 | |
| Hovedtotal | 13 | |
| | | |
| | | |
| | | |
| Bopæl | Antal personer | Sum af indkomst |
| Bornholm | 3 | 1650 |
| Herning | 1 | 750 |
| København | 4 | 150750 |
| Odense | 2 | 400 |
| Aarhus | 3 | 1800 |
| Hovedtotal | 13 | 155350 |

Nedenfor de samme to tabeller, men nu med et aggregeringsniveau, hvor de kan hjemsendes uden risiko for identifikation og uden risiko for, at man får yderligere information om enkeltpersoner, som det var tilfældet ved personen fra Herning i de to første tabeller.

| Bopæl | Antal personer | |
|------------|----------------|-----------------|
| Bornholm | 30 | |
| Herning | 10 | |
| København | 40 | |
| Odense | 20 | |
| Aarhus | 30 | |
| Hovedtotal | 130 | |
| | | |
| | | |
| | | |
| Bopæl | Antal personer | Sum af indkomst |
| Bornholm | 30 | 16500 |
| Herning | 10 | 7500 |
| København | 40 | 1507500 |
| Odense | 20 | 4000 |
| Aarhus | 30 | 18000 |
| Hovedtotal | 130 | 1553500 |

Hvad må IKKE hjemsendes?

Ingen form for mikrodata må hjemsendes

Hverken via hjemsendelsesværktøjet eller på andre måder fx ved at skrive oplysninger fra skærmen, tage et skærmprent eller et billede af skærmen med mobiltelefonen

Det gælder også de mikrodata som brugerne selv har indsendt i forbindelse med et konkret projekt.

Mikrodata defineres som data, der fx indeholder nedenstående:

- Datasæt – eller dele af datasæt - med oplysninger på individ-/virksomhedsniveau - dvs. data, som er knyttet til et enkelt individ eller en enkelt virksomhed. Det gælder også i de tilfælde, hvor man har fjernet den af-identificerede ID-variabel - fx personnummeret eller CVR nummer. Dvs. fx et datasæt på individniveau bestående af baggrundsoplysninger som indkomst, uddannelse og socioøkonomisk status eller beregnede variabler på individniveau, men hvor det pseudonymiserede personnummer for hver person er fjernet. Det samme gælder for virksomheder
- Pseudonymiserede nøglevariabler som personnumre, CVR-numre, arbejdsstedsnumre, adressekoder mv. betragtes altid som mikrodata, da de angiver et unikt nummer, der henviser direkte til en enkelt person eller virksomhed. Pseudonymiserede personnumre mv. må aldrig hjemsendes uanset, om de er knyttet til anden information eller ej

Pas på med programfiler og logfiler

Programmer og logfiler kan også indeholde mikrodata fx:

- Programkode der indeholder mikrodata skrevet ind i en betingelse
- Logfiler der lister observationer. Nogle procedurer lister fx et lille udsnit af det datasæt, der er anvendt til analysen

En hjemsendt logfil, der lister pseudonymiserede personnumre, betragtes altid som et regelbrud – også selv om disse personnumre ikke indeholder anden information.

| |
|---|
| Al arbejde med mikrodata skal foregå på forskermaskinerne |
|---|

Det er tilladt at danne logfiler, der lister pseudonymiserede personnumre – fx som kontrol af et program, men en sådan fil SKAL, i lighed med alle andre filer, der indeholder mikrodata, blive på forskermaskinen og må IKKE hjemsendes.

Brugernes egne mikrodata, som er indsendt til Danmarks Statistik for at indgå i et projekt, er omfattet af de samme regler. Dvs. at det heller ikke er tilladt at hjemsende sine egne mikrodata fra forskermaskinerne.

| |
|---|
| Tabeller med under tre observationer må ikke hjemsendes |
|---|

Indeholder tabellerne færre end tre observationer, skal disse tabeller aggregeres yderligere fx ved at lægge kategorier sammen, alternativt blænde celler med få observationer, før tabellen hjemsendes. Blænding kan fx ske ved at slette cellens indhold, eller erstatte det med en markering såsom ”-”. Hvis man blænder celler, skal man være opmærksom på, at det ikke må være muligt at regne tilbage og finde den værdi, der er blændet - fx hvis der kun blændes én celle, og totalen er angivet, så skal der blændes mindst to celler. Det er derfor bedst at sammenlægge kategorier, hvis muligt, så man undgår denne problematik.

