

GEDCOM Service Programme

ein Satz von Programmen zur

- Analyse
- Validierung
- Korrektur
- Reduzierung
- Listen
- Sortierung
- Umnummerierung
- Duplikate, Merge, Sync

von Gedcom Dateien

Ein Gedcom
Werkzeugkasten



Dazu einige sog. GSP Spezial- und GSP/OFB Dienstprogramme

Diedrich Hesmer, Herrenberg (D) - ofb.hesmer@gmx.de - <http://ofb.hesmer.name/gedcom>

Herrenberg, Sept 2017

© Diedrich Hesmer, Herrenberg

1

GSP (Gedcom Service Programme)

Gedcom Service Programme – Entstehung

- Ursprünglich für das OFB-Programm geschrieben
 - Erhebliche Unterschiede in ged-Dateien bei gleichem Dateninhalt
 - Unterschiedliche „Dialekte“ der Genealogie-Programme
 - Missbrauch von Datenfeldern durch Benutzer
 - Missbrauch von Gedcom Tags durch Programmierer
 - Verwendung „Benutzer-definierte“ Tags (_XXX)
- Viele Anfragen + Wünsche von Anwendern des OFB
 - Fehlerprüfung + Korrektur
 - Einfacherer Datenaustausch mit Kollegen
 - Anpassung der Datei-Strukturen für unterschiedliche Programme

2

GSP Programme

Paket besteht aus 8 Programme

- Grundprinzipien
 - Original ged-Dateien bleiben unverändert
 - Zeiger + Datensatz-Nr. werden unverändert übernommen
 - Fehlerprotokoll- / Logdateien werden erstellt
 - Alle einmal gewählte Einstellungen können in „Steuer“-Dateien zur Wiederverwendung gespeichert werden
 - Letzte 10+ verarbeitete "Steuer"-Dateien werden in eine ini-Datei gespeichert und beim Programmstart erneut geladen
 - Programme über Schaltknöpfe + Reiter steuerbar
keine Menüs
 - "?" öffnet das Handbuch + „Tooltips“ unter Mauszeiger



GSP (Gedcom Service Programme)

Gedcom Start Zentrum

- Start aller GSP
- Programm Update
- USB Transfer
- Einstellungen
- Lizenz Verwaltung
- "Info" zeigt Speicherort der Programmdateien
- Start von "Specials" und "Dienstprogrammen"

Programme erstellen eine neue ged-Datei

Programme erstellen nur Ergebnis-Dateien

Programme benötigen ein Gedcom Profil



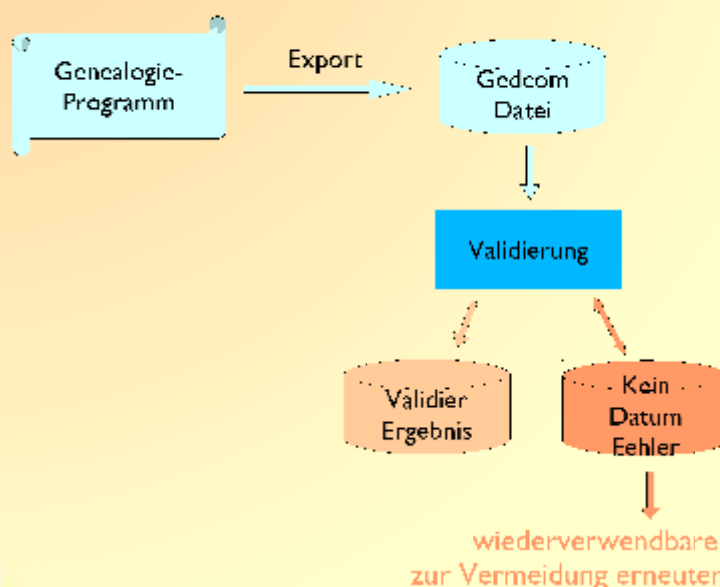
Struktur Analyse

- Analysiert Struktur der ged-Datei
 - Stufenweise Aufbau
 - Aufeinanderfolge der Tags
 - Jeder Datensatz Typ (INDI, FAM, SOUR, NOTE, OBJE, ...)
 - Umfang der Prüfung + Ausgabe durch Optionen einstellbar (>30)

Darstellung ged-Datei	ergibt	Darstellung Struktur Analyse
0 @I150 INDI		n INDI
1 NAME Marion /Afrau/		n INDI.NAME
1 SEX F		n INDI.SEX
1 BIRT		n INDI.BIRT
2 DATE 18 MAY 1888		n INDI.BIRT.DATE
2 F AC Aldorf		n INDI.BIRT.PLAC
2 SOUR Standesamt Aldorf		n INDI.BIRT.SOUR
2 NOTE Geburtsregister ...		n INDI.BIRT.NOTE
3 CONT Fortsetzung Text ...		n INDI.BIRT.NOTE.CONT
1 DEAT		n INDI.DEAT
2 DATE 18 NOV 1976		n INDI.DEAT.DATE
...		n INDI.DEAT
1 RELI evangelisch		n INDI.SOUR
1 SOUR #0		n INDI.BUERGERORT
1 BUEGERORT Hildorf		n INDI.FAMS
1 FAMS 0740		

