

Har du dubletter i din databas?

Svaret på den frågan är sannolikt ja! Efter en tids forskning blir det lätt så att man matar in samma person mer än en gång i sitt program. Men det är ju trots allt så att en person endast har funnits i enkel upplaga i verkligheten, så då bör den normalt sett inte finnas i dubbel upplaga i databasen.

Med version 8.2 av DISGEN tillkom en ny funktion för att hjälpa till att hitta eventuella dubletter i databasen.

Anledningen till att personer förekommer fler än en gång i forskningen är ofta att personen är din ana på mer än en linje, en så kallad anförlust [Bild 1]. Den vanligaste orsaken i övrigt till att man har en person fler än en gång i databasen är att man stöter på en person i olika roller, som barn i en familj och kanske sedan som vuxen ingift i en annan familj. Det är ju inte alls ovanligt när man fors-

kar på sidogrenarna eller när man jobbar med hembygdsvetenskap av något slag, tvärtom!

En kusins man kanske är barnbarns barn till din ana på en helt annan gren. Att släktingar, visserligen ofta ganska avlägset släkt, gifter sig med varandra förekommer oftare än du kanske tror. De ska då kopplas ihop i databasen, inte matas in flera gånger. En annan anledning till dubletter kan vara att man oavsiktligt matat in en person eller en familj fler än en gång.

En dublett kan finnas i din databas utan att du observerat det, kanske p.g.a. att namnet är snarlikt skrivet, men inte exakt lika. Och då hittar du den antagligen inte vid sökningar, beroende på hur du formulerar sökvillkoren eller vilka parametrar du använder när du söker.

Redan tidigare i artikelserien *Lär dig mer om DISGEN ...* har vi i viss mån tagit upp detta med dubletter, eller hur man undviker dem. Läs gärna artiklarna som berör sökning av personer och söklistor och deras användning (artiklarna med löpnummer 11 i Diskulogen nr 71 och nr 18 i Diskulogen nr 77) samt artiklarna om koppling och losskoppling (artiklar-

na nr 19–20 i Diskulogen nr 78 och 79). Läs också senaste artikeln om säkerhetskopiering, anpassad till version 8.2 och 8.2a, som du hittar som artikel nr 30 i Diskulogen nr 85.

Kontroll innan inmatning

Det finns en enkel funktion för att kontrollera om en person redan finns i databasen innan du skriver in honom/henne på nytt. Det beskrivs i DISGENS handledning för 8.2 på sidan 10–11. Knappen för detta hittar du på kortet *Grunddata* men den är gråmarkerad, d.v.s. du kan inte använda den, om du inte har skapat ett Sökindex (se sidan 43 i handledningen).

För att skapa ett sökindex går du in under menyn *Verktyg* och väljer där alternativet *Sökindex*. På kortet *Sökindex* [Bild 2] markerar du i kolumnen *Skapa* genom att klicka för *Alla* och sedan klicka på knappen *Skapa*, och därefter på knappen *Stäng*. Du bör välja att det ska uppdateras efter varje ändring av person, vilket är grundinställningen.

Om du använder kontrollen vid inmatning bör du tänka på att inte skriva in alltför många uppgifter innan du trycker på knappen *Dublett*. Du får upp ett antal