Derudover anvendes for erhvervsstatistik én yderligere fortrolighedsregel for økonomiske variabler (fx omsætning eller værditilvækst), det såkaldte dominanskriterium. Det betyder, at hvis det største eller de to største virksomheder i en tabelcelle tilsammen udgør en dominerende andel, dvs. 85 pct. af værdien i cellen eller derover, så træder dominanskriteriet i kraft, og cellen skal blændes, se evt. eksemplet på side 5: Eksempel på dominansproblem med erhvervsdata.

Særligt omkring Institutionsregistret

Det er fra Institutionsregistret ikke tilladt at hjemsende andre oplysninger på institutionsniveau end institutionsnummeret. Dvs. at man ikke må hjemsende de tilknyttede pseudonymiserede adressekoder og CVR-numre på institutionerne. Det betragter vi som hjemsendelse af mikrodata, også selv om de samme oplysninger er offentligt tilgængelige.

Eksempler på mikrodata

Nedenfor vises eksempler på mikrodata, der skal blive på forskermaskinen.

Tabel 1. viser et eksempel på data, der indeholder mikrodata. Tabellen indeholder oplysninger om 12 personers køn, indkomst, bopæl, alder og en reference til personernes børn. Disse oplysninger henviser direkte til et individ og må ikke hjemsendes.

Tabel 1. Eksempel på mikrodata

| PNR | KØN | INDKOMST | BOPÆL | ALDER | PNRB |
|-----|-----|----------|-----------|-------|------|
| 1 | M | 1500000 | København | 40 | 15 |
| 2 | M | 1500 | Aarhus | 45 | 22 |
| 3 | K | 1000 | Bornholm | 50 | 37 |
| 4 | M | 500 | Bornholm | 50 | 65 |
| 4 | M | 250 | København | 40 | 87 |
| 5 | M | 150 | Aarhus | 20 | |
| 6 | K | 150 | Odense | 25 | |
| 7 | M | 750 | Herning | 35 | 19 |
| 8 | K | 500 | København | 60 | 74 |
| 9 | M | 250 | Odense | 65 | 74 |
| 10 | K | 150 | Aarhus | 20 | |
| 11 | K | 50 | Bornholm | 15 | |
| 12 | M | 0 | København | 10 | |

I Tabel 2 er identifikationsvariablerne pnr og pnrB er fjernet. Men der er stadig tale om mikrodata, fordi oplysningerne kan henføres til et enkelt individ. De skal derfor også blive på forskermaskinen.

Tabel 2. Eksempel på mikrodata

| KØN | INDKOMST | BOPÆL | ALDER |
|-----|----------|-----------|-------|
| M | 1500000 | København | 40 |
| M | 1500 | Aarhus | 45 |
| K | 1000 | Bornholm | 50 |
| M | 500 | Bornholm | 50 |
| M | 250 | København | 40 |
| M | 150 | Aarhus | 20 |
| K | 150 | Odense | 25 |
| M | 750 | Herning | 35 |
| K | 500 | København | 60 |
| M | 250 | Odense | 65 |
| K | 150 | Aarhus | 20 |
| K | 50 | Bornholm | 15 |
| M | 0 | København | 10 |

I tabel 3 indeholder den fra tabel 1 kun det pseudonymiserede personnummer. Pseudonymiserede personnumre er altid mikrodata og må ikke hjemsendes. Det er uden betydning at alle de øvrige variabler er fjernet. Der er stadig tale om individdata, da oplysningerne kan henføres til et enkelt individ.

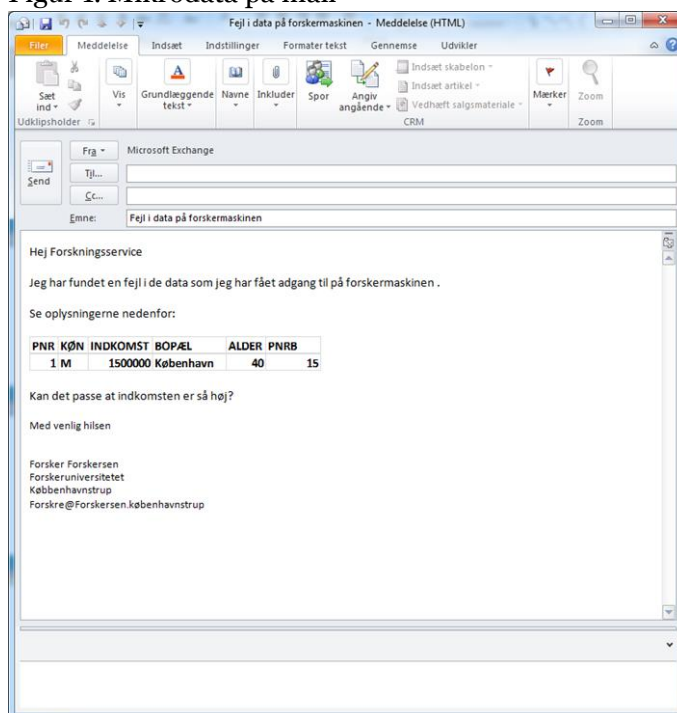
Tabel 3. Eksempel på mikrodata

| PNR |
|-----|
| 1 |
| 2 |
| 3 |
| 4 |
| 4 |
| 5 |
| 6 |
| 7 |
| 8 |
| 9 |
| 10 |
| 11 |
| 12 |

Al mikrodata, uanset form, skal blive på forskermaskinen.

