2026-08-09

Egenkontroll av GEDCOM-fil

Ett hjälpmedel, som släktforskare kan använda för att hålla bra kvalitet på sina egna uppgifter. Man startar egenkontrollen med hjälp av en vanlig webb-läsare, t.ex. Chrome, Edge, Firefox.

Innan man kan starta egenkontrollen skall man ha skapat en GEDCOM-fil, som innehåller de personer och familjer man vill kontrollera.

Egenkontrollen arbetar med GEDCOM-filer, GEDCOM är en ”standard” för alla släktforskningsprogram. Det är ofta mycket enkelt att skapa en GEDCOM-fil, som man sen i egenkontrollen kan ”peka ut” i sin egen filhanterare i sin dator.

Kontrollerna hjälper till att hitta typiska felaktigheter, som ofta smyger sig in när man registrerar sina uppgifter. Typiska fel som dubbletter, personer med fel kön, datum i fel ordning.

Det kan vara tryggt att veta att uppgifterna i GEDCOM-filen inte används till något annat, det är ingen annan än den som läst in GEDCOM-filen, som kan se uppgifterna och resultatlistorna.

Starta egenKontroll

Använd valfri webb-läsare och starta egenKontrollen via länken https://dubblett.dis.se/egenKontroll/ eller indirekt via länken https://rgd.dis.se/

Startsida

Sidan innehåller text som listar de bas-kontroller som görs och lite kort information om de olika resultatlistorna. En länk till inloggningssidan för openRGD finns också på sidan.

Egenkontroll

Det första momentet är att peka ut en GEDCOM-fil.

Rubriken **Välj Gedcom-fil:** har en knapp med texten **Bläddra...** Genom att klicka på den knappen får man upp filhanteraren i sin egen dator. Leta då upp din aktuella GEDCOM-fil i filsystemet. Vanligtvis har en sådan filändelsen '.ged'. Den aktuella GEDCOM-filen pekas ut genom att klicka på den.

Den utpekade GEDCOM-filens namn skall då komma upp efter Bläddra-knappen

Med knappen **Kör egenkontroll** startar man bearbetningen.

Det som då händer är att den utpekade GEDCOM-filen kopieras och bearbetningen startar. Ett antal kontrollprogram bearbetar GEDCOM-filen och skapar ett antal resultatlistor.

Resultatfilerna finner du sedan som ett paket, **egenKontroll.zip** under **Hämtade filer/Downloads** i den egna filhanteraren.

När du dubbel-klickar på **egenKontroll.zip** så öppnas paketet och de enskilda resultatfilerna blir åtkomliga.

Resultatlistorna
Klicka på respektive lista för att öppna listan

Bearbetningen avslutas med ett antal resultatlistor. Inte nödvändigtvis felaktigheter, utan det kan också vara ovanliga eller udda uppgifter.

Varje avvikelse identifierar personen eller familjen som avvikelsen berör. Beteckningen Id är hämtad från GEDCOM-filen och är i alla fall i vissa program direkt sökbar.

Utöver Id listas personens namn och om uppgift finns, även förekommande datum.

Det kan även förekomma andra kommentarer på sådant som hindrar eller försvårar att GEDCOM-filen kan användas i andra syften.

Senare i denna översikt finns lite mer information om barn i mer än en familj och om referensfel.

Tips: Spara gärna resultatlistorna, för att nästa gång undvika ny felsökning på sånt som inte var fel.

Familje-identiteter:
Identiteter för familj är ofta svårare att återfinna, därför skrivs även lite information från någon av personerna i familjen. Utöver Id listas personens namn och om uppgift finns, även förekommande datum. Via den personen brukar man kunna hitta vilken familj som berörs.

Lyckas man inte hitta familjen är enda möjligheten att söka direkt i GEDCOM filen.

En GEDCOM fil är en vanlig textfil och kan läsas med program av typ Anteckning och Notepad.

Tips: Sökningen i GEDCOM filen underlättas om identiteten kompletteras (första tecknet en nolla)
0 @F därefter familje-id och sen **@ FAM** (notera mellanslag bara före första @ och efter andra @)
0 @I därefter person-id och sen **@ INDI** (notera mellanslag bara före första @ och efter andra @)

Då kommer man direkt till rätt familj eller person och kan hitta mer information.

Check_lista.txt:
Check_listan bör alltid kontrolleras, men skall inte ses som fellista utan mer som information

Innehåller flera olika rubriker som beskriver typen av avvikelse.

Varje rad har sedan sin detaljbeskrivning av avvikelsen och information om vilken person eller familj det berör.

Det är ofta lätt att med hjälp av rubriktext och radtext avgöra vad programmet avser för typ av avvikelse.

