

Vejledning til hjemsendelse af filer

Under Danmark Statistiks forskerordning er det tilladt at hjemsende analyseresultater fra forskermaskinen. Denne vejledning beskriver, hvordan I kan anvende vores nye hjemsendelsessystem, som kan hjælpe jer med at identificere filer, der muligvis indeholder mikrodata.

Scanning af hjemsendte filer

Hjemsendelsessystemet skanner alle tilladte filtyper *inden* de hjemsendes. Skanningen tester, hvorvidt filerne indeholder mulige mikrodata.

Bemærk venligst at skanningssystemet kun skanner for mikrodata, der ligger i Forskningsservices grunddata, samt at vi har udeladt filtyper og variabler, der genereret for mange falsk positive advarsler.

Det betyder fx at korte mønstre er udeladt ud fra en forventning om, at de vil give for mange falske positive, fx EJENDOMSNUMMER (som har længde 8)

Forskningservice modtager mange eksterne data, der ligeledes indeholder nøglevariabler. De afidentificerede variabler i eksterne datasæt matcher *ikke* nødvendigvis mønstre i Forskningsservices grunddata.

Skanningssystemet skal derfor kun betragtes som en hjælp og kan ikke erstatte en manuel kontrol af filerne før de hjemsendes.

Nedenfor en oversigt over de mønstre som systemet skanner for og som giver en advarsel, hvis de genkendes af systemet.

| Mønster_id | Variabel | Længde |
|------------|---|--------|
| 1 | Valide PNR-variable Flere ID-variable (familie_id, sags_id, ydernr, OK_NR mv.) | 12 |
| 2 | Valide CVR-variable, SE-variable, JURNR | 10 |
| 3 | PNR-variable med bogstaver | 10 |
| 4 | OPGIKOM-variable | 16 |
| 5 | BOPIKOM-variable | 22 |
| 6 | RECNUM-variable og lign. | 20 |
| 7 | ADRKOD og lign. | 34 |
| 8 | SVPNR og lign. | 14 |
| 9 | MAAL_KEY og lign. | 24 |

Systemet søger/scanner **ikke** diskretioneringsregler. Dvs. systemet identificerer ikke filer, hvis tabel indeholder fx "antal observationer" i hver celle, som er mindre end 3.

Skanning kan ikke erstatte manuel kontrol Skanningen af filer er ment som en hjælp til at understøtte datasikkerheden. Det er vores håb, at det nye hjemsendelsessystem vil minimere antallet af forskere som uoverlagt komme til at hjemtage mikrodata.

Automatisk kontrol kan dog aldrig erstatte jeres manuelle kontrol af de filer, I hjemsender. I skal fortsat selv foretage den manuelle kontrol inden I beslutter jer for at hjemsende en fil.

Systemer reagerer på en mistanke Det er vigtigt at understrege, at systemet reagerer på en mistanke om at en fil indeholder mikrodata. Der kan derfor godt komme en advarsel om en fil, der ved nærmere eftersyn alligevel ikke indeholder mikrodata. Der vil med andre ord være såkaldte "falsk positive" alarmer. Vær venligst opmærksom på, at der fx er forholdsvis mange HTML-filer, der kan give falske alarmer. For at undgå for mange falske alarmer foreslår vi, at I ikke anvender dette format til hjemsendelse af filer.

Mulighed for at se scanningslog Vi har udvidet hjemsendelsessystemet således at du har mulighed for at se scanningslog, når systemet har givet en advarsel. Scanningslog vises og kan gemmes som text-fil og dermed konverteres til andre fil-formater til nærmere eftersyn.

Scanningslog vises med kolonner MoensterId;Moenstertekst;Vaerdi.

- MoensterId og Moenstertekst svarer til Mønster_id og Variabel i den ovennævnte tabel på side 1.
- Vaerdi angiver hvilken/hvilke værdi(er) systemet identificerer som mikrodata

Husk at scanningslog IKKE MÅ HJEMSENDES!

Libre-office format kan scannes Systemet er også blevet udvidet således at Libre-office format (.ods, .odt, .odp) kan scannes og hjemsendes.