Det er heller ikke tilladt at sende mikrodata på mail – fx i forbindelse med fejlfinding. Det gælder også, selv om den sendes til en medarbejder i Forsknings-service eller en medarbejder i et fagkontor.

Figur 1. Mikrodata på mail

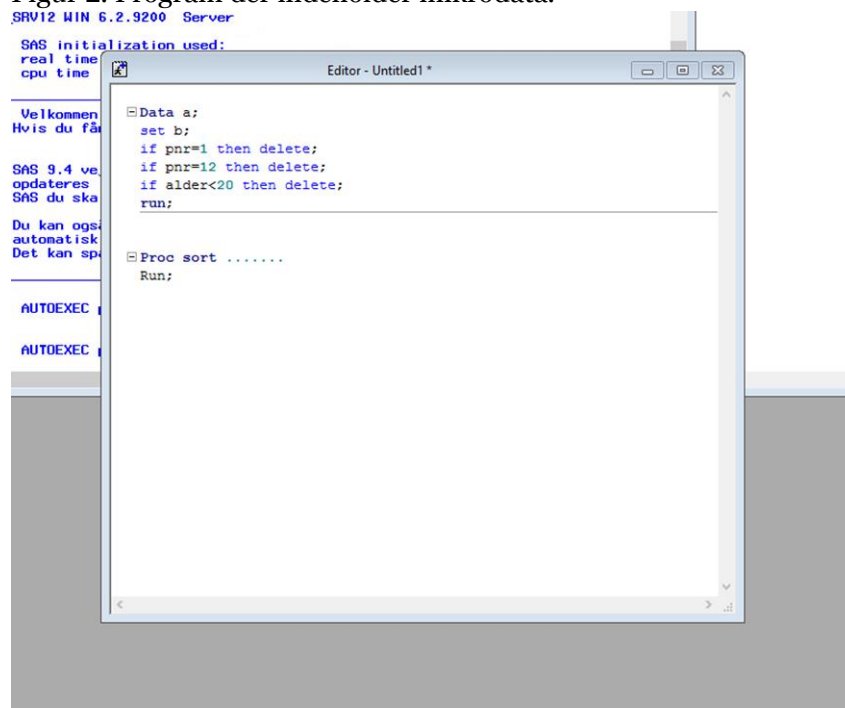


Hvis man konstaterer en fejl i de data, man har fået leveret eller har nogle konkrete spørgsmål til mikrodata, som man gerne vil vise et eksempel på, så skal

man danne en fil og gemme den på forskermaskinen. Herefter skal man orientere FSE om, hvor filen ligger – typisk henviser til den sti, hvor filen er placeret. Herefter kan FSE gå ind på forskermaskinen og se filen. Herved sikres det, at disse filer bliver på forskermaskinen.

Man skal også være påpasselig med at hjemsende programmer fra forskermaskinen, da de også kan indeholde mikrodata, jf. eksemplet nedenfor:

Figur 2. Program der indeholder mikrodata.



I programmet har man kodet pseudonymiserede personnumre direkte ind i programmet. Hvis et sådan program sendes hjem, så er det et brud på reglerne for hjemsendelse. Vær derfor varsom med at programmere sådanne oplysninger direkte ind i programmet. Anvend i stedet nogle generelle regler for at fjerne outliers i et datasæt, fx hvis indkomsten er større eller lig med 150000 etc. jf. tabel 1. Så undgår man at kode personoplysninger direkte i programmet.

Man må godt inkludere mikrodata i programmer, men så skal de blive på forskermaskinen, og man udsætter sig selv og den autoriserede institution for en øget risiko for databrud. Det kunne fx være, hvis programmet senere overtages af en anden bruger og sendes hjem.

Procedurer og sanktioner vedbrud på datafortroligheden

Hvis et brud på datafortroligheden konstateres, så skelnes der mellem hjemsendelse af detaljerede tabeller og hjemsendelse af mikrodata.

Hvis vi konstaterer, at der er hjemsendt mikrodata, så lukker vi for institutionens adgang og følger proceduren beskrevet nedenfor.