Validierung

Erlaubt die Überprüfung von ged-Dateien auf Datum Fehler (einzeln und im Zusammenhang) und diverse potenzielle Probleme



Validierung – 1/2

- **Datum Prüfung** innerhalb einer ged-Datei
 - Prüfung logischer Fehler bei Datum Angaben
 - 10 fixe (1:1) Vergleiche
 - z.B. Geburt nach Tod bzw. Heirat
 - 10 flexible Vergleiche mit Ungenauigkeit in Tagen, Monaten, Jahren
 - z.B. Geburt Kind nach x Jahren seit Geburt von Mutter
 - Spezielle Ereignisse
 - z.B. Abhängig von Volljährigkeit, Geburt, Tod, Heirat
 - Prüfung / Listung von Datum Bereichen (von bis)
 - z.B. fehlende Datum Angaben für * ~ + =
 - Ergebnislisten – mögliche Datum Fehler
 - Validierungsergebnisliste je Datum den "Fehler"
 - Ged Datum Liste + Datei 1:1 der ged-Datei für die Person / Familie
 - „Kein Datum-Fehler“ Datei – Datum Angaben als „korrekt“ definiert



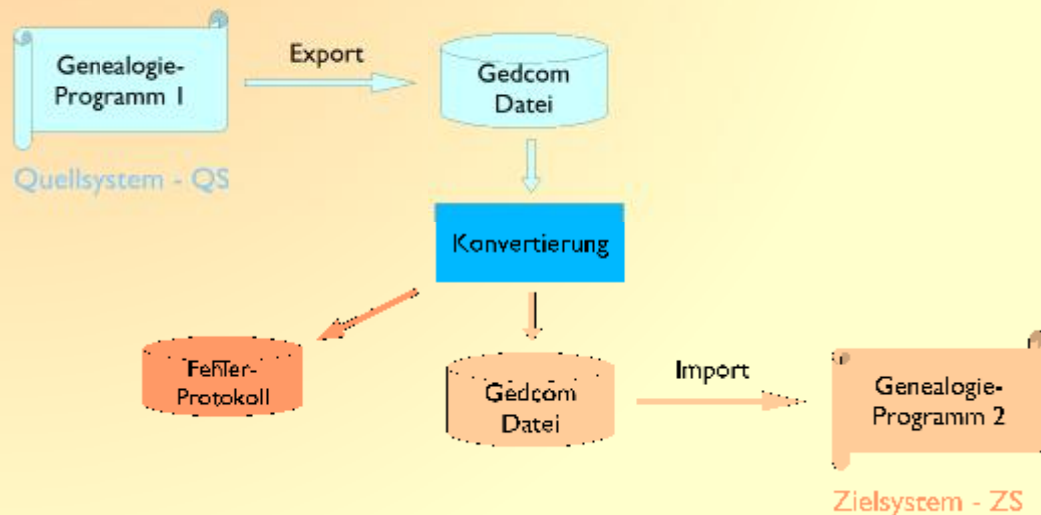
Validierung – 2/2

- **Andere potenzielle Probleme**
 - 8 Fehlerprüfungen
 - z.B. fehlende Tags, Datensätze, Referenzen, Zeiger
 - 27 Warnungsprüfungen
 - z.B. Familie nur 1 Person, nur Kinder, ohne Person
 - z.B. gleicher Nachname von Mann + Frau
 - z.B. vertauschtes Geschlecht bei Mann + Frau / Mutter + Vater
 - z.B. mehrfach identische Datensätze, mehrfache Tags
 - Ergebnislisten – mögliche Fehler & Warnungen
 - Validierungsergebnisliste jedes gefundene Vorkommnis
- Weitere Optionen zur Steuerung des Einlesens der ged-Datei
 - Streichung von Sondertexten in Namen – z.B. NN, N.N, Living, ...
 - Streichung von Sonderzeichen bei Namen
 - Behandlung von hierarchischen Ortsangaben – als PLAC xxx, xxx, , , ,



Konvertierung

Erlaubt die Konvertierung einer ged-Datei in ein Format, dass die Daten vom aufnehmenden Genealogie-Programm (weitestgehend) verlustfrei übernommen werden können