Beroende på allvaret i meddelandet kan det rubriceras som **VARNING** eller som **FEL LISTA**. Fel-lista anger att det bör åtgärdas medan Varning är mer för att ange något sånt som är ovanligt, udda eller avvikande.

Dubbletter.txt – Dubblett-sökning

Söker personer med lika eller snarlika uppgifter. Listan innehåller parvisa dubblett-kandidater.

Om det faktiskt handlar om dubbletter eller om det är falsklarm måste avgöras manuellt.
Programmet anger en sannolikhetsnivå för varje kandidat-par med värde från 1 till 9, där en hög sannolikhetsnivå är tecken på att många uppgifter är lika.

Varje kandidat-par visas i block om 3 rader. Rad 1, Poäng = anger vilken sannolikhetsnivå kandidat-paret fått samt ett hjälptecken, vars betydelse förklaras i botten av listan

Raderna 2 och 3 innehåller de respektive dubblett-kandidaternas GEDCOM-id, förnamn, efternamn och någon händelse, oftast födelseuppgift
.
I vissa registreringsprogram är personidentiteten i GEDCOM-filen samma som i registreringsprogrammet och därmed direkt sökbara.

Namn.txt – Kontroll av namn och tecken i namnfältet

Denna lista kan hjälpa till att hitta personer som är registrerade med fel kön eller är felstavade.

Programmet använder sig av en referens-tabell med vanliga svenska förnamn. Med hjälp av registrerat kön och personens registrerade förnamn listas avvikelser. Namn-tabellen är inte på något sätt komplett så därför kommer lite udda namn, gamla namn, utländska namn ut i listorna. Detta kan skapa en lång listning, trots att namnen är helt korrekta och rättstavade.

Det kan ändå vara värt att kolla igenom dessa då det är ett utmärkt hjälpmedel för att hitta felstavningar.

Rubriken **Snedstreck i namnfältet...** listar personer med tecknet / i namnfältet.
Anledningen är att GEDCOM-standarden använder / för att avskilja efternamnet. Egna / gör att efternamn tolkas fel.

Rubriken **Tecken mindre än/större än...** listar personer med < > i namnfältet
Anledningen är att web-sidor tolkar allt som finns mellan dessa tecken som dold text. Följaktligen syns det inte heller i listan.

Rubriken **Kvinnonamn saknas...** listar personer med namn som kan vara mans-namn eller felstavade.
Rubriken **Mansnamn saknas...** listar personer med namn som kan vara kvinno-namn eller felstavade.

Rubriken **Trippeltecken...** listar personer med tre lika tecken i rad, vilket inte är så vanligt.
Rubriken **Dotterpatronymikon i mans-namn...** listar mans-personer med 'dotter' i efternamnet.
Rubriken **Personen saknar kön...** listar personer utan angivet eller okänt kön (någon namnkontroll har därvid ej kunnat göras).

Namn_2.txt – Lista avvikande och obehandlade namn

Denna lista kan hjälpa till att hitta personer som är felstavade.

Kontrolldatabasen innehåller vanliga moderna svenska förnamn men är på inget sätt komplett. Därför kommer lite udda namn, gamla namn, utländska namn ut i listorna. Detta kan skapa en lång listning, trots att namnen är helt korrekta och rättstavade.

Det kan ändå vara värt att kolla igenom då det är ett utmärkt hjälpmedel för att hitta felstavningar.

Rubriken **Förnamn som kan vara felstavat, bedömts som icke-namn eller som efternamn**
Listar personer med namn som bedömts felstavat, texter som inte hör hemma i namnfältet och även typiska efternamn. Man måste själv avgöra om det är en felregistrering eller om det är så man vill se namnfältet.

Rubriken **Förnamn som inte behandlats, de kan vara helt korrekta men kolla gärna**
Listar personer med namn som inte behandlats i namndatabasen. Databasen är inte på något sätt komplett utan det finns mängder av namn som inte finns med. Dessa kan givetvis vara helt korrekta men det kan också bero på felstavning, så det kan vara bra att göra en ytlig kontroll.

Orter.txt – Ortnamn / Platser...

Alla platser som inte kan identifieras som en svensk församling eller som ett land listas.

Händelserna född, död, vigd bör registreras med en församling där så är möjligt.

För att programmet skall kunna identifiera församling krävs att län är angivet, antingen som **länskod**, som **länsnummer** eller som **text tillsammans med ordet län**.
Som facit används Skatteverkets tabell **Sveriges församlingar**.

För länder kontrolleras mot **ISO-2 landstabell** eller att landskoden är i **2-ställig ISO-kod**.

För länskod, länsnummer och landskod gäller att koden omges av parenteser och placerad efter texten.

För länsbeskrivningar i textform gäller att den är placerad i rätt ordning, församlingsnamn, länsnamn följt av mellanslag och ordet län.