Hvis vi konstaterer, at der er hjemsendt tabeller med mindre end tre observationer i cellerne, så lukker vi IKKE for adgangen, men kontakter brugeren for en redegørelse og vejleder brugeren i reglerne for hjemsendelse af data. Hvis der er tale om gentagne brud på diskretioneringskravet, eller hvis detaljeringsgraden i

de hjemsendte tabeller er meget høj (fx hvis langt de fleste celler indeholder 1-taller), så kan sagen behandles som hjemsendelse af mikrodata.

Sanktioner rammer både den enkelte bruger og den autoriserede institution, der ejer et projekt

Kernen i mikrodataordningerne er, at Danmarks Statistik autoriserer konkrete institutioner til at anvende data til nærmere afgrænsede analyseprojekter. Det er vigtigt, at vi kan have fuld tillid til de institutioner, vi autoriserer. Danmarks Statistik har et godt kendskab til de autoriserede institutioner, mens kendskabet til de enkelte brugere er langt bedre i de autoriserede institutioner end på Danmarks Statistik. Derfor er det afgørende vigtigt, at vi kan stole på, at institutionerne står 100 % inde for de brugere, som de ønsker at give adgang til deres projekter. Institutionerne skal stå inde for, at de brugere, der arbejder på deres projekter, har de nødvendige kompetencer, og at de er godt bekendt med reglerne om datafortrolighed, hjemsendelser mv.

I hver autoriseret institution er der derfor udpeget en ansvarlig chef, der har ansvaret for - og fører tilsyn med - at brugere, der er tilknyttet deres projekter, er bekendt med Danmarks Statistiks datafortrolighedsregler og retningslinjer for hjemsendelse.

Af de aftaler Danmarks Statistik indgår med de autoriserede institutioner og deres brugere, fremgår det bl.a., at en alvorlig overtrædelse af datafortrolighedsreglerne vil medføre, at brugere, der overtræder reglerne, kan udelukkes fra at anvende Danmarks Statistiks mikrodataordninger permanent eller i en periode. Ligeledes gælder det, at den institution, der ejer det konkrete projekt, kan blive udelukket fra at bruge ordningerne. En overtrædelse af datafortrolighedsreglerne kan altså få konsekvenser både for den person, der overtræder datafortrolighedsreglerne og for de resterende projekter og brugere, der er knyttet til den pågældende autoriserede institution, som ejer pågældende projekt. Begrundelsen for dette er, at en alvorlig overtrædelse af datafortrolighedsreglerne, tyder på, at den involverede autoriserede institution ikke har taget sit ansvar tilstrækkeligt alvorligt. Det er et problem for hele mikrodataordningen. Derfor er sanktionen ikke kun rettet mod den enkelte bruger, men også mod den autoriserede institution, der har ansvaret for, at pågældende bruger kender reglerne og efterlever reglerne.

Hjemsendelse af detaljerede tabeller

Ved hjemsendelse af tabeller, hvor vi vurderer, at der er fare for identifikation af enkeltindivider, bliver brugeren, der har hjemsendt filen, samt den autorisationsansvarlige for den institution, som brugeren er tilknyttet, kontaktet og bedt om at sende en redegørelse for hjemsendelsen, samt bekræfte, at filen er slettet. Ligeledes vil der blive bedt om en redegørelse for, hvordan man fremadrettet vil undgå overtrædelser. Ved enkeltstående overtrædelser vil det for brugeren og institutionen normalt ikke få konsekvenser for adgangen til mikrodata, men ved gentagne regelbrud, hvor der er sendt tabeller hjem, hvor der er fare for identifikation af enkeltpersoner, kan denne praksis føre til lukning for institutionens adgang til mikrodata og en videre behandling af sagen svarende til hjemsendelse af mikrodata.

Hjemsendelse af mikrodata

Hvis Danmarks Statistik konstaterer, at en bruger har hjemsendt mikrodata, lukkes der omgående for brugernes adgang til mikrodata og for samtlige projekter, der hører under den autoriserede institution, der ejer det projekt, hvorfra der er hjemsendt mikrodata. Hvis en bruger hjemsender mikrodata på et projekt, der ikke er ejet af vedkommendes egen institution, men af en anden autoriseret institution, lukkes der for den institution, der ejer projektet – ikke for brugerens egen institution. Det datasikkerhedsmæssige ansvar for en bruger påhviler altid den projektejende institution. For den bruger, der har hjemsendt mikrodata, lukkes adgangen til alle projekter uanset, hvilken institution de hører under.

