SEL-GbOv — toelichting

Het oude pakketje: SEL-GbOv.LZH (Ivo Gilisen, sept. 2012)

Ivo Gilisen verdeelde dit in september 2012 als `sel-gbov.lzh` (aangekondigd in de PGggLIMBURG-nieuwsbrief van 27-09-2012; auteur destijds `ivo.gilisen@telenet.be`). Anders dan SEL-HFDL is dit **wél een VTX-programma**: `SEL-GbOv.VTX v1.0` (04-08-2012).

Doel: personen met **eenzelfde geboorte-/doopdatum of overlijdens-/begrafenisdatum** netjes onder elkaar in een lijst zetten — **datumgenoten**, en zo mogelijke **dubbele personen** opsporen. Het sorteert op die datum en markeert/selecteert opeenvolgende personen met dezelfde datum. Opties: *“ook rekening houden met de 1ste letter achternaam”* en *“ook met de plaats (of deel ervan)”*.

Ivo was in zijn eigen REM-commentaar opvallend **eerlijk** over twee zwaktes:

- *“Programma bevat een schoonheidsfoutje: zowel de eerste persoon als de laatste persoon staat steeds op ON, nodig of niet nodig.”*
- Over de 1ste-letter-optie: *“Naar meer letters kan er niet gekeken worden: Huybrecht = Hoebrechts, Tys = Thijs, Lycops = Lijcops … In het geval 1ste letter Achternaam blijven gevallen als* **Kremers = Cremers** *zonder oplossing.”*

Beperkingen van deze aanpak:

1. **Eerste én laatste persoon altijd gemarkeerd** — een cosmetische fout (Ivo's eigen vaststelling): de buitenste records komen steeds in de selectie, terecht of niet.
2. **Naam enkel op 1ste letter** — spelvarianten met een **andere beginletter** worden gemist (*Kremers/Cremers*), en zelfs binnen dezelfde letter pakt 1 letter te grof (*Huybrecht/Hoebrechts*).
3. **Geen voornaam-vergelijking** — op een groot bestand delen heel veel **onverwante** mensen een datum; zonder voornaam is het een ruwe “datumgenoten”-lijst, geen dubbel-detector.
4. **Geen expliciete terugval-keuze** — geboorte→doop / overlijden→begrafenis liep enkel mee via PRO-GEN's veldvervanging, niet als bewuste, zichtbare optie.
5. **Detecteert, corrigeert niet** — het programma selecteert/markeert enkel; samenvoegen/rechtzetten blijft handwerk.

De oplossing — SEL-GbOv.VTX v2.00

Concept **Ivo Gilisen**, herwerking **Luc Klaps** (gereviseerd 2026).

- **Eerste/laatste-fout weg** — twee fasen: eerst **sorteren + RESTART**, dan per gesorteerd record de **vorige met de huidige** vergelijken; het eerste record na de restart zet enkel “vorige” en markeert niets.
- **Naam-niveau 0/1/2/3** — geen / 1ste letter / **klinkt als** / volledige achternaam.
- **Voornaam mee vergelijken** (J/N) — voor échte dubbels.

- **Plaats (deel) mee vergelijken** (J/N).
- **Expliciete terugval** — geboorte valt terug op doop, overlijden op begrafenis als de hoofddatum leeg is.

Het kerngeval: Cremers/Kremers — de soundex-modellen

Precies Ivo's onopgeloste geval. “Klinkt als” via SOUNDEX lijkt de oplossing, maar er zit een val in: PRO-GEN's `SOUNDEX$` kent **drie modellen**, en **model 0 (Amerikaans) én model 8 (Belgisch) behouden de eerste letter** — dus C ≠ K, en *Klaps/Claps*, *Krijns/Crijns*, *Craeghs/Kraegs* blijven uit elkaar. PRO-GEN's eigen ingebouwde “klinkt als”-selectie gebruikt zelfs model 0 en mist die dus óók.

Model 1 (“alle letters gecodeerd”) codeert **ook de eerste letter**, waardoor C en K in dezelfde klankgroep vallen en wél samenkomen. Bewezen op het eigen bestand: de voorwaarde `SOUNDEX$("Klaps",1) = SOUNDEX$("Claps",1)` selecteerde **alle 70.638 personen** (altijd waar). Daarom gebruikt v2.00 **model 1** voor de klinkt-als-optie, zowel in de sorteersleutel als in de vergelijking.

Praktijktest op ±70.000 personen vond o.a. *Boijen/Boyen*, *Goijens/Goyens*, *Smijers/Smeyers*, *Craeghs/Kraegs*, *Vanceer/Vanseer*, *Luyten/Luten* en *Klaps/Claps* samen.

Wat we bouwen — versie v2.00

Versie-afspraak: v1.0 hoort bij **Ivo Gilisen** (origineel 2012); onze herwerking start op **v2.00 (Luc Klaps, gereviseerd 2026)**, zodat de auteursgrens duidelijk is.

v2.00 omvat:

- eerste/laatste-fout weg;
- naam-niveaus 0–3 met klinkt-als via **model 1** (vangt C/K);
- voornaam- en plaats-optie;
- terugval doop/begrafenis;
- BOX-statusscherm met lopende teller;
- selecteren én markeren vanuit de personenlijst (L/A).

Vuistregel: voor het zoeken van **dubbels altijd voornaam = J** (de standaard staat op G/2/J/N); niveau 0/1/2 zónder voornaam zijn enkel grove datumgenoten-overzichten.

Blijft handwerk: het effectief samenvoegen/rechtzetten — een selectie-VTX schrijft niets terug.

Auteurs

Ivo Gilisen v1.0 (nieuwsbrief 27-09-2012) • Luc Klaps v2.00 (gereviseerd 2026) • Johan Mulderij & D.J. Scholte in 't Hoff.