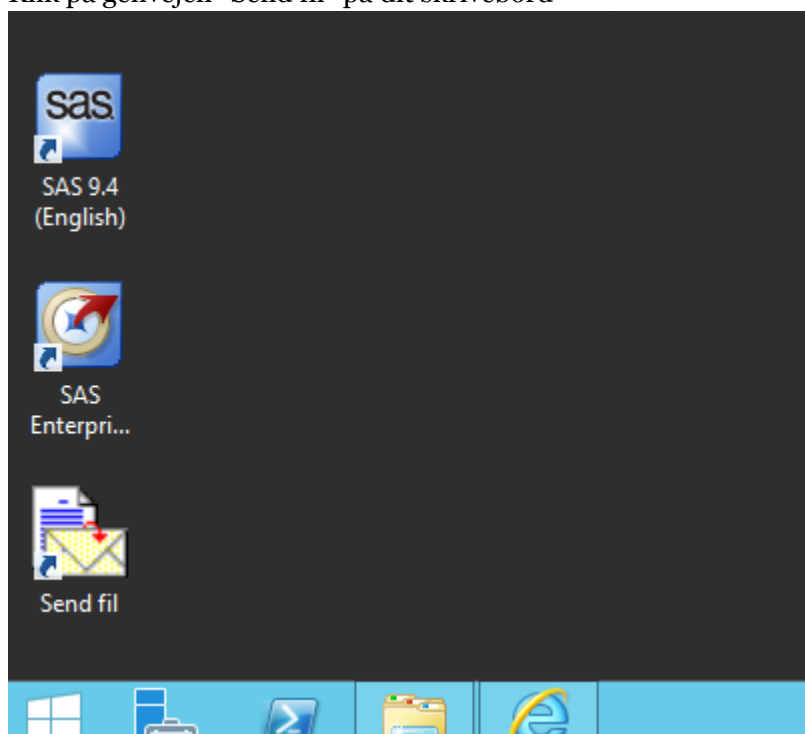
Kend reglerne for hjemsendelse Den automatiske skanning er udelukkende en hjælp til jer. Det er fortsat jeres eget ansvar at sikre, at de hjemsendte filer overholder reglerne for hjemsendelse og at filer er kontrolleret før hjemsendelse. Reglerne er beskrevet i vejledningen: Retningslinjer for hjemsendelse af analyseresultater, der findes under linket: <https://www.dst.dk/da/TilSalg/Forskningsservice/hjemtagelse-af-analyseresultater>

Stikprøvekontrol Forskningsservice foretager løbende stikprøve kontrol af de filer som hjemsendes for at tjekke at reglerne bliver overholdt. Vi beholder en kopi af alle hjemsendte filer i seks måneder og logger alle hjemsendelser inkl. afviste forsøg.

Nedenfor følger en kort brugervejledning i at sende analyseresultater hjem fra forskermaskinen med det nye system.

Send filer hjem

Start program Klik på genvejen "Send fil" på dit skrivebord



Herefter starter en applikation i webbrowseren.

FSEHjemsend

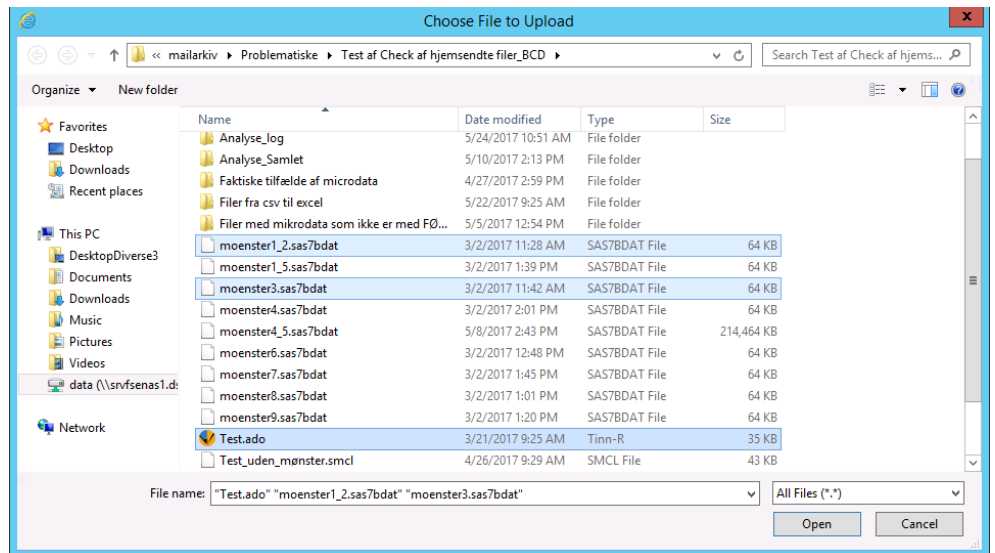
Vælg filer til hjemtagelse via e-mail

Filer Browse...

