Rijksinstituut voor Volksgezondheid
en Milieu
*Ministerie van Volksgezondheid,
Welzijn en Sport*

A. van Leeuwenhoeklaan 9
3721 MA Bilthoven
Postbus 1
3720 BA Bilthoven
www.rivm.nl

Project Start Architectuur
COVID Informatie Rapportage

Document informatie

Auteur	:	5.1.2e 5.1.2e 5.1.2e 5.1.2e, 5.1.2e 5.1.2e en 5.1.2e 5.1.2e (editor)
Datum	:	25 maart 2021
Versie	:	1.0
Status	:	Definitief

Documentversies

Versie	Datum	Opmerkingen
0.1		Initiële versie
0.4		Technologiekeuze (demarcatiepunt) toegevoegd
0.5		Pseudonimiseren, datamanagement en inrichting datavirtualisatie toegevoegd
0.5.5		Pseudonimiseren, datamodel en links naar use case toegevoegd
0.7	31-12-2020	Structuur aangepast, dataroutes en rollen toegevoegd, review commentaar verwerkt
0.8		Keuze pseudonimisering toegevoegd; review commentaar PT, EdN, NV, NvdM verwerkt
0.86	27-1-2021	Keuze technologie pseudonimisering toegevoegd.
0.9	12-2-2021	Defintief concept voor oplevering eerste release BI-CIMS
0.91	24-3-2021	Commentaar EdN verwerkt
1.0	25-3-2021	Goedgekeurd door leadarchitect RIVM

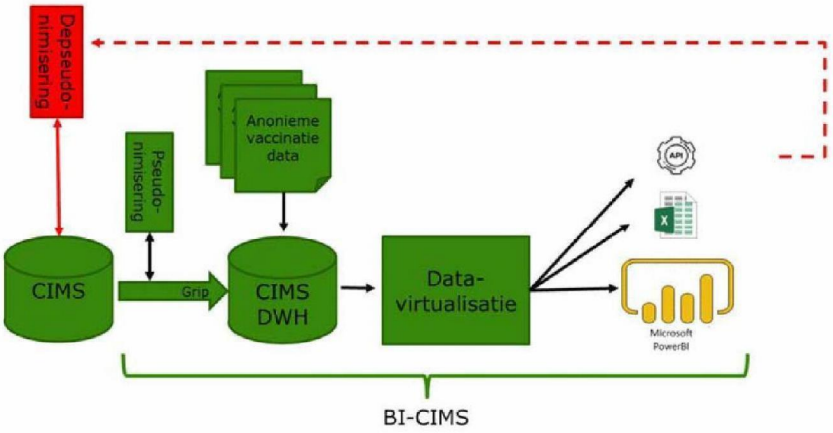
Referenties

Documenttype	Documentnaam/ Verwijzing	Datum
Project Mandaat (*)		
Business Case		
Requirement Analyse		
Projectformulier		
Business Architectuur Klant		

NB de met (*) gemarkeerde velden zijn verplicht.

Managementsamenvatting

Beknopte schets



Kerneigenschappen

Deze PSA beschrijft de architectuur van de rapportagefuncties voor het CIMS-rapportageregister. Door instemming met deze PSA kiest het RIVM voor:

- Uitvoer naar meerdere kanalen in meerdere formaten.
- Inzet van het RIVM-brede platform TIBCO Datavirtualisatie voor eenvoudig en veilig (her)gebruik van data voor onderzoek.
- Tijdreizen mogelijk. Het CIMS-datawarehouse bewaart historie.
- Combineren van gegevens over vaccinaties uit het CIMS-vaccinatieregister en anonieme gegevens in één datawarehouse.
- Separate rapportageomgeving, zodat rapportages geen belasting vormen voor het CIMS vaccinatieregister.
- Pseudonimisering. Alle verwerkingen in BI-CIMS zijn gepseudonimiseerd. Data-analyse vindt plaats met zeer beperkt privacyrisico. Indien aan voorwaarden voor het doorbreken van de pseudonimisering is voldaan, kunnen de persoonsgegevens van individuele gevaccineerden echter benaderd worden.
- Functiescheiding tussen data-analisten, beheer CIMS-DWH, sleutelbeheer en depseudonimisering.
- Audit-trail van alle dataverwerkingen. Datavirtualisatie voorziet hierin.

Inhoudsopgave

Leeswijzer

In hoofdstuk 1 wordt aangegeven wat de scope en doelen van het project zijn. In hoofdstuk 2 worden de gevraagde rapportages vastgelegd. In hoofdstuk 3 wordt het datamodel en bijbehorende -management van de COVID Informatie & Monitoring systeem beschreven. In hoofdstuk 4 worden de koppelingen, rapportage-, en basisapplicaties uitgewerkt. In hoofdstuk 5 worden relevante principes, standaarden en (security) richtlijnen geselecteerd.

1 Inleiding

Deze Project Start Architectuur (PSA) is een projectdocument dat als hulpmiddel bij het project wordt ingezet om veranderingen te faciliteren. De PSA richt zich daarbij op kaders die op dit project van toepassing zijn en de impact van deze kaders op de beoogde verandering. De PSA maakt concreet wat architectuur voor dit project betekent.

1.1 **Doelstelling**

Dit project heeft als doel een functie te realiseren die automatisch rapporten kan leveren voor de monitoring en evaluatie van het vaccineringsprogramma voor COVID-19. Dit betreft voornamelijk rapporten over de opkomst, vaccinatiegraad, effectiviteit en veiligheid. Maar gezien de noodzaak kennis over de ziekte en vaccinaties te ontwikkelen zullen er daarnaast specifieke vragen komen die snel en flexibel beantwoord moeten kunnen worden.

De functie voorziet ook in de mogelijkheid om snel nieuwe bronnen aan te sluiten en resultaten eenvoudig te ontsluiten. Denk hierbij aan pdf-bestanden (Website) en Open data (APIs).

De functie is in staat om bruikbaarheid, integriteit en de veiligheid van de data te waarborgen (Governance). De oplossing borgt data lineage, datakwaliteit, beveiliging van data en ondersteunt master-, reference- en metadatamanagement.

1.2 **Scope, afbakening**

Het Covid-19-vaccinatieregister en de bijbehorende rapportagevoorziening zijn afgeleid van de voorziening voor het Rijksvaccinatieprogramma (RVP). Voor de realisatie van COVID Informatie Rapportage wordt in eerste instantie gebruikt gemaakt van de bestaande middelen van Praeventis. Voor nieuwe functionaliteit (bijv. rapporten) wordt gebruikt gemaakt van Datavirtualisatie.

Tot de scope behoort de verwerking van geaggregeerde voor business analyses en rapportage door DVP. DVP levert gegevens aan andere organisatieonderdelen binnen en organisaties buiten het RIVM. De verwerking buiten DVP behoort niet tot de scope van deze oplossing.

1.3 **Management uitgangspunten**

Deze PSA geeft invulling aan de functionaliteiten beschreven in de Use Case COVID-vaccinatieregister voor wat betreft de daarin genoemde monitoring-functionaliteit.

1.4 Stakeholders

Rol	Belang
Data Eigenaar 5.1.2e	Verantwoordelijk voor het ontsluiten van data naar de functie.
Data Gebruikers 5.1.2e 5.1.2e 5.1.2e 5.1.2e, DVP, EPI	
Business Analisten	Verantwoordelijk voor de analyse van de bedrijfsprocessen
Data Architecten 5.1.2e 5.1.2e 5.1.2e 5.1.2e, 5.1.2e 5.1.2e, 5.1.2e 5.1.2e	Verantwoordelijk voor de kaders waarbinnen de functie wordt gerealiseerd
Data Analisten 5.1.2e 5.1.2e, 5.1.2e 5.1.2e (DVP), 5.1.2e 5.1.2e 5.1.2e (EPI)	Verantwoordelijk voor verwerking van data
Data Beheer 5.1.2e 5.1.2e (DVP), 5.1.2e 5.1.2e (Campus), 5.1.2e 5.1.2e, <Ordina>	Verantwoordelijk voor functioneel en technisch beheer
Informatiemanager 5.1.2e 5.1.2e	Verantwoordelijk voor de informatie stromen naar en van de functie
Hoofd BIS 5.1.2e 5.1.2e5.1.2e 5.1.2e	Verantwoordelijk voor het beheer van de oplossing
Product owner (Datavirtualisatie) 5.1.2e 5.1.2 5.1.2e	Verantwoordelijk voor de levering van Datavirtualisatie functionaliteiten
Leverancier Datavirtualisatie Connected Data Group	Leverancier van Datavirtualisatie functionaliteiten
Leverancier GRIP Grip op Data	Leverancier van de GRIP-functionaliteiten
Projectleider, Scrum Master 5.1.2e 5.1.2e 5.1.2e	Verantwoordelijk voor het realisatieproces van de functionaliteit(en)
Chief Data Officer 5.1.2e 5.1.2e	Verantwoordelijk voor de governance op en het gebruik van gegevens van het RIVM
Data Steward <DVP Data Steward>	Verantwoordelijk voor het beheer van data
Security Officer (CISO) 5.1.2e 5.1.2e	Verantwoordelijk voor de veilige werking van de dienstverlening.
Privacy Officer 5.1.2e 5.1.2e, 5.1.2e 5.1.2e	Verantwoordelijk voor de verwerking van persoonsgegevens.
Enterprise Architect CIO-O RIVM 5.1.2e 5.1.2 5.1.2e, 5.1.2e 5.1.2e	Verantwoordelijk voor de Enterprise architectuur van het RIVM

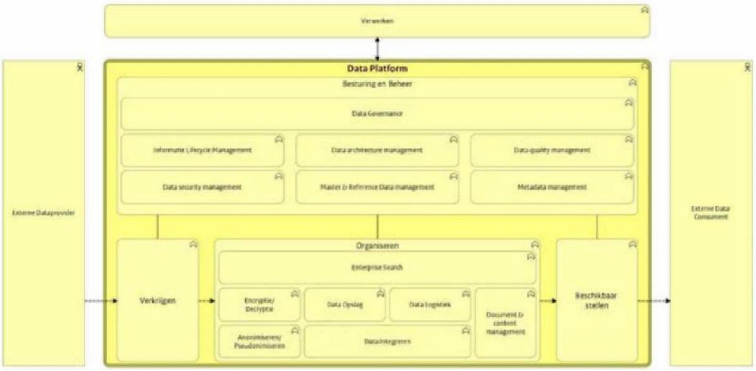
1.5 Algemene Richtlijnen

De volgende kaders en richtlijnen uit de RIVM Data Strategie en de Nederlandse Overheid Referentie Architectuur (NORA) zijn van toepassing:

ID	Enterprise Architectuurprincipe
HP 1	RIVM voegt waarde toe aan data
HP 2	Interne en externe samenwerking
HP 3	Citizen Science
HP 5	Gebruiker Centraal
HP 8	Duurzaam en Circulair
DP 1	Data zijn de grondstof van het RIVM
DP 2	Datamanagement
DP 3	Data zijn open en vrij toegankelijk
DP 5	Standaardiseer en documenteer data
DP 6	Toegevoegde waarde en communicatie
BP 1	Afnemers krijgen data waar ze behoefte aan hebben
BP 2	Afnemers kunnen data eenvoudig vinden
BP 3	Afnemers hebben eenvoudig toegang tot de dienst
BP 4	Afnemers ervaren uniformiteit in de dienstverlening door het gebruik van standaardoplossingen
BP 5	Afnemers krijgen gerelateerde diensten gebundeld aangeboden
BP 6	Afnemers hebben inzage in voor hen relevante informatie
BP 8	Afnemers kunnen erop vertrouwen dat informatie niet wordt misbruikt
BP 9	Afnemers kunnen erop vertrouwen dat de dienstverlener zich aan afspraken houdt
BP 10	Afnemers kunnen input leveren over de dienstverlening

Deze principes en de implicaties voor datavirtualisatie zijn verder uitgewerkt in **Bijlage A: COVID Vaccinatiemodel.**

De principes voor privacybescherming (Privacy-by-design) zijn opgenomen in paragraaf 5.2.



