

# Conversiemethode DB/DC

*Van bron via tussendoel naar zo juist als mogelijk*  
*Willem Mulder, 10 juli 2012*

---

## Wat is conversie

In theorie is een conversie niet meer dan het omzetten van gegevens uit de bron naar een data- en schemaformaat dat bruikbaar is in de doelomgeving. Veelal is voor de werking van het doelsysteem noodzakelijk dat applicatiegebonden parameters toegevoegd worden en er applicatiegebonden parameters van de bron verdwijnen. De toe te voegen parameters worden vastgesteld op basis van algoritmes die toegepast wordt op de te converteren data. Grosso modo kan gesteld worden dat de som van de resultaatgevende gegevens niet zou moeten wijzigen door deze transformatie.

In de praktijk wordt het conversietraject ook gebruikt om de functionele werking van de doelapplicatie te optimaliseren door resultaatgevende gegevens op te schonen. Zo zullen typische (redundante) registraties als gevolg van de functionele werking van de bron ongedaan gemaakt worden, gegevens die niet meer relevant zijn weggelaten worden en gegevens waarvan inmiddels de wenselijkheid tot administratie helder is geworden worden toegevoegd.

Daarnaast komt het uiteraard voor dat de brongegevens onvolkomenheden bevatten die worden hersteld in het conversietraject.

Er gebeuren dus twee dingen:

- transformatie
- correctie

Van beide activiteiten moeten de gevolgen zelfstandig controleerbaar en reproduceerbaar zijn (de correctie zeker gezien Quinto).

## Controleren, reproduceren en traceren

### Controleerbaarheid

#### *Extractie (stap 0)*

De start van iedere conversie is levering van een data-extractie vanuit de bron. Uiteraard moet de juistheid en de volledigheid van deze extractie vastgesteld worden. Deze vaststelling blijft in dit document echter buiten beschouwing. Vanuit implementatie bezien, beschouwen wij de extractie als volledig en juist. Om zeker te weten dat een eventuele nieuwe levering dezelfde groep bevat, stellen wij wel een hash total vast op basis van het somproduct unieke nummer x (maximum van 1 en het opgebouwd pensioen (ouderdomspensioen, nabestaandenpensioen cq wezenpensioen afhankelijk van de status)). Het zou de voorkeur van de organisatie moeten zijn om ook vast te stellen dat de extractie daadwerkelijk de hele betrokken populatie bevat Dit kan door (los van de extractiequery) een hash total direct op de bron te bepalen.

#### *Transformatie (stap 1)*

De gegevens uit de brontabellen worden door middel van queries toegevoegd aan tabellen die nauw aansluiten op het datamodel van Maia. De transformatie is één op één voor wat betreft

unieke registraties in de bron<sup>1</sup>. Zowel redundantie in de bron als onjuiste resultaatgevende gegevens worden overgenomen en toegepast op het model van Maia. Eventuele splitsing van aanspraken<sup>2</sup> zoals deze in Maia geadministreerd gaan worden, worden in deze stap toegepast. In deze stap worden alle Maia-specifieke (stuur)parameters toegevoegd evenals eventuele aanvullende resultaatgevende (bijvoorbeeld basis aanvullende velden).

Er is dus nu een volledige te converteren set aan gegevens ontstaan die één op één aansluit bij de werkelijkheid van de bronadministratie maar dan in Maia formaat.

Die aansluiting is ook controleerbaar omdat zowel in aantalen als in sommen de waarden hetzelfde moeten zijn.

---

<sup>1</sup> Als een deelnemer met deeltijdpensioen in de bron een ingegaan recht heeft en opbouwend recht worden deze in deze fase geconverteerd als een status 50 en een status 20 en (nog) niet 37. Vergelijkbaar worden de aparte registratie van een arbeidsgeschikte en een arbeidsongeschikte strook geconverteerd als 20 en 28 en (nog) niet 25.

<sup>2</sup> Hierbij valt te denken aan bijvoorbeeld hoog/laag en verevende aanspraak. Hoog/laag wordt soms geadministreerd als een tijdelijk hoog en een uitgesteld laag terwijl in Maia een levenslang laag met alleen de delta als tijdelijke aanspraak wordt geadministreerd. Een verevende aanspraak wordt in Maia zelfstandig geadministreerd terwijl in veel pensioenadministratie alleen een aantekening wordt gemaakt dat er verevend moet worden in de uitkeringsfase.

