

1
2
3

4
5
6
7
8
9
10
11
12
13
14
15



Doelarchitectuur secundair datagebruik v0.8

Op weg naar een nationale gezondheidsdata-infrastructuur

16	Inhoud	
17		
18	Colofon	3
19		
20	1. Inleiding	4
21	2. Missie, visie en architectuurprincipe	10
22	3. Huidige situatie	14
23	4. Doelarchitectuur	16
24	5. Beheer doelarchitectuur	44
25	6. Roadmap	46
26		
27	Bijlage 1. Begrippenlijst.....	50
28	Bijlage 2. Afkortingenlijst	60
29	Bijlage 3. Toelichting architectuurprincipes	62
30		

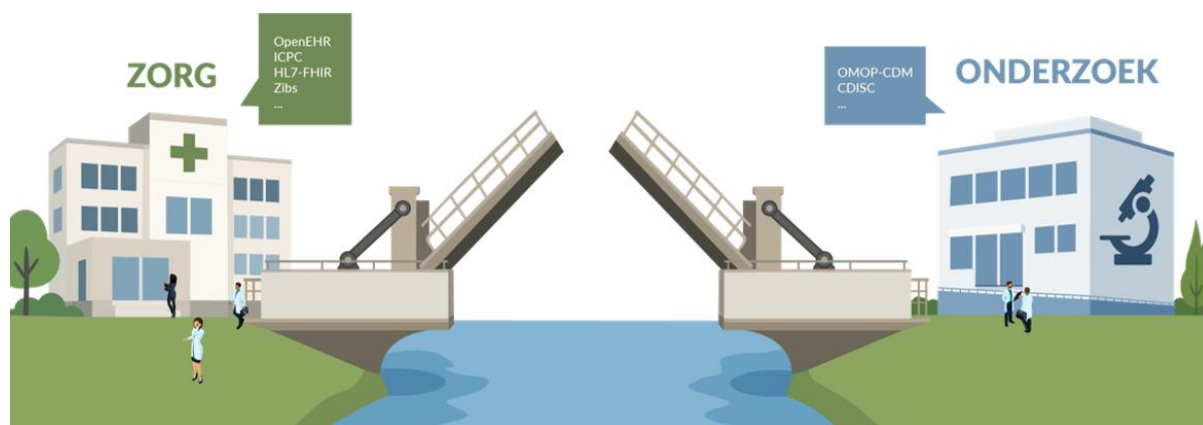
Colofon

Versie	Datum	Toelichting
0.1	2026-03	Eerste opzet door Marieke Groeneveld, Igor Schoonbrood, Rob Cornelisse, Roger Snijder en Mark de Lange
0.2	2026-05	Oplevering hoofdlijnen architectuur door Niels Bolding, Jeroen Beliën en Mark de Lange voor consultatie Health-RI hub en knooppunten
0.3	2026-07	Verwerking commentaren hoofdlijnen architectuur en oplevering doelarchitectuur door Niels Bolding, Jeroen Beliën en Mark de Lange voor consultatie Health-RI hub en knooppunten
0.4 – 0.6	2026-07/08	Verwerking commentaren Health-RI en knooppunten door Niels Bolding, Jeroen Beliën en Mark de Lange
0.7	2026-08	Oplevering doelarchitectuur voor toetsing MT Health-RI
0.8	2026-08	Verwerking commentaren en oplevering doelarchitectuur door Niels Bolding en Mark de Lange voor open consultatie

1. Inleiding

Om de gezondheidszorg zowel efficiënter als betaalbaarder te maken, is de beschikbaarheid van gezondheidsdata essentieel. Dit geldt voor zowel het primaire gebruik, gericht op directe zorg aan burgers, als voor het secundaire gebruik van data buiten de directe zorgverlening. De nationale infrastructuur voor gezondheidsdata ten behoeve van secundair gebruik vormt het fundament van een lerend zorgsysteem en draagt bij aan zowel hoogwaardige zorg als het versterken van de innovatiekracht binnen Nederland.

Health-RI ontwikkelt samen met diverse Nederlandse partners een robuuste infrastructuur voor gezondheidsdata. Hierbij wordt nauw samengewerkt met Cumuluz, de uitvoeringsorganisatie van het Landelijk Dekkend Netwerk, evenals met het IZA-programma voor secundair gebruik. Hiermee slaat Health-RI de brug tussen primaire zorg en secundair hergebruik van gezondheidsdata (figuur 1).



Figuur 1. Brug tussen zorg en onderzoek.

Daarnaast worden er intensieve samenwerkingen aangegaan met kwaliteitsregistraties om een landelijke infrastructuur voor secundair gebruik te realiseren. Met TWIIN wordt gewerkt aan een landelijk afsprakenstelsel voor zowel primair als secundair datagebruik.

Binnen de tijdshorizon van deze doelarchitectuur worden de EHDS, HealthData@EU infrastructuur en de HDAB-NL operationeel. De ontwikkeling van de Nationale gezondheidsdata-infrastructuur secundair gebruik vindt daarom plaats met zowel de huidige als de toekomstige situatie als uitgangspunt. Daarbij wordt bewust gestuurd op continuïteit: functies en voorzieningen die vandaag worden gerealiseerd, zijn waar mogelijk ook bruikbaar binnen het toekomstige Europese kader. Op deze wijze ontstaat een groeipad waarin stap voor stap wordt toegewerkt naar de EHDS-doelsituatie, terwijl investeringen, implementaties en opgedane ervaringen maximaal kunnen worden hergebruikt.

Voor secundair gebruik bevindt het Health-RI afsprakenstelsel zich momenteel in ontwikkeling; dit stelsel wordt voortdurend aangevuld en verbeterd. Jaarlijks wordt de actuele inhoud breed afgestemd met het zorgveld. Naast deze periodieke consultaties bestaat tevens de behoefte om specifieke onderwerpen tussentijds te consulteren. De in dit document beschreven doelarchitectuur voor secundair datagebruik is daar een voorbeeld van, maar ook op andere thema's zullen separate feedbackrondes plaatsvinden. Dit document vormt een eerste stap richting een gezamenlijke doelarchitectuur voor een nationale gezondheidsdata-infrastructuur die meervoudig gebruik van gezondheidsdata faciliteert, met bijzondere aandacht voor secundair datagebruik. De inhoud van dit document kan dan ook in de toekomst worden herzien of uitgebreid.

Bij het opstellen van deze doelarchitectuur zijn de elementen uit het Health-RI afsprakenstelsel als uitgangspunt genomen. Deze elementen zijn vervolgens getoetst aan onder meer:

- de Gezondheidsdata Informatie Stelsel Architectuur (GISA);
- de CumuluZ-doelarchitectuur;
- de European Health Data Space verordening (EHDS) en TEHDAS2- richtlijnen en technische specificaties;
- de specificaties voor het gefedereerde datastation ([1. Inleiding - Datastations voor secundair gebruik in de zorg](#));
- de datastation specificaties uit de Health-RI Data-GitHub ([GitHub - Health-RI/data-station-specification: Specification of composable data stations · GitHub](#));
- voortschrijdende inzichten, mede vanuit de partijen die momenteel nauw betrokken zijn bij Health-RI, zoals de UMC's.

In dit hoofdstuk wordt verder stilgestaan bij het doel en de scope van de doelarchitectuur. Ook wordt het document gepositioneerd en worden de ontwikkelingen geduid. Tot slot wordt ingegaan op de lezersdoelgroep en wordt een leeswijzer gegeven.

1.1. Doel

De in dit document beschreven doelarchitectuur geeft richting aan de ontwikkeling van landelijke en gezondheidssector brede databeschikbaarheid en het hergebruik van gezondheidsdata voor onderzoek, beleid en innovatie¹. De doelarchitectuur biedt een gemeenschappelijk en stabiel kader voor ontwerp- en implementatiekeuzes van zorgorganisaties, leveranciers en andere betrokken partijen en ondersteunt daarmee de ontwikkeling van een samenhangende, duurzame en interoperabele nationale infrastructuur voor secundair gebruik. Het kader geldt zowel vóór als na de operationalisatie van de Health Data Access Body (HDAB). De doelarchitectuur vormt daarmee een fundament voor een soepele aansluiting op de EHDS en voor verdere uitwerking in solution- en softwarearchitecturen, met functies, applicatiecomponenten en koppelvlakken die interoperabele en uitwisselbare implementaties mogelijk maken.

1.2. Scope

De scope van de doelarchitectuur wordt bepaald door de tijdshorizon, functionele reikwijdte, technische inrichting en betrokken partijen.

- **Scope in tijd;** De doelarchitectuur beschrijft de ontwikkeling van de Nationale Gezondheidsdata-infrastructuur Secundair Gebruik over een periode van vijf jaar. Binnen deze periode worden de EHDS, de HealthData@EU infrastructuur en de HDAB-NL operationeel. De architectuur richt zich daarom zowel op de huidige situatie als op de toekomstige situatie waarin de HDAB is geoperationaliseerd. Daarbij verschuiven verantwoordelijkheden rondom datatoegang gedeeltelijk van datahouders naar de HDAB, terwijl functies en voorzieningen zoveel mogelijk gelijk blijven. Hierdoor vormt de vanuit het Nationaal Groeifonds gerealiseerde infrastructuur een duurzaam fundament voor de toekomstige EHDS-situatie.
- **Functionele scope;** De doelarchitectuur richt zich op landelijk en gezondheidssector breed databeschikbaarheid en hergebruik van gezondheidsdata. De architectuur is primair uitgewerkt voor onderzoek, beleid en innovatie, maar is vanuit het principe generiek boven specifiek zo opgezet dat zij ook toepasbaar kan zijn op andere vormen van secundair datagebruik, zoals kwaliteitsregistraties. De daadwerkelijke toepasbaarheid voor deze domeinen wordt in een volgende fase nader onderzocht.

¹ Vanuit principe van Generiek boven specifiek dient de doelarchitectuur voor secundair gebruik in de breedste zin toepasbaar te zijn. De doelarchitectuur is momenteel vooral gericht op onderzoek. Naar verwachting is voor de bredere toepassing nog aanpassing van de doelarchitectuur nodig. Nader onderzoek in een volgende fase dient dit uit te wijzen.

- **Technische scope;** De doelarchitectuur beschrijft op hoofdlijnen de toekomstige inrichting van processen, data, applicaties en koppelvlakken voor secundair datagebruik. De architectuur schrijft daarbij geen specifieke producten, leveranciers of technische oplossingen voor.
- **Organisatorische scope;** De doelarchitectuur beschrijft de verantwoordelijkheden van partijen die betrokken zijn bij secundair datagebruik van gezondheidsdata, waaronder datahouders, datagebruikers en regie voerende organisaties binnen en rondom de gezondheidssector.