Fortsæt Nulstil

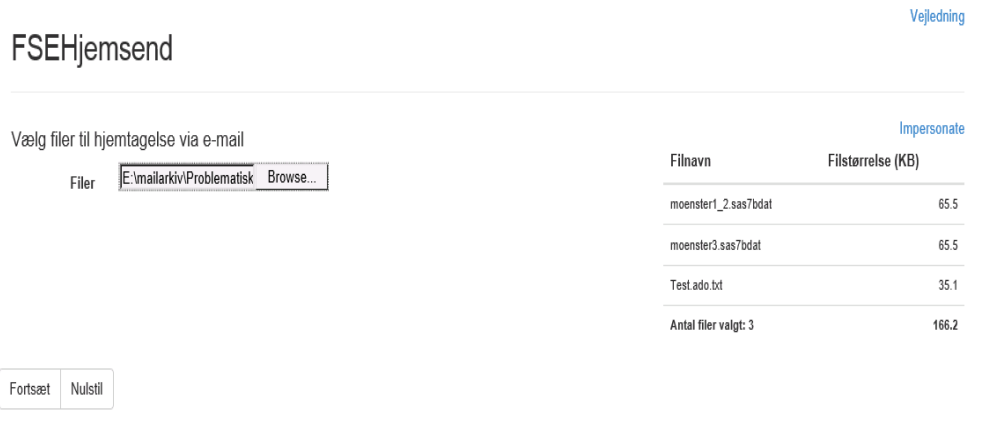
Vælg filer Vælg herefter **Browse...** En stifinder åbner og her får du mulighed for at vælge den eller de filer du gerne vil hjemsende.

Vi anbefaler at I danner et særligt bibliotek under Workdata, fx hjemsend, der kun indeholde de filer der skal kontrolleres og hjemsendes, for at minimere risikoen for at vælge en fil, der indeholder mikrodata.

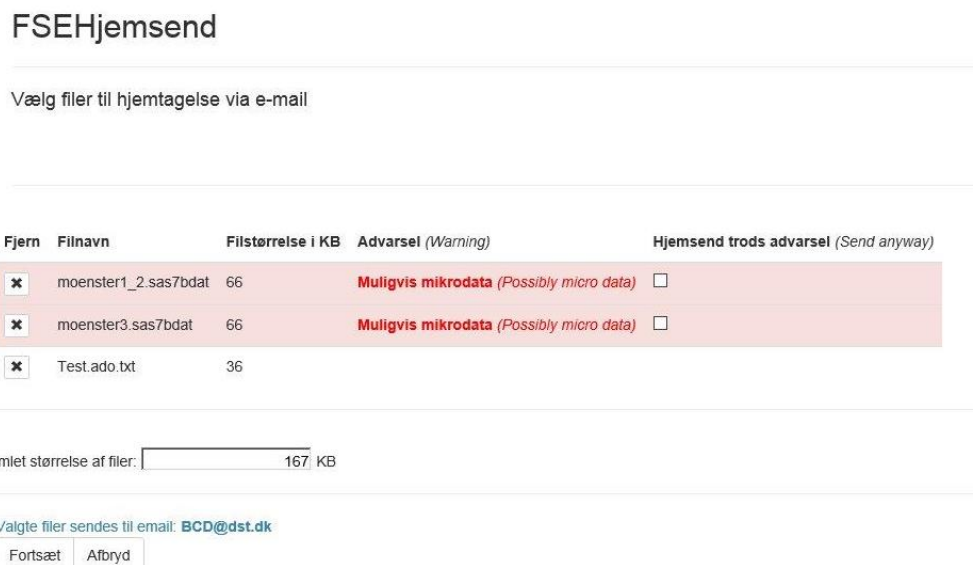


Klik på **Open** når du har markeret de filer du ønsker at hjemsende. Man kan vælge flere filer ved at holde CTRL nede.

Scan filerne Nu er filerne tilføjet og klar til at blive scannet. Klik på **Fortsæt** for at scanne filerne, jf. billedet nedenfor.



Herefter fremkommer nedenstående billede:



Din e-mail vil automatisk blive vist på billedet. Kontakt evt. forskningservice@dst.dk, hvis din e-mail skal rettes.

