

Manual för registerprogrammet Sveriges kyrkoböcker

Innehållsförteckning

MANUAL FÖR REGISTERPROGRAMMET <u>SVERIGES KYRKOBÖCKER</u>	1
INNEHÅLLSFÖRTECKNING	2
INFORMATION OM PROGRAMPAKETET	3
SYSTEMKRAV	4
SUPPORT	4
BAKGRUND	4
ARBETSGÅNG	4
1. ANPASSNING AV DET BEFINTLIGA DATAMATERIALETS FÄLTSTRUKTUR	5
2. TVÄTTNING AV DATAMATERIALET	6
<i>Tvingande åtgärder</i>	7
<i>Rekommenderade åtgärder</i>	7
3. GENERERING AV INDEXFILER FÖR DATABASEN	10
<i>Introduktion</i>	10
<i>Egen text</i>	10
<i>Generering</i>	10
<i>Normering i efterhand</i>	11
<i>Normering i förhand</i>	11
<i>Hur man skapar en bildfil för logotypen</i>	12
<i>Ytterligare tvättning</i>	12
4. INSTALLATION AV SÖKPROGRAMMET	13
TEKNISKA SPECIFIKATIONER (BILAGA)	14
A. FÄLTFORMAT	14
<i>Födda (huvudtabell 1)</i>	14
<i>Vigda (huvudtabell 2)</i>	15
<i>Döda (huvudtabell 3)</i>	16
<i>Kommentarer</i>	17
B. VISNINGSFORMAT	18
C. SÖKNING	19
<i>Födda, sökindex</i>	20
<i>Vigda, sökindex</i>	20
<i>Döda, sökindex</i>	21
<i>Sökindex för allmän sökning (tvärs över alla posttyper):</i>	21
D. GENERERING OCH UTFILFORMAT	22
<i>Utformat</i>	22
<i>Informat</i>	23
<i>Normeringsstabeller</i>	23
<i>Statistikfiler</i>	24
E. VISNING OCH INTERAKTIVITET	25
<i>Huvudskärm</i>	25
<i>Sökskärmar</i>	25
<i>Utskrift</i>	26

Information om programpaketet

Originalidé: *Sam Blixt, Carl Szabad*
Projektansvarig: *Sven Johansson*
Programmerare: *Johan Gidlöf*

Sveriges Släktforskarförbund
Allén 7
172 66 Sundbyberg

Tel: 08-440 75 50
Fax: 08-695 08 24
e-mail: info@genealogi.se
webb: www.genealogi.se/forbund/presprog.htm

Manualens version 2007-01-28.

© Sveriges Släktforskarförbund 2007
ISBN 91-87676-41-9

Systemkrav

Windows 98 eller senare version, rekommendation minst 256 MB RAM, 10 MB hårddiskutrymme (+ datamängden, ca 5.000 poster per MB), 800x600 pixlar skärmupplösning med 16-bitars färg.

Support

Förbundet har utbildat ett antal faddrar som kan hjälpa till med frågor som uppstår under framtagningen av det färdiga registret. Uppgifter om vilka dessa är framgår i Nättidningen RÖTTER på sidan www.genealogi.se/forbund/presprog.htm.

Bakgrund

Behovet av ett särskilt program för presentation av kyrkboksregister har funnits länge. Många föreningar har gjort mer eller mindre lyckade egna konstruktioner, men någon rikstäckande standard har inte funnits tidigare.

2001 framlades ett förslag om att släktforskarförbundet skulle ta fram ett sådant program och en fältstruktur för kyrkboksregister som skulle kunna användas.

Något konkret programmeringsarbete blev dock inte igångsatt innan *Borås Släktforskarförening* lade en motion till riksstämman 2003 med innehållet att förbundet skulle ta fram det föreslagna programmet till riksstämman 2004. Den sommaren förelåg således en första version av programmet.

Därefter har olika betatesters och justeringar av funktionerna vidtagit. Till förtroendemannakonferensen i maj 2005 i Visby kunde således programmet presenteras helt färdigt att använda för alla föreningar som bedriver registrering av ministerialböcker, dvs födelse-, vigsel- och dödböcker.

Arbetsgång

Arbetet med produktionen av en register-CD består av fyra faser, 1) anpassning av fältstrukturen i det befintliga materialet till den som presentationsprogrammet behöver, 2) tvättning av det befintliga datamaterialet, 3) generering av indexfiler via ett särskilt genereringsprogram och 4) skapande av själva CDn. Via nedanstående beskrivning kan du följa arbetsgången i detalj.

1. Anpassning av det befintliga datamaterialets fältstruktur

Registrering av kyrkböcker har pågått i decennier och det finns många idéer om hur en sådan ska gå till. I det program som släktforskarförbundet tagit fram finns ett antal obligatoriska fält och ytterligare ett antal som är anpassade till programmet, men ej obligatoriska, se sidorna 14-16. Därutöver kan man själv bygga på med hur många valfria fält som helst. Dessa presenterar då sitt innehåll, där det finns något, i Övrigt-fältet med den rubrik man själv valt.

Många har tidigare valt att lägga både för- och efternamn i samma fält. Detta går att bryta upp med hjälp av exempelvis Excel, så att de båda olika namntyperna hamnar i sina respektive fält, för att kunna indexeras separat.

Välj hellre fler antal fält i början av tvättarbetet, det är lätt att i ett senare skede slå ihop dem om det blir mer överskådligt.

Tänk på att skilja på anmärkningar som står i källmaterialet och registrerarens anmärkningar. De ska stå i olika fält, om det över huvud taget behövs registreringsanmärkningar. Det finns ett särskilt fält för markering om posten behöver kollas extra, *Status*. Använd gärna det i stället.

Det är viktigt med fält för genus när man ska normera förnamn. Finns inte det redan i materialet, försök lägga in det, om det inte blir ett orimligt arbete.

Har titel och civilstånd registrerats i olika fält bör dessa slås ihop till fältet Titel.

Tänk på att fältet för ålder kan vara angivet på många sätt. Det fältet är därför ett rent textfält och blir inte indexerat för sökning.