Figuur 1 Architectuurkaders

1.6 Referentie Architectuur

ID	Referentie Architectuur	Relevantie
DAMA	Data Management Body of Knowledge	Internationale kaders en richtlijnen voor het inrichten van de informatie-huishouding
NORA	Nederlandse Overheid Referentie Architectuur (NORA)	Kaders en richtlijnen voor het inrichten van de informatie-huishouding van de Nederlandse overheid
EAR	Enterprise Architectuur Rijksdienst (EAR)	Kaders en richtlijnen voor het inrichten van de informatie-huishouding van de Rijksdienst
RIVM	Campus Architectuur	Kaders en richtlijnen voor de RIVM-IV (Campus) informatievoorzieningen

1.7 Wet- en regelgeving

ID	Wet- en Regelgeving/ Norm	Relevantie
BIO	Baseline Informatiebeveiliging Overheid	Een 'In Control Verklaring' (ICV) is verplicht voor het leveren van deze functionaliteit.
AVG	Algemene Verordening Gegevensbescherming	Een 'Data Privacy Impact Assessment' (DPIA) is verplicht voor het leveren van deze functionaliteit.
VIRBI	Voorschrift Informatiebeveiliging Bijzondere Informatie	Rubriceren van gegevens
WPG	Wet Publieke Gezondheid	Registreren van infectieziekten (Osiris-AIZ)
WABVPZ	Wet aanvullende bepalingen verwerking persoonsgegevens in de zorg	Waarborgen voor cliënten bij uitwisseling van medische gegevens
COVID	Tijdelijke wet maatregelen COVID-19	Tijdelijke maatregelen ter bestrijding van de COVID-19 epidemie

1.8 Afbakening en relaties met andere producten en projecten

Afhankelijkheid	Opmerkingen	Relatie
CIMS	Corona Informatie en Monitoring Systeem	Bron van de gegevens
Datavirtualisatie	Creëren van geïntegreerde, virtuele views op data	Platform van de rapportage-functie

Afhankelijkheid	Opmerkingen	Relatie
RIVMdata	Het RIVM-dataregister	Bevat beschrijvingen van binnen RIVM beschikbare datasets

1.9 Verwijzingen

ID	Referentie
ARC	Architectuur richtlijnen SSC-Campus v1.1
ASV	AVG Stappenplan verwerkingsverantwoordelijke
AVG	Algemene Verordening Gegevensbescherming (AVG) 27 April 2016
BIO	Baseline Informatiebeveiliging Overheid
CIMS	PSA Corona Informatie en Monitoring Systeem (CIMS), P.F.W.M. Thuis
DAMA	DAMA International Guide to Data Management Body of Knowledge https://dama.org/content/body-knowledge
DV	Doelarchitectuur Datavirtualisatie, 5.1.2e en 5.1.2e 5.1.2e
GLDV	Leveraging Data Virtualization in Modern Data Architectures (Gartner)
REA	RIVM Enterprise Architectuur
USE	Use Case covidvaccinatie, DVP 195 v0.9, 5.1.2e 5.1.2e 5.1.2e en 5.1.2e 5.1.2e
DOEL	Centraal register COVID-19 vaccinatie, gezamenlijke notitie van RIVM, LAREB & VWS, 22 december 2020

1.10 Afkortingen

ID	Referentie
ANSI	American National Standards Institute
API	Applicatie Programmeer Interface
BIO	Baseline informatiebeveiliging Overheid
DAMA-DMBOK	Data Management Body of Knowledge
DV	Datavirtualisatie
DWH	Datawarehouse
ETL	Extract-Transform-Load
ISO	International Organization for Standardization
IV	Informatie Voorziening
NORA	Nederlandse Overheid Referentie Architectuur
ODBC	Open Database Connectivity
DPIA	Digital Privacy Impact Assessment
RDBMS	Relational Database-Management System
REST	Representational state transfer
RVP	Rijksvaccinatieprogramma
SOAP	Simple Object Access Protocol

ID	Referentie
SQL	Structured Query Language
SOC	Service and Organization Controls
VE	Vaccinatie-effectiviteit
XML	Extensible Markup Language

2 Rapportages

Dit hoofdstuk beschrijft de inmiddels bekende typen rapportages. De benodigde rapportages zijn onder andere beschreven in de Use Case COVID-vaccinatieregister [ref. USE]. In paragraaf 2.1.9 van de Use Case staan eisen opgenomen met referentie **[MONxx]**. In dit hoofdstuk staan verwijzingen naar de Use Case op dezelfde manier genoteerd.

Op businessniveau is het zogenaamde 'Doelendocument' [ref. DOEL] van belang. In het vierde hoofdstuk van dit document staan de rapportagedoelen beschreven, samen met de daarvoor benodigde gegevens.

Dit hoofdstuk wil vanuit de benodigde rapportages vaststellen welke architectuur van de oplossing nodig is. Het is nadrukkelijk niet bedoeld als gedetailleerde functionele specificatie van de individuele producten die BI-CIMS op moet leveren. Hoofdstuk 3 bevat een schets van de informatie in BI-CIMS en het bronsysteem CIMS.

2.1 **Opkomst**

In de Use Case staat dit gedefinieerd als **[MON01]**. Het rapport over de opkomst bevat de volgende cijfers:

- Het aantal cliënten dat in aanmerking komt voor vaccinatie. (Het RIVM kan dit slechts rapporteren voor zover zij over deze informatie beschikt. Doelgroepen op basis van geboortedatum en postcode kan het RIVM onderscheiden. Doelgroepen met medische indicatie of zorgverleners zijn echter niet bekend).
- Het aantal cliënten dat uitgenodigd wordt. (Het RIVM kan dit slechts rapporteren indien de oproep middels CIMS gedaan is of daarin geregistreerd is.)
- Het aantal cliënten dat de 1e prik (vaccinatie) ontvangen heeft
- Het aantal cliënten dat de 2e prik (vaccinatie) ontvangen heeft (optioneel; afhankelijk van vaccin)

Deze cijfers worden dagelijks gerapporteerd en kunnen uitgesplitst worden naar regio, indicatie en vaccin. De eerste indicaties zijn: medisch, beroep en leeftijd.

De cijfers worden teruggekoppeld aan de opdrachtgever en zorgverleners via standaardrapportage, dashboard en/of open data.

2.2 **Vaccinatiegraad**

Het rapport over de vaccinatiegraad bevat de volgende cijfers:

- Het aantal cliënten dat in aanmerking komt voor vaccinatie
- Het aantal cliënten dat de 1e prik (vaccinatie) ontvangen heeft

- Het aantal cliënten dat de 2e prik (vaccinatie) ontvangen heeft (optioneel; afhankelijk van vaccin)

Deze cijfers worden periodiek gerapporteerd en kunnen uitgesplitst worden naar *indicatie*. Het rapport houdt rekening met de wijzigingen in het cliëntrecord gedurende deze periode. Denk hierbij aan nieuwe ingezetenen met of zonder vaccinatiebewijs en cliënten die Nederland verlaten hebben.

De cijfers worden teruggekoppeld aan de opdrachtgever en zorgverleners via standaardrapportage (RIVM-rapport).

2.3 **Effectiviteit**

Het rapport over de effectiviteit bevat de volgende cijfers:
- Effectiviteit tegen sterfte
- Effectiviteit tegen ziekenhuisopname
- Effectiviteit tegen ziekte

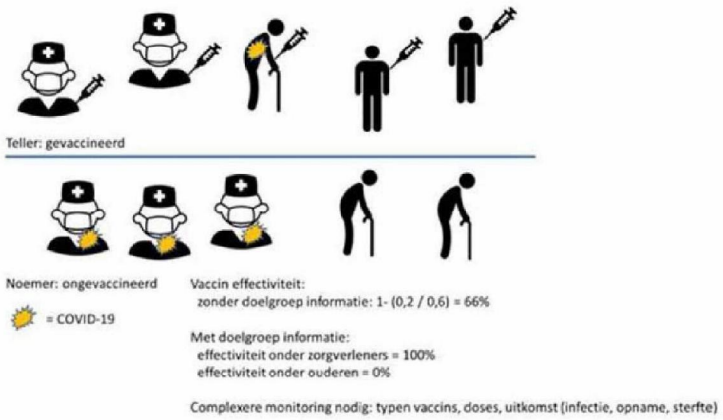
Deze cijfers worden periodiek gerapporteerd en kunnen uitgesplitst worden naar *medische indicatie*, *vaccin* en *leeftijd*.

Manieren om de vaccinatie-effectiviteit (VE) te berekenen:
- Screening methode: formule, waarin percentage van het aantal ziektegevallen* dat gevaccineerd is en de vaccinatiegraad in de bevolking worden gebruikt.
- De verhouding tussen het aantal ziektegevallen* onder de gevaccineerden en het aantal ziektegevallen* onder de ongevaccineerden.

Liefst wil je VE per doelgroep (v.b. medische indicatie) kunnen uitrekenen. Dan moet je voor methode 1 dus de vaccinatiegraad onder de mensen met een medisch indicatie weten en voor methode 2 moet je weten van alle ziektegevallen[1] onder mensen met een medisch indicatie weten of ze wel/niet gevaccineerd zijn.

Zie Figuur 2 met een apart pictogram voor elk van de 3 doelgroepen (zorgmedewerkers, leeftijd, medische indicatie).

[1] met ziektegevallen bedoelen we dan: meldingen van de ziekte in Osiris, opnames in het ziekenhuis vanwege die ziekte of sterfte als gevolg van die ziekte.



Figuur 2 Vaccinatie-effectiviteit

Bronnen voor ziektegevallen:
- Osiris (vragenlijst)
- Nationale Intensive Care Evaluatie[2] (NICE)

Bron sterftecijfer: <CBS-sterfteregistratie>

Dit zijn bronnen zowel binnen als buiten het RIVM. Koppeling met bronnen is eventueel op termijn mogelijk, wanneer hiervoor de juridische grondslag en de data beschikbaar zijn.

De cijfers worden teruggekoppeld aan de opdrachtgever en zorgverleners via standaardrapportage, dashboard en/of open data.

2.4 **Veiligheid**

De rapportage over vaccinatieveiligheid wordt verzorgd door het LAREB. Hier worden meldingen van bijwerkingen verzameld. DVP verstrekt vaccinatiegegevens van de gevaccineerde op verzoek aan bijwerkingen centrum LAREB. Dit is echter geen rapportagefunctie, maar een functie op CIMS zelf.

Daarnaast moet DVP verschillende rapporten kunnen opleveren voor dit doel. Belangrijke voorbeelden hiervan zijn:
- Aantallen vaccinaties per vaccin en batch.
- In geval van terugroepacties: lijsten met gevaccineerden per vaccin/batch. Deze lijsten bevatten persoonsgegevens.
- Andere onderzoeken naar aanleiding van een veiligheidssignaal van bijvoorbeeld CBG, LAREB, EMA, RIVM of een fabrikant.

[2] https://bronnen.zorggegevens.nl/Bron?naam=Nationale-Intensive-Care-Evaluatie

Deze gegevens moeten bijdragen aan het onderzoek naar dit signaal (RVP Richtlijn calamiteiten).

Deze rapportage wordt ook beschreven in de Use Case onder referentie **[MON02]**.

2.5 Algemene functionaliteiten

Naast de hiervoor genoemde inhoudelijke faciliteiten worden de volgende eisen gesteld:

- De Use Case vraagt met **[MON05]** om een algemene querymogelijkheid. Hiermee moeten rapporten gemaakt kunnen worden op de volgende terreinen:
  - Logistiek: distributie en verbruik van vaccins
  - Financiën
  - Managementinformatie (bijv. aantallen af te handelen fouten in vaccinaties)
- **[MON07]** Koppelingen naar individuele records voor surveillance-, bestrijding- en onderzoeksdoeleinden. Dit kent de volgende vormen:
  - Aanleveren van adresgegeven voor het benaderen van cliënten om hen te verzoeken tot deelname aan vervolgonderzoek.
  - Aanleveren van vaccinatiegegevens ten behoeve van vervolgonderzoek. Dit kan zowel op persoonsniveau zijn als gepseudonimiseerd.
  - In deze situatie kunnen onderzoeksgegevens met persoonsgegevens aan een TTP aangeleverd worden. Deze partij relateert op basis van de persoonsgegevens de dataset aan onderzoek data van een andere partij. Vervolgens levert de TTP een op persoon gecombineerde dataset op. De persoonsgegevens zijn uit deze set verwijderd.
- **[MON09]** koppelen op BSN: Onderzoeksgegevens uit verschillende bronnen moeten op basis van BSN of andere persoonsgegevens gecombineerd kunnen worden. Het resultaat zal echter vrijwel altijd gepseudonimiseerd of geanominiseerd zijn.
- **[MON08]** Samengevat: goede afspraken over toegang tot data. De voorgaande eisen laten zien dat tevens persoonsgegevens verwerkt moeten kunnen worden. De eisen hieraan moeten volledig uitgewerkt en beschreven zijn.