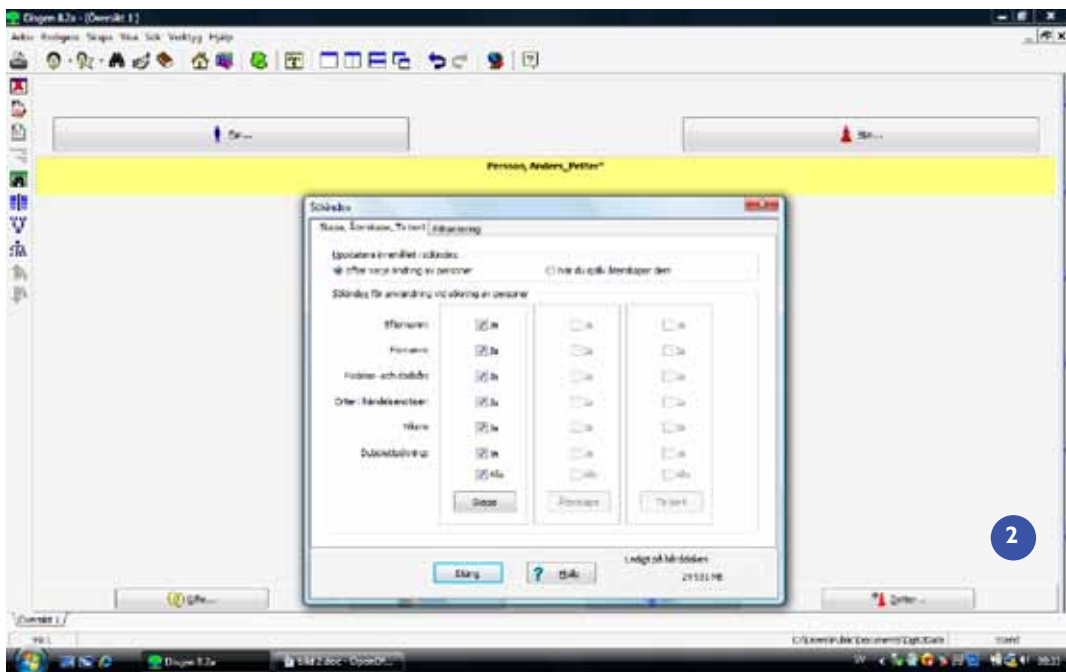
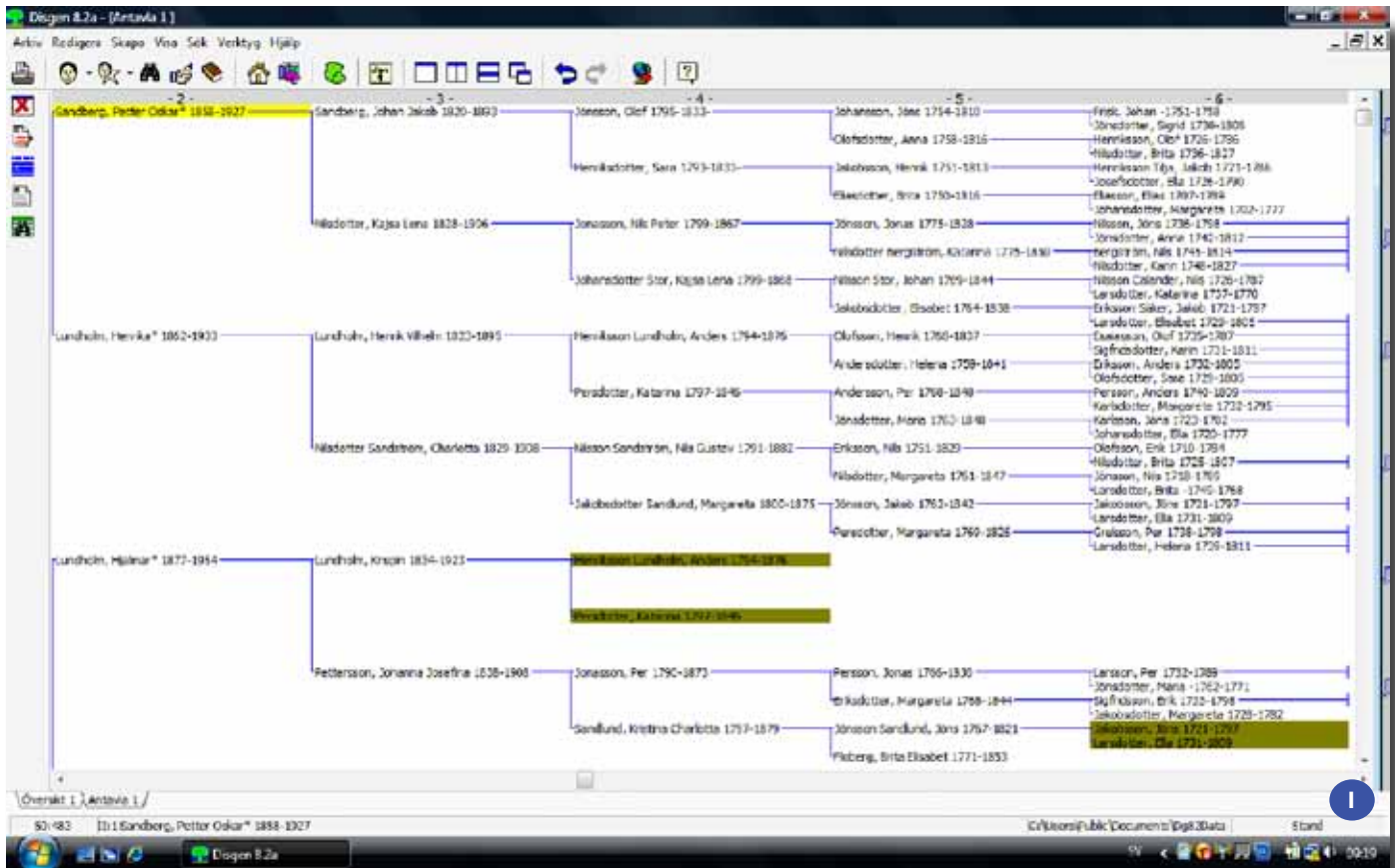


Bild 1. Exempel på anförlost.
Bild 2. Skapa ett sökindex.