Derefter bliver brugeren, der har hjemsendt filen, samt den ansvarlige for institutionens autorisationsaftale kontaktet. De bliver bedt om øjeblikkeligt at slette samtlige filer, der overtræder reglerne for hjemsendelse af mikrodata, samt sende en bekræftelse på, at alle filer, der overtræder reglerne er slettet. Dette gælder både data beliggende på harddiske, mail konti, og hvor disse data ellers måtte være opbevaret.

Danmarks Statistik vil derudover udbede sig en redegørelse for regelbruddet, samt omfanget af hjemsendte filer, der bryder reglerne. Ansvar for redegørelsens udarbejdelse påhviler den autorisationsansvarlige for institutionen. Efterfølgende vil Danmarks Statistik bede institutionen om snarest muligt at sende en plan for, hvordan den fremadrettet vil undgå yderligere regelbrud. Alle adgange vil forblive lukket indtil både redegørelsen og planen er modtaget.

Herefter beslutter Danmarks Statistiks direktion, hvilken sanktion institutionen skal have.

Som udgangspunkt vil sanktionen være, at der lukkes for adgangen til mikrodata for alle brugere og projekter tilknyttet institutionen i minimum en måned, samt for den berørte bruger al adgang til mikrodata i samme periode. Ved gentagne brud vil adgangen blive lukket i en længere periode eller i særlig alvorlige tilfælde permanent. Har man begået et brud, er der en karensperiode på to år. Det vil sige at begår man et nyt brud inden for to år, så vil sanktionen stige, svarende til tabellen nedenfor. Fx vil et nyt brud indenfor karensperioden kunne betyde, at institutionen bliver lukket i to måneder. Efter to år forældes et brud, og det vil ikke længere tillægges betydning ved efterfølgende sanktioneringer.

Hvis brugeren selv gør Danmarks Statistik opmærksom på, at vedkommende har hjemsendt mikrodata, vil det blive betragtet som en formildende omstændighed, og det vil muliggøre at dataadgangen vil kunne åbnes igen, så snart Danmarks Statistik har fået klarhed over overtrædelsens omfang og karakter. Disse brud er ikke omfattet af en karensperiode. Dvs. at der ses bort fra disse brud ved fastlæggelse af en sanktion, hvis et nyt brud opdages af Danmarks Statistik.

En vejledende oversigt over sanktioner fremgår af oversigten i Tabel 1.
Tabel 1. Vejledende oversigt sanktioner ved regelbrud.

| | | Sanktion mod institution | | | Sanktion mod bruger | |
|--|---|--|--------------------------------------|--------------------------------------|---|---|
| | | 1. gang for institution | 2. gang for institution | Gentagne gange for institution | 1. gang | Flere gange |
| Hjemsendelse af individdata | | | | | | |
| a. | Teknisk uheld, uden vilje til at gøre dette | 1 måneds karantæne for institution (kan være en formildende omstændighed hvis institutionen fx selv underretter om fejlen) | 2 måneders karantæne for institution | 3 måneders karantæne for institution | 1 måneds karantæne fra alle forskningsprojekter (kan være en formildende omstændighed hvis institutionen fx selv underretter om fejlen) | 3 måneders karantæne fra alle forskningsprojekter |
| b. | Bevidst handling, ville se på dem (fejlsøgning) | 2 måneders karantæne for institution | 3 måneders karantæne for institution | 3 måneders karantæne for institution | 2 måneders karantæne fra alle forskningsprojekter | Permanent udelukkelse |
| c. | Bevidst, forsøg på at identificere | 3 måneders karantæne for institution | 3 måneders karantæne for institution | Permanent udelukkelse | Permanent udelukkelse | Kan ikke forekomme |
| Password eller adgang videregivet af autoriseret person | | | | | | |
| a. | Sjusk | 1 måneds karantæne for institution | 2 måneders karantæne for institution | 3 måneders karantæne for institution | 1 måneds karantæne fra alle forskningsprojekter | 3 måneders karantæne fra alle forskningsprojekter |
| b. | Med vilje | 3 måneders karantæne for institution | 3 måneders karantæne for institution | Permanent udelukkelse | 3 måneders karantæne fra alle forskningsprojekter | Permanent udelukkelse |

Orienter os snarest hvis der sker brud på datafortroligheden

Hvis der ved en fejl sendes mikrodata eller detaljerede tabeller hjem så kontakt Danmarks Statistik hurtigst muligt.

Orienter altid snarest Danmarks Statistik ved et regelbrud

Beskriv i mailen, hvornår data er sendt hjem, samt omfanget af hjemsendelsen. Hvis der er hjemsendt mikrodata vil adgangen til mikrodata fortsat blive lukket, og der skal sendes en redegørelse og plan for forebyggelse, men det betragtes som en formildende omstændighed, at institutionen selv rapporterer regelbruddet.