2.6 Eisen aan CIMS

De volgende eisen uit de Use Case onder Monitoring dienen ingevuld te worden door CIMS:

- **[MON06]** vraagt om een snelle invoer van data zodat ook snel rapportages gemaakt kunnen worden. Dit is allereerst een eis aan CIMS, zijn koppelingen en de gekoppelde systemen/organisaties.

2.7 Patronen

De hierboven beschreven uitvoerformaten laten een aantal verschillende patronen zien. De onderstaande tabel beschrijft dit patronen, onderverdeeld naar de vraag of persoonsgegevens verwerkt worden en waar de data vandaan komt. In hoofdstuk 4 worden de verschillende werkwijzen voor deze patronen beschreven.

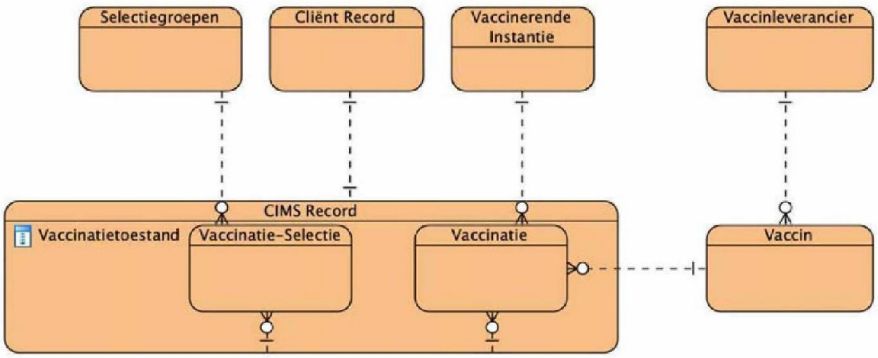
Bron Resultaat	Alleen CIMS	CIMS plus andere bronnen binnen RIVM	CIMS plus externe bronnen
Gepseudo-nimiseerd	• Opkomst • Vaccinatie-graad • Management-informatie • Verbruik per batch • Datakwaliteit (gedeeltelijk)	• Vaccinatie-effectiviteit • Vaccinaties relateren aan ander onderzoek	• Specifieke onderzoeken gematched via via TTP
Met persoons-gegevens	• Recall-lijsten • Benaderen voor onderzoek		

Tabel 1 Patronen voor dataverwerking

3 Data

3.1 **Datamodel**

Het COVID Informatie en Monitoring Systeem (CIMS) is gebaseerd op het volgende logische datamodel:



Figuur 3 Datamodel

Hieronder worden de verschillende entiteiten/objecttypes nader beschreven. De PSA van CIMS [ref. CIMS] beschrijft deze applicatie in meer detail.

3.1.1 Cliëntrecord

Het Cliëntrecord bevat alle benodigde persoonsgegevens van de te vaccineren of gevaccineerde persoon.

De verwerking in BI-CIMS is gepseudonimiseerd. Uitgangspunt is dat de verwerking in BI-CIMS (tweeweg) gepseudonimiseerd is. Anonimisering of één weg-pseudonimisering is niet mogelijk, in verband met de eisen om later voor nader onderzoek terug te kunnen grijpen op de brondata of de gegevens te combineren met gepseudonimiseerde informatie uit andere bronnen. Belangrijkste reden is echter het maken van lijsten met gevaccineerden bij terugroepacties.

Bij gepseudonimiseerde verwerking worden alleen persoonsgegevens overgenomen die nodig zijn voor de geaggregeerde rapportages. De exacte set wordt in het ontwerp vastgesteld. Als voorbeeld denken we aan de volgende:

- Pseudoniem
- Geboortejaar
- Geslacht

- Cijfers Postcode of een andere aggregatie van locatie. Het 'Doelendocument' kiest voor Gemeente.
- Selectiecriterium

3.1.2 Selectiegroepen/Vaccinatie-Selectie

Deze entiteit ondersteunt het oproepen voor vaccinaties. Hiermee worden groepen vastgelegd die vervolgens met functionaliteit van CIMS kunnen worden opgeroepen.

Voor de eerste doelgroepen zal de oproep echter vanuit de systemen van zorgverleners gebeuren. Oproepinformatie zal dus naar verwachting niet in CIMS worden vastgelegd.

Deze gegevens worden in principe geheel overgenomen uit CIMS, voor zover dit de pseudonimisering niet doorbreekt.

3.1.3 Vaccinaties

Deze gegevens worden in principe geheel overgenomen uit de CMIS-DB. Behalve de feitelijke vaccinaties wordt ook de beoordeling in het licht van het vaccinatieschema meegenomen. Dit bepaalt de geldigheid van de vaccinaties.

3.1.4 Vaccins

Deze gegevens worden in principe geheel overgenomen uit de CMIS-DB.

3.1.5 CIMS-record

Deze gegevens worden overgenomen, voor zover deze niet zodanig persoonsgebonden zijn dat ze de pseudonimisering onderbreken. Belangrijk onderdeel van het CIMS-record is de vaccinatietoestand van de client zoals die in het licht van het vaccinatieschema is bereikt.

3.1.6 Geanonimiseerde gegevens

In CIMS zijn alleen gegevens opgenomen van cliënten die hiertoe expliciet toestemming hebben gegeven. Het RIVM moet een zo volledig mogelijk beeld van de vaccinatiecampagne opbouwen en ontvangt daarom ook informatie over vaccinaties waar de client die toestemming weigerde. Hiervoor worden aparte datastromen ingericht die op basis van geanonimiseerde gegevens een vergelijkbare dataset aanlevert. De geanonimiseerde gegevens zijn degene genoemd hierboven beschreven als gepseudonimiseerde gegevens bij Clientrecord, met uitzondering van het pseudoniem. Deze gegevens worden opgeslagen in het CIMS Datawarehouse (CIMS-DWH).

3.1.7 Filtercriteria

Datasets moeten gefilterd kunnen worden op een aantal criteria. Hieronder staat beschreven welke data hiervoor beschikbaar is. Deze lijst

is niet uitputten. Ook andere gegevens uit het datamodel kunnen hiervoor gebruikt worden.

- Locatie. Dit kan op verschillende niveaus, waaronder woonplaats, gemeente, veiligheidsregio, GGD-regio: dit gebeurt op basis van postcode en een vertaaltabel van postcodes naar de verschillende regioindelingen. Deze tabellen worden van het CBS betrokken.
- Bewoners zorginstellingen: hiervoor zijn nog geen gegevens beschikbaar. Mogelijk dat dit met behulp van adreslijsten van zorginstellingen te realiseren is. Als dit kan, zal naar verwachting tevens het volledige adres beschikbaar moeten zijn om het GBA-adres van de burger te matchen met deze lijst.
- Medische indicatie: dit gebeurt op de waarde Selectiecriterium = 1 (Medisch) uit de aangeleverde vaccinatiegegevens.
- Zorgmedewerkers: dit gebeurt op de waarde Selectiecriterium = 2 (Beroepsgroep) uit de aangeleverde vaccinatiegegevens. Voorwaarde is dat alle aanleverende partijen dezelfde definitie van 'Beroepsgroep' definiëren.
- Leeftijd: dit kan met behulp van het geboortejaar bepaald worden.
- Vaccinsoort en -batch: deze informatie is onderdeel van CIMS en volledig bruikbaar.
- Zorgverlener: deze informatie is onderdeel van CIMS en volledig bruikbaar (voor zover dit geen persoonsgegevens betreffen).

3.2 Datamanagement

3.2.1 Data lineage

De oplossing moet 100% data lineage kunnen weergeven, zowel gezien vanuit de bronkant als vanuit het afgiftepunt (presentatie voor eindgebruikers). Hiermee wordt bedoeld dat data gevolgd kan worden vanaf het ontstaansmoment tot aan het afgifte moment en vice-versa. Er moet willekeurig kunnen worden geraadpleegd om de datastroom op elk punt in kaart te brengen. Dit vergroot de transparantie en levert te allen tijde het vereiste inzicht voor alle stakeholders. Data in deze lineage is alleen beschikbaar voor daartoe geautoriseerde rollen.

3.2.2 Datakwaliteit

In BI-CIMS mag data niet worden aangepast. Een uitgangspunt is dat de kwaliteit van data in de bron (CIMS) moet worden opgelost. Wel moet de oplossing in staat zijn, om data defecten te signaleren en te rapporten. Bij het inrichten van het systeem moet worden besloten of data die niet voldoet aan de gestelde eisen, moet worden doorgegeven. Hierbij dient rekening te worden gehouden met het volledigheidsprincipe.

3.2.3 Beveiliging van data

De toegang tot data is rolgebaseerd. Binnen de omgeving worden de volgende rollen onderkend:

Rol	Omschrijving
DVP Data Eigenaar	Governance op bestaande en nieuwe datasets; waaronder toekennen van rollen en rechten op gegevens
RIVM Sleutelbeheerder	Beheerder van het sleutelmateriaal van de pseudonimisering. Beheerder vernieuwt periodiek het sleutelmateriaal en borgt de veilige bewaring hiervan.
DVP depseudonimiseerder	Depseudonimisering van datasets en aanvulling met NAW-gegevens in die gevallen dat hier toestemming voor is.
DVP Data Scientist	Analyseren en verwerken van gegevens op het gebied van COVID-vaccinaties; definiëren (coderen) van views op gegevens (rapportages)
EPI Data Scientist/Gebruiker	Analyseren en verwerken van gegevens op het gebied van COVID-vaccinaties ten behoeve van infectieziektebestrijding
DVP Data Steward	Een ondersteunende rol om Data Scientists te helpen op het gebied van metadata, datakwaliteit en data governance binnen een aangewezen gebied.
DVP Functioneel Beheer	Bewaken en wijzigen van de gegevens(stromen); onderhoud gebruikersgroepen en rechten op groepen.
IV Applicatie Beheer	Bewaken, probleem oplossen en upgraden van de applicaties (Database engines, Datavirtualisatie, etc.)

Tabel 2 Rollen binnen COVID Informatie Management Rapportages

Data scientists krijgen via Datavirtualisatie toegang tot data middels een rollenpatroon dat wordt gekoppeld aan het RIVM Identity & Access managementsysteem (IAM) en aanvullende policies in Datavirtualisatie. Ze kunnen data manipuleren in datavirtualisatie en verwerken met PowerBI

3.2.4 Privacy van data

In paragraaf 3.1.1 is aangegeven dat persoonsgegevens zeer beperkt verwerkt worden. Identificerende gegevens worden als volgt behandeld:

- Het CIMS Clientnummer wordt twee-weg gepseudonimiseerd. Herleiding tot de brongegevens in CIMS is na toestemming hiermee mogelijk.
- Het BSN wordt twee-weg gepseudonimiseerd. Herleiding tot de brongegevens in CIMS is na toestemming hiermee mogelijk.
- Technische sleutels die direct of indirect kunnen verwijzen naar een specifieke client worden één-weg gepseudonimiseerd. Gegevens in CIMS kunnen bijgewerkt worden aan de hand van deze pseudonimisering. Herleiding naar brongegevens is echter niet mogelijk.

In de toekomst kan aanvullende versleuteling van de twee-weg gepseudonimiseerde sleutels worden toegepast, indien deze in datasets naar buiten worden verstrekt.

