

## Instruktion till dig som ska skicka datafiler till Socialstyrelsen

1. Kontrollera att din handläggare har lagt in dig i Sofi
2. Dubbelkolla att filformaten accepteras
3. Skapa en metadatafil / beskriv datamaterialet som ska inskickas
  - Fem eller mer filer? → metadatafil behövs
  - Datamaterialet 1GB eller större? → metadatafil behövs
4. Komprimera och kryptera datamaterialet med lösenord
5. Lösenordet skickar du via e-post till [registerservice@socialstyrelsen.se](mailto:registerservice@socialstyrelsen.se).
  - Glöm inte att ange ärendenummer i e-postmeddelandets ämnesrad!

### Vad är Sofi?

Sofi är Socialstyrelsens filuppladdningstjänst. Du behöver tillgång till Sofi och din handläggare ansvarar för att lägga in ditt namn, din e-post och ditt mobilnummer i systemet. Efter att du fått bekräftat att du är inlagd i systemet kan du logga in och ladda upp filer i de ärenden som du har behörighet att ladda upp data till.

### Vilka filformat kan skickas in?

För att kunna garantera en korrekt inläsning av datamaterialet måste datafilerna inkomma i något av följande format:

- SAS-format (.sas7bdat)
- Textfil (.txt)
- csv-fil *character-separated values* (.csv)

#### Filnamn ska **inte**

- börja på en siffra
- innehålla mellanslag
- innehålla bindestreck
- vara längre än 28 tecken

Vi accepterar **inte** filer i Excelformat (.xls) eller andra format förutom metadatafilen.

### När behövs en metadatafil?

Metadatafilen behövs för .txt och .csv filer enligt nedan. Skickar du SAS-dataset behövs ej metadata.

- Om du skickar **fem eller fler datafiler** till Socialstyrelsen
- Om den totala storleken på datamaterialet är **1GB eller större** innan komprimering och kryptering

För att vi ska kunna läsa in de inskickade datafilerna korrekt och snabbt behöver vi få metadata om datafilerna. Vi kräver att metadata (hädanefter *metadatafil*) skickas in som en enda Excel fil enligt bifogad metadata-mall.

Vid frågor om metadatafilen kan du mejla [registerservice@socialstyrelsen.se](mailto:registerservice@socialstyrelsen.se). Ange i e-postmeddelandets ämnes rad "Fråga om metadata".

## Hur ska metadatafilen formateras?

En metadatafil skickas in som en enda Excel-fil enligt bifogad metadatamall och innehåller två flikar:

1. Information om varje fil i leveransen; och,
2. Information om varje kolumn i varje fil.

Följande information måste finnas i **första fliken** (OBS en rad per fil):

filename	namn på fil inklusive filändelse: <i>.sas7bdat</i> eller <i>.txt</i> eller <i>.csv</i>
encoding	teckenkod i datafilen, <i>UTF-8</i> eller <i>ISO-8859-1</i> (eller annan ISO-teckenkod) <ul style="list-style-type: none"> <li>• <i>UTF-8</i> stöder de flesta språk och inkluderar även Latin1.</li> <li>• <i>ISO-8859-1</i> (även kallad Latin1) kodar tecken ur det latinska alfabetet och stödjer västeuropeiska språk.</li> </ul>
separator	tecken som avgränsar kolumner. Kan vara <i>komma</i> , <i>semikolon</i> , <i>tabb</i> eller <i>pipe</i> . Du kan även specificera en annan symbol (såsom t.ex. <i>⌘</i> ) men tänk på att denna symbol inte får återfinnas i själva datamängden i strängvariabler
end_of_line	kan vara <i>CRLF</i> , <i>LF</i> eller <i>CR</i>
Nrows	antal rader med data i aktuella filen (exklusive raden med variabelnamnen, s.k. header)

Följande information måste finnas i **andra fliken** (OBS en rad per kolumn per fil):

filename	namn på fil inklusive filändelse: <i>.sas7bdat</i> eller <i>.txt</i> eller <i>.csv</i>
variable	Variabelnamn
position_in_file	i vilken kolumn finns variabeln i filen
Type	t.ex. numerisk, sträng, datum
Length	variabelns längd i antal bytes eller tecken

Ett exempel på hur mallen ska se ut när den är ifylld hittas i den bifogade filen *metadata\_mall\_exempel*.