1.3. Positionering

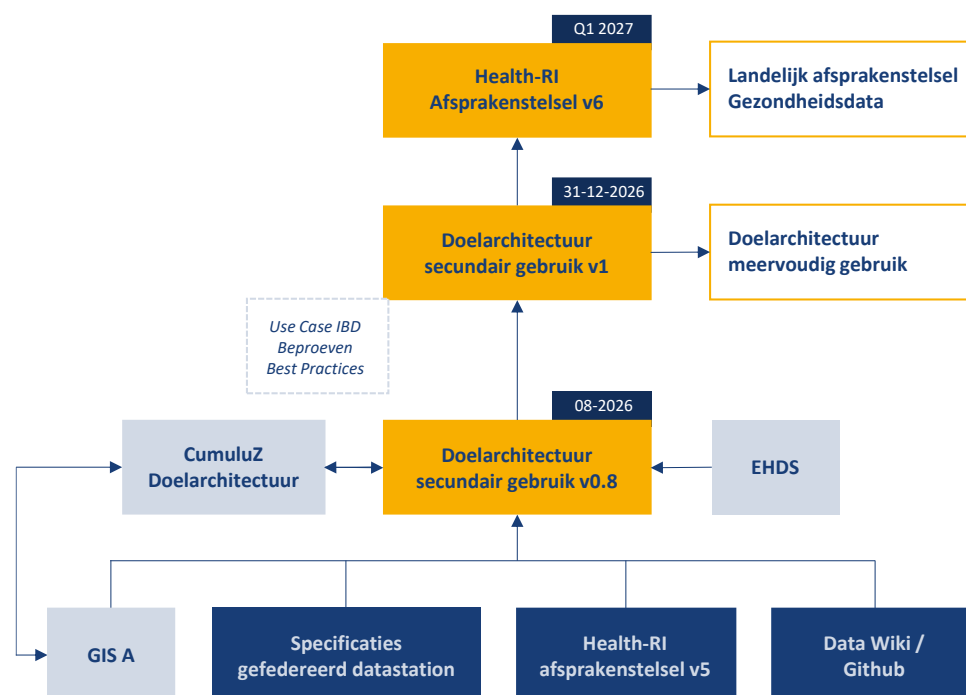
Deze doelarchitectuur staat niet op zichzelf, maar maakt onderdeel uit van een bredere ontwikkeling naar één samenhangende nationale en Europese gezondheidsdata-infrastructuur. De doelarchitectuur maakt deel uit van het Health-RI afsprakenstelsel en sluit aan op bestaande nationale en Europese afsprakenstelsels, architecturen en infrastructuren voor databeschikbaarheid. Daarmee draagt zij bij aan de ontwikkeling van één samenhangend gezondheidsinformatiestelsel waarin zowel primair als secundair gebruik van gezondheidsdata wordt ondersteund.

Binnen Nederland vormen de Nationale Visie & Strategie, het Landelijk Afsprakenstelsel (LAS), de GISA, het Landelijk Dekkend Netwerk (LDN) en de CumuluZ-doelarchitectuur belangrijke kaders. Op Europees niveau vormen de EHDS en de HealthData@EU infrastructuur de belangrijkste randvoorwaarden voor grensoverschrijdende interoperabiliteit en datagebruik. Daarnaast sluit het LAS aan op domeinoverstijgende ontwikkelingen zoals de Interbestuurlijke Datastrategie (IBDS), het Federatief Datastelsel (FDS) en de inrichting van data spaces, waaronder de European Open Science Cloud.

Ontwikkeling	Hoofdfunctie
EHDS / HealthData@EU	Europees kader voor grensoverschrijdend datagebruik
FDS/ IBDS	Landelijke domein overstijgende afspraken voor databeschikbaarheid
GISA	Stelselarchitectuur voor het gezondheidsinformatiestelsel
LAS	Landelijke afspraken voor gezondheid databeschikbaarheid
LDN / CumuluZ	Landelijke infrastructuur voor databeschikbaarheid binnen de zorg
NVS	Nationale visie van VWS op het gezondheidsinformatiestelsel en de strategie om daar te komen.

De Nationale gezondheidsdata-infrastructuur secundair gebruik positioneert zich hierbij als de architecturale invulling voor secundair gebruik van gezondheidsdata binnen het bredere nationale gezondheidsinformatiestelsel en de toekomstige Europese Health Data Space. Daarbij wordt maximaal aangesloten op bestaande afspraken, standaarden, voorzieningen en organisatie van verantwoordelijkheden-structuren, zodat een samenhangende infrastructuur ontstaat voor zorg, onderzoek, beleid en innovatie.

De CumuluZ-doelarchitectuur richt zich primair op het beschikbaar maken van data voor het primaire zorgproces in het kader van het LDN. Health-RI is verantwoordelijk voor de gezondheidsdata-infrastructuur voor secundair gebruik. Samen met CumuluZ wordt toegewerkt naar één samenhangende doelarchitectuur voor zowel primair als secundair hergebruik van gezondheidsdata. Deze doelarchitectuur vormt daarvoor een eerste aanzet en sluit aan op de CumuluZ-doelarchitectuur.



Figuur 2. Van doelarchitectuur versie 0.8 naar inbedding in afsprakenstelsel

Figuur 2 toont dat verschillende bronnen zijn gebruikt om tot versie 0.8 van de doelarchitectuur voor secundair gebruik van gezondheidsdata te komen. In augustus 2026 is een open feedback ronde gestart, die loopt tot en met eind september. De commentaren worden gebruikt om tot een 1.0 versie te komen. Deze versie wordt eind 2026 gepubliceerd. Zowel de CumuluZ doelarchitectuur als de doelarchitectuur secundair gebruik worden geïntegreerd binnen de architectuur voor meervoudig gebruik die door het ministerie van VWS wordt vastgesteld in het kader van de GISA². De doelarchitectuur voor secundair gebruik zal in het eerste kwartaal van 2027 ook worden ingebed in het afsprakenstelsel van Health-RI.

1.4. Ontwikkelingen

In deze paragraaf wordt ingegaan op de belangrijkste ontwikkelingen die een drijfveer vormen voor de totstandkoming van een Nationale gezondheidsdata-infrastructuur secundair gebruik. De volgende ontwikkelingen worden onderkend:

1. Europese wetgeving.
2. Toenemende internationale samenwerking en grensoverschrijdend onderzoek.
3. Groeiende behoefte aan datagedreven onderzoek.
4. Opkomst van AI en data-intensieve innovaties.

Hierna wordt kort ingegaan op elk van deze ontwikkelingen.

1.4.1. Europese wetgeving

De EHDS vormt de belangrijkste juridische aanjager voor de ontwikkeling van de Nationale gezondheidsdata-infrastructuur secundair gebruik. De verordening verplicht lidstaten om voorzieningen in te richten voor secundair gebruik van gezondheidsdata ten behoeve van onderzoek, innovatie, beleid en AI-ontwikkeling. Daarbij stelt de EHDS eisen aan onder meer datatoegang, metadata, interoperabiliteit, datakwaliteit, beveiligde verwerkingsomgevingen en grensoverschrijdende gegevensuitwisseling via de HealthData@EU infrastructuur.

Binnen de tijdshorizon van deze doelarchitectuur worden de EHDS en de bijbehorende HealthData@EU infrastructuur operationeel. Daarmee verschuift de context van een nationaal

² Meer informatie over GISA is te vinden op [1. Introductie GISA - GISA Consultatie - Confluence](#).

georganiseerd stelsel naar een Europees ecosysteem. Voor Nederland betekent dit dat de ontwikkeling van een Nationale gezondheidsdata-infrastructuur secundair gebruik niet alleen voortbouwt op de reeds ingezette ontwikkelingen vanuit onder meer het Nationaal Groeifonds, maar ook noodzakelijk is om aan Europese verplichtingen te voldoen en als volwaardig onderdeel van de HealthData@EU infrastructuur te functioneren.

Een van de belangrijkste veranderingen betreft de data-aanvraagfunctie. Waar datahouders nu verantwoordelijk zijn voor de inhoudelijke beoordeling van aanvragen, verschuift deze verantwoordelijkheid voor aanvragen die onder de EHDS vallen naar de HDAB. Deze verandering raakt niet alleen de toegangsfunctie, maar de gehele keten van datatoegang, waaronder autorisatie, gebruiksregistratie, outputcontrole en de vrijgave van resultaten vanuit beveiligde verwerkingsomgevingen.

Naast de EHDS beïnvloeden ook de AI Act, Medical Device Regulation (MDR), In Vitro Diagnostic Regulation (IVDR), Data Governance Act (DGA) en de Algemene Verordening Gegevensbescherming (AVG) het secundair gebruik van gezondheidsdata. Samen vormen zij het juridische kader waarbinnen de Nationale gezondheidsdata-infrastructuur secundair gebruik zich ontwikkelt.

1.4.2. Toenemende internationale samenwerking en grensoverschrijdend onderzoek

Veel gezondheidsvraagstukken vragen om samenwerking over landsgrenzen heen. De EHDS ondersteunt dit door nationale infrastructuren via de HealthData@EU infrastructuur met elkaar te verbinden. Hierdoor kunnen datasets uit verschillende lidstaten beter worden gevonden en de data daaruit gebruikt voor onderzoek. De nationale gezondheidsdata-infrastructuur secundair gebruik draagt daarmee niet alleen bij aan nationale, maar ook aan Europese databeschikbaarheid.

1.4.3. Groeiende behoefte aan datagedreven onderzoek

De hoeveelheid beschikbare gezondheidsdata groeit sterk. Onderzoekers hebben steeds vaker behoefte aan het combineren van data uit verschillende organisaties en domeinen om nieuwe inzichten te verkrijgen. De nationale gezondheidsdata-infrastructuur secundair gebruik ondersteunt dit door data beter vindbaar, toegankelijk en herbruikbaar te maken.