3.2.5 Master-, reference- en metadatamanagement

Er kan onderscheid worden gemaakt in de verschillende soorten data:

Master-datamanagement
Als definitie voor Masterdata wordt gebruikt:

- *Master Data Management (MDM) is het geheel aan processen en applicaties waarmee waardevolle kerngegevens (behalve transacties) in de organisatie worden beheerd.*

Het betreft dus de kerngegevens van het systeem die op 1 plaats worden verzameld en beheerd.

In BI-CIMS worden de volgende gegevens aangemerkt als Masterdata:

- Cliënten. Deze data wordt ontvangen uit BRP, COA, Probas en PIVA.
- Locaties
- Vaccinaties
- Vaccins

De overige masterdata wordt onderhouden door de beheerorganisatie CIMS. De Masterdata zal worden gebruikt voor COVID-Rapportages, LAREB- en AIZ-registers.

Referentie-datamanagement
Met referentiedata wordt data bedoeld die gebruikt kan worden om de bestaande data aan te vullen met een aantal vastgelegde entiteiten. Deze entiteiten worden vaak extern vastgelegd en onderhouden. Een voorbeeld in deze oplossing is de dataset:

- BIG- en AGB-register
- Zorg-AB (Adresboek)

Meta-datamanagement
Deze data worden gebruikt om data zodanig te beschrijven dat alle informatie beschikbaar is voor de juiste interpretatie van data tijdens de data-analyse. Daarnaast kan data snel en consistent worden toegevoegd om bijvoorbeeld meerdere doelgroepen te benoemen. Een voorbeeld in deze oplossing is de dataset:

Open-data
Van data die niet is geclassificeerd zal de metadata worden gepubliceerd op RIVMdata.

4 Applicaties

In paragraaf 2.7 staan de verschillende typen rapportages beschreven. In dit hoofdstuk staat beschreven hoe deze gerealiseerd gaan worden. Gegeven de hiervoor geschetste opzet, beschrijven we op basis van de gevraagde rapportages hieronder de wijze waarop ze gerealiseerd zullen worden.

Bron / Resultaat	Alleen CIMS	CIMS plus andere bronnen binnen RIVM	CIMS plus externe bronnen
Gepseudo-nimiseerd	- Pseudonimisering bij opname in CIMS-DWH - CIMS-informatie CIMS-DWH naar DV - Combinatie met anonieme data in DV - Rapportage via API of PowerBI	- Beide bronnen in pseudonimisen met dezelfde sleutel[3] - CIMS-informatie CIMS-DWH naar DV - Andere bron ontsluiten in DV - Combineren data in DV - Rapportage via API of PowerBI	- Maken dataset zoals beschreven bij 'Alleen CIMS' - Toevoegen persoonsgegevens in depseudonimiserings-functie - TTP matched alle betrokken bronnen - TTP pseudonimiseert of anonimiseert resulterende dataset
Met persoons-gegevens	- Rapportage zoals hierboven - Toevoegen persoonsgegevens in depseudonimiserings-functie		

4.1 Applicatie per gevraagde rapportage

4.1.1 Opkomst

De rapportage over opkomst kan het BI-CIMS leveren indien het oproepen via CIMS gebeurt of in CIMS geregistreerd wordt. In dat geval volgt de rapportage volledig de opzet zoals beschreven in deze PSA.

[3] Als alternatief kan gekozen worden voor het combineren van geaggregeerde informatie uit verschillende bronnen. Deze methode is echter doorgaans minder nauwkeurig.

Oproepen door het RIVM zijn voorzien in de latere fases van de vaccinatiecampagne.

4.1.2 Vaccinatiegraad

Het RIVM kan rapporteren over de vaccinatiegraad volgens de opzet die is beschreven in deze PSA.

4.1.3 Vaccinatie-effectiviteit

Deze rapportage betreft een combinatie van de gegevens uit CIMS over vaccinaties en gegevens uit andere bronnen over ziektemeldingen en overlijden. Deze gegevensbestanden moeten bij voorkeur op persoonsniveau aan elkaar gerelateerd worden. Voor het RVP worden vaccinatiegegevens handmatig geleverd aan EPI. Voor Covid19 is automatisering dringend gewenst. Daarvoor is het echter allereerst noodzakelijk dat éénduidige identificerende gegevens aanwezig zijn om de cliënten te identificeren. Op dit moment zijn de juridische mogelijkheden daarvoor nog niet aanwezig.

Vervolgens gebruiken zowel CIMS als Osiris pseudonimisering met dezelfde sleutel. Daarna is het direct in Datavirtualisatie mogelijk de gegevens uit verschillende bronnen te relateren.

4.1.4 Vaccinatieveiligheid

De rapportage over vaccinatieveiligheid wordt verzorgd door LAREB. Deze organisatie ontvangt informatie over de uitgevoerde vaccinaties wanneer mogelijke bijwerkingen op deze vaccinatie zijn gerapporteerd. Deze aanlevering is een functie op CIMS en wordt daarom in deze PSA niet verder uitgewerkt.

Wel zijn er de rapportagefuncties nodig om zowel rapportages zonder persoonsgegevens (bijv. aantallen vaccinaties per batch) als mét persoonsgegevens (bijv. terugroeplijsten) te maken.

4.2 **Koppelingen**

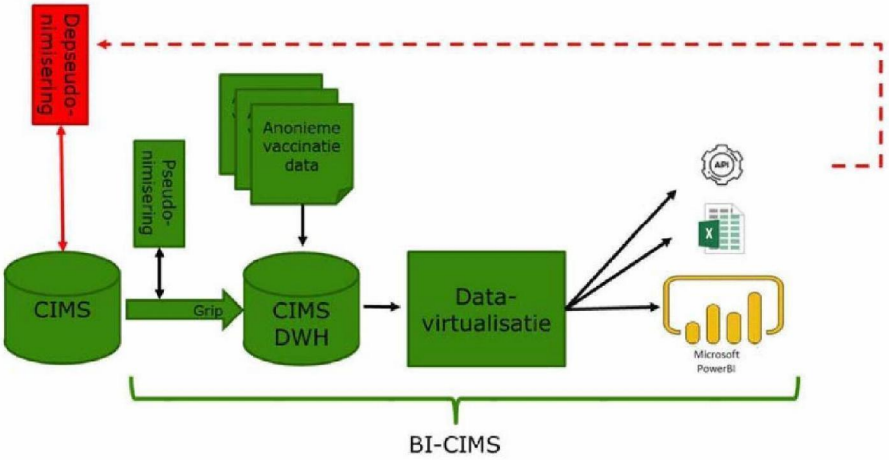
4.2.1 RIVMdata

De metadata van de rapportages zal worden gepubliceerd op RIVMdata.

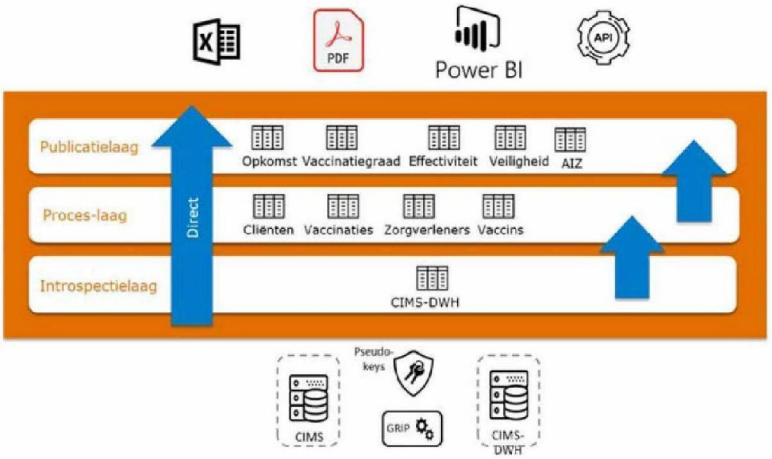
4.3 **Basisapplicaties**

Om te voorzien in de informatiebehoefte van de COVID-informatierapportage wordt het technische platform met bijbehorende componenten beschreven in de vorm van een technische infrastructuur.

Qua componenten ziet dit er als volgt uit:



Figuur 4 Opzet in componenten



Figuur 5 Dataroutes

De registratie van data vindt plaats in het CIMS. Met behulp van GRIP, ETL-tool, vindt er een datatransitie plaats richting het datawarehouse genaamd CIMS-DWH. Bij binnenkomst in CIMS-DWH wordt de data gepseudonimiseerd. In het CIMS-DWH vindt er een opbouw plaats van historie over data. Daarnaast is CIMS-DWH de bewaarplaats voor anoniem aangeleverde vaccinatiegegevens.

Rapportage verloopt via Datavirtualisatie. Dit biedt goede mogelijkheden informatie te ordenen en de combineren. Daarnaast biedt

Datavirtualisatie andere mogelijkheden voor ontsluiting, afhankelijk van de specifieke behoefte.

Voor het presenteren van data wordt onder andere gebruik gemaakt van Power BI. De Power BI-rapporten worden on-premises aangeboden via een web portaal.

4.3.1 Pseudonimisering

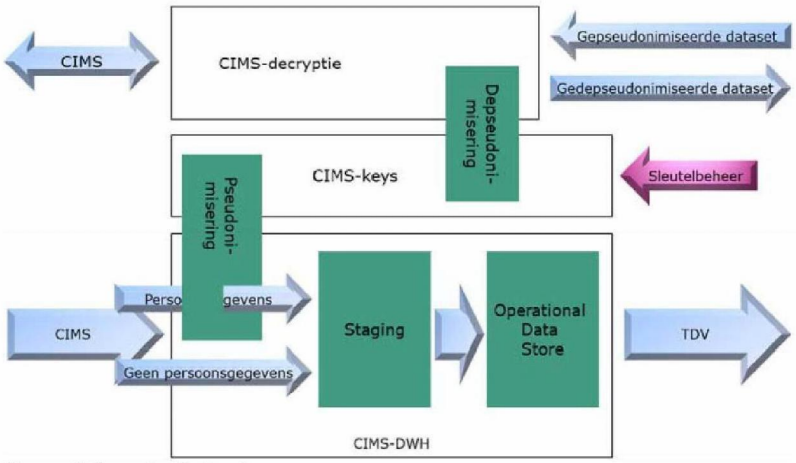
Pseudonimisering vindt plaats bij de ingang van BI-CIMS. De gehele verwerking in BI-CIMS vindt gepseudonimiseerd plaats. Hierbij wordt de databron ontsloten in DWH, waarna een masker over de data wordt geplaatst, waarin:

- BSN en CIMS-clientnummer worden gepseudonimiseerd. Hiervoor wordt AES-256 encryptie gebruikt.
- Overige identificerende gegevens worden gehashed. Hiervoor wordt SHA-256 hashing gebruikt.
- Gegevens benodigd voor filteren naar een hoger aggregatieniveau worden gebracht.

Keuze technologie

In bijlage C zijn de opties voor het implementeren van pseudonimisering uitgewerkt. Vanuit de beschikbare technologieën in het project is gekozen voor gebruik van een separaat schema in BI-CIMS-database. Deze oplossing kan in d toekomst omgezet worden naar een RIVM-brede voorziening.

Feitelijke pseudonimisering vindt plaats bij de ingang van de staging-omgeving van het CIMS-DWH. De ETL-tooling filtert persoonsgegevens tijdens het inlezen uit de CIMS-database. Hashing en encryptie worden daarbij gedaan met functies uit het CIMS-keys-schema. Bij het uitwerken van de oplossing moet tevens invulling gegeven worden aan de aandachtspunten voor beveiliging punten genoemd in paragraaf 5.3.



Figuur 6 Pseudonimisering

Het sleutelmateriaal is geïmplementeerd in het schema CIMS-keys. De sleutelbeheerder heeft hierin toegang voor het plaatsen en vervangen van sleutel. Deze handelingen vinden plaats onder het een vier-ogen-principe. De sleutels moeten daarnaast elektronisch en op papier worden opgeslagen in een kluis. Hierin wordt tevens de historie van gebruikte sleutels bewaard.

Vernieuwen van sleutels

Bij het werken met versleutelingsmechanismes op basis van SALT-keys of encryptiecertificatie, is periodieke vernieuwing wenselijk. Indien de sleutel gecompromiteerd raakt, kan de pseudonimisering doorbroken worden. Het pseudonimiseringmechanisme moet daarom de mogelijkheid hebben de pseudoniemen te vernieuwen. Tegelijkertijd moet de metadata van BI-CIMS informatie bevatten over de gehanteerde sleutel (datum, volgnummer, e.d.).