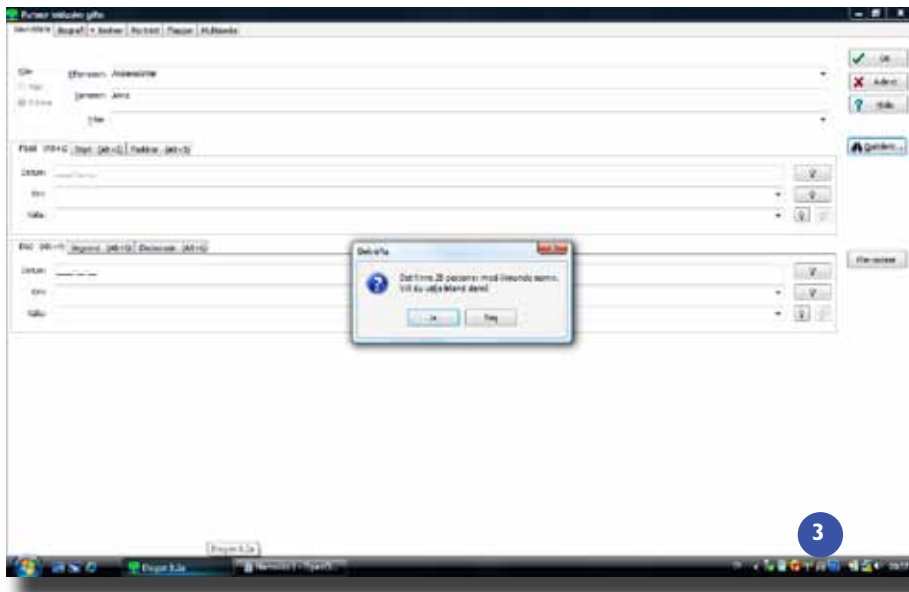


Bild 3. Verktöget Dubblett visar personer som kan finns redan.



personer att välja mellan [Bild 3] för att se om personen kan finnas redan, men om du inte väljer någon av dem så förloras de uppgifter du skrivit in på kortet *Grunddata* och du får skriva in dem på nytt.

Dubblettkontrollen

Det nya verktyget för dubblettkontroll gör en genomgång av ditt inskrivna material och försöker bedöma vilka personpar i databasen som kan vara dubbelregistreringar. Du når det när du i menyn *Sök* väljer alternativet *Sök efter dubletter...* Det kan alltså användas när som helst, inte som en funktion när du ska skriva in en ny person. Det beskrivs även i handledningen på sidan 42–43, varifrån denna text har hämtats.

Automatisk jämförelse

Kontrollverktyget använder en så kallad Soundex-kod och jämför personer i din databas för att få fram personer som har liknande uppgifter och därför skulle kunna vara dubletter. Personerna jämförs två och två och poängsätts.

För att starta sökningen när du har valt *Sök efter dubletter ...* klickar du på

knappen *Sök i hela arkivet* i det nyöppnade fönstret. Om du använder flera flockar och bara vill välja att söka i några av dem, så markerar du bara dessa i flocklistan genom att klicka och *hålla ner Ctrl*. När du valt klart klickar du på *Sök i valda flockar*. Om du redan har gjort en dubblettsökning under ditt arbetspass och har stängt ner sökningen men inte har stängt ner DISGEN sedan dess, så kommer ditt tidigare sökresultat fram när du väljer *Sök efter dubletter* [Bild 4].

Jämför utpekade dubletter

När sökningen är klar, vilket kan ta lite tid beroende på hur stort ditt material är och hur snabb din dator är, får du en tvåspaltig lista över möjliga dubblettpar [Bild 5]. Programmet säger inte att de med säkerhet är dubletter. När du markerar ett personpar och sedan klickar på knappen *Visa* i sökfönstrets egen menyrad så öppnas i bakgrunden familjeöversikterna för de två personerna i paret [Bild 6]. Klicka på knappen *Flytta/Koppla* så läggs familjeöversiktens sida vid sida och du kan lätt jämföra personerna. Du kan även få fram notiserna för dem genom att klicka på knappen *Notiser*. Du

kan hela tiden välja att visa eller dölja träfflistan, genom att klicka på knappen *Visa lista*.