Sidnummer för källan ska stå i det fält som är avsett, liksom källan ska anges som den volym materialet är hämtat ur. Om inte den är inlagd tidigare går det bra att komplettera i efterhand i Excel. Volymbeteckning och tidsomfattning framgår exempelvis på NAD-CDn, alternativt i NAD på nätet.

När fältstrukturen väl är definierad är det dags att gå till nästa fas, tvättarbetet.

2. Tvättning av datamaterialet

Kvaliten på ett register är i första hand beroende av hur noggrant arbetet är genomfört vid inmatningen. Fel som begåtts där kan du i viss mån rätta i den här arbetsfasen, dock naturligtvis inte felläsningar. Det måste till en korrekturläsning för att rätta den typen av fel.

För att den här processen ska bli så effektiv som möjligt beskrivs den steg för steg. Vissa åtgärder är tvingande medan andra bara underlättar sökning eller ger ett snyggare resultat i presentationen.

Manualen för tvättarbetet bygger på att Microsoft Excel används, men även andra liknande program som Access eller Works kan användas. Upp till 50.000 poster kan hanteras åt gången i Excel (teoretiskt 65.536, men då fungerar inte programmet bra längre), så om registret innehåller fler poster skapas flera Excel-ark som sedan i slutändan läggs ihop i ett annat program.

Om de olika registren ligger i ett antal mindre Excelfiler rekommenderas att de först läggs ihop till en fil, då tvätten vanligtvis inte tar längre tid för en stor än för en liten fil. Tänk då bara först på att fältstrukturen innan sammanslagningen måste vara identisk i filerna!

Först skapas rubriker för varje kolumn genom att en rad läggs till överst i registret, med de namn för fälten som framgår under den tekniska specifikationen för informat på sidorna 14-16, beroende på om tvätten avser ett födelse-, vigsel- eller dödboksregister. Ordningen av fälten spelar ingen roll, bara samtliga obligatoriska fält i beskrivningen finns med. Då undviker man kommentar om detta i den felrapport som sedermera kommer att skapas av genereringsprogrammet. Läs rubrikraden med *Läs fönsterrutor* under *Fönster*.

Om det finns fält som inte förekommer i fältförteckningen på sidorna 14-16 läggs en rubrik in som passar det fältet på samma översta rad som de tidigare inskrivna rubrikerna.

I ett födelseregister kan faddrar läggas in och blir då sökbara via Övrigt-fältet.

På samma sätt kan eventuella föräldrar till brud och brudgum göras sökbara i ett vigselregister.

I ett dödsregister kanske man skapat det egna fältet Runa. Innehållet i det fältet kommer då att presenteras i Övrigt-fältet på en egen rad med texten *Runa: xxxxxxxxxxxxxx*

Tänk på att med jämna mellanrum under arbetet med tvätten spara ner databasen under nytt namn, så att eventuella misstag under arbetets gång inte innebär att arbetet måste göras om från början.

Den som använde den något annorlunda fältbeskrivningen från 2001 behöver inte känna det arbetet som bortkastat. De tillägg *dp* (döpt), *bg* (begravd) och *ly* (lyst) som den gången rekommenderades att lägga till efter datumen där födelse-, död- och vigseluppgifter saknades kommer vid genereringen att placera uppgifterna i rätt fält ändå.

Tvingande åtgärder

; (semikolon) ska helst inte förekomma i datamaterialet eftersom det är en fältseparator och kan således åstadkomma att uppgifter hamnar i fel fält. Använd *Sök och ersätt* med *Bekräfta* för att ändra till kolon eller komma, beroende på sammanhanget.

Om du absolut vill ha med dem, läs under Teknisk specifikation sid 22 hur man då förfar.

ÿ (y med prickar) är en postseparator och måste också ersättas om den finns. Tecknet har inte heller något eget värde utan betyder efter sammanhanget y eller ij. Exempel: *Bÿ* blir *by* och *wÿk* blir *wijk*.

Dubbla blanksteg kan störa sökningen. Ersätt dubbla blanksteg med enkla utan att bekräfta. Den här rutinen görs om till inga träffar längre finns. Om det visar sig att registreraren använt många mellanslag efter varann kan det vara lämpligt att först kolla vad avsikten varit. Prova då att söka på dubbla mellanslag med bekräfta.

Rekommenderade åtgärder

? (frågetecknen) är inget sökbart tecken och bör bytas. Det lär inte heller förekomma i källmaterialet. Eventuell markering av en osäker post görs i stället i fältet *Status*, med ett *E*, eller med en kommentar i fältet *Reganm* (registeranmärkning). Sökning på frågetecknen i Excel görs med ~? (tilde och frågetecknen).

* kan sökas i Excel via ~* (tilde och stjärna) om man vill ersätta det tecknet.

() (parenteser) torde inte förekomma i originalmaterialet och byts således ut, vanligtvis mot ingenting.

Ovana registrerare kan ha haft svårt med accenter. Sök på i tur och ordning 'e, e', `e, e`, ´e, e´, è och ersätt med é. è med grav accent torde inte förekomma i någon utsträckning, men gör granskningen med *Bekräfta* för varje ändring för säkerhets skull. Det kan finnas anledning även att söka på de tre olika accenterna var för sig också.

+ (plustecken) ska normalt inte förekomma, men kan ha kommit in av misstag i Excel-arket som en cellreferens. Sök på + och rätta eventuella fel.

” (citattecken) förekommer normalt inte i källmaterialet och bör därför tas bort. Det stör dessutom sorteringen. Om det finns anledning att det absolut måste vara kvar ska det då i stället dubbeltecknas, se också Teknisk specifikation, sid 23. Det är även viktigt att det i så fall verkligen finns två citattecken i fältet och inte bara ett.

NN i namnfält förekommer inte i källmaterialet och tas därför bort via *Sök och ersätt*, hela cellinnehållet, utan *Bekräfta*.

