|  |
| --- |
| Datafabriek CBBS |



Functioneel Ontwerp

Datamart Medische beoordeling CBBS

Datafabriek

Versie 0.1

Datum: 16 -2-2021

# Versiebeheer

|  |  |  |  |
| --- | --- | --- | --- |
| Versie | Datum | Status | Korte beschrijving aanleiding / wijziging |
| 0.1 | 16-02-2021 | Concept | Eerste conceptversie |
|  |  |  |  |
|  |  |  |  |
|  |  |  |  |
|  |  |  |  |
|  |  |  |  |
|  |  |  |  |

Versiebeheer: *Erik Paardekooper*

Opgesteld door: *Erik Paardekooper*

|  |  |  |  |
| --- | --- | --- | --- |
| Afgestemd met: | Status | Datum | Naam / Contactpersoon |
| CEC | Concept |  | Yvonne Vergunst |
| OAA | concept |  | Roepesh Serpal |
|  |  |  |  |

|  |  |  |  |
| --- | --- | --- | --- |
| Besproken met: | Status | Datum | Naam / Contactpersoon |
| CEC | Concept |  | Yvonne Vergunst |
| OAA | concept |  | Roepesh Serpal |
| Datafabriek |  |  | Mossaab Agdi |

Vertrouwelijkheid:

De lezer/gebruiker van dit document wordt geacht de inhoud daarvan vertrouwelijk te behandelen, tenzij uit de toelichting of bronvermelding blijkt dat de informatie als openbaar kan worden beschouwd.

# Document Informatie

In dit document wordt de standaard voor het Functionele Ontwerp (FO) neergezet.

Het FO beschrijft de requirements en de logische datastromen die ten grondslag liggen aan de te realiseren informatieproducten en informatiegebieden, alsmede eventuele de ETL-jobs in het DIM.

Dit FO betreft een afstemmingsdocument voor het laten aansluiten bestaat voornamelijk uit:

* Een beschrijving van de huidige situatie
* Een beschrijving van de requirements
* Een beschrijving van de migratie stappen

Het Functioneel Ontwerp vormt dus de basis voor, en is input voor, het technische ontwikkelwerk.

Voor meer detail over het maken van een Functioneel Ontwerp en de gebruikte standaarden, zie de documenten “[Standaarden modelleren en FO voor het DIM](https://samenwerken.sharepoint.uwv.nl/projecten/uwv/DFDatafabriek/werkmappen/_Fase%20IV/P101%20-%20%20MetaData%20Standaarden/Archief/Modelleren%20in%20IDA%20standaarden.docx)” en “[WoW Modelleren en FO maken in IDA](https://samenwerken.sharepoint.uwv.nl/projecten/uwv/DFDatafabriek/werkmappen/_Fase%20IV/P101%20-%20%20MetaData%20Standaarden/Archief/WoW%20Modelleren%20en%20FO%20maken%20in%20IDA.docx)”.

Inhoudsopgave

[Versiebeheer 2](#_Toc63951116)

[Document Informatie 3](#_Toc63951117)

[1 Inleiding 5](#_Toc63951118)

[1.1 Doelen 5](#_Toc63951119)

[1.2 Achtergrond 6](#_Toc63951120)

[2 Requirements 7](#_Toc63951121)

[2.1 Integration platform architecture 7](#_Toc63951122)

[2.2 Databases 7](#_Toc63951123)

[2.3 Delta / Full 8](#_Toc63951124)

[2.4 Data lifecycle 8](#_Toc63951125)

[2.5 Source systems 8](#_Toc63951126)

[2.6 Migration 9](#_Toc63951127)

[2.7 Horizontal and vertical lineage 9](#_Toc63951128)

[2.8 Actions and activities 9](#_Toc63951129)

# Inleiding

Dit Functioneel Ontwerp (FO) bevat de functionele beschrijving van de informatiebehoefte van DMAP.

In dit hoofdstuk wordt achtereenvolgens de achtergrond, het doelstelling en de scope van de gegevensaanlevering vanuit DIM naar DMAP beschreven.

## Doelen

Het FO zal worden geschreven voor de volgende doeleinden:

* Vastlegging en beschrijving van de BI omgeving.
* Bron voor toekomstige aanpassingen.
* Communicatie met de opdrachtgever.
* Communicatie met de aanleverende divisies.
* Bron voor ontwikkelaars.
* Dient als testbasis en als input voor het testplan.
* De acceptatie van het product.
* Naslag voor de beheerders.

## Achtergrond

### DIM architectuur

DIM is a new AVG-compliant and governed Enterprise DataWarehouse. In FUGEM/RLO the sensitivity of data is defined.

The data architecture is split into two separate parts:

1. a staging/bronzone, where the source data is processed and administered.

This data is only available via an Ontkoppelview (OKV). This OKV is split into a masked and an unmasked version.

1. a integrationzone/bedrijfszone where information products for end-users (Afnemers) are created.

A gegevensvenster (GGV) is a basic information product and is a view on an OKV table. Based on the end-users “doelbinding”, the end-user will be granted access to the masked or unmasked version, never to both!