Genom att klicka på knappen *Notiser* får du alltså fram notiskorten för de båda personerna i familjeöversiktterna som du ska jämföra. Klicka sedan på symbolen i huvudmenyn för flera fönster bredvid varandra, så får du en fyrdelad bild och kan jämföra notiserna [Bild 7]. Observera att när du förflyttar dig i träfflistan med pilarna *Nästa* respektive *Föregående* så byts personerna i familjeöversiktterna, men inte i notiserna, så du får klicka på knappen *Notiser* igen när du vill jämföra notiser för ett nytt dubblettpar.

I mitt exempel har jag ingen särskild text inskriven för någon av de två Olof, men det syns att de ligger i olika flockar och att de har forskats fram vid olika tidpunkter. Den ene har föräldrar, den andre har fru och barn och jag har alltså träffat på dem i olika roller.

Beslutet är ditt

Du måste själv avgöra om personerna är dubletter eller inte. Du får sannolikt många träffar som inte är dubletter utan bara har snarlika namn, t.ex. sys-

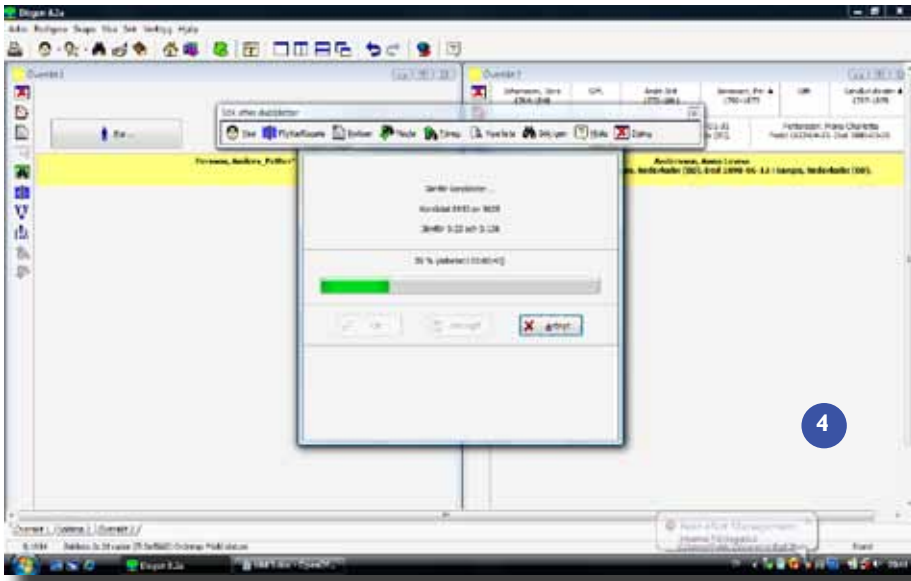


Bild 4. Sök efter dubletter i ditt arkiv.

kon har fått samma förnamn. Det är ett trubbigt verktyg eftersom du inte kan styra hur DISGEN beräknar de möjliga dubletterna, och poängsättningen. Du kan få högre poäng för två syskon med samma namn men med olika födelseår om det finns föräldrar inmatade till båda personerna, än vad du får för två personer med samma namn och samma födelsedag och födelseort, som sannolikt är samma person men där det finns föräldrar till den ena personen och hustru och barn till den andra personen. Men det är ett hjälpmedel för att kunna hitta personer som av någon anledning har blivit inmatade flera gånger.

Ett troligt dubblettpar som däremot inte hittats med sökningen, inte ens när jag begränsat till att söka i endast den flock där de båda finns inlagda, är två personer där stavning av både förnamn och efternamn skiljer sig åt [Bild 8]. Det paret fick jag dock träff på med den vanliga söklistan (se artikel 18 i Diskulogen nr 77, sidan 31). Poängberäkningen i Soundex-systemet bakom funktionen ger detta dubblettpar lägre poäng än andra träffar och de hamnar utanför träfflistan. Du kan tyvärr inte påverka anta-

let träffar vid dubblettkontrollen genom notering av att vissa möjliga dubblettpar inte är dubletter för att undvika att de kommer upp vid varje sökning.