Sök på mellanslag före skiljetecken som komma och punkt, och ta bort dessa via *Sök och ersätt*. De kommer annars att stå med som valbara namn i plocklistorna.

Enbart blanksteg i ett fält ersätts med ingenting. Gör *Sök och ersätt*, välj *Matcha hela cellinnehållet*.

Sök och ersätt återigen dubbla mellanslag med ett. Ofta har nämligen tvättåtgärderna skapat nya dubbla mellanslag.

Det går alltså att söka på i stort sett alla kombinationer av tecken, exempelvis ”....-.-..” (som vissa registerprogram skapar). De byts ut mot ingenting. Fält som består enbart av ensamma bindestreck eller punkter rensas. *Sök och ersätt* hela fältet. Sök gärna på stavfel som uppträder. En registrerare gör inte sällan samma fel flera gånger.

Spara Excelarket. Spara det sedan igen som en skv-fil, dvs semikolonseparerad textfil. Stäng filen. Öppna Utforskaren och leta rätt på skv-filen. Döp om suffixet på filen till *txt*.

Du har nu fått en textfil som ska öppnas igen i Excel för fortsatt tvättarbete. Åtgärden att spara den som en textfil för att sedan öppna den igen är till för att få bort oönskade formateringar som ofta stör sorteringen i nästa moment.

Öppna textfilen i Excel. I Textimportguiden som kommer upp ska du välja Semikolon som avgränsare och som kolumndataformat Text i samtliga fall. Du har nu en ny version av databasen som är helt oberoende av den tidigare formateringen i de olika registrerarnas datorer.

Markera ett fält i kolumn A. Blockmarkera hela databasen med Ctrl-A. Tryck på Sortera. I början och slutet av den sorterade databasen letar du efter konstigheter i kolumn A. Det kan ha uppstått något fel vid registreringen, som att fel uppgifter hamnat i det fältet. Den här sorteringen avslöjar genast sådana missar. Det kan gälla text som hamnat i ett datumfält eller vice versa. Det kan också ha skapats blanksteg i början på en rad som man då får ta bort manuellt. Rätta de fel du hittar och gör om proceduren tills det ser bra ut.

Fortsätt sedan på samma sätt med kolumn B och sedan resten av kolumnerna som har något datainnehåll. Kolla extra att det verkligen bara är förnamn i förnamnsfältet, efternamn i efternamnsfältet, osv.

Spara ner databasen som semikolonseparerad textfil. Tänk på att inte ha med citationstecken som textseparator om det valet finns.

Om er databas är större än 50.000 poster, så att fler Excelfiler behövs, gör du om proceduren med nästa fil. Tänk bara på att när den är tvättad och klar ska du den här gången ta bort den översta raden som du lagt in med fältbeteckningar. I den slutliga databasen ska det bara finnas en sådan rad. Under arbetets gång med de olika Excelfilerna måste den dock finnas med för att hålla ihop databasen. Om något fält helt saknar uppgifter kommer sorteringen annars att bli felaktig.

Du har nu fått en databas som är tvättad så långt det går på rimligt sätt. Under genereringsfasen kommer ytterligare kontroller med fellista att skapas, men genom ovanstående åtgärder blir den sannolikt minimal.

När allt material av den typen (födelse-, vigsel- eller dödboksregister) är tvättat läggs textfilerna ihop, antingen i ett databasprogram som exempelvis Microsoft Access eller helt enkelt i textprogrammet Anteckningar. Fältnamnen ska befinna sig överst i den slutliga filen. Spara ner allt som en textfil med namnet *ssfppxxxiny.txt*, där *ssfpp* betyder Sveriges Släktforskarförbunds Presentationsprogram, *xxx* är det unika nummer som varje förening fått sig tilldelat för sitt register, *in* betyder indatafiler och *y* ersätts med 1, 2 eller 3 beroende på om textfilen innehåller födelse-, vigsel- eller döduppgifter.

3. Generering av indexfiler för databasen

Introduktion

Starta installationsprogrammet *Presgeninst.exe* som ligger på CDn. Se till att ingen av förbundets övriga CD är igång då det i så fall stör installationen.

Välj mappar för programinstallationen (*SSFpresgen*) samt för datainnehållet (*SSFpresdata*). I den senare mappen skapas tre undermappar av installationsprogrammet:

- *hdbas* (där de genererade databasfilerna kommer att placeras)
- *infiler* (där du placerar de tre färdiga textfilerna med födelse-, vigsel- och dödbregistren)
- *normering* (där du kan lägga in normeringstabeller, se sidan 11)

När installationen är färdig, lägg in de tre textfilerna *ssfppxxxxiny.txt*, som skapats under tvättprocessen, i undermappen *infiler*, där *xxxx* är föreningens unika nummer och *y* ska vara 1, 2 eller 3 beroende på om filen innehåller ett födelse-, vigsel- eller dödbregister.

Egen text

Föreningen kan skriva en egen text om registret, exempelvis hur arbetet gått till, vilka som deltagit, kvalitetsdeklaration eller annat valfritt. Den ska skapas i rent textformat, läggas i mappen *infiler* och heta *ssfppxxxxtext.txt*, där *xxxx* ersätts med föreningens tilldelade nummer. Texten kommer att presenteras under knappen *OM...* tillsammans med andra, av programmet skapade statistikuppgifter.

Generering

Starta genereringsprogrammet (ikonen med blixten).

Skriv in det fyrsiffriga nummer föreningen fått sig tilldelat under Databasnummer.

Skriv in namnet på den blivande CDn (ex. *Söderbo härads släktforskarförening*).

Skriv in underrubrik (ex. *Födelse-, vigsel- och dödbregister för Söderbo härad 1688-1930*).

I nedersta vänstra fältet kan man ange en mapp där tillfälliga filer läggs under genereringen. Annars läggs de direkt under C:. Detta blir bara aktuellt för stora databaser, över 200.000 poster, och filerna raderas dessutom automatiskt efter att genereringen avslutats.

Skriv namn och sökväg till föreningens logotyp i bmp-format, se närmare på sidan 10.

I rutan till höger kommer loggfilen att visas under själva genereringen.