Programmet kan även ge alternativa förslag på giltiga församlingar som kan vara till hjälp. Men ibland är det bara snarlika namn eller kan vara helt felaktigt.

Orter_X.txt – Ortnamn / Platser...

Innehåller orter med referenser till personer och familjer

En ort lista med referenser till personer och familjer kan tendera att bli väldigt lång och svåröverskådlig.

Den komprimerade listan Orter.txt kan användas som översikt och när man bestämt att ändra en ort söker man upp orten i denna lista och får då referenser till personer eller familjer som berörs av just den orten.

Utöver ort listas identitet, personens namn och om uppgift finns, även förekommande datum.

Källor.csv – Saknade källor

En csv-fil är inte så läsvänlig om man tittar på den med mänskliga ögon, utan är avsett att öppnas i ett matris-program, t.ex. **Excel** eller liknande.

Endast händelserna **född, död** och **vigd** kontrolleras och listas.

Som kolumnavskiljare används **semikolon,** eftersom vissa fält ibland innehåller kommatecken. Inställning för teckenuppsättning bör vara Unicode (UTF-8) och språk bör vara Svenska.

Listningen saknar rubrik men innehåller:
Id, efternamn, förnamn, händelse, ort, datum i GEDCOM-format och ev. -information.

Om Id refererar till Familj finns normalt lite personuppgifter som information.
Se information om **Familje-id** under huvudrubriken **Resultatlistorna.**

GEDCOM fil – Konverterad GEDCOM till UTF-8 format

En GEDCOM fil kan skapas med olika teckenformat, t.ex. UTF-8, ANSI, ANSEL.

För användaren har detta normalt ingen större betydelse och beror hur man valt teckenformat i sitt släktforskningsprogram.

egenKontrollen använder teckenformatet UTF-8, som är det modernaste och mest vanliga.

Kontrollerar man en GEDCOM fil med annat teckenformat, så konverteras GEDCOM filen innan bearbeteningen.

Har GEDCOM filen konverterats till UTF-8, bifogas den tillsammans med resultatlistorna. Avser man att använda GEDCOM filen till annan än denna kontrollen, bör man i så fall använda den bifogade GEDCOM filen med textformat UTF-8.

Barn i mer än en familj

I ett genealogiskt släktträd är regeln att ett barn bara kan ha en biologisk relation.

Numera hanterar många program olika relationer, d.v.s. både en biologisk relation och en eller flera icke-biologiska.

Det här behöver då inte alls betraktas som ett fel. Har man ett program som rätt hanterar adoptiv- och foster-relationer behöver man bara se till att man registrerat uppgifterna på rätt sätt.

Men det är viktigt att komma ihåg att om man skapat GEDCOM-filen för något annat ändamål så måste man försäkra sig om att det mottagande programmet hanterar icke-biologiska relationer på samma sätt. Tyvärr har olika program också valt lite varierande delar av GEDCOM-standarden för detta, så det händer att det ändå skapar fel hos mottagaren.

Relationer.txt – Informationslista med saknade relationskopplingar

Listar icke kompletta relationer som bör kompletteras om möjlighet finns. En normal relation ska normalt ha minst två personer, t.ex. Make-Maka, Pappa-Barn, Mamma-Barn.

Rubriken **Familj som helt saknar individ** avser en relation helt utan personer.
Rubriken **Familj med endast en individ** avser en relation som inte uppfyller minikravet.
Rubriken **Individ som saknar familjekoppling** avser en orelaterad person utan familjekoppling.

Om Id refererar till Familj finns normalt lite personuppgifter som information.
Se information om **Familje-id** under huvudrubriken **Resultatlistorna.**

Datum.txt – Informationslista opreciserade datum

Listar sådant som inte är felaktigt men som kan kvalitetsförbättras om information finns.
Rubriken **Notering om opreciserade datum,** typiskt sådana är när man angivit t.ex. omkring, före, efter eller mellan. Troligtvis har man inte mer detaljerad information så därför har dessa fått en egen lista.

Referensfel

Formella fel i GEDCOM-filen, t.ex. att kors-referenser inte stämmer eller saknas. Det är ett allvarligt fel som gör att GEDCOM-filen troligen inte blir accepterad av andra program.

Dessa fel beror normalt inte på släktforskaren, utan det är programmet man använder som gör fel. Möjligen på grund av att man inte registrerat sitt data så som programmet kräver eller att man själv försökt göra någon justering direkt i GEDCOM-filen.

Denna typ av fel kan normal användaren ofta inte lösa, man behöver mycket god kunskap i GEDCOM-standarden, hur GEDCOM använder referenserna och hur man kan rätta felen.

Felsignalerna som egenkontrollen har möjlighet att ge är heller inte speciellt informativa. Programmet kan mest bara konstatera att det inte stämmer.