Orienteres der hurtigt om et regelbrud, betragtes det som en formildende omstændighed og kan afkorte den periode institutionen er lukket ned ved hjemsendelse af mikrodata

Som nævnet lukker vi som udgangspunkt ikke institutionen ned, hvis vi bliver orienteret om hjemsendelse af detaljerede tabeller, dvs. tabeller der ikke opfylder diskretionskriterierne.

Gode råd

Nedenfor følger en liste med gode råd til, hvordan man undgår brud på reglerne om datafortrolighed og datasikkerhed:

For brugeren, der har adgang til mikrodata:

1. Kontroller altid samtlige filer, der skal hjemsendes
 - a. Kontakt altid Forskningsservice i tvivlstilfælde
2. Begræns hjemsendelsen af filer til det mest nødvendige.
 - a. Er det fx nødvendigt at hjemsende programmer og logfiler?
– de opbevares sikkert under forskerordningen
 - b. Er det nødvendigt at hjemsende foreløbige resultater?
3. Få gerne kontrolleret indholdet af filer af en kollega der også har adgang til projektet, før du sender filer hjem

For den ansvarlige for autorisationen:

1. Lav en plan der sikrer, at alle brugere er orienteret om reglerne vedrørende datafortrolighed og datasikkerhed under forskerordningen
 - a. Herunder også at både den sidst tilkomne og den udenlandske bruger har tilstrækkelig viden om reglerne!
 - b. At der i miljøet findes en kontaktperson ved tvivlstilfælde
2. At alle autoriserede brugere, du har ansvar for og har godkendt tilknytningsaftaler for, kender proceduren ved et regelbrud

3. At alle autoriserede brugere er bekendt med de "fælder", der kan være med at arbejde med mikrodata – fx at logfiler kan indeholde mikrodata
4. Lav evt. kontrolforanstaltninger ved hjemsendelse af resultater
5. Begræns evt. antallet af brugere pr. projekt, der må sende output hjem
6. Jævnligt at kontrollere at kun de relevante brugere har adgang til mikrodata til institutionens projekter

Eksempler på overtrædelser

Nedenfor findes nogle generelle eksempler på forskellige typer af databrud, samt hvordan de kan undgås. Eksemplerne baserer sig hovedsageligt på de typer databrud vi har set siden 2014, hvor vi strammede vores sanktionspolitik, og eksemplerne vil blive udvidet, hvis vi ser nye typer af databrud.

| Type af databrud | Beskrivelse | Hvordan undgås det? |
|---|--|--|
| Afidentificerede PNR sendt på mail til en FSE medarbejder i forbindelse med fejlretning | Bruger sender en liste med afidentificerede personnumre (variablen PNR) på mail til FSE kontaktperson for at lette processen i forbindelse med fejlretning. | Mikrodata til fejlretning skal altid gemmes på et drev på forskermaskinen. Skriv til FSE-kontaktpersonen, hvor på forskermaskinen filen ligger. Herfra kan FSE-medarbejderen se filen. |
| Hjemsendelse af programmer der indeholder mikrodata | En bruger hjemsender et stort program fra forskermaskinen for at være sikker på, at det ikke går tabt. Hjemsendelsessystemet kommer med en advarsel, men brugeren ignorerer den og er overbevist om, at det er en falsk positiv advarsel. Det viser sig i vores kontrol af filen, at programmet indeholder mikrodata, da der i programmet er kodet personnumre på outliere, der skal fjernes fra datasættet. | Der er som udgangspunkt ikke grund til at sende programmer hjem. De ligger sikkert på forskermaskinen. De kan også af en FSE-medarbejder blive kopieret fra et projekt til et andet, hvis programmet skal bruges på et andet projekt. Det er OK at hjemsende programmer fra forskermaskinen, hvis man er sikker på, at de ikke indeholder mikrodata. Hvis man får en advarsel ved hjemsendelse af et program, anbefales det at få en anden medarbejder, der har adgang til projektet, til at se programmet igennem for mikrodata. |