När alla uppgifter är ifyllda, tryck på *Starta!* Programmet börjar då granska och indexera de textfiler du tidigare skapar. Processen kan följas i det högra fältet. Se till att inga andra program är igång samtidigt, då den här processen kräver mycket datakraft!

När bearbetningen är klar kan du bläddra upp och ner i loggtexten för att granska hur resultatet blev. Första gången brukar det bli ett antal påpekanden, som behöver åtgärdas. Listan skapas även som en textfil *ssfppxxxlogg.txt* (där xxxx betyder föreningens unika nummer) under mappen *\normering* och kan således den vägen skrivas ut.

Normering i efterhand

Normering kan göras på fyra olika begrepp, där de filer man vill använda ligger under mappen *\normering* och har följande namn:

- Efternamn (filen *ssfppnormenamn.txt*)
- Förnamn män (filen *ssfppnormfnamn.txt*)
- Förnamn kvinnor (filen *ssfppnormfnamnk.txt*)
- Orter (filen *ssfppnormort.txt*)

Filerna skapas av genereringsprogrammet, där personnamnsfilerna får fyra semikolonseparerade fält och ortnamnsfilen fem.

Filerna öppnas företrädesvis i Excel varefter normerade namnformer fylls i i den tomma andra kolumnen. För ortnamnsfilen ska ortnamnet stå i den andra och socknen i den tredje kolumnen. Fältet därefter innehåller antalet träffar för den angivna ursprungsformen av namnet och det sista fältet den databasfil namnet är hämtat ur.

Normering i förhand

Har du redan under registreringsarbetet fyllt i ett extra normeringsfält för för-, efter- och/eller ortnamn kan du skapa en egen normeringsfil ur din databas som då ersätter den fil genereringsprogrammet skapade enligt förra stycket, exempelvis på följande vis.

Gör en kopia av Excel-arket eller Accessdatabasen. Öppna den. Ta bort alla fält utom de två för onormerat och normerat namn. För ortnamn behövs då de tre fälten onormerat, normerat och socken i normerad form. Spar innehållet som en semikolonseparerad textfil med det namn och på den plats som anges under föregående rubrik. Den befintliga filen ersätts alltså av den nya, manuellt skapade.

Det kommer med den här metoden att bli många dubbletterposter vilket inte påverkar programmets funktion, men dubbletterna går att ta bort via funktioner i såväl Excel som moderna versioner av Access.

Tänk på att normeringen bara görs till ett namn. Det går alltså inte att i det färdiga programmet söka antingen på *Margareta* eller *Greta* och då få träff på alla varianter av dessa ursprungligen samma namn. Träff får man på de varianter som är kopplade antingen till det ena eller till det andra namnet.

Hur man skapar en bildfil för logotypen

Logotypen ska föreligga i bmp-format med en storlek på högst 80x80 pixlar.

Den som har sin logotyp i digitalt format och har tillgång till ett bildbehandlingsprogram, exempelvis PhotoShop, kan själv anpassa bilden efter tillgängligt utrymme. Starta programmet och skapa en ny fil med den storlek som passar logotypen, dock högst 80x80 pixlar med 72 pixlar/tum upplösning. Öppna även bildfilen med logotypen. Markera den, kopiera den och klistra in den i den fil du först skapade.

Så här långt är det inte så svårt, men för att klara av följande moment bör man vara van vid PhotoShop. Bakgrunden ska nämligen göras helt grön (RGB-färg 0,255,0) för att logotypen ska kunna bli genomskinlig. Annars lägger den sig som en ruta över bakgrundsbilden i programmet.

Om du nu förstorar din bild (Ctrl+) så kan du bättra på den manuellt pixel för pixel med hjälp av verktygen till vänster, om den inte blir bra med en gång. Det är dock ett knepigt jobb som du bör testa på en kopia av filen först.

Ytterligare tvättning

När genereringsprogrammet kört igenom datamaterialet är nästa moment att rätta de fel som den här vägen kom fram.

Loggfilen kan ses i högra rutan i genereringsprogrammet. Den ligger som textfil *ssfppxxxlogg.txt* i mappen *\normering* och kan således också skrivas ut.

Vissa tecken som normalt inte förekommer i ett register presenteras särskilt med frekvens och kan ersättas mot andra. Detta görs i den ursprungliga Excel-filen som sedan återigen får sparas som textfil och döpas om till lämpligt namn. Fortsätt att korrigera de fel som påtalats i loggfilen och gör sedan en ny generering tills du är nöjd.

Nu är databasmaterialet i sådant skick att du kan gå in i fas 4, dvs att testa sökprogrammet.

4. Installation av sökprogrammet

I mappen *ssfpresdata* ligger filen *presinst.exe*. Det programmet skall tillsammans med mappen *\hdbas* brännas till CDn när databearbetningen väl är färdig.

Starta programmet och välj en mapp på hårddisken där du vill installera sökprogram och data.

När installationen är klar startar du programmet *SSPpres.exe* som ligger i den mappen.

Kolla nu att logon ser bra ut, att rubrikerna ligger snyggt samt att den egna fritexten ser bra ut under knappen *OM....*. Kolla även olika utskrifter av valfritt urval.

Via plocklistorna i sökfönstret kan du nu ytterligare granska att det verkligen bara ligger förnamn i förnamnsfältet, efternamn i efternamnsfältet, ortnamn i ortnamnsfältet, mansnamn i mannens fält i vigselregistret och vice versa, inga underliga stavningar, osv. Notera de avvikelser du hittar eftersom de måste korrigeras i den tidigare tvättade databasfilen, som sedan får genereras om igen.

Du kan också använda dig av de olika textfilerna under mappen *\normering* exempelvis vid kontroll av plocklistorna. För innehållet i filerna, se sid 20-21. Du kommer säkert att hitta mansnamn i fältet för kvinnliga förnamn och vice versa. Korrigera genusbeteckningen i originalfilen och gör om genereringen tills allt stämmer.