Konvertierung - 1/2

- Änderung der Kodierung der ged-Datei
 - ASCII, ANSEL, (ANSI) → UTF-8, (ANSI)
- Strukturänderungen fest codiert / nicht veränderbar
 - 13 Entfernen bestimmter Zeilen bzw. Textteile in Zeilen
 - 16 Fest vorgegebene mehrzeilige Änderungen (Korrekturen)
 - 12 Datum Konvertierung in Standard „15 MAR 1820“
 - NOTE- und SOUR Datensätze → eingebettete Texte
 - Spezialänderungen wie z.B. Beziehungen für „Brother's Keeper“ in Standard Form
- Textänderungen – variabel definiert
 - 9 Entfernen von Einzelzeilen + Zeilenblöcken
 - 2 Datum Text Änderungen
 - 6 Textänderungen (Zeilenanfang, allgemein, PLAC Texte)



Konvertierung - 2/2

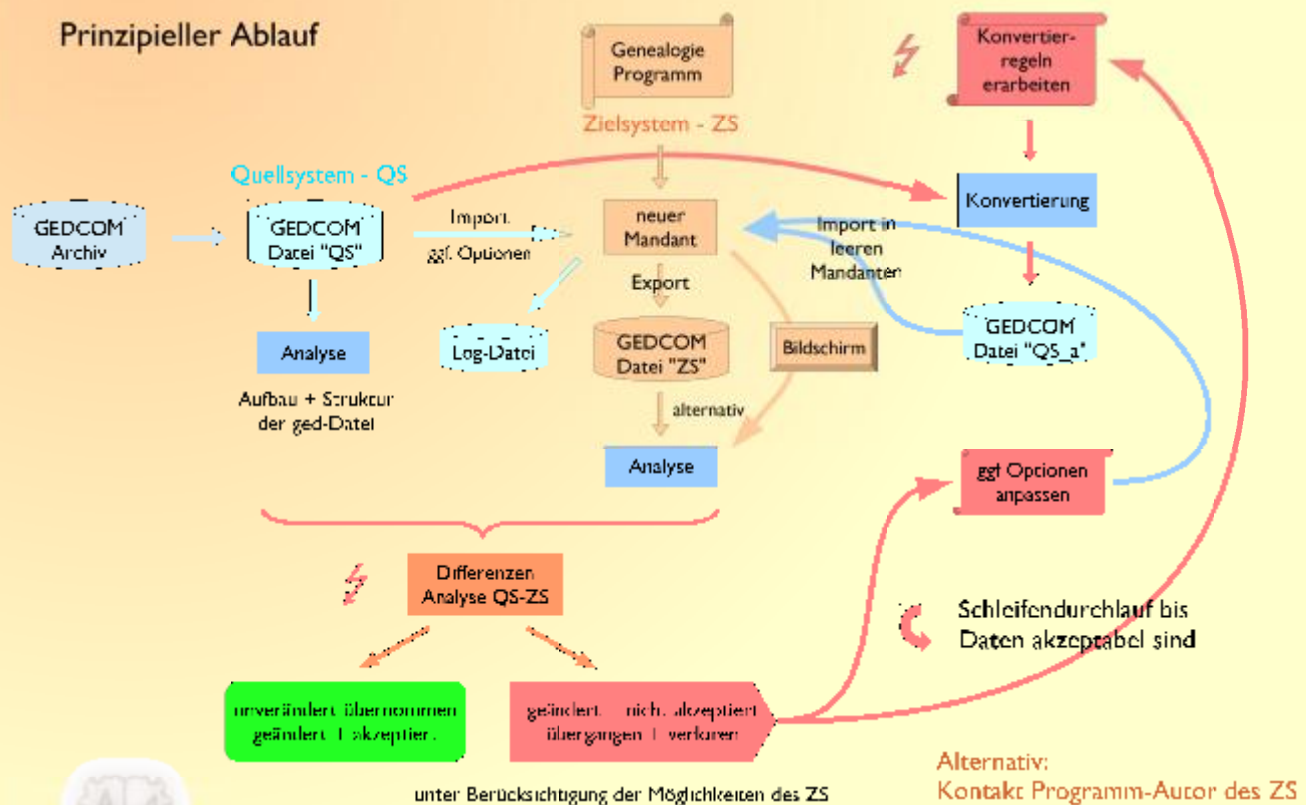
- "Verschiebe Text" Option
 - Variable Strukturänderungen bis zu 5 Zeilen
 - Verschieben bis zu 5 Merkmalstexte
 - Streichen von Merkmalstexten
 - Änderung von Tag-Namen, Stufen-Nr. und Tag-Reihenfolge
 - Streichen, Einfügen und Verschieben von Zeilen
- "Verschiebe Tags" Option
 - Verschieben von Tag-Gruppen unter andere Tags bis zur Stufe 4
- "Transfer Text" Option
 - Verschieben von Daten zwischen Personen- ↔ Familien-Datensätzen
- **Achtung:**
 - Mit großer Vorsicht zu verwenden – gute Gedcom Kenntnisse notwendig
 - Kann bei Fehlern zur „Schrottung“ der Zielsystem ged-Datei führen
 - Reihenfolge der Verarbeitung beachten
 - Dateivergleich durchführen → WinMerge, DiffMerge