Träfflistan kan inte skrivas ut i sig, men du kan ta en skärmdump (med hjälp av tangenten Print Scrn/ Prt Scr eller Print Screen på din dators tangentbord) av en del av träfflistan och klistra in i t.ex. Word och skriva ut den. Du kan också välja ut en intressant person att föra till en söklista, genom att som vanligt i familjeöversikten ta fram popupmenyn genom att högerklicka.

Kopplingar

Om du då hittar dubletter som du vill ta bort, så är det hanteringen i artiklarna om koppling som kan vara till hjälp. Du ska inte radera en person utan att ha försäkrat dig om att det är en dublett, och du måste ta tillvara uppgifter i t.ex. noter som bara står hos den ena av personerna. Men dubblettkontrollen tar inte bort någonting för dig, så var inte rädd för att använda den.

Lycka till med dina sökningar! 🍀

I Diskulogen har vi sedan 2004, Diskulogen nr 64, skrivit fördjupningsartiklar på olika nivåer om DISGEN. Några av de äldre artiklarna gäller endast för DISGEN 8.0 och 8.1, men flera av artiklarna som skrevs för de versionerna av DISGEN kan ändå användas av dig som använder DISGEN 8.2a. Du hittar en förteckning över samtliga artiklar, liksom artiklarna i sig, på adressen www.dis.se/disgen/articles. När det gäller dubletter och sökningar så är det särskilt artiklarna med löpnummer 11 (Diskulogen nr 71) och 18–20 (Diskulogen nr 77–79) du kan använda dig av, även om du har DISGEN 8.2a.