The main goal of DIM is to replace DWH 3.0.

### DMAP architecture

Just like DIM, DMAP is an Enterprise Datawarehouses, but is not AVG-compliant and based on a Microsoft SQLServer environment. DMAP is an analytics environment and is (temporary) created to fulfil the informational needs of the Handhaving & K&S (Klan ten Service).

Technical DMAP maintenance is not yet part of DataFabriek, but will be soon.

DMAP gets the data from DWH 1.0, DWH 2.0 and DWH 3.0 and because these data warehouses will be replaced by DIM, it needs to get the data from DIM.

Product owner Handhaving: Ewoud Vegt (Manager Handhaving / Datafabriek stuurgroep)

Product owner K&S: Ben van der Hee (DIAscoop stuurgroep)

### Differences DMAP / DIM at a glance

|  |  |  |  |
| --- | --- | --- | --- |
|  | DMAP | DIM | More details |
| Data Integration platform | Microsoft SQLServer ETL | IBM Information Server platform | 3.1 |
| Environments | P en P’(accent) | O,T,A en P | 3.1 |
| Databases | MS SQLServer (16TB) | Oracle | 3.2 |
| Source data process | Delta  Pull | Delta & Full  Push / Scheduled runs | 3.3 |
| Source systems | DWH 1.0, 2.0, 3.0 | GINA, RESAFasa, etc | 3.5 |
| Data lifecycle – history | Ja | Ja | 3.4 |
| Data lifecycle - Physical deletes | Nee | Ja | 3.4 |
| Maskering sensitieve data | Actief | Vooraf | 3.6 |
| Registration Primary keys (PKs) | Id\_hist | Id\_hist |  |
| Horizontal and vertical Lineage | Not available | Available | 3.7 |

Tabel 1 Voorbeeld overzicht ingangsdocumentatie

| **Ingangsdocument** | **Datum** | **Auteur** |
| --- | --- | --- |
| Requirementsdocument | xx-xx-xxxx | <auteur> |
| GEB | xx-xx-xxxx | <auteur> |
| Juridische documenten | xx-xx-xxxx | <auteur> |
| Interface-beschrijvingen | xx-xx-xxxx | <auteur> |

# Requirements

## Integration platform architecture

DMAP (MicroSoft) requires DIM (IBM) data stored in the Oracle database. Both platforms are independent from eachothre.

DMAP Production will use data from DIM Production (PDIM).

DMAP P-Accent will use data from DIM Acceptance (ADIM).

DMAP does not require access to DIM O & T environments.

DMAP will have to open ports to the Oracle databases ADIM and PDIM. DMAP will have to be whitelisted into the Oracle databases. 2

## Databases

DMAP will have to retrieve the DIM data by querying the DIM Oracle database.

On DIM Acceptance and DIM Production there will be created an user account for DMAP.

DMAP will get unmasked data from DIM Production and from DIM Acceptance.

DMAP will have read access to logging tables to know that source data is processed in DIM.

The DMAP GEB is already approved by the functionaris Gegevensbescherming

Per source, DMAP needs official GLA approval to have access and use this data. Initially DIM-GINA approval will be requested. (Borgmeyer, Arjen)

If DMAP databases are not written in Oracle, DMAP will have to change their queries to Oracle queries (concerns functions, datatime formats, datatypes).

The column names are equal to the source columns. DIM added some additional administration columns that can be used. Schema names and table names will have to be changed in the DMAP queries.

## Delta / Full

DIM is usually processing delta loads from the source systems. If this delta does not contain source delete indicators, periodically a full load is processed to correct this in the DWH.

If the source does not provide a complete/correct set of data, DIM will not process the data delivery as a whole. In this case, the data will not be available for DMAP also.

DMAP can get access to logging tables to know if data is processed.

## Data lifecycle

DIM follows data lifecycle of the source systems.

DIM will process a logical delete first, before it will be physical deleted after 5 (?) years period.

According AVG, data must be physically deleted when it exceeds an particular age. In DIM this will take place. (Needs to be defined how DMAP can handle this situation)

DIM build up history. In the GGV a start\_date, end\_date and change indicator will be provided. This change indicator is a I or U, not a D!

DMAP is only interested in the current state. This can be accomplished by only processing the DIM data with the end\_date ’31-12-9999’, but this will not include the source deletes.

## Source systems

DMAP uses data from these sourcesystems: SMF, CBBS, IMF, ResaFasa, Excasso, UZS, GINA, WWO, BAS,Peoplesoft HRM and UPA

## Migration

DMAP will migrate from DWH 1,2 or 3 sources to DIM sources, source by source, starting with GINA.

## Horizontal and vertical lineage

DIM fullfills in the needs of supplying horizontal and vertical lineage. DMAP does not have this functionality. Afnemers of DMAP can get access to the IGC console to search on Business Terms and Categories associated to the data that DIM supplies to DMAP. This way technical data is supplied including context.

## Actions and activities

* Request GEB / GLA for use of DIM GINA data
* Grant DMAP access to DIM GINA data
* Data/impact analyses on DMAP changes
* Data/impact analyses on Data Lifecycle situations