Fjerne filer fra listen

Hvis du ønsker at fjerne filer til hjemsendelse kan filerne fjernes via X-ikon under ”**Fjern**”. Filerne forsvinder herefter fra oversigten. Fjernes krydserne ud for de to filer med advarsler fås nedenstående billede:

FSEHjemsend

Vælg filer til hjemtagelse via e-mail

| Fjern | Filnavn | Filstørrelse i KB | Advarsel (Warning) | Hjemsend trods advarsel (Send anyway) |
|--------------------------|--------------|-------------------|--------------------|---------------------------------------|
| <input type="checkbox"/> | Test.ado.txt | 36 | | |

Samlet størrelse af filer: KB

Valgte filer sendes til email: forskerpost@dst.dk

Send filerne

Hjemsend filen ved at klikke på **Fortsæt**. Du skulle gerne modtage filen kort efter.

Hvis du vil starte forfra, så tryk på **Afbryd**.

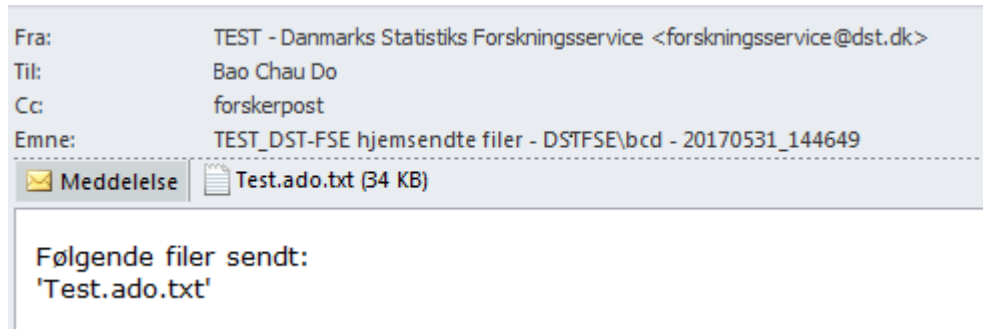
Sender du filen hjem får du nedenstående besked på skærmen:

FSEHjemsend

Sendt til: BCD@dst.dk

Følgende filer sendt:
'Test.ado.txt'

Eksempel på mail Efter man har hjemsendt filer, vil man få følgende mail



Advarsel ved identificerbare data

Hvis scanningsystemet mistænker at der er identificerbare data i filerne, vil du få en advarsel. Her har du mulighed for at tjekke filerne igen og fortsætte, hvis du er sikker på at der ikke er mikrodata i filerne eller afbryde hjemsendelsen.

FSEHjemsend

Vælg filer til hjemtagelse via e-mail

| Fjern | Filnavn | Filstørrelse i KB | Advarsel (Warning) | Hjemsend trods advarsel (Send anyway) |
|-------------------------------------|----------------------|-------------------|--|---------------------------------------|
| <input checked="" type="checkbox"/> | moenster1_2.sas7bdat | 66 | Muligvis mikrodata (Possibly micro data) | <input type="checkbox"/> |
| <input checked="" type="checkbox"/> | moenster3.sas7bdat | 66 | Muligvis mikrodata (Possibly micro data) | <input type="checkbox"/> |

Samlet størrelse af filer: 132 KB

Valgte filer sendes til email: forskerpost@dst.dk

Fortsæt Afbryd

Scanningslog til eftersyn Når systemet giver en advarsel, har du mulighed for at se scanningslog, som viser hvilke mønstre og hvilke værdier systemet tror, at det er tale om mikrodata, ved at trykke på "Vis scanningslog".

FSEHjemsend

Vælg filer til hjemtagelse via e-mail

| Fjern | Filnavn | Filstørrelse i KB | Advarsel (Warning) | Hjemsend trods advarsel (Send anyway) | Vis scanningslog (Scanningslog må IKKE hjemsendes) |
|-------------------------------------|----------------------------|-------------------|--|---------------------------------------|--|
| <input checked="" type="checkbox"/> | test med PNR og CVRRR.xlsx | 10 | Muligvis mikrodata (Possibly micro data) | <input type="checkbox"/> | Vis scanningslog |

Samlet størrelse af filer: KBValgte filer sendes til email: BCD@dst.dk

Scanningslog kan åbnes og gemmes som text-fil til nærmere eftersyn.

Husk at scanningslog IKKE MÅ HJEMSENDES!

Skal filerne ikke sendes hjem klikkes på **Afbryd**.

Hvis du er helt sikker på at en fil, på trods af advarslen og eftersynet af scanningslog, IKKE indeholder mikrodata, så kan den hjemsendes ved at klikke på boksen under **Hjemsend trods advarsel** og herefter på **Fortsæt**.