1.4.4. Opkomst van AI en data-intensieve innovaties

AI is zowel een belangrijke gebruiker als een aanjager van de Nationale gezondheidsdata-infrastructuur secundair gebruik:

- Net als onderzoek, beleid en innovatie is de ontwikkeling van AI een vorm van secundair gebruik van gezondheidsdata. AI-toepassingen zijn afhankelijk van toegang tot grote hoeveelheden kwalitatieve data voor training, validatie en doorontwikkeling en kennen daarmee grotendeels dezelfde behoeften als andere datagebruikers.
- Tegelijkertijd maakt AI de waarde van databeschikbaarheid sneller zichtbaar. Waar onderzoek en innovatie vaak een lange doorlooptijd kennen, kunnen AI-toepassingen relatief snel tot concrete resultaten leiden. Hierdoor neemt de urgentie toe om gezondheidsdata beter beschikbaar, vindbaar en herbruikbaar te maken.
- Daarnaast kan AI bijdragen aan de infrastructuur zelf. AI kan ondersteuning bieden bij het vinden van datasets, het vastleggen van data aan de bron, het transformeren en harmoniseren van data en het selecteren van relevante datasets voor specifieke toepassingen. Daarmee kan AI zowel de kwaliteit als de bruikbaarheid van gezondheidsdata vergroten.

Het momentum rond AI versnelt daarmee de ontwikkeling van het bredere ecosysteem voor onderzoek, beleid en innovatie. Tegelijkertijd bestaat het risico dat AI-initiatieven eigen, project specifieke data-oplossingen ontwikkelen, wat leidt tot verdere versnippering. Het is daarom essentieel dat AI-toepassingen aansluiten op de generieke Nationale gezondheidsdata-infrastructuur

secundair gebruik. Alleen zo kunnen schaalbaarheid, herbruikbaarheid en duurzaamheid van het lerend gezondheidszorgsysteem worden geborgd.

1.5. Lezersdoelgroep

Dit document is geschreven voor architecten die betrokken of belanghebbend zijn bij databeschikbaarheid voor secundair datagebruik, maar bevat ook relevante informatie voor andere doelgroepen, zoals bijvoorbeeld leveranciers. Het primaire hoofddoel van deze architectuur is om de keuzes die vanuit Health-RI de komende periode gezamenlijk met het veld op detailniveau gaan maken, vast te leggen. Daarmee wordt een handelingskader geboden:

- voor softwareleveranciers en dienstverleners om oplossingen te implementeren;
- voor gezondheidsorganisaties om te begrijpen hoe zij optimaal kunnen participeren in de Nationale gezondheidsdata-infrastructuur secundair gebruik.

1.6. Leeswijzer

In hoofdstuk 2 wordt ingegaan op de missie, visie en de basisprincipes, waarna in hoofdstuk 3 op hoofdlijnen de huidige situatie omtrent hergebruik van gezondheidsdata wordt geschetst. Hoofdstuk 4 beschrijft vervolgens de eigenlijke, op de toekomstgerichte doelarchitectuur (processen, data, applicaties en verantwoordelijkheden). In hoofdstuk 5 wordt het beheer van de doelarchitectuur toegelicht. Afgesloten wordt met de roadmap op weg naar de realisatie van de doelarchitectuur in hoofdstuk 6.

In bijlage 1 is een uitgebreide lijst met begrippen en hun duiding opgenomen en in bijlage 2 een afkortingenlijst. Voor de duiding van de begrippen wordt zoveel als mogelijk gebruik gemaakt van bestaande duidingen. Echter, er zijn veel begrippen die op verschillende plekken net een iets andere betekenis hebben. Het advies is dan ook om eerst de begrippenlijst door te nemen aangezien begrippen in andere contexten een net iets andere betekenis kunnen hebben. Twee belangrijke begrippen in dit document worden hier kort uitgelicht:

- Er wordt gesproken over data, niet over gegevens. Het begrip informatie wordt daarnaast gebruikt waar data betekenis heeft voor mensen in de context waarin zij dit nodig hebben. Dit is in lijn met de GISA.
- De term burger wordt gebruikt voor de persoon die deelneemt in het zorgproces omdat deze zorg wil krijgen, krijgt of gekregen heeft. Onder de term burger vallen nadrukkelijk ook patiënt en cliënt. Een burger kan een patiënt of cliënt zijn, maar dat hoeft niet. Een bredere term dan cliënt of patiënt is nodig omdat de doelarchitectuur niet alleen over zorg gaat, maar ook over welzijn en gezondheid, en dus ook over mensen die geen zorg ontvangen omdat ze niet ziek zijn.

Tot slot wordt in bijlage 3 op de architectuurprincipes een uitgebreide toelichting gegeven inclusief rationale en implicaties.

2. Missie, visie en architectuurprincipe

Voordat dieper wordt ingegaan op de huidige en toekomstige situatie, wordt in dit hoofdstuk stilgestaan bij de missie en visie van Health-RI en worden de gezamenlijk overeengekomen architectuurprincipes toegelicht. Ook wordt ingegaan op de ontwikkelingen die een drijvende kracht vormen achter de noodzaak om tot een Nationale gezondheidsdata-infrastructuur secundair gebruik te komen.

2.1. Missie en beoogde maatschappelijke waarde van secundair gebruik van data

De missie van Health-RI is om de gezondheid van burgers te verbeteren door hergebruik van gezondheids- en life sciences data met een geïntegreerde data-infrastructuur voor datagedreven onderzoek, beleid en innovatie. Het creëren van een geïntegreerde data-infrastructuur voor gezondheids- en life sciences data vergroot niet alleen de kwaliteit van de gezondheidszorg. Het draagt ook bij aan meer en sneller onderzoek, waardoor de innovatiekracht van Nederland wordt vergroot. Het eenvoudiger en beter kunnen delen en hergebruiken van gezondheidsdata versnelt onderzoek, beleid en innovatie. Dit geeft niet alleen economische groei, maar is helpt ook de gezondheid van burgers.

2.2. Visie

De kwaliteit van de huidige gezondheidszorg is het resultaat van decennia van onderzoek, beleid en innovatie en het vermogen om te leren van de kennis, inzichten en data uit de praktijk. Deze continue wisselwerking tussen zorg en onderzoek vormt de kern van het lerend gezondheidszorgsysteem: een systeem waarin data uit de dagelijkse zorg worden benut om nieuwe inzichten te genereren, die vervolgens worden omgezet in betere behandeling, beleid en preventie. Op dit moment staat de zorg onder druk door een toenemende zorgvraag en een afnemende zorgcapaciteit. Daarom is het nodig om dit lerend vermogen structureel te versterken. Een Nationale gezondheidsdata-infrastructuur secundair gebruik maakt het mogelijk om data veilig, herbruikbaar en schaalbaar in te zetten voor onderzoek, beleid en innovatie, en vormt daarmee een essentiële randvoorwaarde voor toekomstbestendige, toegankelijke en betaalbare zorg.

2.3. Architectuurprincipes

In de NVS en GISA van het ministerie van VWS worden drie fundamenteën van het gezondheidsinformatiestelsel gehanteerd:

1. Databeschikbaarheid: Gegevens veilig en eenvoudig delen voor betere en passende zorg;
2. Vertrouwen: Zorg dat mensen en systemen bouwen op betrouwbare data en veilige apps (trust en confidence);
3. Regie: Samen sturen op landelijke standaarden en duidelijke afspraken.

In het verlengde van deze drie fundamenteën zijn er in de NVS vier en in de GISA 7 leidende principes vastgesteld:

1. NVS: Data is beschikbaar voor alle betrokkenen in het zorgnetwerk van een burger en voor de burger zelf
2. NVS: Data is beschikbaar voor secundair gebruik, met minimale registratielast voor zorgverleners
3. NVS: Data en functionaliteit worden los van elkaar georganiseerd om flexibiliteit en hergebruik te ondersteunen
4. NVS: Het gezondheidsinformatiestelsel creëert een open markt die innovatie stimuleert
5. GISA: Het GIS is mens- en functioneel gericht
6. GISA: Het GIS is samenhangend en adaptief
7. GISA: Het GIS heeft een open federatieve architectuur
8. GISA: Gezondheidsdata wordt eenmalig vastgelegd bij de bron en beschikbaar gesteld voor meervoudig gebruik

9. GISA: Het GIS borgt onderling vertrouwen

10. GISA: Het GIS waarborgt de integriteit en kwaliteit van data

11. GISA: Het GIS gebruikt bestaande passende oplossingen

Voor een verdere toelichting op deze principes wordt verwezen naar de [NVS](#), respectievelijk de GISA.

Health-RI omarmt deze nationale principes en heeft in aansluiting hierop nog enige aanvullende architectuurprincipes geformuleerd. Hiermee wordt richting en structuur gegeven aan de inrichting van de Nationale gezondheidsdata-infrastructuur secundair gebruik. Om de leesbaarheid van dit document te behouden, zijn in dit hoofdstuk uitsluitend de definities van de aanvullende principes opgenomen. De volledige uitwerking van elk principe inclusief rationale en implicaties is opgenomen in Bijlage 3. Architectuurprincipes.

De volgende aanvullende principes worden door Health-RI gehanteerd:

1. **Data nabij de bron waar mogelijk:** Gezondheidsdata voor secundair gebruik (zoals wetenschappelijk onderzoek, beleid en innovatie) worden bij voorkeur niet permanent gecentraliseerd in één grote landelijke database. In plaats daarvan blijven de gegevens waar mogelijk staan in de bronsystemen van de datahouder of worden ze naar een regionale/lokale voorziening gebracht dicht bij de bron, totdat ze nodig zijn. In de Nationale gezondheidsdata-infrastructuur secundair gebruik wordt er wel rekening mee gehouden dat soms verplaatsing, kopiëren of centrale opslag van data nodig is. Dit vindt echter alleen plaats wanneer dit noodzakelijk is voor de beoogde toepassing, aantoonbare meerwaarde biedt of vanuit wet- en regelgeving vereist is.
2. **Alleen delen analyseresultaat:** Wanneer het functioneel, juridisch en technisch mogelijk is, wordt alleen de analyse toegelaten tot de data (data visiting). Als het mogelijk is worden na de analyse de resultaten gedeeld. De resultaten bevatten normaliter alleen statistische gegevens en dus geen persoonsgegevens. Dit sluit aan bij het AVG-principe van Privacy by Design, waarbij de meest privacy vriendelijke verwerkingsvorm wordt toegepast die de beoogde doelstelling ondersteunt.
3. **Verantwoordelijkheid en regie bij de datahouder:** De datahouder is en blijft verantwoordelijk voor de rechtmatigheid, de kwaliteit en de beveiliging van de gezondheidsdata, én beslist over de beschikbaarstelling voor secundair gebruik binnen de kaders van de wet en landelijke afspraken. Wel wordt de regie van de datahouder beperkt op het moment dat de operationele HDAB aan een datagebruiker via vergunningverlening permissie heeft gegeven om data beschikbaar te krijgen en te gebruiken. De datahouder dient in dat geval de data beschikbaar te stellen. Hiertegen kan slechts bezwaar worden aangetekend bij een vermeende overtreding van intellectueel eigendomsrechten.
4. **Decentrale autonomie over eigen gezondheidsdata-infrastructuur:** (Onafhankelijke) organisaties binnen het brede ecosysteem van gezondheid, preventie en zorg behouden het recht om binnen de geldende wet- en regelgeving hun eigen (ICT-)infrastructuur bronsystemen en databeheer in te richten t.b.v. secundair gebruik, mits zij voldoen aan de (inter)nationale vastgestelde afspraken voor interoperabiliteit en gegevensuitwisseling. De architectuur beschrijft dan ook functionaliteit, geen specifieke oplossingen die invulling geven aan functionaliteit. De decentrale autonomie op de inrichting en implementatie wordt gerespecteerd.
5. **Gestandaardiseerde ontsluiting:** Toegang tot gezondheidsdata en de bijbehorende metadata verloopt zo veel mogelijk via gestandaardiseerde, open, internationaal geaccepteerde en goed gedefinieerde koppelvlakken en informatie-architectuurstandaarden.
6. **Hergebruik als uitgangspunt:** Data, datasets en onderzoeksresultaten worden vanaf het begin ontworpen voor hergebruik.
7. **Generiek boven specifiek:** Oplossingen en voorzieningen worden generiek ingericht, zodat ze herbruikbaar zijn over gezondheidszorg, onderzoek, beleid en innovatie, tenzij een specifieke casus aantoonbaar een uitzonderingspositie vereist. De doelarchitectuur stuurt dan ook actief