Het hash total somproduct “unieke nummer x opgebouwd pensioen” zou, behoudens afrondingsverschillen in de berekening van het product, volledig aan moeten sluiten. Daarnaast wordt een vergelijking geproduceerd tussen bron en tussendoel per statusgroep per geslacht van de aantallen unieke verschijningen en de sommen van opgebouwde en te bereiken aanspraken per hoofdtype dekking.

#### *Correctie (stap 2)*

Als eerste worden de door te voeren correcties geanalyseerd en per correctie wordt een tabel aangemaakt. Bij de correctie valt te denken aan uitsluiting (deelnemers zonder enige aanspraak), samenvoegen van registratie (bij partieel arbeidsongeschikten, deeltijdpensioen, etc.) en afwijkingen van persoonsgegevens die volgen uit de vergelijking met GBA-gegevens. Uiteraard kunnen correcties het gevolg zijn van onjuiste of nog niet geregistreerde gebeurtenissen, zoals dubbel of niet afgetrokken bijzonder nabestaanden pensioen, gemiste scheidingen.

Al deze correctietabellen bevatten de gegevens die noodzakelijk zijn om een correctie door te voeren aangevuld met de elementen voor de berekening van het hash total. Per correctietabel zal een update-query aanwezig zijn. Het uitsluiten en samenvoegen zal bij bijna alle conversies moeten gebeuren en zijn dus redelijk standaard. Overige correcties zullen sterk verschillen per conversie.

Per correctietabel moet een hash total bepaald kunnen worden. De som van de hash totals over de correctie(s) en het hash total van het eindresultaat moet overeenkomen met het hash totaal van de extractie (stop 0) en de stand direct na transformatie (stap 1).

Na deze bewerking kan een vergelijking geproduceerd worden tussen het de stand direct

na transformatie (stap 1) en de te converteren stand (stap 2) per statusgroep per geslacht van de aantallen resulterende polissen en de sommen van opgebouwde en te bereiken aanspraken per hoofdtype dekking. Van de verschillen kan per groep op het overzicht een specificatie worden gemaakt. Zinnige informatie over de wijziging kan verkregen worden door in de eigenschappen van zowel de tabel als de correctievelden een adequate omschrijving op te nemen.

### Converteren (stap 3)

De doeltabellen worden met behulp van de outputmodule omgezet in tekstbestanden (spatie gescheiden) met door Progress OpenEdge bruikbare opmaak van getallen en data (de zogenaamde .d bestanden). Deze bestanden worden ingelezen in Maia en resulteren in een conversiemutatie. Eventuele fouten die ontstaan bij het inlezen moeten worden opgelost, totdat een foutloze omzetting in een conversiemutatie is bereikt. Het verslag van het aanmaken van de conversiemutaties wordt opgeslagen in de projectmap voor controledoeleinden.

Hierna kunnen de conversiemutaties worden verwerkt. Ook bij het verwerken kunnen foutmeldingen getoond worden. Het verwerkingsverslag van een uiteindelijke foutloze

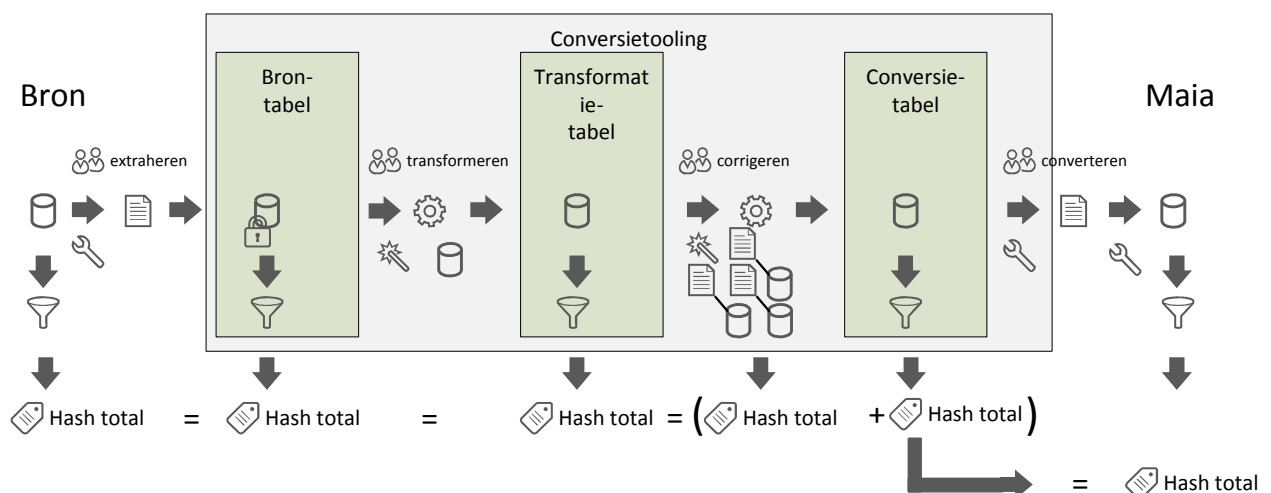
verwerking wordt opgeslagen in de projectmap voor controledoeleinden.