Var noggrann med tvätten av datamaterialet! Det är så mycket tid och arbete nedlagt på själva registreringen att det är värt den extra tid det tar med tvättarbetet. Dela gärna upp registren på olika personer och byt med varann sen. Man ser alltid olika fel.

När omgenerering skett och eventuella andra brister rättats är registret klart för produktion.

Kopiera programmet *presinst.exe* samt undermappen *\hdbas* från mappen *\ssfpresdata* till en brännbar CD-skiva. Den innehåller nu allt som behövs för att installera databasen med sökprogrammet på en användares dator. Installationen hos användaren sker sedan som tidigare beskrivits genom att programmet *SSFpres.exe* körs. Då installeras även de dll-filer som behövs för programmets funktion på användarens dator.

Vill du ändra det namn som skapas för ikonerna går det bra att göra det i filen *presinst.ini* som ligger i undermappen *\hdbas*. Annars skrivs automatiskt de båda rubrikerna *hrub* och *urub* ut som namn.

Sedan är det bara att skapa en etikett och omslag till fodralet.

För att ta fram mindre upplagor av CD-skivor finns ett företag som heter Polyvox AB, Datavägen 4, 175 43 Järfälla, som använder en teknik med bränning, som är lönsam för upplagor upp till några hundra exemplar. För upplagor över 500 exemplar är det mer ekonomiskt att pressa CD-skivorna.

Tänk på att inte ta fram för stora upplagor! När registren blir mer allmänt tillgängliga kan det initiera mer arbete med registrering, och även rättning, så att en ny utgåva kanske snart är aktuell igen.

Tekniska specifikationer (bilaga)

A. Fältformat

I genereringsprogrammet hämtas källregistret från en hårddiskcatalog eller CD, analyseras, komprimeras och läggs som färdig databas på en annan katalog. I processen används också normeringslistor – enkla textfiler med två kolumner. Förutom själva databasen matar genereringsprogrammet ut frekvenslistor på samtliga fält som blir söklistor i databasen. Dessutom fyller programmet på normeringslistorna med onormerade ("missade") begrepp med frekvens och filnamn för den indatafil där begreppet först dök upp. Dessa fylls manuellt på med nya normerade begrepp, programmet körs om osv tills man är nöjd.

Varje källregister kommer att bestå av tre textfiler – en för vardera posttyp – och (ej obl) fyra olika normeringslistor – förnamn män, förnamn kvinnor, efternamn och ort. Mer om genereringsprocessen i kap. D.

Födda (huvudtabell 1)

Obligatoriska fält:

Forskx	Församling
Lan	Län
Kaella	Källuppgift (arkivbildarnummer, volym, littera osv).
Fodat	Födelsedatum
Fnamn	Förnamn (mellanslagseparerade)
Farfnamn	Faderns förnamn
Farenamn	Faderns efternamn
Morfnamn	Moderns förnamn
Morenamn	Moderns efternamn

Valfria fält som programmet är förberett att hantera:

LokalID	Varje enskild posts ID
Forskod	SCB-kod el dyl.
Sida	Sida i originalvolymen
Status	E = ej säker, K = korrläst, annat eller tomt visas inte
Enamn	Barnets efternamn om det förekommer
Ort	Ort
Genus	Kön (M, K, tom)
Bord	Äktenskaplig börd (U, Ä, tom)
Dfodd	Dödfödd (tom/ej tom)
Dopdat	Dopdatum
Faralder	Faderns ålder
Fartitel	Faderns yrke/titel

Farort	Faderns boendeort
Moralder	Moderns ålder
Mortitel	Moderns yrke/titel
Morort	Moderns boendeort
Antfod	Antal födda samtidigt (2, 3, etc)
Paritet	Turordning på barnet (1, 2, 3, 4, etc)
Dopvittnen	Dopvittnen
Ovrigt	Övriga uppgifter i källan
Reganm	Registeranmärkning

Vigda (huvudtabell 2)

Obligatoriska fält:

Forskx	Församling, klartext
Lan	Län
Kaella	Källuppgift (arkivbildarnummer, volym, littera osv).
Mfnamn	Mannens förnamn
Menamn	Mannens efternamn
Kfnamn	Kvinnans förnamn
Kenamn	Kvinnans efternamn

Valfria fält som programmet är förberett att hantera:

LokallID	Varje enskild posts ID
Forskod	SCB-kod el dyl.
Sida	Sida i originalvolymen
Status	E = ej säker, K = korrläst, annat eller tomt visas inte
Vigddat	Vigseldatum
Lysndat	Lysningsdatum (kan vara flera, avdelade med komma).
Mtitel	Mannens yrke/titel
Mort	Mannens boendeort
Malder	Mannens ålder
Ktitel	Kvinnans yrke/titel
Kort	Kvinnans boendeort
Kalder	Kvinnans ålder
Ovrigt	Övriga uppgifter i källan
Reganm	Registeranmärkning

Döda (huvudtabell 3)

Obligatoriska fält:

Forskx	Församling, klartext
Lan	Län
Kaella	Källuppgift (arkivbildarnummer, volym, littera osv).
Fnamn	Förnamn
Enamn	Efternamn

Valfria fält som programmet är förberett att hantera:

LokallID	Varje enskild posts ID
Forskod	SCB-kod el dyl.
Sida	Sida i originalvolymen
Status	E = ej säker, K = korrläst, annat eller tomt visas inte
Dodat	Dödsdatum
Begrdat	Begravningsdatum
Fodat	Födelsedatum
Alder	Ålder
Titel	Yrke/titel
Genus	Kön (M, K, tomt)
Ort	Boendeort
Dorsak	Dödsorsak
Anf1titel	Anförvant 1 titel
Anf1fnamn	Anförvant 1 förnamn
Anf1enamn	Anförvant 1 efternamn
Anf1ort	Anförvant 1 ort
Anf1relat	Anförvant 1 relation
Anf1genus	Anförvant 1 genus (M, K, tomt)
Anf2titel	Anförvant 2 titel
Anf2fnamn	Anförvant 2 förnamn
Anf2enamn	Anförvant 2 efternamn
Anf2ort	Anförvant 2 ort
Anf2relat	Anförvant 2 relation
Anf2genus	Anförvant 2 genus (M, K, tomt)
Ovrigt	Övriga uppgifter i källan
Reganm	Registeranmärkning

Kommentarer

Tanken med *LokalID* är att numret skall vara en stabil identifierare av en post (som inte ändras vid omsortering el dyl), och att det kommer från källregistret. Det kan då användas för kvalitetskontroll, felrapporter osv oavsett dataversion. Numret kan vara en kombinationskod, ett ”syntetiskt personnummer”, ett enkelt löpnummer el dyl.