- op uniformiteit en samenhang om gezamenlijk te convergeren naar landelijke (en Europese) doelarchitectuur, gedeelde standaarden en centrale/federatieve voorzieningen.
8. **Proces- en gebruikersgericht:** Richt de doelarchitectuur, voorzieningen en organisatie van verantwoordelijkheden zodanig in dat deze optimaal aansluiten op de feitelijk werkprocessen en behoeften van zowel de datahouders als de datagebruikers., zonder onnodige administratieve of technische barrières. Ontwerpbeslissingen worden hierbij gestuurd door de behoeften van gebruikers en werkprocessen, niet door technologie of architectuur als doel op zich.
 9. **Systeembenadering:** Architectuurbeslissingen worden genomen vanuit een integraal perspectief, met expliciete afweging van impact op het totale stelsel. Keuzes moeten dus bijdragen aan het grotere geheel.
 10. **Internationaal boven nationaal boven lokaal:** Internationale standaarden, afspraken en ontwikkelingen zijn leidend boven nationale keuzes ("pas toe of leg uit"). Nationale afspraken zijn op hun beurt leidend boven lokale implementaties. Afwijkingen van internationaal of nationaal vastgestelde standaarden zijn alleen toegestaan wanneer daarvoor een aantoonbare noodzaak bestaat en de afwijking expliciet wordt gemotiveerd. Waar harmonisatie nodig is, wordt deze bij voorkeur georganiseerd op nationaal niveau in plaats van door individuele partijen afzonderlijk.
 11. **Omgaan met diversiteit:** De doelarchitectuur ondersteunt variatie in datatypen, datakwaliteit, volwassenheid en organisatorische context van datahouders, zodat uiteenlopende datahouders kunnen participeren en data beschikbaar kunnen stellen op een wijze die past bij hun mogelijkheden en situatie. Hierbij wordt rekening gehouden met verschillen tussen bijvoorbeeld gecureerde onderzoekscollecties, real world data, ad-hoc dataverzamelingen en verschillende typen zorg- en kennisinstellingen.
 12. **Maximale benutting van bestaande data (incl. legacy):** Bestaande data, inclusief historische en legacy-data, worden waar mogelijk actief hergebruikt voor onderzoek, beleid en innovatie. De doelarchitectuur stimuleert dat bestaande gegevens niet opnieuw worden verzameld wanneer deze reeds beschikbaar zijn en geschikt kunnen worden gemaakt voor secundair gebruik. Daarbij worden kwaliteit, herkomst, context en juridische randvoorwaarden expliciet meegewogen.
 13. **Wat kan wel:** Richt beslissingen primair op wat mogelijk is binnen de geldende randvoorwaarden, in plaats van op wat nog niet mogelijk is. Accepteer dat oplossingen of voorzieningen niet vanaf het begin volledig of perfect hoeven te zijn, mits zij voldoende waarde bieden om verantwoord mee te starten en verder te ontwikkelen.

2.4. Federatief versus centraal model

Bij onderzoek kan de analyse van data op meerdere manieren worden georganiseerd: in een centraal model of in een federatief model:

- Bij een centraal model worden data van verschillende datahouders, zoals bijvoorbeeld zorgaanbieders of onderzoeksinstellingen, samengebracht in één enkele geharmoniseerde dataset in één centrale beveiligde verwerkingsomgeving (BVO). Datagebruikers maken vervolgens gebruik van deze centrale omgeving voor het uitvoeren van analyses. De wijze waarop de analyse plaatsvindt kan variëren van data access, waarbij gebruikers toegang krijgen tot de data binnen een beveiligde verwerkingsomgeving, tot data visiting, waarbij analyses naar de beschikbare data worden gebracht en uitsluitend gecontroleerde resultaten worden vrijgegeven en waar nodig worden geconsolideerd met resultaten van andere deelnemende datahouders. Een centraal model maakt analyses relatief eenvoudig en snel, omdat alle data op één plek beschikbaar zijn. Nadelen zijn dat het verzamelen (en soms ook actueel) houden van de data veel inspanning vraagt en dat het voor datahouders lastiger is om controle te houden over hun eigen data. Ook vormt bij het verzamelen van data op één plek misbruik een groter risico.
- Bij een federatief model blijven de data bij de bron. Analyses worden uitgevoerd binnen de omgeving van de datahouder. Afhankelijk van de toepassing kan daarbij sprake zijn van data access of data visiting. Doordat de data niet hoeft te worden getransporteerd naar een centrale BVO behouden datahouders de regie over hun data en is het eenvoudiger om data actueel te

houden. Gevoelige (persoons-)gegevens hoeven niet centraal te worden opgeslagen. Dit is in lijn met het principe van dataminimalisatie vanuit privacywetgeving. Daar staat tegenover dat federatieve analyses technisch complexer zijn, afhankelijk zijn van goede data-interoperabiliteit en standaardisatie en soms minder efficiënt zijn dan werken met één centrale dataset.

Een centraal model biedt dus meer eenvoud en snelheid voor analyses, terwijl een federatief model meer nadruk legt op datasoevereiniteit en samenwerking tussen organisaties. De keuze tussen beide modellen hangt af van de onderzoeksdoelen, de wettelijke kaders en de mate waarin deelnemende organisaties bereid zijn data centraal beschikbaar te stellen. Daarnaast zullen ook combinaties van centraal en federatief voorkomen in de praktijk. De doelarchitectuur omvat dan ook beide modellen.

3. Huidige situatie

Health-RI neemt sinds 2022 het voortouw bij het vaststellen van een generieke set afspraken over de noodzakelijke functies voor databeschikbaarheid, ontwikkeld door en voor datagebruikers (zoals onderzoekers en innovatoren) en datahouders (zoals een zorgaanbieder, preventie-instelling, publieke gezondheidsorganisatie of kennisinstelling). Nieuwe ontwikkelingen, zoals de EHDS, de groeiende behoefte aan data-intensief onderzoek en de opkomst van AI-toepassingen, versterken de behoefte aan een dergelijke infrastructuur waarmee gezondheidsdata sneller, beter, veiliger en op grotere schaal beschikbaar kunnen worden gemaakt.

In de afgelopen jaren is binnen Health-RI, onder meer via het Obstakel Verwijder Traject (OVT), samen met stakeholders uit de gezondheidszorg, onderzoek, overheid, burger- en patiënten vertegenwoordiging en het bedrijfsleven gewerkt aan het wegnemen van een breed spectrum van maatschappelijke, organisatorische, juridische, procesmatige en technische obstakels rondom de huidige databeschikbaarheid. Het OVT levert hierbij waardevolle inzichten op in de belemmeringen die partijen in de praktijk ervaren en de maatregelen die nodig zijn om deze weg te nemen. Niet al deze obstakels raken direct de doelarchitectuur; veel vraagstukken liggen buiten de scope van architectuur en vragen bijvoorbeeld om bestuurlijke, juridische of maatschappelijke oplossingen. De hierna beschreven uitdagingen zijn daarom geen één-op-één vertaling van de binnen het OVT geïdentificeerde obstakels, maar een architecturale duiding van een aantal belangrijke kenmerken van de huidige situatie die relevant zijn voor de ontwikkeling van de Nationale gezondheidsdata-infrastructuur secundair gebruik. Deze uitdagingen vormen daarmee de belangrijkste procesmatige en technische aandachtspunten waarop de doelarchitectuur richting geeft. Het betreffen de volgende uitdagingen:

- Gezondheidsdata zijn versnipperd en moeilijk vindbaar
- Beperkingen bij het koppelen van gezondheidsdata.
- Administratieve lasten beperken duurzame databeschikbaarheid
- Gezondheids(metadata) zijn niet uniform gestructureerd volgens standaarden
- Ontbreken van context
- Beperkte uniforme inrichting

Deze uitdagingen beïnvloeden de efficiëntie, schaalbaarheid en betrouwbaarheid van secundair datagebruik. Health-RI heeft inmiddels forse stappen genomen in het wegnemen van deze uitdagingen. De voortgang hierop is meegenomen bij hierna gegeven toelichting op de uitdagingen.

1. Gezondheidsdata zijn versnipperd en moeilijk vindbaar

Nederland beschikt over een sterke uitgangspositie met hoogwaardige elektronische patiënten-/cliëntendossiers, registraties, cohortstudies en biobanken. Gezondheidsdata zijn echter verspreid over duizenden zorgorganisaties, onderzoeksinstituten, registraties en andere datahouders incl. burgers en patiënten zelf. Iedere datahouder hanteert eigen systemen, procedures, organisatie van verantwoordelijkheden en interpretaties/toepassing van wet- en regelgeving. Hierdoor bestaat geen uniforme manier om datasets te vinden, aan te vragen, te combineren en veilig te analyseren. Onderzoekers moeten veelal per datahouder afzonderlijke afspraken maken over toegang, datalevering, privacy, beveiliging en gegevensuitwisseling, wat leidt tot lange doorlooptijden en projectmatige werkwijzen. Hoewel verschillende organisaties beschikken over professionele voorzieningen voor privacybescherming, metadata registratie en beveiligde analyseomgevingen, zijn deze voorzieningen vaak lokaal georganiseerd en beperkt op elkaar aangesloten. Met name kleinere zorgorganisaties beschikken niet altijd over vergelijkbare mogelijkheden. De grootste uitdaging ligt daardoor niet in het ontbreken van gezondheidsdata, maar in het efficiënt vinden, toegankelijk maken, combineren en hergebruiken ervan over de grenzen van organisaties heen.