Metadatafilen måste beskriva samtliga inskickade datafiler. Om den inte är ifyllt korrekt (d.v.s. att metadata inte stämmer överens med det inskickade datamaterialet, t.ex. om du anger fel *encoding* eller fel antal rader) kommer det bli fel vid inläsningen hos oss. Vi kommer då att kontakta dig och be dig komplettera med rätt metadata. Att skicka rätt metadata är viktigt för att vi ska kunna säkerställa oss att datafilerna blir korrekt behandlade av oss, och för att inläsningen ska gå snabbt.

## Att tänka på

### Personnummer

Personnummer måste anges med 12 siffror, i formatet ÅÅÅÅMMDDNNNN.

### Unika identifierare

*Unika identifierare* som är inte projektspecifika och som inte behövs för samkörningen ska tas bort från datafilerna. Unika identifierare kan till exempel vara namn, adresser, radnummer, primära/sekundära/främmande databasnycklar som kopplar ihop datafiler (exempelvis vårdtillfälle-id, tumör-id). Unika identifierare kännetecknas av att de avslöjar personens identitet, direkt eller indirekt. Om det finns i datamaterialet unika identifierare som måste ingå i leveransen till beställaren, behöver ni informera om dem i *beskrivning av datamaterialet*.

Unika identifierare som inkommer från andra registerhållare t.ex. kvalitetsregister Kvalitetsregister och andra registerhållare som inkommer med data till Socialstyrelsen måste pseudonymisera unika identifierare. Det betyder att unika identifierare ska vara projektspecifik. Variabler och unika identifierare som är inte relevant för projektet ska tas bort från data.

### *Innan data laddas upp via Socialstyrelsens filuppladdningstjänst Sofi*

Dubbelkolla om datafilerna är enligt instruktionerna och kontrollera det följande:

- Skickar jag 5 eller fler datafiler till Socialstyrelsen?
- Är den totala storleken på datamaterialet 1GB eller större innan komprimering och kryptering?

Om svaret är *ja* på en av frågorna och datamaterialet består av datafiler i andra format än SAS (*.sas7bdat*), behöver du fylla i den bifogade metadatamallen och följa ovanstående instruktioner om metadata. Ladda upp den ifyllda metadatafilen i SOFI tillsammans med övriga data.

### *Kryptera datamaterialet med lösenord*

Komprimera och kryptera datamaterialet som du vill skicka till Socialstyrelsen med lösenord. OBS! Kryptera endast mappen där de underliggande filerna finns med lösenord. Se till att inte kryptera enstaka filer, bara mappen som innehåller filerna. För kryptering kan du exempelvis använda winzip (filändelse *.zip*) eller 7-zip (filändelse *.7z*). För dig som använder Mac kan datamaterialet krypteras med program som iZip. Både 7-zip och iZip är gratis.

### *Ladda upp datamaterialet och beskrivningen via Sofi*

Innan du har tillgång till tjänsten så måste du kontakta din handläggare som då lägger in ditt namn, din e-post och ditt mobilnummer i systemet. Efter att du fått bekräftat att du är inlagd i systemet kan du logga in och ladda upp filer i de ärenden som du har behörighet att ladda upp data till.

### *Skicka lösenord via e-post*

Lösenordet skickar du via e-post till [registerservice@socialstyrelsen.se](mailto:registerservice@socialstyrelsen.se). Ange ärendenummer i e-postmeddelandets ämnesrad.

### **Frågor?**

Mejla till [registerservice@socialstyrelsen.se](mailto:registerservice@socialstyrelsen.se). Glöm inte att ändra ämnet!

Fråga om	Ämnesrad
Indata	Fråga om indata
Metadata	Fråga om metadata