Uppgifter om församling och län bör vara samtida men stavningsnormerade. Om moderna län används skall detta vara enhetligt genomfört i källregistret. Det finns inget flaggfält för att markera samtida/modern geografi, så man bör inte blanda. *Forskod* kan exempelvis vara SCB-kod enligt Sveriges församlingar genom tiderna. Koden kommer att visas som en sekundär källuppgift i denna version av presentationsprogrammet men kan senare komma att användas för kartvisning/sökning (beroende på kvalitet).

Alla datum lagras internt (och söks) i formen *ååååmmdd*. I källregistret får dessutom förekomma formerna *dd/mm åååå*, *åååå*, *åååå-mm-dd*, men inte exempelvis ”13 januari 1885”, ”13/5 –86”, ”1887?” el dyl. Osäkerhet skall inte markeras i fältet utan som en E-flagga i statusfältet. I dödsdatumfältet kan hängas på *bg* (för begravd) – datumet kopieras då till fältet för begravningsdatum. På samma sätt fungerar *dp* i födelsefältet (kopieras till dopdatumfältet) samt *ly* i vigseldatumfältet (kopieras till fältet för lysningsdatum). I lysningsfältet kan flera datum skrivas in om de separeras via komma. Det blir dock bara det första som blir sökbart via index.

I ”ålder”-fälten kan finnas diverse uppgifter, beroende på källans art och år: födelseår, ålder i år eller t o m år, månader och dagar. Dessa fält kan inte enkelt datumstandardiseras enligt ovan – man får nöja sig med fritextsökning.

Vid sökning av döda söker man anförvanternas normerade för- och efternamn i listor, medan övriga uppgifter söks fritt.

B. Visningsformat

Av vardera huvudposttypen skapas en visningspost i fritextformat, på samma sätt som i ”Sveriges befolkning 1890” och andra produkter. Till varje post hör en rad (i vigselregistret två) i en träfflista med personens namn och relevant datum för händelsen. Finns inget dödsdatum används begravningsdatum, och finns inget födelsedatum används dopdatum. Listan kan sorteras på namn eller datum.

Posterna ställs upp med primäruppgifterna avseende ”huvudpersonen” överst (eller de två huvudpersonerna, i vigselregistret). På fliken allra längst upp visas Lokal-ID (eller postnummer inom parentes) samt en röd boll (för status E, ej komplett), en grön boll (status K, korrekturläst) eller tomt (ingen statusinformation).

Längre ned visas primäruppgifter om övriga personer (föräldrar, anförvanter, vittnen).

Längst ned visas uppgifter om primär- och sekundärkälla: län, församling, församlingsidentitet, volym, sida samt och anmärkningar från registreraren. Här läggs även de icke-standardiserade fält som automatiskt läggs till Övrigt-fältet vid genereringen.

Allra sist visas eventuella normerade former av namn och orter.

C. Sökning

Till varje posttyp skapas ett antal söklister (index). Här finns fyra varianter:

- Söklista med frekvensvisning, där man dels kan välja enskilda rader, dels fritextsöka.
- Dold söklista för datum och kodvärden
- Dold söklista med fritextsökbara värden
- Ren frisöklista (för löptext, ej uppräkningsbara värden)

I söklister med frekvensvisning eller med fritextsökbara värden kan man söka med * och ”-tecknen för trunkering eller exakta värden. Fler sökvärden avgränsas med semikolon (motsvarar boolesk eller-operation). Sökningen *Johan*;*”*Andersson*” ger sålunda alla Johansson, Johansdotter osv men även alla som heter exakt Andersson.

I datumfält söks på enskilda, automatiskt högertrunkerade, värden eller intervaller. *188509* ger träff för september 1885, medan *188509-188604* ger träff på datum mellan 1/9 1885 och 30/4 1886, och *1900-19100630* ger träff mellan 1/1 1900 och 30/6 1910, och *18* ger träff på hela 1800-talet.

Sökmöjligheterna är koncentrerade på individsökning i ett antal enkla, skarpa, fält – geografi, datum och primära personuppgifter. I övrigt ges frisökningsmöjlighet.

Däremot rekommenderas att man inte frekvenssöker på uppgifter som sällan förekommer.

Fritextsökning är normalt helt fri, men hittar aldrig uppgifter ”på tvären” över fältgränser – varje fält är s a s en egen ”mening” och programmet söker endast i hela meningar. Om jokertecknet * avslutar ett sökord söks ordet endast i början av ett fält – bra vid exempelvis datumsökningar. På samma sätt ger ett sökord inom citationstecken endast träff på exakt fältinnehåll.

Födda, sökindex

Textfilerna i undermappen `\normering` med namnen `ssfppxxxxfrekvyyy.txt`, där `yyy` ersätts med nedanstående siffror, innehåller följande:

101	Forskx	frekvens
102	Lan	frekvens
103	Kaella	frekvens
104	Fodat	datum
106	Fnamn	frekvens (norm)
107	Enamn	frekvens
108	Ort	frekvens
109	Morenamn	frekvens (norm)
110	Morfnamn	frekvens (norm)
111	Morort	frekvens (norm)
112	Mortitel	frekvens
113	Farenamn	frekvens (norm)
114	Farfnamn	frekvens (norm)
115	Farort	frekvens (norm)
116	Fartitel	frekvens
117	Dopvittnen	ren fritext
118	Ovrigt	ren fritext
119	LokalID	kod

Födelse- och dopdatum söks i samma sökindex ("född eller döpt ett visst datum").