2. Beperkingen bij het koppelen van gezondheidsdata

Het combineren van gezondheidsdata uit verschillende bronnen is een belangrijke voorwaarde voor veel vormen van secundair gebruik. In de praktijk is het echter niet mogelijk om gegevens uit verschillende registraties en systemen eenduidig aan dezelfde persoon te relateren. Hierdoor kunnen beperkingen ontstaan bij het samenstellen van volledige en betrouwbare datasets voor onderzoek, innovatie en beleidsontwikkeling.

3. Administratieve lasten beperken duurzame databeschikbaarheid

Het registreren, beheren en beschikbaar stellen van gezondheidsdata vraagt tijd, capaciteit en middelen van organisaties. Wanneer deze lasten als hoog worden ervaren, kan dit ertoe leiden dat gegevens niet volledig worden vastgelegd, niet langdurig beschikbaar blijven of minder geschikt zijn voor secundair gebruik. Hierdoor komt de duurzame beschikbaarheid van gezondheidsdata onder druk te staan.

4. Gezondheids(meta)data zijn niet uniform gestructureerd volgens standaarden

Gezondheidsdata en metadata worden binnen organisaties vastgelegd volgens verschillende informatiemodellen, coderingen, definities en standaarden. Hoewel standaarden zoals SNOMED CT, LOINC, HL7 FHIR, openEHR en DICOM steeds vaker worden toegepast, is de implementatie niet uniform. Hierdoor verschillen betekenis, structuur en kwaliteit van gegevens tussen systemen en organisaties. Het combineren, vergelijken en hergebruiken van datasets vraagt daardoor veel tijd en inspanning voor interpretatie, harmonisatie en kwaliteitscontrole voordat gegevens daadwerkelijk kunnen worden ingezet voor onderzoek, innovatie en beleidsontwikkeling.

5. Ontbreken van context

Gezondheidsdata zonder context hebben vaak onvoldoende betekenis voor secundair gebruik. Een dataset kan technisch compleet lijken, terwijl belangrijke informatie over de omstandigheden waaronder een gegeven is ontstaan ontbreekt. Daardoor kan een onderzoeker het gegeven niet goed interpreteren of verkeerd gebruiken. Een laboratoriumwaarde van bijvoorbeeld 5,8 mmol/L zegt op zichzelf iets, maar voor hergebruik kan het van belang zijn om te weten of de patiënt nuchter was, wanneer het monster is afgenomen en met welke meetmethode de waarde is bepaald. Hoewel de context veelal essentieel is, ontbreekt in veel systemen de mogelijkheid om deze gestructureerd vast te leggen. Gevolg is dat de context ook niet digitaal beschikbaar worden gesteld voor hergebruik. Aanpassing van systemen om context (automatische) vast te leggen, is nodig om dit op te lossen, maar hiertoe wordt vooralsnog nauwelijks actie ondernomen.

6. Beperkte uniforme inrichting

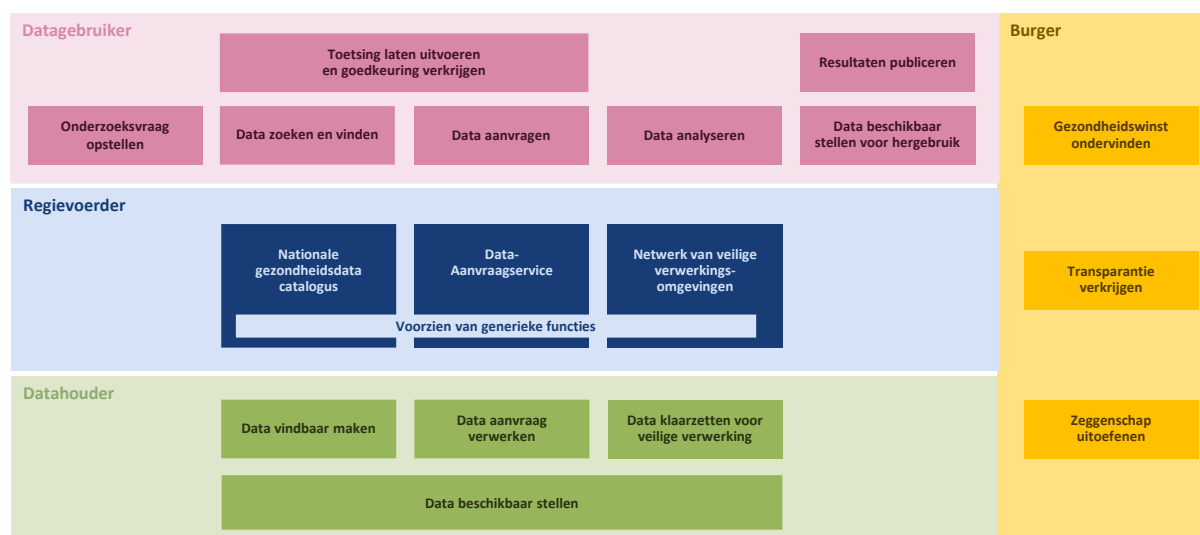
Processen, verantwoordelijkheden, voorzieningen en dienstverlening voor secundair gebruik zijn niet overal op een uniforme en samenhangende wijze ingericht. Hierdoor zijn veel werkzaamheden nog handmatig, organisatie specifiek en projectmatig van aard. Dit beperkt de efficiëntie, herhaalbaarheid en schaalbaarheid waarmee gezondheidsdata beschikbaar kunnen worden gesteld voor onderzoek, innovatie en beleidsontwikkeling.

4. Doelarchitectuur

Dit hoofdstuk bevat de kern van dit document, de doelarchitectuur voor de Nationale gezondheidsdata-infrastructuur secundair gebruik zelf. Eerst wordt kort stilgestaan bij de hoofdelementen van en de hoofdactoren bij de gezondheidsdata-infrastructuur secundair gebruik. Daarna volgen achtereenvolgens de proces-, de data- en als laatste de applicatie-architectuur. Tot slot wordt geschetst welke partijen verantwoordelijk zijn voor het realiseren en beheren van welke applicaties / functies.

4.1. Hoofdlijn nationale gezondheidsdata-infrastructuur secundair gebruik

Dit overzicht laat op hoofdlijnen zien hoe de nationale gezondheidsdata-infrastructuur secundair gebruik partijen ondersteunt bij het veilig en verantwoord hergebruiken van gezondheidsdata. In onderstaande figuur is dit beeld weergegeven:



Figuur 3. Hoofdactoren en belangrijkste functies

In de figuur worden vier hoofdactoren onderscheiden: de datagebruiker, de regievoerder, de datahouder en de burger. Bij iedere actor zijn de belangrijkste functies weergegeven die bijdragen aan dit proces. De regievoerders vervullen hierin een verbindende rol door zelfstandig en in samenhang functies aan te bieden voor vindbaarheid, toegang en veilige verwerking, ondersteund door generieke functies voor interoperabiliteit, vertrouwen en samenwerking. Datahouders zijn verantwoordelijk voor het beschikbaar stellen van data, terwijl burgers zowel inzicht, transparantie en zeggenschap behouden over het gebruik van hun gegevens, als profiteren van de maatschappelijke waarde die met hergebruik van gezondheidsdata wordt gerealiseerd.

In de figuur is van links naar rechts het hoofdproces voor secundair data gebruik ten behoeve van onderzoek, beleid en innovatie weergegeven:

- het vinden van relevante data,
- het aanvragen van toegang tot data,
- (het door de datahouder) beschikbaar stellen van data voor analyse binnen een beveiligde verwerkingsomgeving,
- het uitvoeren van de analyse
- het herbruikbaar maken van resultaten en deze beschikbaar stellen voor hergebruik.

De gehanteerde kleuren en indeling vormen de vaste beeldtaal van deze doelarchitectuur en worden in de verdere uitwerking van dit document opnieuw gebruikt om de samenhang tussen actoren, functies en architectuuronderdelen herkenbaar en consistent weer te geven.

In lijn met de opdracht van het Nationaal Groeifonds aan Health-RI worden momenteel op landelijk niveau de volgende hoofdelementen uit de figuur gerealiseerd:

- In de **ationale gezondheidsdata catalogus** kunnen datagebruikers bestaande gezondheidsdatasets zoeken en vinden. Datahouders kunnen beschrijvingen van gezondheidsdatasets (ook wel: “metadata”) publiceren in de catalogus aan de hand waarvan de datagebruiker de catalogus kan doorzoeken;
- De **data-aanvraagservice** (in dit document uitgebreid tot Nationale gezondheidsdata toegangsservice omdat in de toekomst onder meer ook workflow management en contractbeheer hieronder vallen) heeft als doel het faciliteren van aanvragen voor het hergebruik van data uit gezondheidsdatasets die in de catalogus worden aangeboden. Via een aanvraagformulier kan de datagebruiker de gewenste data-aanvraag indienen. De aanvraag wordt geregistreerd. De data-aanvrager wordt daarbij ondersteund door de centrale helpdesk. Daarna wordt de aanvraag doorgezet naar de datahouder, die zorg draagt voor het verdere proces van de verwerking van de data-aanvraag tot en met het uiteindelijk klaarzetten van de data voor veilige verwerking in een geselecteerde beveiligde verwerkingsomgeving³;
- Health-RI biedt voor de verwerking van data een BVO aan. In een BVO kan data verwerkt worden en zijn passende maatregelen op het gebied van systeembeveiliging en gebruikerstoegang geïmplementeerd om de data en de privacy van de data subjects te beschermen. Er zijn verschillende vormen van veilige verwerking die Health-RI met een BVO ondersteunt, waaronder het door de datahouder laten klaarzetten van data in de BVO voor analyses door de datagebruikers en het door de datahouders beschikbaar stellen data aan de BVO voor het uitvoeren van gefedereerde analyses door de datagebruikers;
- Om een generieke infrastructuur voor zowel primair als secundair gebruik te realiseren, is het belangrijk om **generieke functies** in te richten. Deze functies dienen zowel in de zorg (primair) als daarbuiten (secundair) toepasbaar te zijn. De genoemde generieke functies dragen bij aan het verminderen van versnippering en het bevorderen van interoperabiliteit en herbruikbaarheid van data. Het ministerie van VWS heeft een regierol bij het ontwikkelen van deze generieke functies. De Nationale gezondheidsdata-infrastructuur secundair gebruik sluit aan bij de keuzes die in dat proces worden gemaakt en zal meedenken over vereisten voor generieke functies voor secundair gebruik.