| Type af databrud | Beskrivelse | Hvordan undgås det? |
|---|--|---|
| Hjemsendelse af logfil med mikrodata | En bruger skal se kørslerne af sit program grundigt igennem og sender derfor programmets logfil hjem. Der kommer en advarsel, som brugeren vælger at tolke som en falsk positiv advarsel, da det jo kun er en logfil, der er sendt hjem. | Der er flere eksempler på, at logfiler fra forskellige programmer kan indeholde mikrodata – fx at logfilen automatisk printer en lille stikprøve af filen og medtager denne i loggen. Det er tilladt at hjemsende Logfiler, men kontroller dem grundigt før hjemsendelse. |
| Udlån af password | En bruger har glemt sin mobiltelefon og har i øvrigt en meget stram deadline. Han låner derfor loginoplysninger og mobilnummer fra en medarbejder, der også har adgang til projektet. | Uanset hvor travlt man har, så er det aldrig tilladt at låne andres adgang til forskermaskinen. Det password man får er strengt personligt og må ikke overdrages til andre. |
| Hjemtagelse af egne indsendte mikrodata | En bruger har indsendt sit eget datasæt med mikrodata, som skal kombineres de øvrige data på projektet. Da det er brugerens egne data, tænker brugeren, at det er OK, at de bliver sendt hjem igen via hjemsendelsessystemet. | Der er ingen mikrodata, der må hjemsendes fra forskerordningen. Det gælder både data fra Danmarks Statistik, og de data brugeren selv indsender. |

| Type af databrud | Beskrivelse | Hvordan undgås det? |
|---|--|---|
| Kopiere resultater fra skærmen | En bruger har dannet en tabel, som hurtigt skal med til et møde om projektet. Brugeren skriver resultaterne ned fra skærmen og tager dem med til mødet. | Det er ikke tilladt på nogen måde at kopiere resultater eller andet fra de skærme, der har adgang til forskermaskinerne. Alle de resultater, som man skal bruge fra forskermaskinen, skal hjemsendes via hjemsendelsessystemet. |
| Tage billeder med mobiltelefon af skærmen | En bruger har nogle spørgsmål til sin vejleder og beslutter sig for at tage et billede af skærmen med mikrodata og sende billedet i en sms til vejlederen for bedre at kunne forklare problemet. | Det er ikke tilladt at tage billeder af skærmen med sin mobiltelefon, når man er koblet til forskermaskinen. Gem i stedet filen på forskermaskinen og skriv til vejlederen hvor den ligger. Så kan vejlederen selv gå ind på forskermaskinen og se filen. Det kræver, at vejlederen har adgang til projektet. |
| Skærmdeling | To brugere sidder ikke på samme sted og vil dele en skærm, så de begge kan se på et datasæt. | Det er ikke tilladt at skærndele på afstand, når man har adgang til forskermaskinerne. Man må kun dele en skærm, hvis man sidder i det samme lokale, og begge har adgang til projektets data. |

Bilag 1 – Brugerftale

Brugerftale for *[Navn på bruger]*

er indgået om etablering af direkte elektronisk adgang til udvalgte datasæt under mikrodataordningerne:

Der gælder følgende bestemmelser:

1. De datasæt, der gives adgang til, er fortrolige i henhold til Forvaltningslovens §27, stk. 3 og Straffelovens §152.
2. Arbejde med mikrodata under mikrodataordninger hos Danmarks Statistik, må kun ske i overensstemmelse med Danmarks Statistiks regler herfor. Dette kan enten ske ved opkobling fra den institution, som er tildelt autorisation, eller via en hjemmearbejdsplads i overensstemmelse med retningslinjer fastsat af Danmarks Statistik.
3. En pc, der er opkoblet til Danmarks Statistik, må ikke overlades til andre personer, og forbindelsen skal, når pc'en forlades, enten være lukket helt ned eller "disconnected", dvs. beskyttet mod uautoriseret brug.
4. De af Danmarks Statistik udleverede passwords er personlige og må ikke udlånes eller meddeles til andre.
5. Data, såvel som afledte datasæt, der indeholder mikrodata, må ikke hjemtages hverken direkte eller indirekte.
6. Alle hjemsendelser af resultater fra projekter under mikrodataordninger (tabeller, analyseresultater mv.) må kun ske i overensstemmelse med de af Danmarks Statistik fastsatte retningslinjer og metoder. Danmarks Statistik foretager en logning af disse overførsler.
7. Der må ikke hjemsendes fortrolige data, herunder data på individ- eller virksomhedsniveau - mikrodata og alt output skal være aggregeret i en sådan grad, at der ikke er fare for direkte eller indirekte identifikation af personer eller virksomheder. Der må ikke foretages forsøg på en sådan identifikation.
8. Publiceringer fra projektet må ikke indeholde oplysninger, der kan identificere enkeltpersoner eller enkeltvirksomheder.
9. Danmarks Statistik kan få tilsendt publiceringer fra projektet til orientering
10. Ved udlevering af fysisk "token" skal afleveres til Danmarks Statistik ved aftalens ophør.