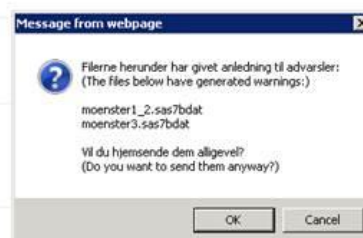
Herefter fremkommer endnu en meddelelsesboks hvor man skal klikke **OK**, jf. billedet nedenfor.

FSEHjemsend

Vælg filer til hjemtagelse via e-mail

| Fjern | Filnavn | Filstørrelse i KB | Advarsel (Warning) | Hjemsend trods advarsel (Send anyway) |
|-------------------------------------|----------------------|-------------------|--|---------------------------------------|
| <input checked="" type="checkbox"/> | moenster1_2.sas7bdat | 66 | Muligvis mikrodata (Possibly micro data) | <input checked="" type="checkbox"/> |
| <input checked="" type="checkbox"/> | moenster3.sas7bdat | 66 | Muligvis mikrodata (Possibly micro data) | <input checked="" type="checkbox"/> |
| <input type="checkbox"/> | Test.ado.txt | 36 | | |

Samlet størrelse af filer: KBValgte filer sendes til email: BCD@dst.dk



Begrænsninger

| | |
|---------------------------|--|
| Tekstbaserede filetyper | <p>Tekstbaserede filer i følgende kategorier bliver skannet i begrænset omfang. ”txt;htm;html;mht;do;doh;ado;tab;tex;log;lst;spv;xml;r;sas;sps;xlm;smcl”</p> <p>Fx hvis en fil i de nævnte filetyper indeholder en liste af numeriske værdier (hel-tal), der er adskilt af komma eller punktum, kan systemet ikke fange værdier, som kunne være mikrodata, og I bliver derfor ikke advaret herom. Derfor skal I som nævnt også kontrollere de filer som I ønsker at hjemsende manuelt, inden I beslutter jer for at hjemsende filerne.</p> |
| Diskrektioneringsregler | <p>Systemet søger/scanner ikke diskrektioneringsregler. Dvs. systemet identificerer ikke filer, hvis tabel indeholder fx ”antal observationer” i hver celle, som er mindre end 3.</p> |
| Max 3 MB pr. fil | <p>Det er tilladt at hjemsende filer, der er mindre end 3.000 KB. Sender du en fil som er større end det tilladte, får du følgende besked:</p> <p>Filer til overførsel må ikke overstige 3.000 KB. Den samlede filstørrelse må ikke overstige 30.000 KB.</p> |
| Max 30 MB i alt | <p>Den samlede mængde filer i hver mail kan maksimalt have en størrelse på 30.000 KB</p> <p>Har du behov for at sende mange store filer hjem, skal du derfor sende flere mails. Det er <u>ikke</u> længere tilladt at hjemsende komprimerede filer.</p> |
| Kun godkendte ekstensions | <p>Det er tilladt at hjemsende filtyper, der er godkendt af Danmarks Statistiks Forskningservice.</p> |
| Tilladte filtyper | <p>En liste over filtyper, det er tilladt at hjemsende, kan findes i ”Hjemsendelse af filer - tilladte filtyper” under https://www.dst.dk/da/TilSalg/Forskningservice/hjemtagelse-af-analysere-sultater</p> <p>Kun filer med ekstensions, der er specificerede, hjemsendes. Øvrige er ikke tilladte og hjemsendes ikke.</p> <p>Filer, der er tilladte at hjemsende, men hvor scanningen af filen fejler af en eller anden grund, bliver hjemsendt. Man får advarslen: ”Kunne ikke scannes”</p> |
| Ikke-tilladte filetyper | <p>Følgende filekstensions er ikke tilladte at hjemsendes og som ikke kan scannes: csv, dat, db, gph, msg, raw, sas7bcat, spo, ster og xps, rda og rdata</p> <p>Sender du en filekstension, som ikke er tilladt, får du følgende besked: ”Filetype ikke tilladt”</p> |
| Grafikfiler | <p>Grafikfiler scannes ikke men må gerne hjemsendes, hvis de har den specificerede ekstension. Man får følgende besked: ”Grafik, scannes ikke”</p> <p>Det er ikke tilladt at ændre på fil-ekstension i forhold til indholdet af filen.</p> |

Hvis man ønsker at få en ny filtype tilføjet til denne liste, skal man henvende sig til Jørn Korsbø Petersen JKP@dst.dk

Hjælp

Ved yderligere spørgsmål til hjemsendelse af filer er du velkommen til at kontakte os på tlf. 3917 3140 eller på mail forskningservice@dst.dk