Deze hoofdelementen worden de komende periode maximaal compliant gemaakt aan de eisen van de EHDS. De BVO's en de toegangsfuncties tot de data van de datahouder zullen ook op lokaal / regionaal / sectoraal niveau blijven bestaan en verder worden ontwikkeld. Hierbij is een goede aansluiting op de landelijke voorzieningen essentieel.

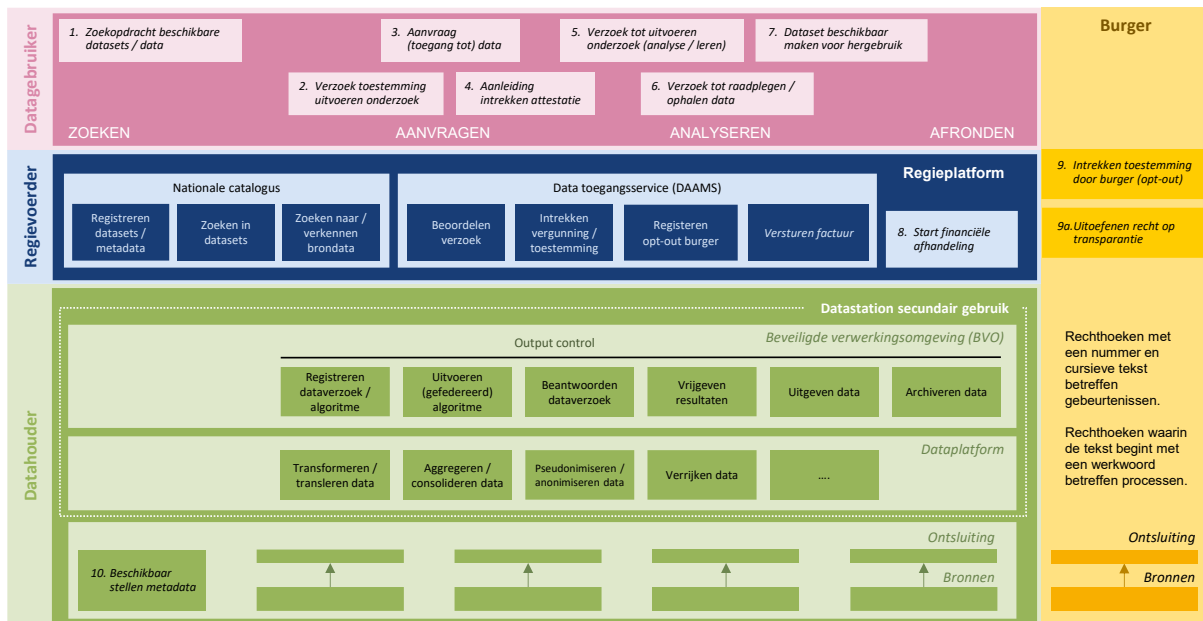
4.2. Business- en procesarchitectuur

In deze paragraaf wordt een nadere verdieping gegeven van de procesarchitectuur binnen de context van hergebruik van gezondheidsdata voor onderzoek, beleid en innovatie. De leidraad hierbij wordt gevormd door de hoofdstappen in het kernproces van de datagebruiker (onderzoeker, beleidsmaker, innovator). De procesarchitectuur beschrijft op hoofdlijnen hoe een datagebruiker data lokaliseert (1 Zoeken), autorisatie verkrijgt (2 Aanvragen), data beschikbaar krijgt en analyseert (3 Uitvoeren) en het onderzoek afrondt (4 Afronden). Aanvullend op deze hoofdstappen wordt in dit hoofdstuk ook ingegaan op de rol van de burger (5) en processen bij het overig secundair gebruik van gezondheidsdata (6).

Binnen de procesarchitectuur worden 4 uitvoerende actoren onderscheiden: de datagebruiker, de datahouder, de regievoerder en de burger. In het volgende figuur worden de gebeurtenissen met daaropvolgende processen per actor weergegeven:

³De term *beveiligde verwerkingsomgeving* komt ook voor de Engelstalige variant *SPE, Secure Proces Environment*.

625



626

627

628

Figuur 4. Gebeurtenissen en processen secundair gebruik

629

In de volgende paragrafen worden de gebeurtenissen en processen uit figuur 4 nader toegelicht. Om onderscheid te maken tussen wat altijd gebeurt en wat alleen in specifieke situaties voorkomt, zijn de paragrafen opgebouwd uit vier onderdelen:

632

- **Gebeurtenissen** beschrijven de aanleidingen of triggers die een processtap activeren.
- **Proces** beschrijft het standaardproces en de activiteiten die in beginsel altijd plaatsvinden.
- **Mogelijke scenario's** beschrijven varianten, uitzonderingen of aanvullende stappen die afhankelijk van de context kunnen voorkomen, maar geen onderdeel vormen van het standaardproces.
- **Verschil met primair gebruik** licht toe hoe het proces voor secundair datagebruik zich verhoudt tot de processen die binnen het zorgdomein voor primair gebruik gebruikelijk zijn.

639

640

Lokale autonomie blijft daarbij het uitgangspunt: datahouders en -gebruikers bepalen zelf hoe deze stappen worden ingericht, passend bij hun eigen processen, organisatie van verantwoordelijkheden en technische omgeving, zolang wordt voldaan aan de gestelde gemeenschappelijk kaders voor interoperabele en uitwisselbare implementaties.

644

645

4.2.1. Zoeken

646

Gebeurtenissen

647

0. Start voorbereiding onderzoek; formuleren onderzoek, formuleren datavraag, uitvoeren van een haalbaarheidsanalyse om beschikbaarheid en bruikbaarheid van data te bepalen, ondersteund door catalogus en verkenner.

649

1. Zoekopdracht beschikbare data; bepalen welke datasets en bronnen beschikbaar zijn voor de specifieke toepassing.

651

2. Verzoek toestemming uitvoeren onderzoek; binnen eigen kennisinstelling / -omgeving laten toetsen van voorgenomen datagebruik op onder meer juridische, ethische en organisatorische aspecten.

654

NB: Toetsing of toegang tot data kan worden verleend, vindt plaats bij proces voor afhandeling van de aanvraag.

657

Proces (weergegeven in figuur 5)

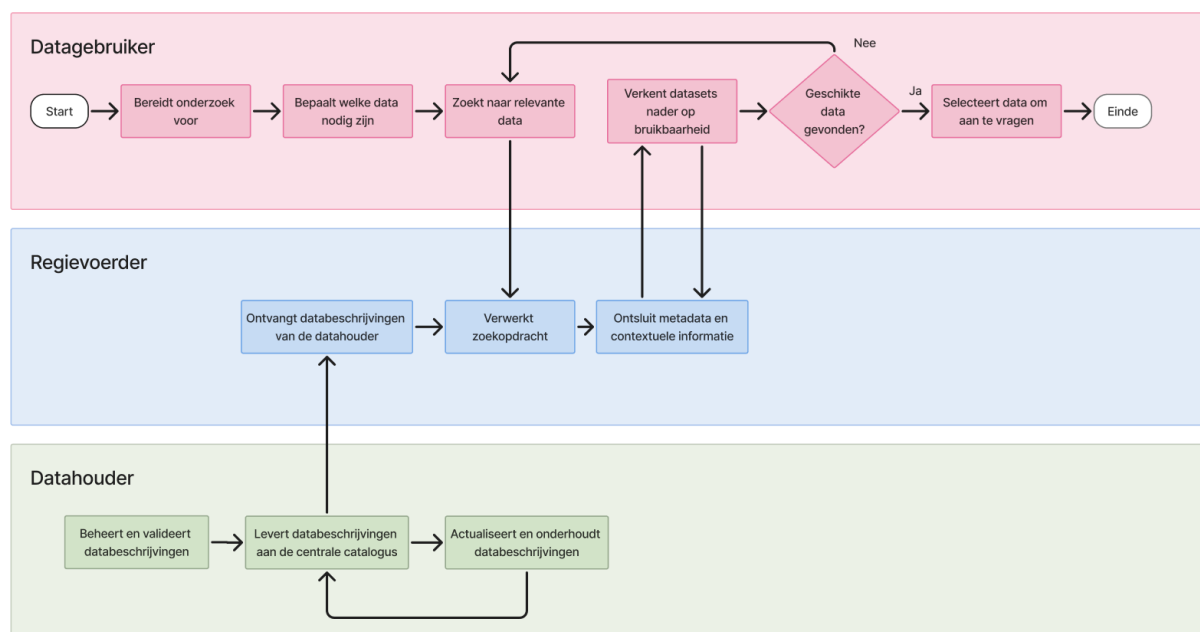
658

- **Datagebruiker;** Bereidt het onderzoek voor en zorgt dat dit organisatorisch is goedgekeurd; bepaalt welke data nodig is en zoekt in de nationale gezondheidsdata catalogus naar relevante

659

datasets. Ook kan gebruik worden gemaakt van de verkenner om relevante data te vinden. Het formuleren en het zoeken naar data betreft veelal een iteratief proces tussen datagebruiker en datahouder.

- **Beheerder Nationale gezondheidsdata catalogus (Regievoerder);** De catalogus ontsluit metadata over nationaal beschikbare data en biedt contextuele informatie om de bruikbaarheid te beoordelen.
- **Datahouder;** Beheert en valideert databeschrijvingen en levert deze in de vorm van metadata aan de centrale catalogus, zowel via handmatige of geautomatiseerde invoer, als via geautomatiseerde data omschrijving service (FAIR Data Point).
NB: In het Quantus-project wordt gesteld dat de catalogusbeheerder de datakwaliteits- en toepasbaarheidslabels beoordeelt en monitort ook of datahouders tijdig hun metadata updaten. Dit is echter nog een punt van discussie.
- **Datagebruiker;** Zoekt relevante data
- **Datagebruiker;** Verkent data en datasets nader op bruikbaarheid en selecteert geschikte datasets voor het onderzoek. Tijdens het verkennen verkrijgt de datagebruiker geen toegang tot de data. De data zijn dan ook niet in te zien.
- **Datagebruiker;** Kan in de verkenningsfase iteratief zoeken en verkennen, waarbij de stappen van lokaliseren en beoordelen van data meerdere malen kunnen worden doorlopen in het kader van een haalbaarheids- of exploratieve analyse in nauwe samenspraak met de datahouder(s).
- **Datagebruiker;** Selecteert data om aan te vragen



Figuur 5. Zoeken

Mogelijk scenario's

- **Datagebruiker;** Zorgt dat het voorgenomen data (her)gebruik organisatorisch is goedgekeurd, waarbij – afhankelijk van de aard van het onderzoek – toetsing kan plaatsvinden door bijvoorbeeld een METC, wetenschapscommissie of biobank-/datatoegangscommissie, gericht op de proportionaliteit en toelaatbaarheid van het voorgenomen datagebruik
- **Datagebruiker;** Beschikt reeds vooraf over kennis van potentieel relevante databronnen en benut deze stap primair om deze bekende datasets verder te verkennen en te selecteren voor het onderzoek.
- **Datagebruiker en datahouder;** Veelal zijn er gesprekken nodig tussen de datagebruiker en de datahouder om een dieper begrip van de vraag en de (context) van de data is nodig voor het uit

te voeren onderzoek.