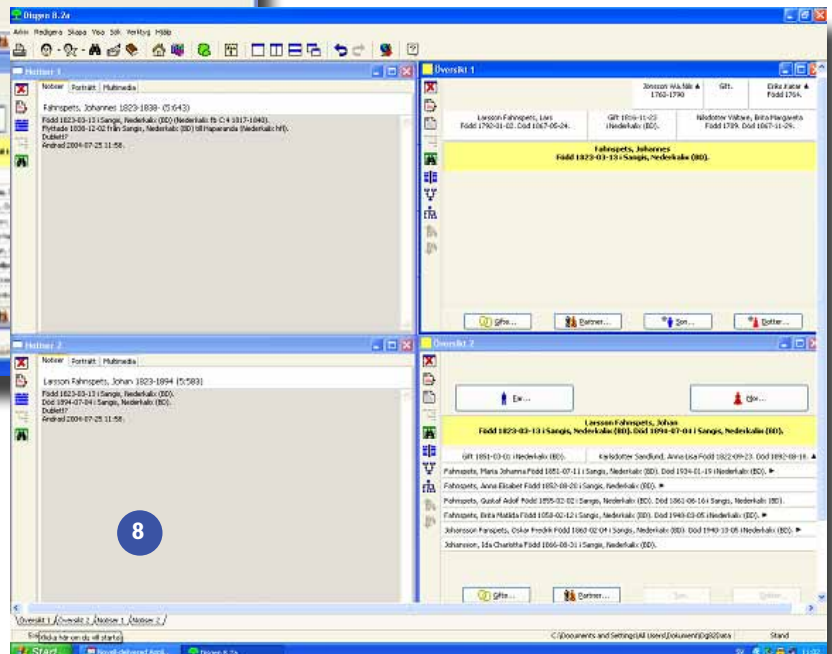
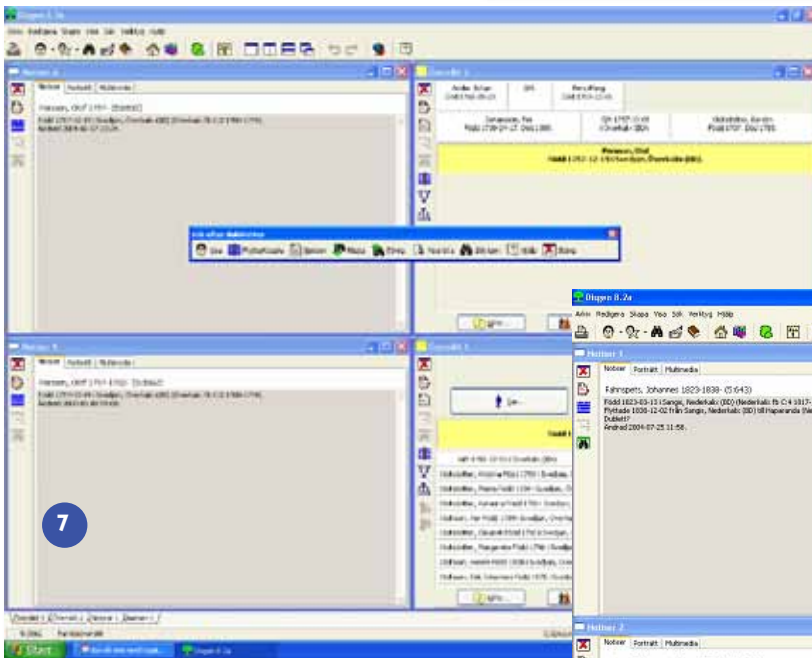
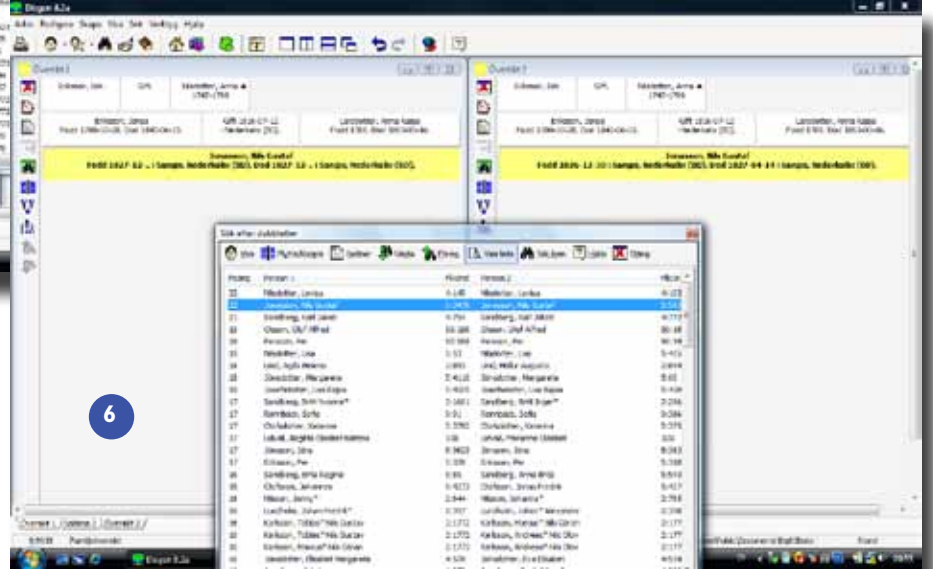
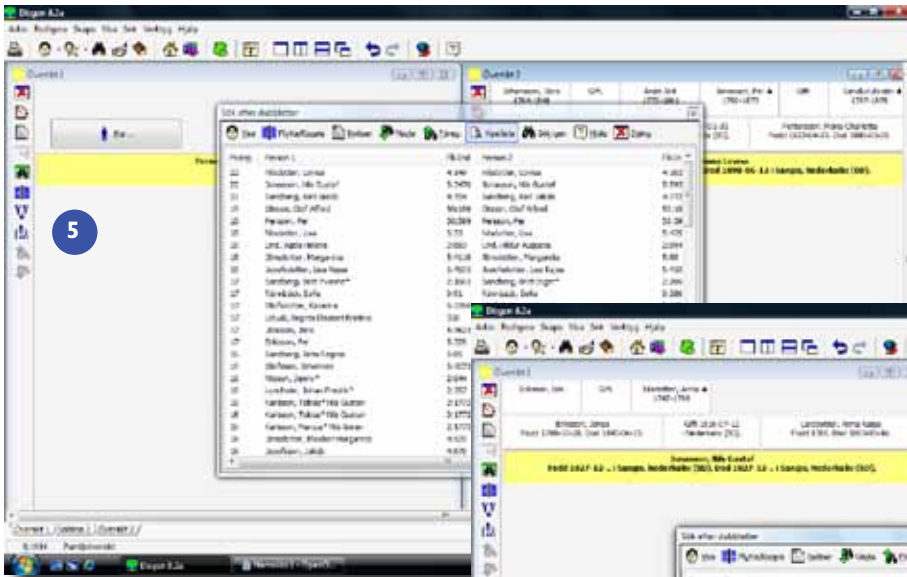


Bild 5. En lista över möjliga dubblettpar.
 Bild 6. Jämförelse av familjööversikter för två personer.
 Bild 7. Jämförelse av noter.
 Bild 8. Samma namn med olika stavning uppfattas inte som dubletter av programmet.