Decryptie

Ten behoeve decryptie is er nog een schema. Dit schema is geautoriseerd tot de decryptiefuncties in CIMS-keys. Daarnaast is dit schema ook gerechtigd tot CIMS na decryptie persoonsgegevens toe te kunnen voegen aan een dataset.

4.3.2 GRIP (ETL)

GRIP is een datawarehouse ETL-tool om grip op je data te krijgen. Vanuit de Programmeer-Editor kan de ontwikkelaar en beheerder maar ook de business-analist op eenvoudige wijze GRIP-commando's uitvoeren op de geselecteerde database. Met deze commando's is een complete data-migratie te realiseren. Andere toepassing zijn bijvoorbeeld (near-realtime) datawarehouses.

De programmeertaal is GRIP in combinatie met SQL. Met de GRIP-routines worden de processen van de ETL-functies gecodeerd. In ANSI-SQL worden de databewerkingen geprogrammeerd. De SQL wordt als view opgeslagen in de databases van waar de data is opgeslagen.

4.3.3 CIMS Datawarehouse

Het COVID Informatie en Monitoring Systeem (CIMS) bevat:

- Een centrale database met alle actuele vaccinatiegegevens, de CIMS-DB; deze wordt gebruikt voor het opslaan en raadplegen van gegevens, het bepalen van de geldigheid van de vaccinatie de vaccinatietoestand en eventuele vervolgstappen.
- Een datawarehouse met een kopie en de historie van de gegevens, het CIMS-DWH; deze heeft als belangrijkste taken:
    - Beschermen van CIMS-DB tegen overbelasting door (onverwachte) rapportage-taken
    - Bewaren historie via journalling.

Het CIMS-DWH wordt gevoed vanuit de volgende bronnen:

- CIMS, het vaccinatieregister.
- Anonieme vaccinaties. Deze worden niet verwerkt in CIMS, omdat de client geen toestemming heeft gegeven voor verwerking van zijn persoonsgegevens door het RIVM. De anonieme gegevens zijn echter wel gewenst, om een zo volledig mogelijk beeld van de vaccinatiecampagne te krijgen.

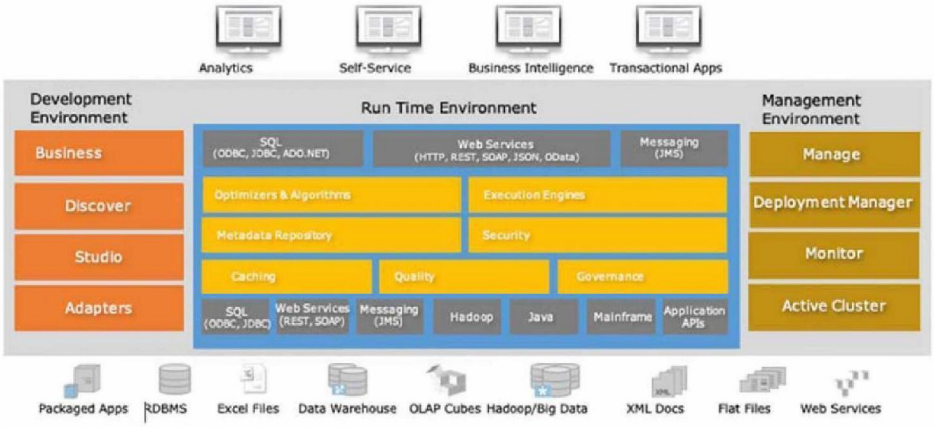
Het CIMS Datawarehouse bestaat uit één Oracle-database maar bevat meerdere schema's in het data warehouse. Ondanks dat dit maar één database betreft hebben de schema's verschillende functionaliteiten. De STG (staging area) bevat een gepseudonimiseerde kopie met dezelfde structuur vanuit CIMS. Deze stap wordt uitgevoerd om het bron CIMS-systeem te ontlasten. Het schema ODS volgt na de STG en zorgt ervoor dat de historie over data wordt opgeslagen. In Figuur 4 is de STG niet apart getekend, aangezien dit een tijdelijke verblijfplaats is, ondersteunend aan het ETL-proces.

4.3.4 Datavirtualisatie

Het schema ODS is het koppelpunt met Datavirtualisatie en wordt derhalve gezien als primair bronsysteem voor de datavoorziening. Dit is ook het punt waar de lineage start van Datavirtualisatie. Vanuit Datavirtualisatie zal gewerkt worden met een vast toegangsaccount tot de ODS-schema. Dit toegangsaccount heeft raadpleegrechten op de relevante tabellen en/of views en kan alleen worden gebruikt door de Datavirtualisatie software. Dit account staat los van de rechten op de data die eindgebruikers hebben.

Een Active Cluster bestaat uit een groep DV-instanties die met elkaar zijn verbonden via een systeemnetwerk. Active Cluster garandeert dat alle knooppunten identiek zijn, met hun metadata, activiteit en configuratie

automatisch gesynchroniseerd op de achtergrond. Dit betekent ook een uitbreiding van de onderliggende servers en hun besturingssystemen.



Figuur 7 Datavirtualisatie infrastructuur

5 Security & Privacy richtlijnen, Principes & Standaarden

In deze paragraaf wordt een opsomming gegeven van de principes, standaarden en (security)richtlijnen ten aanzien van het architectuuraspect, dat voor het project relevant zijn.

5.1 **Security uitgangspunten**[4]

Bij het ontwikkelen van het product dient rekening te worden gehouden met de onderstaande security principes.

Het uitgangspunt hierbij is Security-by-Design, oftewel beveiliging als onderdeel van het ontwerp. Dit principe heeft als uitgangspunt de beveiligingsmaatregelen bij de totstandkoming van het product te bepalen en deze hiermee integraal te implementeren.

Dit betekent dat de architectuur vanaf de basis is ontworpen om veilig te zijn. Beveiliging is hiermee geen losstaand punt, maar onderdeel van het product zelf. Uitgangspunten zoals in dit hoofdstuk beschreven zijn hiermee van toepassing. Het vulnerability assessment faciliteert in advies hierover.

Het vier ogen principe wordt toegepast op wijzigingen in de gedeelde infrastructuur. Dit betekent dat een expert de voorbereidingen en uitvoering van een wijziging op deze infrastructuur mede beoordeeld.

De volgende uitgangspunten zijn van toepassing:

ID	Security & Privacy Architectuurprincipe
SP 1	Minimaliseer de attack-surface
SP 2	Least-privilege principe
SP 3	Defence-in-depth principle
SP 4	Vertrouw externe koppelpunten niet zomaar
SP 5	Faciliteer Functie scheiding
SP 6	Voorkom security-by-obscurity
SP 7	Houd het eenvoudig
SP 8	Zorg voor juiste implementatie en oplossingen

Er worden twee beveiligingsmodellen ontworpen en geïmplementeerd in Datavirtualisatie:

Bronbeveiligingsmodel

- Gebruik indien mogelijk het verbindingsprotocol van de bronfabrikant (inclusief SSO),

[4] Zie https://www.owasp.org/index.php/Security_by_Design_Principles en BIO Versie 1.

- Gebruik een lokaal Active Directory-account of een eigen RDBMS-cliënt account om u aan te melden,
- Gebruik alleen speciale bronaccounts met bepaalde privileges voor noodzakelijke objecten,
- Indien mogelijk, veilige verbindingen tussen gegevensbron en datavirtualisatie,
- Toegang tot 'introspected' objecten binnen DV kan worden beveiligd door het Applicatie beveiligingsmodel.

Publicatiebeveiligingsmodel

- Ontwerp groepen en beleid voor authenticatie en autorisatie per virtuele laag,
- Gebruik eenmalige aanmelding,
- Toegang op objectniveau en gegevenstoegang zijn gescheiden en moeten worden verleend via de vereiste Active Directory-groepen,
- Alle toegang en voor authenticatie en autorisatie via Active Directory.

Implicatie voor Datavirtualisatie

Het dataverkeer tussen Datavirtualisatie en de Database (DST) schema is altijd encrypted, volgens de in de markt geldende normen.

De toegang tot Datavirtualisatie (Studio) is encrypted en het autorisatie en authenticatie proces verloopt via de Active Directory van het RIVM

In Datavirtualisatie (Studio) wordt gewerkt met een autorisatiematrix, op objectniveau en dataniveau voor elke gebruiker (geen eindgebruikers, analisten en ontwikkelaars) om rechten te geven op objecten en data.

Binnen Datavirtualisatie worden beveiliging policies ontworpen, welke garanderen dat gebruikers van data altijd moeten voldoen aan bepaalde autorisatieniveaus voordat zij data kunnen raadplegen. Deze autorisatieniveau's kunnen op kolom en rijniveau worden ontworpen.

In een eerste schets zijn de eerste virtuele ensembles[5] te onderkennen:

Vaccinatie
Data over het proces van het vaccineren zelf. Gelinkt aan ensembles

- Cliënten
- Vaccins
- Zorgverleners

Daarnaast gelinkt aan meta data tabel Doelgroep en Referentie tabel Locatie
Cliënten
Data over de cliënten, inclusief de kenmerken die 1:n zijn. Gelinkt aan ensembles

[5] Een ensemble is verzameling van data entiteiten inclusief de relaties tussen deze entiteiten; het geheel beschrijft een bedrijfsconcept

- Vaccinatie

Vaccins
Data over de vaccins, inclusief de kenmerken die 1:n zijn. Gelinkt aan ensembles
- Vaccinatie

Zorgverlener
Data over de zorgverleners, inclusief de kenmerken die 1:n zijn. Gelinkt aan ensembles
- Vaccinatie

5.2 **Privacy uitgangspunten**

Privacy-by-design houdt in dat u als organisatie al tijdens de ontwikkeling van producten en diensten (zoals informatiesystemen) ten eerste aandacht besteedt aan privacy verhogende maatregelen, ook wel privacy enhancing technologies (PET) genoemd. Ten tweede houdt u rekening met dataminimalisatie: u verwerkt zo min mogelijk persoonsgegevens, dat wil zeggen alleen de gegevens die noodzakelijk zijn voor het doel van de verwerking.

We vertalen dit naar **Verstandig Gebruik** en **Passende Bescherming**[6].

Verstandig Gebruik
1 Verzamel alleen wat je nodig hebt (Grondslag)
2 Gebruik gegevens alleen waarvoor je ze gevraagd hebt (Doelbinding)
3 Kies bij configuratie standaard de privacy vriendelijke variant
4 Let op de kwaliteit van data
5 Maak geen onnodige kopieën
6 Gooi weg wat je niet langer nodig hebt

Passende Bescherming
7 Sla gescheiden op
8 Beperk de toegang
9 Verwerk geaggregeerd
10 Pas versleuteling, pesudonimisering of anonimisering toe

Aangaande "Verstandig Gebruik" gaat het daarmee dus ook om Grondslag en Doelbinding (1 en 2). De Grondslag bepaalt of de gegevens verwerkt mogen worden, de Doelbinding bepaalt waarvoor de gegevens worden gebruikt.

Tijdens de pilot is geen grondslag en doelbinding bepaald, aangezien er geen persoonsgegevens worden verwerkt en gebruikt.

Er zijn tijdens de pilot diverse ontwerpbeslissing genomen die zich deels hebben vertaald naar maatregelen waarmee het Verstandig Gebruik en

[6] Uit het boekje Privacy-by-Design door [redacted], VKA.

zorgen voor Passende Bescherming van persoonsgegevens kan worden toegepast in vervolgprojecten.