Vigda, sökindex

Textfilerna i undermappen `\normering` med namnen `ssfppxxxxfrekvyyy.txt`, där `yyy` ersätts med nedanstående siffror, innehåller följande:

201	Forskx	frekvens
202	Lan	frekvens
203	Kaella	frekvens
204	Vigddat	datum
206	Kfnamn	frekvens (norm)
207	Kenamn	frekvens (norm)
208	Kort	frekvens (norm)
209	Ktitel	frekvens
210	Mfnamn	frekvens (norm)
211	Menamn	frekvens (norm)
212	Mort	frekvens (norm)
213	Mtitel	frekvens
214	Ovrigt	ren fritext
216	LokalID	kod

Vigsel- och lysningsdatum söks i samma sökindex ("vigd eller lyst ett visst datum").

Döda, sökindex

Textfilerna i undermappen *\normering* med namnen *ssfppxxxxfrekvyyy.txt* där *yyy* ersätts med nedanstående siffror innehåller följande:

301	Forskx	frekvens
302	Lan	frekvens
303	Kaella	frekvens
304	Dodat	datum
306	Fodat	datum
307	Fnamn	frekvens (norm)
308	Enamn	frekvens (norm)
309	Ort	frekvens (norm)
310	Titel	frekvens
311	Anf	fritextsökbar (kombination av alla anförvantsfält)
312	Ovrigt	ren fritext
313	LokalID	kod
314	Anförvants efternamn	frekvens
315	Anförvants förnamn	frekvens

Döds- och begravningsdatum söks i samma sökindex ("död eller begravd ett visst datum").

Sökindex för allmän sökning (tvärs över alla posttyper):

404	Fritx	ren fritext, allt
-----	-------	-------------------

Detta "allfält" kan också användas för att söka på uppgifter som inte finns indexerade för delregistren, exempelvis källa *Töllsjö C:3* samt vissa datum och åldrar.

D. Generering och utfilformat

Utfomat

Presentationsprogrammets databasmodell bygger på ett antal genererade databastabeller som installeras eller läses från CD. Dessa ligger fysiskt i över hundra filer och borde ta mellan 150 och 200 byte per post, inklusive allt. Filerna namnges enligt:

SSFPPxxxxh1, h2, h3	Huvudfiler
SSFPPxxxxl1, l2, l3	Träfflistor
SSFPPxxxxo1, o2, o3	Filer för sortering
SSFPPxxxxsy	Sökindex
SSFPPxxxxnrm	Tabell över normerade begrepp
SSFPPxxxxgrfk	Specialgrafik ("branding")
SSFPPxxxxtext	Text från utgivaren samt genererad databasstatistik.
SSFPPtext	Allmän text (om SSF och hjälptexter)
SSFPPgrfk	Allmän grafik

SSFPP står för "Sveriges Släktforskarförbund, presentationsprogrammet".

xxxx är produktens unika löpnummer (nollutfyllt). Med produkt menas en viss urskiljbar databas med sin egen bakgrundshistoria, upphovsrättsinnehavare m m. Dessa nummer administreras av Sveriges Släktforskarförbund. Filen SSFPPxxxxgrfk är inte obligatorisk (den innehåller en komprimerad logotyp för utgivaren). Samtliga filer har dessutom olika suffix beroende på deras funktion hos programmet.

yyy är sökindexnummer enligt kap. C ovan.

Filerna lagras och komprimeras på ett sådant sätt att det (i praktiken) är omöjligt att packa upp dem och använda rådata utanför presentationsprogrammet. Däremot finns inga inbyggda kopieringsskydd, exempelvis koppling av filerna till en viss dator eller CD-medium.

Vid massdistribution bränns den katalog `\hdbas` som genereringsprogrammet skapar och installationsfilen `presinst.exe` och andra filer (kanske pdf-version av hjälptexterna) på CD:n. Skivan kan sedan kopieras i valfritt antal exemplar eller skickas till tryck.

Informat

Filerna levereras i semikolonseparerat format. Poster åtskiljs med radbrytning och fält med semikolon. Citationstecken får omgärda ett fält, och måste göra det om fältet innehåller semikolon, men avstå gärna. Citationstecken inuti fält måste dubleras. Exempel på ett register med tre poster:

"Andersson; A:son";Jonas;18850910
Pettersson;Petter;18860320
Pettersson ""på byn"";Nils;18900220

Översta raden (första posten) skall innehålla kolumnrubriker som är identiska med de formella fältnamnen ovan, exempelvis ENAMN för efternamn eller FORSKX för församlingen i klartext. Genereringsprogrammet kan då läsa kolumnerna i valfri ordning. Vill man ha med egna fält som inte finns i den gemensamma specifikationen får man skapa ett eget fältnamn (exempel: RUNA). Dessa fält läses in och hängs på övrigt-fältet, på egen rad med fältnamnet som rubrik (exempel: *Runa: xxxxxxxxxxxxxxx*). De egna fälten blir sökbara via fritextsökningen men inte via namn-, datum- eller ortsökningarna.

Genereringsprogrammet tar inte hänsyn till stora/små bokstäver i fältnamnen.

Ny rad markerar ny post, så radbrytning får inte förekomma i enstaka fält. Kommentarer i originalkällan, anmärkningar osv skrivs som separata meningar i ett enda stycke. Om radbrytning absolut måste förekomma i ett (fritext-)fält, ersätts radbrytningen med ett kodtecken (pipetecknet, |), som sedan automatiskt byts tillbaks mot radbrytning vid genereringen.

I övrigt finns inga restriktioner på antal tecken per fält. I genereringsprogrammets logg visas "konstiga tecken" = höga eller sällan använda ASCII-koder som oftast indikerar felaktig inmatning eller överföring.

Ett fullständigt källregister finns på sidorna 14-16.