Særligt for Projektdatabasemanagere:

11. Projektdatabasemanageren varetager kontakten med Forskningservice i forbindelse med institutionens projekter.
12. Projektdatabasemanageren har adgang til projektdatabasen, men ikke til institutionens øvrige projekter. Undtaget er projekter, hvor projektdatabasemanageren ligeledes er bruger på et projekt.
13. Der skal indsendes en indstilling til godkendelse i forbindelse med nye projekter på basis af projektdatabasen.
14. Det er projektdatabasemanagerens ansvar at danne udtræk af databasen til nye og eksisterende projekter.
15. Til det enkelte projekt må der kun dannes udtræk med indhold, der afspejler indgået aftale med Forskningservice.
16. Projektdatabasen kan ikke indeholde data fra Lægemedeldatabasen med mindre, der foreligger en særlig godkendelse fra Lægemedelstyrelsen.
17. Der må ikke udføres analyser på projektdatabasen, og der må ikke hjemtages eller på nogen anden måde videregives data fra projektdatabasen.
18. Projektdatabasen og øvrigt projektdata skal være adskilt ved forskellige projektnumre, og der må ikke gøres forsøg på at samkøre data fra forskellige projekter.
19. Forskningservice overfører data fra projektdatabasen til institutionens projekter.
20. Projektdatabasemanageren sikrer, at alle brugeres adgang til et projekt, som indeholder projektdatabase efter gammel ordning er ophørt senest seks (6) måneder efter aftalens indgåelse.

21. Forskningsservice skal orienteres, hvis projektdatabasemanageren fratræder.

Tilladelse til at dele oplysninger i brugeradministrationssystemet FSE-BOA

22. Jeg tillader, at nedenstående oplysninger, som jeg har afgivet, deles med andre brugere i FSE-BOA.

- a. For medarbejdere i Danmarks Statistik tillader jeg at dele nedenstående oplysninger

- Navn
- Mail
- Kontakttelefon (ej obligatorisk at udfylde)
- Mobilnummer til SMS token (password)
- Projektadgange
- Tilknytningsaftaler

- b. For brugere, der er tilknyttet samme autoriserede institution, tillader jeg at dele nedenstående oplysninger:

- Navn
- Mail
- Kontakttelefon (ej obligatorisk at udfylde)
- Projektadgange
- Tilknytningsaftaler

- c. Jeg tillader at dele nedenstående oplysninger med andre autoriserede institutioners administratorer, institutionsansvarlige, samt disses stedfortrædere:

- Navn
- Mail
- Kontakttelefon (ej obligatorisk at udfylde)

Oplysningerne er nødvendige for, at en bruger med en tilknytningsaftale kan blive tilknyttet flere autoriserede institutioner.

Bilag 2 – Tilknytningsaftale

Tilknytningsaftale mellem [indsæt navn på brugeren] og [indsæt navn på institut]

1. Den ansvarlige person, som har underskrevet autorisationsaftalen på den autoriserede danske institution, godkender og tager ansvaret for, at den tilknyttede bruger **[indsæt navn på brugeren]** overholder alle gældende regler for omgang med mikrodata
2. Det pålægges den danske autoriserede institution at orientere den tilknyttede bruger om reglerne for brug af mikrodata, herunder de gældende diskretionsregler, samt reglerne vedrørende hjemtagelse af data
3. Den tilknyttede brugers adgang til mikrodata fra udlandet skal gå gennem den danske autoriserede institution og viderestilles i overensstemmelse med reglerne for hjemmearbejdsplads

Særligt vedrørende udenlandske brugere

4. Den autoriserede danske institution udpeger en kontaktperson, der varetager al kontakt mellem den tilknyttede bruger og Danmarks Statistik
5. Alle faktureringer vedrørende den tilknyttede bruger sendes til og betales af den pågældende autoriserede danske institution på de faktureringsvilkår, som i øvrigt gælder for institutionen

Særligt vedrørende virksomhedsdata

6. Personer ansat ved konsulentfirmaer må ikke få adgang til virksomhedsdata. Der kan under særlige omstændigheder søges dispensation, såfremt projektet ejes af en offentlig myndighed eller en interesseorganisation. Kontakt Forskningsservice vedrørende dispensation.

Personer ansat ved andre enkeltvirksomheder må ikke få adgang til virksomhedsdata.

En overtrædelse af reglerne i denne aftale vil medføre, at den pågældende institutions brugere udelukkes fra at anvende nogen af Danmarks Statistiks mikrodataordninger i en begrænset periode, der fastsættes af Danmarks Statistik - eller permanent. Reglerne for udelukkelse af mikrodataordningerne fremgår af Danmarks Statistiks hjemmeside under linket:

<https://www.dst.dk/da/TilSalg/Forskningsservice/hjemtagelse-af-analyseresultater> - Retningslinjer for hjemsendelse af analyseresultater