GSP (Gedcom Service Programme)

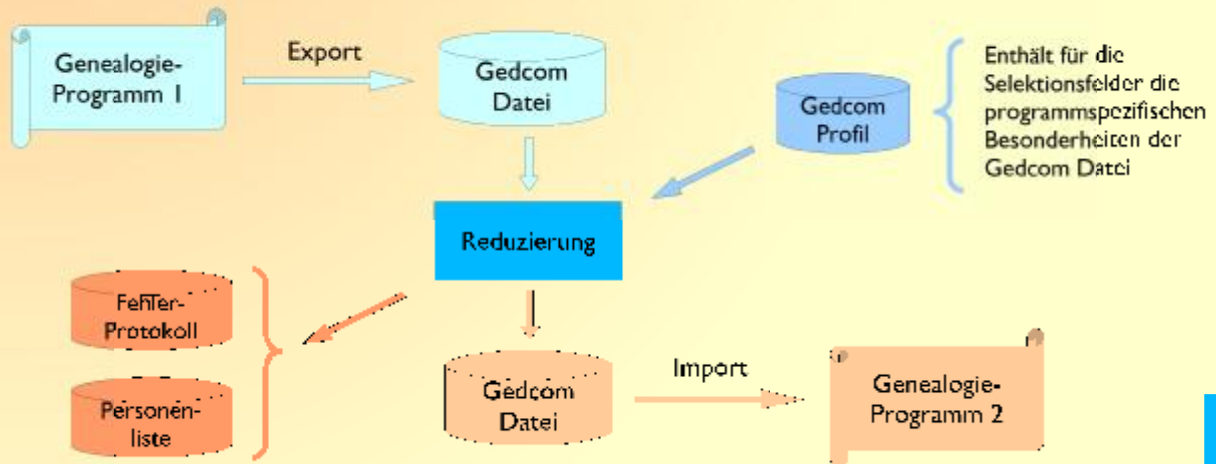
Import von GEDCOM Dateien

Prinzipieller Ablauf



Reduzierung

Erlaubt die Reduzierung einer ged-Datei durch selektive Auswahl einer Teilmenge um die Daten von Einzelpersonen oder ganzen Familienverbänden an andere Forscherkollegen übermitteln zu können



Reduzierung - 1/2

- Selektion von Personen
 - Personenauswahl (20 Selektionsparameter) u/o
 - Ortsauswahl (11 Ortsgruppen)
- Selektives Zufügen von
 - Partner / Kinder / Eltern / ...
 - Vorfahren / Nachkommen
 - Zeugen / Paten / ...
 - Medien – mit verschiedenen Optionen
- Optionales Zufügen von Daten für je 2 gefundene zusammenhängende Personen
 - Familiendaten (FAM Datensätze) bei Partnern / Kinder
 - Beziehungsdaten (ASSO Strukturen) bei Zeugen / Paten / ...

Reduzierung - 2/2

- Auswahl der zu übertragende Datenfelder für gefundene Personen
- Datenschutz Funktionen stehen zur Verfügung
 - Standard DS über Berechnung von Jahren
 - Individueller DS über spezielle Kennzeichen in der ged-Datei
- Benötigt „Gedcom Profil“
 - Definiert wie welche Felder in der ged-Datei kodiert sind
 - Zu definieren für ... – sofern verwendet – ... mit "GSP Profil Verwaltung"
 - Taufe
 - Beruf
 - Familienname
 - Lebens-/Wohnort
 - Bürgerort
 - Beziehungen



GSP (Gedcom Service Programme)

Listen

Erlaubt die Ausgabe von bis zu 10 Feldern je Person aus einer ged-Datei als html- und csv-Datei durch selektive Auswahl der Daten