Normeringsstabeller

SSFPPnormenamn.txt	Efternamn
SSFPPnormfnamnk.txt	Förnamn kvinnor
SSFPPnormfnamnm.txt	Förnamn män
SSFPPnormort.txt	Orter

Normeringslistorna är semikolonseparerade i två kolumner: kolumn 1 är onormerat och kolumn 2 normerat. De ord som inte hittas vid normeringen läggs upp sist i listorna, med fyra kolumner: onormerat ord, (tomt), frekvens och källfil. Den tomma kolumnen fylls sedan på med normeringar (i Excel el dyl).

En del av ett normeringsregister kan alltså se ut som ...

Joh:son;Johansson
J:son;Johansson
Johansson;Johansson

... medan den (ännu) icke normerade delen kan se ut som ...

Jo:sn;;5;ssfpp0382in1
Joansson;;46;ssfpp0382in3

Genusfältet samt skillnaden M/K i vigselregistret används i genereringsprogrammet för att välja förnamnsnormeringslista. Förnamnsuppgifter utan kön blir alltså inte normerade.

Normeringslistorna skapas normalt ur specifika register, men man kan ändå tänka sig att en och samma lista skall kunna användas till flera körningar. Därför är inte dessa filer nummerkodade som de övriga. Det blir upp till ”fadderföreningarna” eller SSF att se till att normeringslistorna hålls isär och, om möjligt, samkörs på ett organiserat sätt.

Statistikfiler

Vid varje generering skapas dels en logg (ses på skärmen) och dels ett antal frekvensfiler över samtliga fält där det skapas sökindex.

Frekvensfilerna namnges *ssfppxxxfrekvyyy.txt*, där *yyy* motsvarar sökindexets nummer. Ordet läggs i första kolumnen och frekvensen i den andra. Endast sökindex med frekvens och datumfält skapar en statistikfil.

E. Visning och interaktivitet

Huvudskärm

Presentationsprogrammet är uppbyggt av en träfflista till vänster och en post till höger. Längst upp till vänster läggs en logotyp för presentationsprogrammet. Klickar man på denna får man information om Sveriges Släktforskarförbund och om presentationsprogrammet, inkl programversion. Bredvid huvudlogotypen läggs föreningens egen rubrik med mellanstor stil, samt en underrubrik med mindre stil. Möjlighet finns att lägga en logotyp till vänster om rubriken.

Hjälptexter nås med frågetecknen i hörnet.

Längst ned till vänster läggs fyra sökknappar – Sök födda, Sök vigda, Sök döda och Sök allt – samt utskrift och ”Om”. Om-knappen visar text om registret, inkl kvalitetsdeklaration, och utgivaren samt statistik över de ingående registren, med totalantal poster, postantal och tidsperiod per källa.

Programmet kan användas på skärmar med minst 800x600 upplösning.

Programmet har tre träfflistor, en per huvudregister, som väljs med flikssystem i överkant. Man kan alltså ha tre helt olika sökurval – i de tre huvudregistren – och bläddra oberoende i dem. I söklistan kan man använda rullhjul, piltangenter eller rullpil till höger. Med två knappar ovanför tabellen väljs namn- eller datumsortering. Observera att datumläget kan gå långsammare på en äldre dator med mycket stora register (hundratusentals poster).

Posten visas till höger och ställs upp som beskrivs i kap. B ovan. Rullning sker med rullpil. Ovanför finns knappar för att kopiera hela posten till urklippet och att skriva ut den. Man kan även markera delar av texten som i vanlig Windows-text och kopiera den med högerklick (meny, kopiera), alternativt använda Ctrl-C/Ctrl-Ins.

Sökskärmar

Tryck på sökknapparna visar sökskärmarna. Dessa fungerar som i tidigare applikationer (Sveriges dödbok, 1890 osv). Samtliga ifyllda fält måste passa in för att posten skall bli utvald. Fält som inte har någon information i databasen gråmarkeras och kan inte väljas.

I sökskärmen för födda finns sökrutor med listknapp för församling, län, källa, förnamn, ort, moders förnamn, efternamn, titel och ort samt faders förnamn, efternamn, titel och ort. Dessutom finns inmatningsrutor för födelse-/dopdatum, dopvittnen och övriga uppgifter. De två senare söks fritt.

I sökskärmen för vigda finns sökrutor med listknapp för församling, län, källa, mannens och kvinnans förnamn, efternamn, ort och titel samt inmatningsrutor för vigsel-/lysningsdatum och övriga uppgifter (som söks fritt).

I söskärmen för döda finns sökrutor med listknapp för församling, län, källa, förnamn, efternamn, titel, dödsorsak, ort samt anförvants för- och efternamn. Dessutom finns inmatningsrutor för dödsdatum, födelsedatum, anförvanter (alla fält, sammanslagna) samt övriga uppgifter. De två senare söks fritt.

I alla tre skärmarna kan man söka på lokal ID-nummer. Detta matchas exakt mot den kod som finns i basen. Sökningen genomförs genom att trycka på knappen till vänster om rutan (inte med den vanliga sökknappen). Posten med det valda lokal-ID visas omedelbart, alternativt en miss-ruta.

Man kan också trycka på knappen ”normering”. Då visas en tabell över alla normerade begrepp, med tre kolumner för onormerad form, normerad form samt frekvens (i alla tre tabellerna). Med knapparna längst upp väljs normeringstyp. Finns inga normeringar gråmarkeras respektive knapp.

I Allsökningen finns endast en ruta – för total fritextsökning i hela registret, tvärs över alla tre posttyperna.

Utskrift

De valda posterna kan skrivas ut i text- eller tabellformat. Tabeller skrivs ut med grunduppgifter med fältrubriker, på liggande papper.

Posterna kan sorteras i namn-, datum- eller källans ordning, dvs efter källreferens och därefter datum.

Inmatningsrutorna används för att välja utskriftsintervall, om antalet poster är för stort för att skrivas ut på en gång. Maximalt 10.000 poster kan skrivas ut på en gång.

För att starta utskriften när ovanstående val är utförda, tryck på *STARTA*-knappen.

Knappen *STÄLL IN* används för att välja skrivare, papper eller andra parametrar.