PP 1 - Verzamelen alleen wat je nodig hebt (Grondslag)	
Beschrijving:	Uitgangspunt is om alleen data vast te leggen die werkelijk nodig is voor het beantwoorden van de informatievraag.
Toepassing:	

PP 2 - Gebruik gegevens alleen waarvoor je ze gevraagd hebt (Doelbinding)	
Beschrijving:	Uitgangspunt is om alleen data te gebruiken voor het beantwoorden van de informatievraag die daarvoor is gedefinieerd.
Toepassing:	

PP 3 - Kies bij configuratie standaard de privacy vriendelijke variant	
Beschrijving:	Onderdeel van de functie is de configuratie waarin doelbinding en grondslag worden gekoppeld aan de informatievraag, zodat alleen data uitsluitend daarvoor gebruikt kunnen worden.
Toepassing:	

PP 4 - Let op de kwaliteit van data	
Beschrijving:	Datakwaliteitsdimensies: nauwkeurigheid, volledigheid, tijdigheid, bruikbaarheid, relevantie en betrouwbaarheid van data
Toepassing:	

PP 5 - Maak geen onnodige kopieën	
Beschrijving:	Voorkomt extra onderhoud en verhoogt consistentie van data
Toepassing:	

PP 6 - Gooi weg wat je niet langer nodig hebt	
Beschrijving:	Verwijderen van virtuele views waardoor data niet meer toegankelijk is.
Toepassing binnen ontwerp:	

PP 7 - Sla gescheiden op	
Beschrijving:	Sla gescheiden op
Toepassing:	

PP 8 - Beperk de toegang	
Beschrijving:	Expliciet gebruikers en groepen toegang toewijzen
Toepassing:	

PP 9 - Verwerk geaggregeerd	
Beschrijving:	Voorkom het delen van persoonsgerelateerde data door deze te aggregeren.
Toepassing:	

PP 10 - Pas versleuteling, pseudonimisering of anonimisering toe	
Beschrijving:	Ter voorkoming van ongewenst delen van persoonsgerelateerde gegevens.
Toepassing:	Pseudonimisering en anonimisering zijn ontwerpkeuzes, waarbij alle data on-demand verwerkt wordt.

Voorafgaand aan de datavirtualisatie-intake zal een Privacy Impact Assessment moeten worden uitgevoerd.

5.3 **Pseudonimisering**

Belangrijk onderdeel van de privacybeschermende maatregelen zijn is het gebruik van pseudonimisering. De keuze voor pseudonimisering is beschreven in paragraaf 3.2.4. De vorm van realisatie binnen CIMS – gegeven het feit dat hiervoor nog geen RIVM-brede voorziening beschikbaar is – is beschreven in paragraaf 4.3.1.

Voor een zorgvuldige werking van de pseudonimisering is het noodzakelijk de omgeving waarin dit gebeurd goed af te schermen. Hiervoor zijn onder andere de volgende maatregelen nodig:

- Scheiding tussen de rollen zoals aangegeven in paragraaf 3.2.3.
- Toekennen van functies en autorisaties aan de rollen zodat daadwerkelijk scheiding ontstaan. Daarbij moet in ieder geval onderscheid bestaan tussen de mogelijkheid om:
    - Gegevens te versleutelen
    - Gegevens te ontsleutelen
    - Sleutelbeheer
- Autorisaties zijn expliciet vastgelegd.
- Sleutelbeheer vindt alleen plaats onder vier ogen.
- Procedures voor sleutelbeheer en ontsleutelen zijn expliciet vastgelegd.
- Sleutelmateriaal is beveiligd opgeslagen. Het staat uitsluitend in het CIMS-keys-schema van de productieomgeving en wordt onleesbaar gemaakt ('wrapping').
- Sleutels worden tevens in een kluis bewaard, elektronisch en op papier. Hierbij wordt tevens de volledige historie van sleutels bewaard.

5.4 **Implicaties voor Datavirtualisatie**

Data die binnenkomt binnen Datavirtualisatie zal met een kolom doelbinding worden verrijkt. De kolom doelbinding wordt aan elke virtuele tabel toegevoegd en kent een numerieke waarde en een betekenis (meta data)

Waarde doelbinding:

Standaard: 0
Doelbinding 1 1
Doelbinding 2 2
Enz.

Standaard 0 betekent dat voor de data geen bijzondere regels gelden ten aanzien van de AVG. Dit betekent dus niet dat de data vrij toegankelijk is.

Doelbinding 1:

Dit is een beschrijving van het type data waar dit over gaat. Doelbinding is afgegeven voor het gebruik onder bepaalde omstandigheden door daartoe bevoegde personen. Dit geldt bijvoorbeeld voor de dataset cliënten. Voor de eerste rapportages zijn deze detail gegevens nog niet relevant maar wellicht in een latere fase wel. De persoon die met deze data aan de slag wil, zal dus moeten voldoen aan de regels gesteld rondom deze specifieke doelbinding. Binnen Datavirtualisatie zal deze data niet worden opgeslagen maar alleen ten tijde van de verwerking worden geraadpleegd. Dit gebeurt bij vragen aan het systeem die in de kern detaildata nodig hebben om te komen tot een aggregaat. Bijvoorbeeld hoeveel cliënten uit een bepaalde wijk van een bepaalde leeftijd..., etc.

De precieze inrichting hiervan wordt nader ontworpen. De twee-weg pseudoniemen (Clientnummer en BSN) vragen echter sowieso aandacht. De beschikbaarheid van deze gegevens bepalen of er sprake is van gepseudonimiseerde of geanonimiseerde uikomsten. Aparte autorisaties hiertoe zijn nodig.

De oplossing zal als volgt worden ingericht:

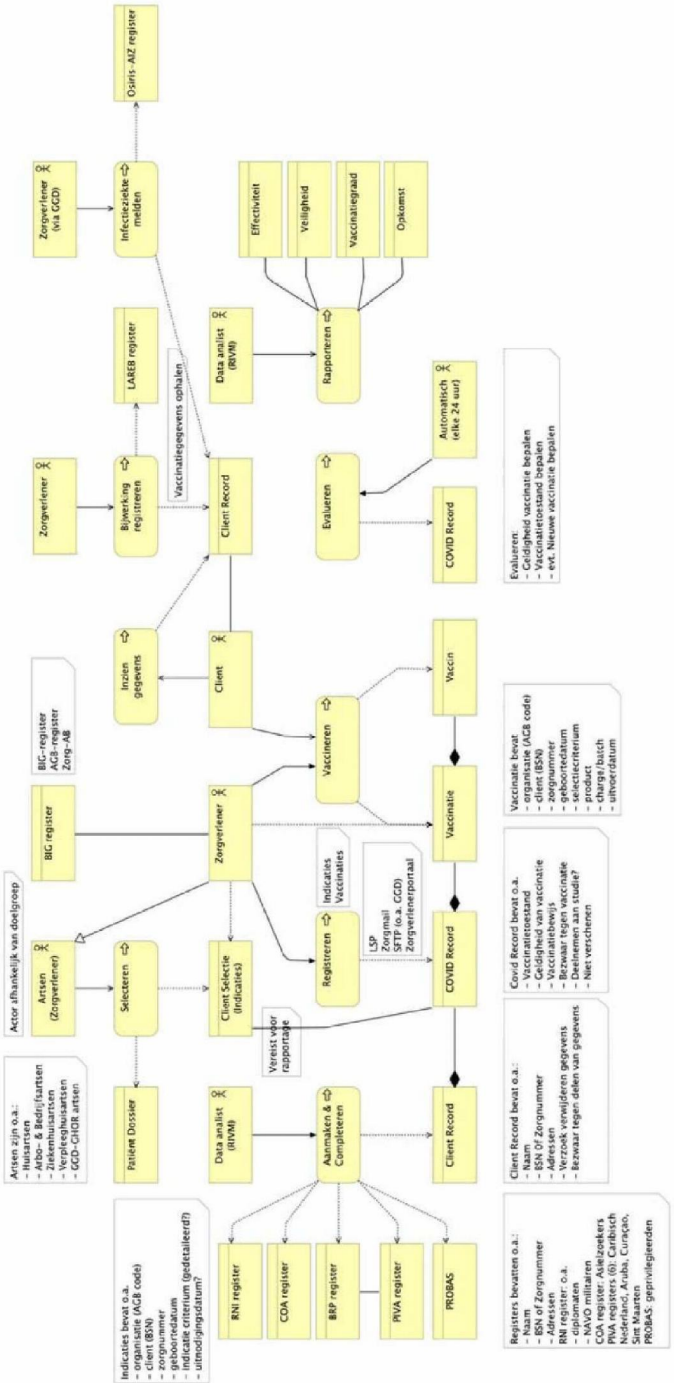
Publicatie
Hier wordt het virtuele sterschema gepresenteerd. Als bron wordt hiervoor de Proces laag gebruikt

Proces
Hier wordt gemodelleerd volgens het zogenaamde Ensemble Logical Model principe. Hier worden de entiteiten van de Introspectie laag gegroepeerd en logisch samengebracht. Daar waar een relatie tussen virtuele ensembles is te herkennen, zal deze worden vastgelegd. Hierbij wordt per Ensemble aangegeven wat de reikwijdte van een dataset kan zijn.

Introspectie
Deze laag bevat de connectie naar het bronsysteem en leest de meta data van de tabellen uit de database waar het recht op heeft via het account.

De beveiligings-policies alsmede de AVG-regels worden voor eindgebruikers direct toegepast op data objecten in de Publicatie laag. Dit is de enige laag waar eindgebruikers toegang toe hebben.

6 Bijlage A: COVID Vaccinatiemodel



7 Bijlage B: Architectuurkaders

HP 1 - RIVM voegt waarde toe aan data	
Beschrijving:	Onze informatievoorziening ondersteunt het toevoegen van waarde. In ons onderzoek, het RIVM-onderzoek, gebruiken we data en aan die data voegen we waarde toe, dit is een kern onderscheidend vermogen.
Toepassing binnen ontwerp:	Er zijn altijd minimaal 2 databronnen In elke project wordt vooraf beschreven wat de waarde van de gecombineerde data is, waarom combineren waarde vertegenwoordigt en voor wie.
Relatie tot kader:	RIVM IV Principe

HP 2 - Interne en externe samenwerking	
Beschrijving:	Het werk van het RIVM wordt steeds vaker uitgevoerd in samenwerking met anderen. Partner-organisaties, binnen en buiten Nederland, tot aan Citizen Science waar de burger in betrokken is en ook aan mee werkt.
Toepassing binnen ontwerp:	Zonering, Role Based Access, Column Based access Afnemers van de informatie worden vooraf beschreven Zie NORA-principe 8
Relatie tot kader:	RIVM IV Principe

HP 3 - Citizen Science	
Beschrijving:	Om de herkenbaarheid, het draagvlak en daarmee de afstand tussen wetenschap en burgers te verkleinen, zetten we actief in op burgerparticipatie en inspraak.
Toepassing binnen ontwerp:	
Relatie tot kader:	RIVM IV Principe

HP 5 - Gebruiker Centraal	
Beschrijving:	Intern zijn het de RIVM-medewerkers die de informatievoorziening gebruiken als middel om hun taken goed te kunnen verrichten. Maar we hebben ook externe gebruikers van onze informatievoorziening. Ook voor hen geldt dat we ze de bij bouw of doorontwikkeling van voorzieningen nadrukkelijk betrekken.
Toepassing binnen ontwerp:	In elke project wordt vooraf beschreven wat de waarde van de gecombineerde data is, c.q. waarom combineren waarde vertegenwoordigt en voor wie Afnemers van de informatie worden vooraf beschreven.