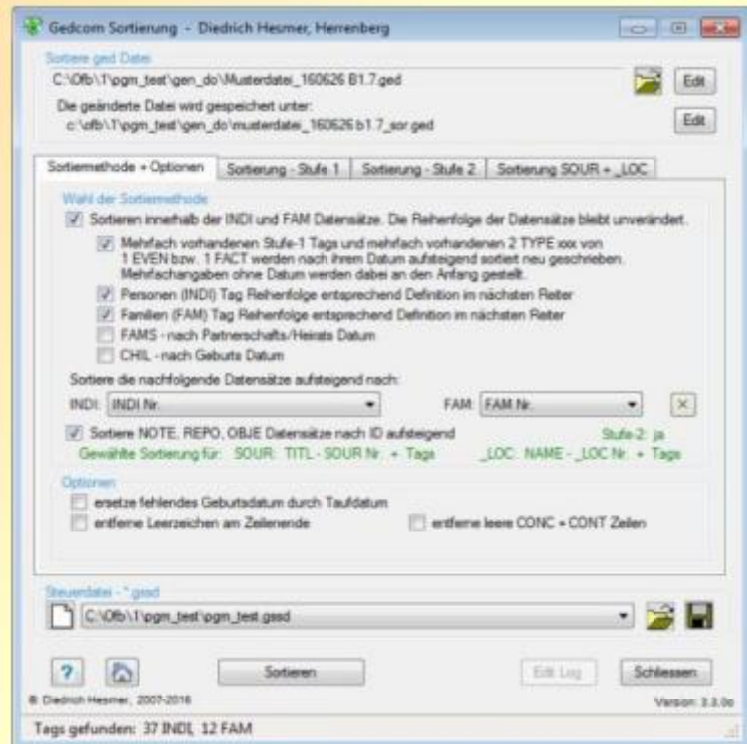
- Selektion von Personen
 - Personenauswahl (20 Selektionsparameter) u/o
 - Ortsauswahl (11 Ortsgruppen)
 - Zusatzauswahl (4 Tags)
- Selektives Zufügen von
 - Partner / Kinder / Eltern
 - Vorfahren / Nachkommen
 - Zeugen / Paten / ...
- Auswahl der auszugebende Datenfelder/Tags für gefundene Personen
- Datenschutz Funktionen stehen zur Verfügung
- Benötigt „Gedcom Profil“

Siehe auch "GSP Reduzierung"



Sortierung

- Gleiche Ereignisse in zeitlicher Folge
 - Berufe, Wohnort, ...
- Umsortierung innerhalb
 - INDI Datensätze
 - FAM Datensätze
- Umsortieren ganzer
 - INDI Datensätze
 - FAM Datensätze
- FAMS nach oo Datum
- CHIL nach * Datum
- weitere Optionen