NB: De noodzaak om met elkaar in gesprek te gaan, geldt voor alle fasen van dit proces.

- **Datahouder**; wordt verplicht om de beschrijving van de data waarover hij de beschikking heeft, op te nemen in de catalogus.

Verschil met primair gebruik

- In het primair gebruik wordt data doorgaans per individuele cliënt / patiënt uit het bronsysteem geraadpleegd en geprepareerd volgens afgesproken standaarden. Bij secundair datagebruik ligt de focus op datasets met data van meerdere personen, die voldoen aan bepaalde selectiecriteria.

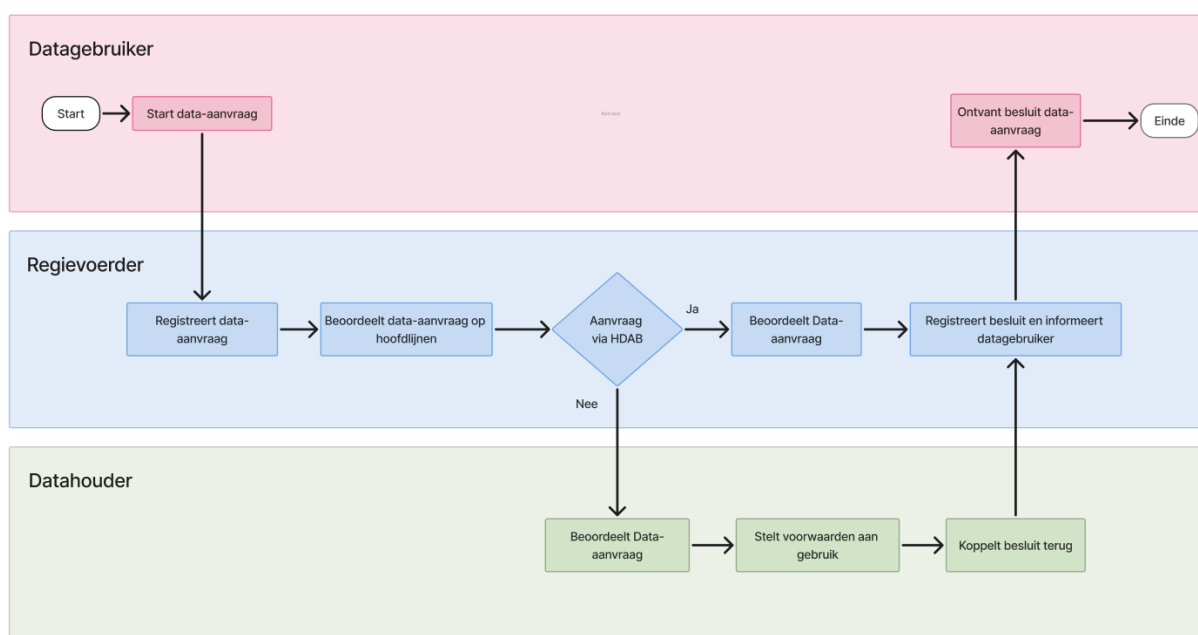
4.2.2. Aanvragen

Gebeurtenissen

3. Aanvraag toegang tot data; verkrijgen van geautoriseerde toegang tot datasets en BVO's via identificatie, authenticatie en autorisatie op basis van goedgekeurde aanvraag.
4. Aanleiding intrekken goedkeuring/ vergunning; herroepen van goedkeuring/ vergunning wanneer niet langer wordt voldaan aan gestelde voorwaarden, bijvoorbeeld door non-compliance, incidenten, wijzigingen in context of op verzoek van een van de betrokken partijen.

Proces (weergegeven in figuur 6)

- **Datagebruiker**; Start een aanvraag in de **ationale gezondheidsdata toegangsservice** met aanvraagformulier waarin doel, dataset(s) en beveiligde verwerkingsomgeving beschreven wordt.
- **Beheerder nationale gezondheidsdata toegangsservice** (Regievoerder); Registreert en beoordeelt de aanvraag op hoofdlijnen
 - Voordat **HDAB** operationeel is : **Datahouder** beoordeelt in detail de aanvraag en stelt voorwaarden voor gebruik, waarbij wordt getoetst of het gevraagde datagebruik in verhouding staat tot de onderzoeksvraag, de vraagstelling redelijk is en in overeenstemming met geldende wet- en regelgeving kan worden uitgevoerd. Onderdeel van de voorwaarden voor gebruik zijn veelal ook de onderlinge financiële afspraken.
NB: Binnen de Nationale gezondheidsdata-infrastructuur secundair gebruik wordt gestreefd naar standaardisatie van de te stellen voorwaarden.
 - Als **HDAB** operationeel is en aanvraag via de **HDAB** verloopt, dan vindt de beoordeling door de **HDAB** plaats en mag een **Datahouder** alleen bij vermeende overtreding van intellectueel eigendomsrechten bezwaar aantekenen tegen het voor onderzoek beschikbaar stellen van de data.
- **Beheerder nationale gezondheidsdata toegangsservice** (Regievoerder); Legt het besluit vast en informeert de **Datagebruiker**.



Figuur 6. Aanvragen

Mogelijke scenario's

- **Nationale gezondheidsdata toegangsservice** (Regievoerder) kan gedurende het aanvraagproces iteraties van vraag en antwoord tussen **Datagebruiker** en **Datahouder** faciliteren, met evt. wijzigingen in de aanvraag administratie.
- **Datagebruiker** kan voor een onderzoek data van meerdere **Datahouders** nodig hebben
- Ten aanzien van de duiding van de data die aangevraagd wordt, kan deze data zich in een te identificeren dataset bevinden, of een extract zijn van een grotere databron, zoals een EPD.
- **Datagebruiker** kan verderop in het proces, als analyse loopt, wijziging willen doorvoeren, van bijvoorbeeld onderzoeksvraag of betrokken onderzoekers. Hierbij dient beoordeeld te worden of er een nieuw aanvraagproces opgestart dient te worden of dat de bestaande aanvraag kan worden aangepast.
- In de situatie onder de EHDS kan inafwijking van de standaardaanvraag voor toegang tot data kan een datagebruiker een aanvraag indienen voor een **statistisch antwoord** op een wetenschappelijke vraag (Data Request). Deze aanvraag kan grotendeels hetzelfde pad volgen, waarbij de nadruk ligt op toetsing van het resultaat om te waarborgen dat geen tot personen herleidbare gegevens worden verstrekt. Hierdoor kan een vereenvoudigd of minimaal goedkeuringsproces volstaan.
- Intrekken goedkeuring / vergunning;
 - Voordat HDAB operationeel is; **Datahouder** kan, op basis van nieuwe inzichten, gewijzigde omstandigheden of niet-naleving van voorwaarden, een eerder verleende goedkeuring heroverwegen en indien nodig intrekken.
 - Als **HDAB** operationeel is en aanvraag via **HDAB** verloopt; **HDAB** kan een afgegeven datavergeving wijzigen intrekken overeenkomstig het EHDS-kader.

Verschil met primair gebruik

- Bij primair gebruik beschikt de zorgverlener, vanuit de eigen rol binnen een gezondheidszorgorganisatie en de relatie tot de patiënt/burger, over een directe grondslag voor het gebruik van data. Voor secundair gebruik van data hebben de gebruikers in eerste instantie geen toegang tot de data, omdat zij geen rol hebben in de behandeling van de betreffende patiënten waar de data van afkomstig zijn. Daarom wordt vooraf beoordeeld of de aanvraag aan de voorwaarden voor ontsluiting voldoet (doelbinding). Denk hierbij aan de een wettelijke grondslag voor gegevensverwerking, privacy waarborgen en een beoordeling van ethische aspecten.

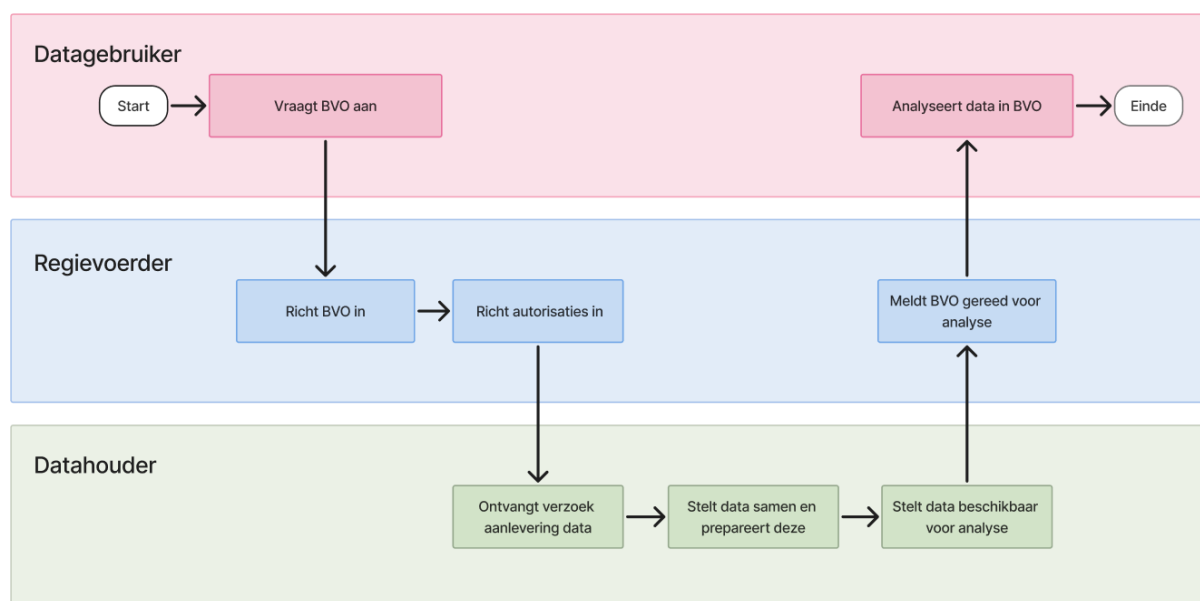
4.2.3. Analyseren

Gebeurtenissen

5. Verzoek tot uitvoeren onderzoek; Verkrijgen toegang tot BVO's en uitvoeren van onderzoek / analyse, waarbij data na voorbereiding gecontroleerd beschikbaar wordt gesteld aan geautoriseerde gebruiker.
6. Verzoek tot raadplegen/ ophalen van de data; ontvangen van data binnen de BVO, uitvoeren van analyses en vrijgave van resultaten na controle door de datahouder.