Relatie tot kader:	RIVM IV Principe

HP 8 - Duurzaam en Circulair	
Beschrijving:	Bij alle sourcing (inhuur van mensen en inhuur & gebruik van middelen) streven we naar een duurzame inzet. Duurzame inzet geeft niet meer belasting dan (strikt) noodzakelijk voor het uitvoeren van de taken.
Toepassing binnen ontwerp:	DV wordt niet ingezet/gebruikt om data persistent op te slaan
Relatie tot kader:	RIVM IV Principe

DP 1 – Data zijn de grondstof van het RIVM	
Beschrijving:	Data vormen de basis voor beslissingen, rapportages en kennis die het RIVM deelt. Om ervoor te zorgen dat het RIVM zijn uitstekende positie in het huidige kennislandschap kan blijven behouden en daarmee toegevoegde waarde blijft creëren, is het noodzakelijk om data als bedrijfsassets te zien en te benaderen. Data, ICT en IV staan op de strategische agenda van het RIVM, het zijn bedrijfsassets en die moeten worden gemanaged.
Toepassing binnen ontwerp:	
Relatie tot kader:	RIVM Datastrategie

DP 2 – Datamanagement	
Beschrijving:	Goed datamanagement is een voorwaarde voor het betrouwbaar vergaren van data en beheren van de datakwaliteit. Wij kiezen hiervoor het DAMA-raamwerk. Dit is een algemeen gebruikt model voor datamanagement dat bestaat uit kennisvelden die verbonden worden door data governance. Dit kader sluit aan bij Rijksbrede kaders en afspraken. In projecten kiezen wij voor de besturingsrol en houden toezicht op het toepassen van standaarden, het gebruik van beschikbare registers en toepassing van datamanagement principes in project start architecturen. Data Governance is het RIVM brede besturingsproces dat zorgdraagt voor eenduidige vastlegging en toepassing van beleidsprincipes, rollen, taken, verantwoordelijkheden en bevoegdheden voor het totale datamanagement-proces. Datamanagement betreft de ontwikkeling, uitvoering en supervisie van plannen, beleid, programma's, procedures en activiteiten binnen het RIVM die de waarde van gegevens gedurende de volledige levenscyclus waarborgen en verbeteren.

	Dit met als doel om de waarde van gegevens maximaal uit te nutten waarbij het RIVM de controle heeft over kosten en kwaliteit van datamanagementactiviteiten. Rollen, taken, verantwoordelijkheden en bevoegdheden voor het totale informatieleveringsproces zijn eenduidig vastgelegd en toegewezen.
Toepassing binnen ontwerp:	Governance m.b.t. databronnen, data routes en gevirtualiseerde datasets is beschreven en ingericht
Relatie tot kader:	RIVM Datastrategie

DP 3 – Data zijn open en vrij toegankelijk	
Beschrijving:	Data is de grondstof van het RIVM. Vrije uitwisseling van data is een voorwaarde om onderzoek te stimuleren en data te verzamelen. Onderzoekers zijn gebaat bij data die betrouwbaar zijn, burgers zijn gebaat bij data die anoniem zijn. De toegang tot en het gebruik van data verloopt zonder drempels, ongeacht waar deze zich bevindt of waar de gebruikers zich bevinden. (Onderzoek)gegevens worden vrijelijk beschikbaar gesteld volgens open standaarden, tenzij zwaarwegende bezwaren dat verhinderen (privacy, wetgeving, overheidsbelang en dergelijke). Overigens wil 'vrijelijk' niet zeggen kosteloos, het beschikbaar stellen heeft een kostprijs. Wie de kosten betaalt (opdrachtgever, afnemer) kan per situatie verschillen
Toepassing binnen ontwerp:	
Relatie tot kader:	RIVM Datastrategie

DP 5 – Standaardiseer en documenteer data	
Beschrijving:	De belangrijke databronnen voor het RIVM vormen de kern van de organisatie. Die databronnen noemen we kernregistraties die breed ingezet kunnen worden voor het RIVM. Dit kan data zijn over de klanten van het RIVM (bedrijfsvoering) maar ook belangrijke registraties zoals het Doetinchem cohort, de gezondheidsmonitor, het infectieziekten surveillance systeem, interventiedatabase, meetstations en nog veel meer. Deze kern registraties worden in applicaties, systemen en ketens gebruikt. INSPIRE-data zijn een goed voorbeeld van de werking van dit principe. De bestaande registers (bijvoorbeeld het Nationaal Georegister (NGR)) en het RIVM-register (RIVMdata) geven weer welke data beschikbaar is en wat de kwaliteit en bruikbaarheid is.

	Standaardisatie bevordert uitwisselbaarheid en documentatie maakt data waardevol.
Toepassing binnen ontwerp:	Per project wordt de architectuur bijgewerkt, gedocumenteerd welke databronnen zijn aangesloten, wie daarvan de eigenaar is, etc. We gebruiken (inter)nationale (meta)data standaarden als die er zijn, denk aan Inspire, SNOMED-CT, LOINC, etc.
Relatie tot kader:	RIVM Datastrategie

DP 6 – Toegevoegde waarde en communicatie	
Beschrijving:	Het RIVM biedt veel en heeft een sterke positie in de samenleving. Data zijn vaak niet direct zichtbaar maar wel als grondstof belangrijk voor kennisproducten. Het RIVM werkt waar mogelijk aan de principes van Open data: 'Open waar mogelijk, gesloten als het moet'. Door specifiek te kiezen waar de toegevoegde waarde groot is en dit breed te communiceren versterkt het RIVM zijn positie binnen de samenleving. Vanuit het principe 'Open by Design' wordt in een vroeg stadium al gekeken naar het businessmodel of het verdienmodel. Dit is geen vast omkaderd standaardmodel, dit kan per situatie verschillen.
Toepassing binnen ontwerp:	De gevirtualiseerde datasets worden via RIVMdata bekend gemaakt
Relatie tot kader:	RIVM Datastrategie

BP01 - Afnemers krijgen de data waar ze behoefte aan hebben	
Beschrijving:	
Toepassing binnen ontwerp:	Zie HP 1 en HP 5
Relatie tot kader:	NORA Proactief

BP02 - Afnemers kunnen de data eenvoudig vinden	
Beschrijving:	
Toepassing binnen ontwerp:	De gevirtualiseerde datasets worden bekend gemaakt via RIVMdata Datasets zijn voorzien van voldoende, bij voorkeur gestandaardiseerde metadata
Relatie tot kader:	NORA Vindbaar

BP03 - Afnemers hebben eenvoudig toegang tot de dienst	
Beschrijving:	
Toepassing binnen ontwerp:	Zie HP 5
Relatie tot kader:	NORA Toegankelijk

BP04 - Afnemers ervaren uniformiteit in de dienstverlening door het gebruik van standaardoplossingen	
Beschrijving:	
Toepassing binnen ontwerp:	Standaard protocollen worden gebruikt voor de toegang tot data (JBDC, Jason, OData, etc.)
Relatie tot kader:	NORA Standaard

BP05 - Afnemers krijgen gerelateerde diensten gebundeld aangeboden	
Beschrijving:	
Toepassing binnen ontwerp:	Allereerst moet worden onderzocht hoe afnemers gebruikmaken van een dienst en hoe zij deze associëren met andere diensten. Op basis daarvan kunnen dienstverleners gaan samenwerken. Gezamenlijk moeten zij bepalen hoe zij hun diensten willen bundelen. Implicatie is dus, dat Datavirtualisatie bijna altijd een onderdeel zal zijn van een project, waarin de gegenereerde data wordt gepresenteerd aan/namens de afnemer.
Relatie tot kader:	NORA Gebundeld

BP06 - Afnemers hebben inzage in voor hen relevante informatie	
Beschrijving:	
Toepassing binnen ontwerp:	Eerst moet worden beschreven wat voor de afnemer(s) relevant is. Elke gevirtualiseerde dataset wordt gemeld in RIVMdata.
Relatie tot kader:	NORA Transparant

BP08 - Afnemers kunnen erop vertrouwen dat informatie niet wordt misbruikt	
Beschrijving:	
Toepassing binnen ontwerp:	Zie HP 4
Relatie tot kader:	NORA Vertrouwelijk

BP09 - Afnemers kunnen erop vertrouwen dat de dienstverlener zich aan afspraken houdt	
Beschrijving:	
Toepassing binnen ontwerp:	Er wordt een SLA voor DV opgesteld. Dit betekent dat de governance goed beschreven moet worden per project. Er worden afspraken m.b.t. kwaliteit en beschikbaarheid van de brondata gemaakt met de eigenaren daarvan. Er worden duidelijke acceptatiecriteria beschreven m.b.t. kwaliteit en beschikbaarheid van de gevirtualiseerde data.
Relatie tot kader:	NORA Betrouwbaar

BP10 - Afnemers kunnen input leveren over de dienstverlening	
Beschrijving:	
Toepassing binnen ontwerp:	Er wordt in elk project een duidelijke feedbackcirkel gedefinieerd en gehanteerd.
Relatie tot kader:	NORA Ontvankelijk

AP01 – Diensten zijn herbruikbaar	
Beschrijving:	De opzet van de dienst anticipeert op onvoorziene afnemers en gebruik.
Rationale:	Toepassing van dit principe maakt de dienst interoperabel en bruikbaar voor een zo groot mogelijke groep afnemers. Dit draagt bij aan een hoger rendement van de dienst.
Toepassing binnen ontwerp:	Toegang tot DV wordt op basis van AD-groepen toegekend en is daarmee voor eenieder met toegang als dienst beschikbaar te maken.
Relatie tot kader:	RIVM Architectuur

AP06 – Gebruik standaard oplossingen	
Beschrijving:	De dienst maakt gebruik van standaard-oplossingen.
Rationale:	Dit principe is erop gericht dat afnemers de overheid in haar dienstverlening zo veel mogelijk zullen ervaren als één organisatie. Dit vraagt om standaardisatie en uniformiteit in dienstverlening en ondersteunende processen. Organisaties hoeven minder zelf te ontwikkelen en het rendement van oplossingen neemt toe.
Toepassing binnen ontwerp:	DV is een solution bestaande uit 2 standaard softwarecomponenten: DV Studio en DV Server. De dienst kan toegang geven tot bestaande servers of er kan een eigen server worden ingericht. Er is een standaard installatie- en ingebruikname proces ingericht.
Relatie tot kader:	RIVM Architectuur
AP08 – Gebruik open standaarden	
Beschrijving:	De dienst maakt gebruik van open standaarden.
Rationale:	Het gebruik van open standaarden bevordert de interoperabiliteit, bijvoorbeeld in het berichtenverkeer. Open standaarden kunnen door alle partijen vrijelijk worden gebruikt. Er zijn geen door private partijen afgedwongen beperking aan het gebruik.
Toepassing binnen ontwerp:	DV ondersteunt de volgende open standaarden voor het uitleveren van data: SOAP, REST, OData, JSON. ANSI SQL wordt als standaard gebruikt voor het zowel het genereren van queries door DV als voor het handmatig opvoeren van queries en scripts.
Relatie tot kader:	RIVM Architectuur

AP17 – Informatie-objecten systematisch beschreven	
Beschrijving:	De aan de dienst gerelateerde informatieobjecten zijn, uniek geïdentificeerd, in een informatiemodel beschreven.
Rationale:	Samenwerking tussen en binnen organisaties is alleen goed mogelijk wanneer de betrokkenen de relevante informatieobjecten kunnen toepassen, hergebruiken en duurzaam archiveren. Een systematische beschrijving van informatieobjecten, hun semantiek en onderlinge structuur is nodig om de informatie eenduidig te kunnen interpreteren en digitale uitwisseling mogelijk te maken. De unieke identificatie is nodig om ervoor te zorgen dat mensen en machines op een eenduidige manier naar informatieobjecten kunnen verwijzen zodat ze vindbaar en van elkaar onderscheidbaar zijn.
Toepassing binnen ontwerp:	DV maakt gebruik van een repository waarin alle objecten en de relaties tussen de objecten gedefinieerd en beschreven zijn. Deze informatie wordt tevens gesynchroniseerd met InData. In InData wordt aanvullende informatie vastgelegd per DV-object, waarmee deze informatie voor zowel gebruikers als ontwikkelaars beschikbaar komt.
Relatie tot kader:	RIVM Architectuur

AP18 – Ruimtelijke informatie via locatie	
Beschrijving:	De dienst ontsluit ruimtelijke informatie locatiegewijs.
Rationale:	Locatiegewijze ontsluiting maakt ruimtelijke informatie voor afnemers begrijpelijk en toegankelijk. Ruimtelijke samenhang is bovendien een belangrijke basis voor bundeling en het proactief aanbieden van diensten. De transparantie van de overheid wordt erdoor bevorderd. Ook zijn informatie-objecten op basis van de locatie eenvoudig buiten de beoogde toepassing te hergebruiken. Informatie over een locatie kan zowel betrekking hebben op de geografische positie, de vorm (geometrie) als op een verwijzing waar elders deze positie en vorm te vinden zijn (bv. postcode, woonplaats). Het ontsluiten kan zowel geschieden door het tonen van de geometrie, als op basis van de verwijzing.
Toepassing binnen ontwerp:	DV-objecten kunnen locatie informatie omvatten, dit zijn metadata-kenmerken van DV-objecten die door ontwikkelaars worden gerealiseerd. Deze objecten zijn deelbaar met anderen, mits de toegang daartoe is geautoriseerd binnen DV.
Relatie tot kader:	RIVM Architectuur