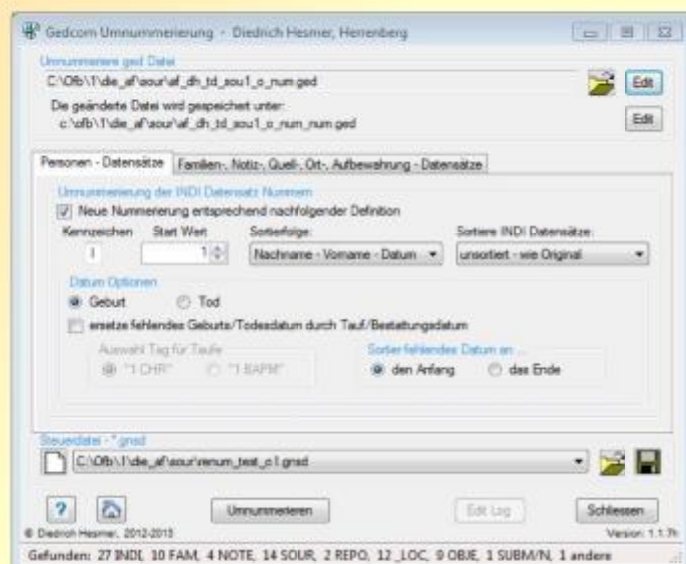


© Diedrich Hesmer, Herrenberg

22

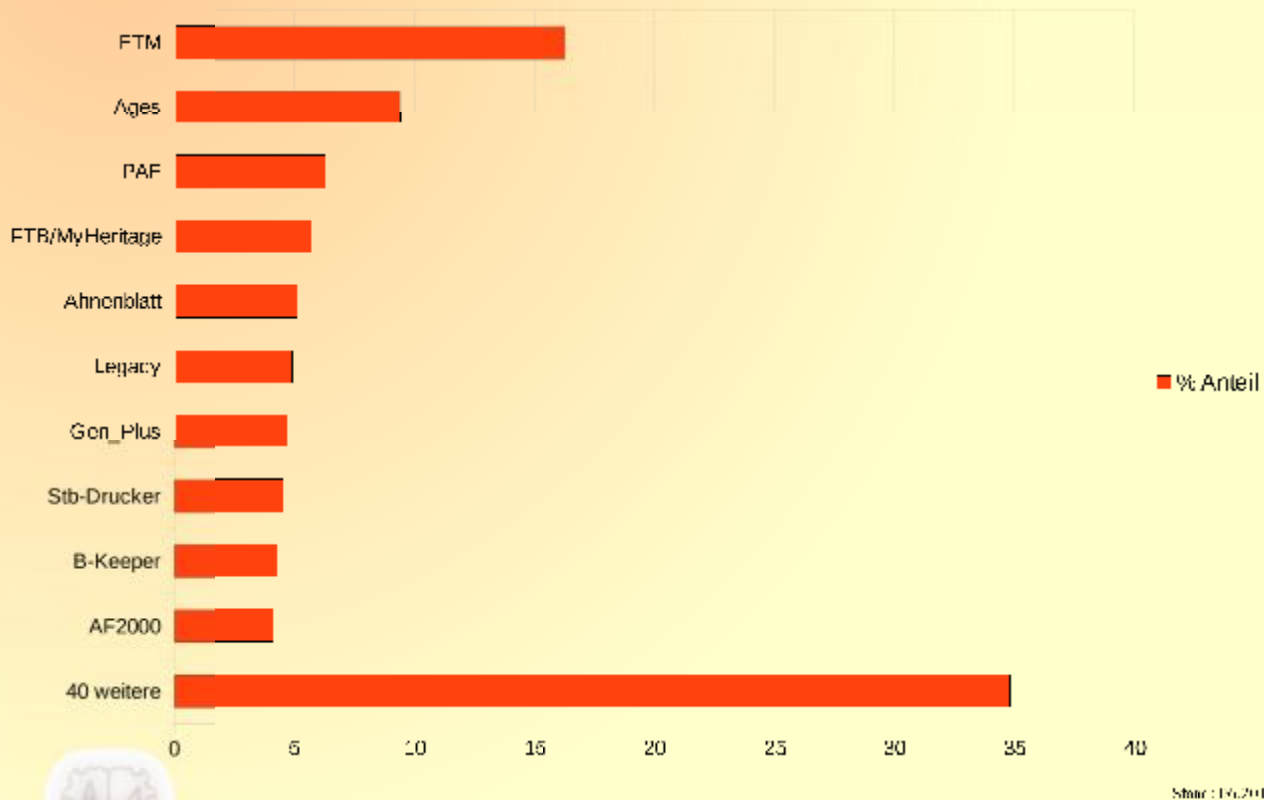
Umnummerierung

- Umnummerierung von INDI, FAM, NOTE, SOUR, _LOC, REPO Nrn.
 - Beliebige Startnummer + Buchstabe
 - Diverse Sortierfolgen (nur bei INDI)
- Sortierung der INDI Datensätze
- Verschiedene Datum Optionen



23

GSP Anwender



GSP Spezial

Programme sind auf Grund spezieller Anforderungen von Anwender-Gruppen entstanden und liefern eine Lösung für spezielle Gedcom Probleme.

6 Programme, davon 3 von allgemeinem Interesse

- Chars2Utf8
 - Konvertierung in UTF-8 Zeichensatz
 - Einzeldateien oder Batch Lauf vieler Dateien
- CorSexHusbWife
 - Prüft Übereinstimmung von SEX mit HUSB + WIFE
 - Erlaubt Korrektur von fehlerhaften Zuordnungen für SEX und Austausch von HUSB ↔ WIFE
- ConvLvl2To1
 - Konvertiert definierte Stufe-2 Tags zu Stufe-1 Tag
 - Bsp: "I NAME mit 2 SOUR" in "I NAME und 1 SOUR"
 - nicht jedes Programm versteht "2 SOUR" → Datenverlust



GSP/OFB Dienstprogramme

4 Programme zur Unterstützung bei Verarbeitung von ged-Dateien mit GSP- und OFB-Programmen

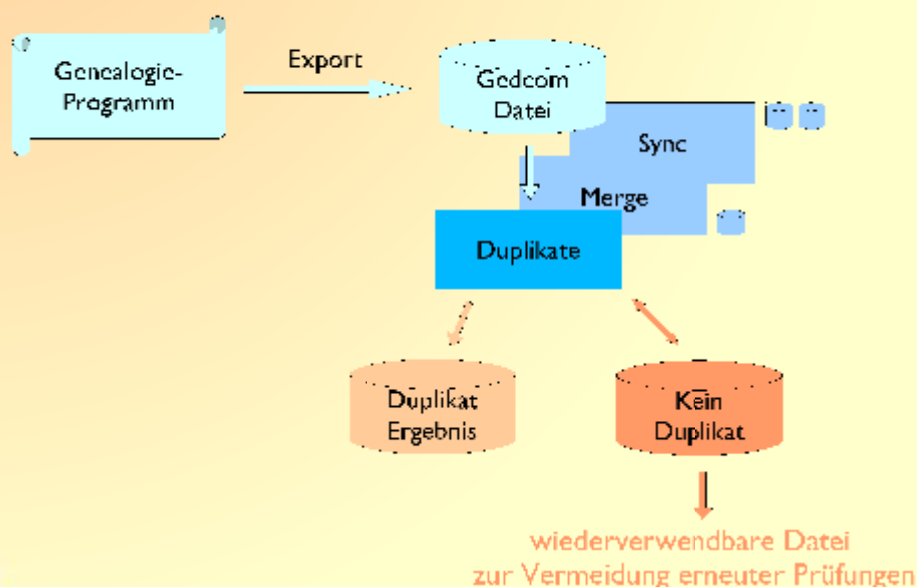
- GedShow - Anzeigen Datensätze der ged-Datei
 - Doppelklick auf Ref-Nr zeigt den Datensatz in separatem Fenster
 - Rechte Maustaste Klick auf Ref-Nr zeigt alle referenzierte DS
- BatchRun - Sequentielle Abarbeitung von Programmen + Befehlen
 - GSP- und OFB-Pgm, Browser, MS Word, Excel, ...
 - Dateien kopieren, umbenennen, löschen, ...
- PersonSearch - Suchen von Personen
 - Suche Personen mit definierten Merkmalen, Orte, Zeiten
 - Bsp: "wer war 1870 Bürgermeister in Buxtehude"
- GedTagSearch - Suchen von Gedcom Tag Kombinationen
 - Sucht in ged-Dateien definierte Tag-Text-Muster
 - Bsp: in FAM alle DS mit "2 HUSB und 3 AGE *" oder in INDI alle DS mit "1 SEX U"