Ter afronding van deze stap wordt een vergelijking geproduceerd tussen de te converteren stand (stap 2) en de geconverteerde stand (deze 3) per statusgroep per geslacht van de aantallen resulterende polissen en de sommen van opgebouwde en te bereiken aanspraken per hoofdtype dekking. Als het aanmaken van de beginstand in Maia goed is verlopen, komen de te converteren en de geconverteerde stand uiteraard volledig overeen.

Na het verwerken van de conversiemutatie wordt doorgaans een zogenaamde nul-mutatie uitgevoerd om de aanspraken door te rekenen op basis van de formuleset van Maia. Deze stap blijft bij de conversie verder buiten beschouwing. De berekening is meer om vast te stellen dat Maia goed rekt en minder bedoeld om vast te stellen dat de conversie goed gaat. Ook de verschillenlijst valt in deze categorie, nuttig voor de gehele implementatie maar minder relevant voor de daadwerkelijke conversie.

### Reproduceerbaarheid

Alle queries die noodzakelijk zijn voor het creëren van de transformatie en de correctie worden opgenomen in macro's. De query wordt door middel van de documenteerfunctie (vanaf Access



**Figuur 1: Schematische weergave conversieproces**

2007) opgeslagen als pdf in de projectmap.

Van de correctietabellen wordt het doel van de correctie, de oorzaak die de correctie noodzakelijk maakt en de verwijzing naar het bronbestand opgenomen in het logboek. Van de ingelezen tabel wordt een pdf overzicht gemaakt en opgeslagen in de projectmap zodat ook de gebruikte data reproduceerbaar blijft.

## Traceerbaarheid

Om de controle vanaf uniek brongegeven tot polis en vice versa te vereenvoudigen komen er twee exportfuncties. De eerste levert de ontwikkeling van bron naar te converteren stand (als deze wordt gehaald) inclusief samenvoegingen, splitsingen en correcties en de andere levert de omgekeerde ontwikkeling vanuit de te converteren stand terug naar de bron(en).

Nog te bepalen:

- Aparte polis verevend: meenemen in stap 1 met hetzelfde unieke nummer en dan telling daar op corrigeren? Carla suggereert het toevoegen van een voorloopcijfer voor het unieke nummer.
- Er zijn vaak meerdere brontabellen, een uniek nummer in de bron hoeft dan op tabel niveau niet meer uniek te zijn!
- Het produceren van verwerkingsverslagen is geen sterk punt van Maia. Ze raken "kwijt" als de verbinding met de client weg valt en bij verweking vanaf meerdere clients zijn ook niet lekker samen te voegen. Als wij ze willen hebben voor controle doeleinden, moeten er goede werkinstructies komen voor het verwerken van de conversiemutatie.
- Gaan wij bij de controle op de conversie de 3-stand vergelijken met de conversietabel of de 4-stand? Voordeel van de 3-stand is dat je afwijkingen als gevolg van de Maia-

formule set negeert en het verschil tussen de 3- en 4-stand via de verschillenlijst kan vaststellen. In een ideaal situatie heb je uiteraard geen verschillen tussen de 3- en de 4-stand en zou je heel stoer moeten gaan voor het vergelijken van het eindresultaat van de conversie.