AP23 – Automatische dienstverlening	
Beschrijving:	De dienst wordt na bepaalde signalen automatisch geleverd.
Rationale:	Het gebruik en gemak van de dienst neemt toe wanneer deze geleverd wordt zodra er een signaal is dat een afnemer behoefte heeft aan de dienst. De afnemer hoeft in dit geval de dienst niet zelf aan te vragen. Dergelijke signalen kunnen ook van andere organisaties afkomstig zijn. Veel signalen zijn locatie gebonden. Overigens heeft dit principe niet alleen betrekking op ICT: ook medewerkers kunnen 'automatisch', volgens afspraken en protocollen handelen.
Toepassing binnen ontwerp:	DV kan worden opgenomen in workflow omgevingen, om op basis van events te worden geactiveerd voor toegang voor gebruikers of ontwikkelaars.
Relatie tot kader:	RIVM Architectuur

AP30 – Verantwoording dienstlevering mogelijk	
Beschrijving:	De wijze waarop een dienst geleverd is, kan worden verantwoord.
Rationale:	Dienstverleners moeten individuele leveringen van diensten kunnen verantwoorden, naar aanleiding van bijvoorbeeld klachten van afnemers, accountantscontroles en gerechtelijke procedures. Met het oog op informatiebeveiliging moet het mogelijk zijn om vast te stellen wie welke handelingen heeft verricht op een ICT-voorziening, of welke fouten zijn opgetreden. Om dit mogelijk te maken, moeten de voor verantwoording relevante informatieobjecten worden vastgelegd. De waarde (en definitie van die waarde) van deze informatie-objecten moeten op een bepaald moment in de tijd gereconstrueerd kunnen worden
Toepassing binnen ontwerp:	DV registreert alle activiteiten (inloggen, raadplegen, verwerken, publiceren etc.) op detailniveau in de repository.
Relatie tot kader:	RIVM Architectuur

AP40 – Onweerlegbaarheid	
Beschrijving:	De onweerlegbaarheid van berichtenuitwisseling wordt gegarandeerd door wederzijdse authenticatie en door versleuteling van elektronische handtekeningen.
Rationale:	Bij diensten met rechtsconsequenties is onweerlegbaarheid van groot belang. Onweerlegbaarheid houdt in dat de afzender of ontvanger van een bericht niet kunnen ontkennen het bericht respectievelijk verstuurd, dan wel ontvangen te hebben. Hiervoor is wederzijdse

	authenticatie en controle op de integriteit van het bericht nodig.
Toepassing binnen ontwerp:	DV ondersteunt wederzijdse authenticatie en versleuteling van elektronische handtekeningen.
Relatie tot kader:	RIVM Architectuur

AP41 – Beschikbaarheid	
Beschrijving:	De beschikbaarheid van de dienst voldoet aan de met de afnemer gemaakte continuïteitsafspraken.
Rationale:	De continuïteitsafspraken zijn gemaakt op basis van de afbreukrisico's die afnemers lopen bij uitval. De processen van afnemers kunnen spaak lopen met financiële en maatschappelijke schade en het vertrouwen in betrouwbaarheid van de dienst kan afnemen.
Toepassing binnen ontwerp:	DV kan worden geïmplementeerd in een Active Cluster (3 nodes) opstellen waarmee een uptime van 99,999 % kan worden gegarandeerd. TVD op 1 node kent een uptime van 99,5 %.
Relatie tot kader:	RIVM Architectuur

AP42 – Integriteit	
Beschrijving:	De dienstverlener waarborgt de integriteit van gegevens en systeemfuncties.
Rationale:	De gebruiker van een gegeven moet erop kunnen vertrouwen dat hij het correcte, complete en actuele gegeven ontvangt.
Toepassing binnen ontwerp:	DV onderhoudt diverse certificaten waaronder HITRUST, FedRAMP, SOC 1 en 2, ISO 27001.
Relatie tot kader:	RIVM Architectuur

AP43 – Vertrouwelijkheid	
Beschrijving:	De dienstverlener verschaft alleen geautoriseerde afnemers toegang tot vertrouwelijke gegevens.
Rationale:	De gebruiker moet erop kunnen vertrouwen dat gegevens niet worden misbruikt.
Toepassing binnen ontwerp:	
Relatie tot kader:	RIVM Architectuur

AP44 – Controleerbaarheid	
Beschrijving:	De dienstverlener zorgt ervoor dat de beoogde toegang tot gegevens en de juiste werking van zijn systemen continu alsook achteraf te controleren is.
Rationale:	Het gebruik en gedrag van de dienst moet voldoen aan de gestelde regels. Om te borgen dat dit gebeurt, moet continu worden gemonitord. Om de juistheid van uitkomsten van het systeem aan te kunnen tonen, moet gelogd worden.
Toepassing binnen ontwerp:	Toegekende toegang tot DV is volledig transparant en controleerbaar door geautoriseerde gebruikers.

Relatie tot kader:	RIVM Architectuur

AP101 – Gegevens zijn Vindbaar	
Beschrijving:	Metadata en data zijn vindbaar voor zowel mensen als machines.
Rationale:	• Data kan opgevraagd worden door middel van een gestandaardiseerd protocol • Toegang tot data kan worden beperkt met authenticatie en autorisatie • De metadata is toegankelijk, onafhankelijk van data
Toepassing binnen ontwerp:	Via de DV repository en InData zijn alle beschikbare object binnen de autorisaties van de gebruikers en ontwikkelaars vindbaar.
Relatie tot kader:	RIVM Architectuur

AP102 – Gegevens zijn Toegankelijk	
Beschrijving:	Zowel mensen als machines hebben eenvoudig toegang tot metadata en data. Authenticatie en autorisatie worden ingezet wanneer de eigenaar de toegang tot metadata en data wil beperken.
Rationale:	• Data kan opgevraagd worden door middel van een gestandaardiseerd protocol • Toegang tot data kan worden beperkt met authenticatie en autorisatie • Metadata is toegankelijk, onafhankelijk van data
Toepassing binnen ontwerp:	DV wordt toegankelijk op basis de rechten binnen de AD-groepen van de gebruikers en ontwikkelaars.
Relatie tot kader:	RIVM Architectuur

AP103 – Gegevens zijn onderling uitwisselbaar	
Beschrijving:	Metadata en data zijn onderling uitwisselbaar en kunnen gemakkelijk geïntegreerd worden in nieuwe datasets.
Rationale:	• Maak gebruik van metadata vocabulaires die gebaseerd zijn op open industrie standaarden • Maak gebruik van unieke identificatie van data en referenties naar data
Toepassing binnen ontwerp:	Door het gebruik van open standaard zijn alle gegevens binnen DV uitwisselbaar.
Relatie tot kader:	RIVM Architectuur

AP104 – Gegevens zijn herbruikbaar	
Beschrijving:	Metadata en data kunnen hergebruikt worden in nieuwe toepassingen.
Rationale:	• Data is herleidbaar (data provenance) • De voorwaarden waaronder data beschikbaar gesteld worden zijn bekend • Data maakt gebruik van domein-relevante (metadata)standaarden

Toepassing binnen ontwerp:	DV ondersteunt het hergebruik van data en metadata door middel van uitwisseling via de open standaarden en exporteren van metadata.
Relatie tot kader:	RIVM Architectuur

8 Bijlage C: Pseudonimiseringsoplossing

Pseudonimisering is een kerneigenschap van deze architectuur. Deze bijlage onderbouwt de keuze voor de implementatietechnologie. Hierbij is gekeken naar drie oplossingen. Twee hiervan staan binnen de omgeving ter beschikking. Eén daarvan heeft tevens aansluiting van een externe dienst.

	Database (ETL)		Datavirtualisatie (Code)		(Zorg)TTP	
Functionaliteit	*Pseudonimiseren met (Oracle)Database functies en ETL-programmatuur.* *Inlezen gegevens met ETL-tool.* *Sleutels pseudonimiseren in separaat schema.* *Pseudoniemen uitlezen via functie op dit schema.* *Depseudonimiseren en persoonsgegevens toevoegen via bevraging sleutelkast vanuit derde separaat schema.*		*Pseudonimiseren met (TIBCO)Datavirtualisatie functies (incl. SALT-Keys)* *Pseudoniemen uitlezen in virtuele view op persoons-gegevens.* *Scheiding via autorisaties.*		*Pseudonimiseren sleutels wordt uitbesteed aan een Trusted Third Party (ZorgTTP)* *Persoonsgegevens worden via API aangeboden aan TTP* *Pseudoniemen worden geretourneerd via API.* *Rollen: pseudonimiseren, de-pseudonimiseren en bevragen.*	
Architectuur	+/-	Maakt gebruik van bestaande middelen (hergebruik voor nieuwbouw) Sleutelkast of standaard versleutelingstechniek (AES-256) Kennis van GRIP in kleine kring aanwezig. Onderliggende taal (PL/SQL) is toekomstvast. GRIP-tool bedoeld voor ETL, niet pseudonimisering.	+	Maakt gebruik van voorkeur architectuur-bouwsteen (state of the art techniek) Standaard versleutelingstechniek (AES-256)? Datavirtualisatietool bedoeld voor dataverwerking, niet pseudonimisering.	+	Wordt uitgevoerd door een externe dienst (schaalvoordeel) Standaard versleutelingstechniek (AES-256) Regelt alleen versleuteling. Daarnaast tooling voor dataverwerking nodig bij depseudonimisering (i.c. GRIP, Datavirtualisatie of een alternatief).
Beveiliging	+	Scheiding van gegevens via separaat schema in Database. Sleutelopslag in de database. Toegang tot deze sleutels via Role-Based-Access (autorisatie in database). Bescherming van sleutels met 2MFA. (toegang tot RIVM Infrstructuur). Gegevenstransport over beveiligd RIVM DC Netwerk.	+	Scheiding van gegevens via separate map in Datavirtualisatie. Sleutelopslag op Datavirtualisatie server. Toegang tot deze sleutels via Role-Based-Access. Bescherming van sleutels met 2MFA (toegang tot RIVM Infrastructuur). Gegevenstransport over beveiligd RIVM DC Netwerk.	+/-	Scheiding van gegevens via separate dienst bij een externe partij. Dit heeft vanuit privacy-perspectief grote meerwaarde. Sleutelopslag op ZorgTTP-server. Toegang tot deze sleutels via Role-Based-Access. Toegang tot deze sleutels met gebruikersnaam/wachtwoord. Gegevenstransport over publiek Internet.

	Database (ETL)		Datavirtualisatie (Code)		(Zorg)TTP	
Beschikbaarheid	+	Database redundant uitgevoerd? Beheer in consignatiedienst.	+	Draait op redundant uitgevoerd platform. Beheer in consignatiedienst.	-	Enkelvoudig uitgevoerd. Beheer in kantooruren?
Inzetbaarheid	-	Pseudoniemen kunnen binnen CIMS opgevraagd worden (bijv. BSN).	+/-	Pseudoniemen kunnen RIVM-Breed ingezet worden.	+	Pseudoniemen kunnen binnen zorgdomein ingezet worden.
Realisatietijd	+	Toevoegen data- en autorisatieschema aan bestaande database. Haalbaar binnen huidige planning.	+/-	Toevoegen view- en autorisatieschema aan datavirtualisatie. Extra inspanning nodig bij het omleggen van gegevensstromen. Ambitieuze planning.	-	Extra inspanning nodig bij het omleggen van gegevensstromen. Ambitieuze planning.
Huidige stand van zaken en geplande doorlooptijd		Oplossing is technisch gereed en kan worden geactiveerd		TDV inrichting 4-8 dagen, Oracle TDV koppeling en configuratie 4-8 dagen in gunstigste geval. Daarna testen. Totaal circa 3 weken uitloop.		Configuratie uitdenken en implementeren minimaal 3 weken. Aanbesteding niet meegerekend