GSP (Gedcom Service Programme)

Duplikate / Merge / Sync

Erlaubt die Überprüfung einer oder zweier ged-Dateien auf doppelte Personen (Duplikate), deren Verschmelzung oder Synchronisierung



Duplikate / Merge / Sync

- **Duplikat Prüfung** innerhalb einer oder zwischen 2 ged-Dateien
 - Vergleiche je zweier Personen von
 - Nachnamen + Vornamen
 - Geburt / Taufe + Tod / Bestattung + Heirat
 - Datum + Ort
 - Vergleichsmethoden
 - Namen + Ortsnamen
 - Kölner Phonetik → deutsche Aussprache
 - Soundex → englische Aussprache
 - 1 : 1
 - Datum mit Zeitfenster
 - Weitere Optionen zur Verfeinerung der Auswahlkriterien
 - Ergebnislisten – mögliche Duplikate
 - Duplikat-Ergebnisliste in Kurzfassung die 2 Personen
 - Ged Duplikat Liste + Datei 1:1 der ged-Datei für die 2 Personen
 - „Kein Duplikat“ Datei – Personenpaare als „nicht Duplikat“ definiert



Duplikate / Merge / Sync

- **Nutzungshinweise**
 - Verwendung unterschiedlicher Einstellungen
 - Je mehr Optionen gewählt um so kleiner Anzahl Duplikate
 - Anfangs nur Nachname + Datum
 - Später Vornamen + Ortsnamen zu/abwählen
 - Johannes, Johann + Hans führen zu unterschiedlichen Ergebnissen, ähnliches gilt für Abkürzungen wie bei Joh., Wilh., ...
 - Bei Vergleich von 2 Dateien sollte die kleinere als 1. Datei genommen werden
- **Ref: Testberichte**
 - "Ein Weg zur Dubletten Bereinigung" – Computergenealogie 1/2010
 - Auch verfügbar im Download-Bereich der OFB Homepage



Merge / Sync

- **Verschmelzung** von Personen + Familien innerhalb 1 ged-Datei
- **Synchronisierung** von Personen + Familien zwischen 2 ged-Dateien
 - Alle gefundene Duplikate stehen zur Verfügung
 - Die jeweiligen INDI- und FAM-Datensätze werden in der Gedcom Form angezeigt
 - Merge: 6 Fenster – je 3 für Personen und Familien
 - links: Original, rechts: Duplikat, Mitte: zusammengeführte Daten
 - Sync: 4 Fenster – je 2 für Personen und Familien
 - links + rechts je 1 Datei
 - Farbliche Darstellung von gleichen, unterschiedlichen + fehlenden Daten
 - Einzelne Zeilen oder Zeilengruppen können kopiert, ersetzt, verschoben, gelöscht oder verkettet werden
 - Halbautomatische Synchronisierung möglich
 - Einzelne Datensätze können in einem Edit-Fenster manuell beliebig geändert werden



©Diedrich Hesmer, Herrenberg

34

GSP (Gedcom Service Programme)

Merge

The screenshot shows the 'Gedcom Merge' application window. It is divided into three main sections for data comparison and merging:

- Personendaten (Left Panel):** Displays data for a person with ID @I1748. Fields include NAME (Susanna /Ackermann/), SEX (F), BIRT, PLAC (Ueberach), DEAT, PLAC (Ueberach), _HEIRATHNAME (Rückert), FAMS (F708), and children (HUSB @I1758, WIFE @I1748, CHIL @I19528 [Susanne /Rickert/ * 5 JUL 1791], MARR, and a note about the source).
- Personendaten (Middle Panel):** Displays data for a person with ID @I1748. Fields include NAME (Susanna /Ackermann/), SEX (F), BIRT, PLAC (Ueberach), DEAT, PLAC (Ueberach), _HEIRATHNAME (Rückert), FAMS (F708), and children (HUSB @I1758, WIFE @I1748, CHIL @I19528 [Susanne /Rickert/ * 5 JUL 1791], MARR, and a note about the source).
- Personendaten (Right Panel):** Displays data for a person with ID @I18. Fields include NAME (Susanna /Ackermann/), SEX (F), BIRT, DATE (6 NOV 1760), PLAC (Ueberach), SOUR (U-Pb-003), DEAT, DATE (22 NOV 1822), PLAC (Ueberach), SOUR (U-Pb-003), LOC (St.468), RELI (rk römisch-kathol.), NOTE (Alter: 62 0 16), COM (Kind Nr. 5), _LEBENSSTADT (Ueberach +Ueberach), and LOC (St.1958).
- Familiendaten (Bottom Section):** Displays data for a family with ID @F708. Fields include FAMS (F708), HUSB (HUSB @I1758), WIFE (WIFE @I1748), CHIL (CHIL @I19528 [Susanne /Rickert/ * 5 JUL 1791]), MARR, and a note about the source.

The interface includes various buttons for navigation and editing, such as 'Optionen', 'Kopiere', 'Lösche', 'Ersetze', 'Verketze', 'Edit', 'Verketze mit "Neu"', 'Verketze mit "nach Text"', 'Verketze mit "vor Text"', 'Edit / Anzeigen', 'Auswählen', 'Speicher Neu', 'Speicher Ged', 'Abbruch', and 'Schließen'. The status bar at the bottom indicates 'in Basis Datei - Anzahl Personen: 397 Familien: 80'.

35

Sync

Gedcom Sync

Personendaten ged-1: @I32@ EG PK P K

0 @I32@ INDI
 1 NAME Elisabetha /Hönche/
 2 GIVN Elisabetha
 2 SURN Hönche
 1 SEX F
 1 BIRT
 2 DATE Abt 12 May 1809
 2 PLAC Dieburg
 2 SOUR #S35@
 1 CHR
 2 DATE 12 May 1809
 2 PLAC Dieburg
 2 LOC #L195@
 2 SOUR #S35@
 1 DEAT
 2 DATE 11 Aug 1885
 2 PLAC Dieburg
 2 SOUR #S35@

24 Z
 11 Z

Familiendaten ged-1: @F12@ Auswahl: @F12@

0 #F12@ FAM
 1 HUSB #I32@
 1 WIFE #I32@
 1 MARR
 2 DATE 9 Feb 1833
 2 PLAC Dieburg
 2 SOUR #S47@
 2 NOTE #NF12@
 1 NOTE Kirchliche Trauung in St.Gallus, Urberach: 24.08.1978
 2 COM1 Civile Trauung am 31.12.1978
 2 COM2 _EUGEN Johann Eder, Spertshausen + Erwin Gehlen, Urber

Personendaten ged-2: @I32@ NoDup: 4

0 @I32@ INDI
 1 NAME Elisabetha /Hönche/
 2 GIVN Elisabetha
 2 SURN Hönche
 1 SEX F
 1 BIRT
 2 DATE 12 May 1809
 2 PLAC Dieburg
 2 SOUR #S41@
 1 CHR
 2 DATE 12 May 1809
 2 PLAC Dieburg
 2 SOUR #S41@
 1 DEAT Y
 1 RELI rk römisch-kathol.
 1 NOTE #NI32@
 1 EVEN _LEBENSORT
 2 TYPE _LEBENSORT

22 Z
 9 Z

Familiendaten ged-2: @F12@ Auswahl: @F12@

0 #F12@ FAM
 1 HUSB #I102@
 1 WIFE #I32@
 1 CHIL #I1021@
 1 MARR
 2 NOTE XXXXXXXXXXXX
 3 COM1 CCCCCCCCCC
 2 PLAC Dieburg
 1 NOTE Kirchliche Trauung in St.Gallus, Urberach: 24.08.1978

Edit Dates in Texteditor
 Optionen Todo Org1 Org2 Neu1 Neu2 Todo Vlog Stop

Anzeigen Datensatz
 Org1 Org2

Verkettete "von" Texte
☒ nach "zu" Text ...mit
☐ vor "zu" Text ...mit

Markieren Speicher 1/2
 Speicher Ged
 Abbruch Schliessen

Speicherort neue Dateien: "C:\06b\1\af\hps_af\sync\Ged-1-LY-b_syn.ged" + "...ged-2-LY-b_syn_syn.ged" 2

in Datei ged1 | 2 - Anzahl Personen: 25 | 34 Familien: 9 | 15

Aktionen
 Verarbeiten:
☒ Person
☐ Familie
 Die markierten Zeilen in ged1/2 ...
 AutoCopy
 Kopieren
 Löschen
 Ersetzen
 Verketten
 entspr. Definition unten
 Edit Zeile
 Edit Feld
 Symbole
 Zelle gleich
 Tag gleich
 Zelle fehlt
 Zelle blockiert
 nicht getestet