Proces (weergegeven in figuur 7)

- **Datagebruiker**; Initieert proces voor beschikbaar stellen van data in een BVO.
- **Datahouder**; Ontvangt vraag om data aan te leveren voor betreffende toepassing, controleert autorisaties a.d.h.v. **nationale gezondheidsdata toegangsservice**, stelt data samen en prepareert deze (opschonen, transformeren, pseudonimiseren, minimaliseren).
- **Beheerder BVO's** (Regievoerder); Richt BVO's in met benodigde analyse gereedschappen.
- **Datahouder**; Draagt data over naar / stelt deze beschikbaar aan een BVO.
- **Datagebruiker**; De betrokken onderzoekers worden geïdentificeerd, geauthenticeerd en geautoriseerd op basis van de data uit de centrale data-aanvraagvoorziening (Regievoerder). Vervolgens verkrijgen de betrokken onderzoekers op die basis gecontroleerde toegang tot de BVO's, waar zij de benodigde analyses uitvoeren.



Figuur 7. Analyseren

Mogelijke scenario's

- Sommige datagebruikers willen niet dat de data geïnterpreteerd worden door deze op te schonen en te transformeren naar gewenste codering/ modellering. Zij werken liever of noodzakelijkerwijs met originele data.
- Er kunnen uitzonderlijke toepassingen zijn die geen de-identificatie verlangen
- De BVO kan zijn
 - Een BVO bij de datahouder, centraal (bijvoorbeeld een nationale of regionale BVO binnen de Nationale gezondheidsdata-infrastructuur secundair gebruik), of bij de datagebruiker.
 - Een gefedereerde data-analyse omgeving bij de datahouder, centraal (bijvoorbeeld de centrale BVO van Health-RI), of bij de datagebruiker.
- In bepaalde onderzoeken dienen datasets van meerdere datahouders te worden samengebracht, waarbij data wordt gestapeld of op persoonsniveau wordt gekoppeld om integrale analyses mogelijk te maken. Een voorbeeld hiervan is onderzoek naar medicijngebruik,

waarbij datasets vanuit het ziekenhuis EPD, het apotheek EPD en het PGO van de patiënt worden samengebracht.

Verskil met primair gebruik

- In het primair gebruik wordt data doorgaans per individuele patiënt uit het bronsysteem geraadpleegd en geprepareerd volgens afgesproken standaarden. Bij secundair datagebruik ligt de focus op datasets met data van meerdere personen, die voldoen aan bepaalde selectiecriteria, waarbij de individuele identiteit doorgaans minder relevant is. Afhankelijk van de toepassing kan de **Datagebruiker** behoefte hebben aan data in bewerkte, gestandaardiseerde vorm, of juist in een meer ruwe, niet-geïnterpreteerde vorm. Daarnaast kan ervoor gekozen worden om niet de dataset zelf te verplaatsen, maar analyses bij de bron uit te voeren en uitsluitend geaggregeerde resultaten beschikbaar te stellen.
- Secundair datagebruik vraagt om een bredere en verder geharmoniseerde dataset dan primair gebruik, inclusief aanvullende datapunten die voor onderzoek, beleid en innovatie relevant zijn.

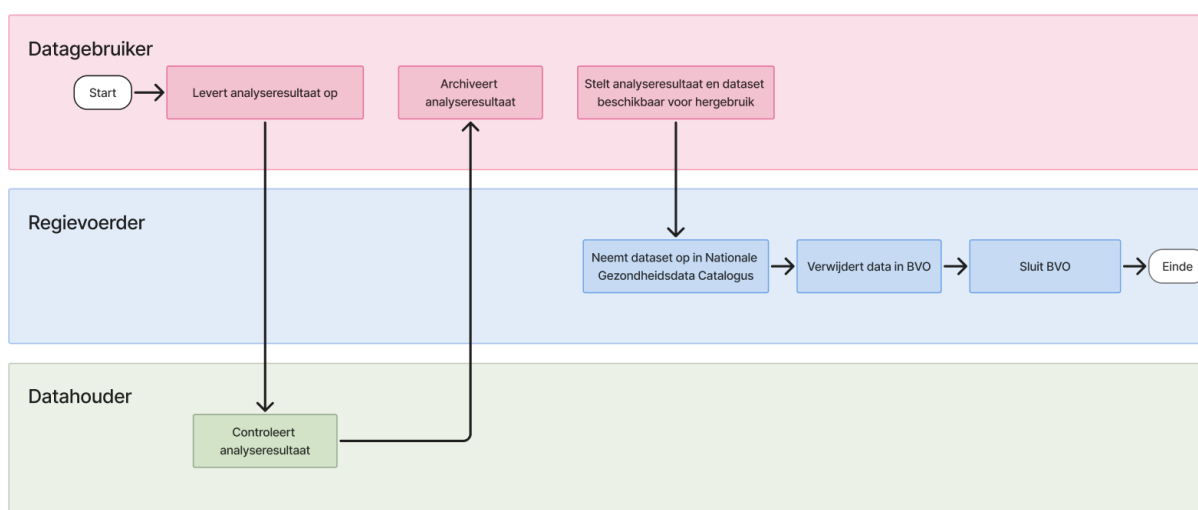
4.2.4. Afronden

Gebeurtenissen

7. Dataset beschikbaar maken voor hergebruik; beschikbaar stellen van dataset voor meervoudig gebruik binnen het ecosysteem, inclusief toetsing, inrichting van levering en publicatie van metadata en persistente identificatie.
8. Start financiële afhandeling: Afhandeling van financiële afspraken, zoals deze voorafgaand aan de uitvoering van het onderzoek zijn overeengekomen en vastgelegd. Er kan eventueel sprake zijn van nacalculatie op basis van daadwerkelijk gebruik (o.a. CPU-capaciteit) en/of op basis van verdeling tussen betrokken partijen van kosten en opbrengsten na afloop.

Proces (weergegeven in figuur 8)

- **Datagebruiker**; Laat het analyseresultaat controleren.
 - Voordat **HDAB** operationeel is dan wordt de output control uitgevoerd door de **datahouder(s)**.
 - Als **HDAB** operationeel is en aanvraag via **HDAB** gelopen is dan wordt de output control uitgevoerd door de **HDAB**.
- **Datagebruiker** archiveert het resultaat. Eventueel kan het analyseresultaat na toetsing, beschikbaar worden gesteld voor hergebruik, bijvoorbeeld door opname in een register van afgeronde of opgeslagen data met verwijzing naar het bijbehorende onderzoek.
- **Beheerder Nationale gezondheidsdata catalogus** (Regievoerder): Neemt dataset op in de catalogus.
- **Beheerder BVO's** (Regievoerder); Verwijdert data en sluit de BVO's na afronding.
- **Datagebruiker**; Levert verrijkingen terug aan de datahouder.
- **Regievoerder**; Binnen de verschillende stappen van het kernproces kunnen kosten in rekening worden gebracht door betrokken partijen. Dit kan onder andere betrekking hebben op het afhandelen van het aanvraagproces (bijvoorbeeld door een centrale data-aanvraagvoorziening), de werkzaamheden van de datahouder voor het selecteren en prepareren van data, en het gebruik van een BVO.



Figuur 8. Afronden

Mogelijke scenario's

- Datagebruikers creëren in het kader van onderzoek ook vaak data, bijvoorbeeld door het verzamelen van antwoorden op uitgezette vragenlijsten of door het verzamelen van data door een datagebruiker tijdens het zorgproces. Door het aanmelden van de dataset bij de catalogus kan worden geborgd dat ook deze data beschikbaar is voor toekomstig hergebruik.
- Uiteraard is het in rekening brengen van kosten geen verplichting. De exacte inrichting van de financiële afhandeling, waaronder de wijze van calculatie (bijvoorbeeld op basis van gebruik van infrastructuur of capaciteit), facturering en verdeling van inkomsten tussen betrokken partijen, dient de komende periode nog verder uitgewerkt te worden.
- Standaardisatie van voorwaarden voor secundair gebruik van gezondheidsdata, waaronder financiële afspraken, is wenselijk. Schept duidelijkheid voor de datagebruiker en creëert de mogelijkheid om procesondersteunende applicaties te convergeren.

4.2.5. Burgerproces

Gebeurtenissen

9. Intrekken toestemming door burger; wijzigen of intrekken/herstellen van toestemming voor gebruik van data conform geldende rechten en afspraken.

9a. Uitoefenen recht op transparantie.

Proces

- Zeggenschap; De burger behoudt regie over het gebruik van data over de burger via een generieke voorziening voor het vastleggen en beheren van zeggenschap. Via deze voorziening kan de burger aangeven in hoeverre data gebruikt mag worden voor secundaire doeleinden. Er is hierbij sprake van opt-out, dus standaard wordt uitgegaan van toestemming voor secundair datagebruik.
- Transparantie; Daarnaast wordt transparantie richting de burger gefaciliteerd via een transparantieportaal, waarin inzicht wordt geboden in welke onderzoeken worden uitgevoerd en in welke onderzoeken data over de betreffende burger is gebruikt, inclusief doel en betrokken partijen.
NB: Of op deze wijze invulling gegeven dient te worden aan transparantie dient nog nader onderzocht te worden.
- Nevenbevindingen; De EHDS stelt dat bevindingen in de onderzoekscontext die mogelijk relevant zijn voor de gezondheid van de burger (significante bevindingen), moeten worden teruggekoppeld aan de burger via de betrokken zorgverlener, conform geldende afspraken en verantwoordelijkheden binnen de zorgverlening. De datagebruiker communiceert eventuele

nevenbevindingen aan de HDAB. De HDAB koppelt dit vervolgens terug aan de datahouder, die op zijn beurt de burger informeert.

Mogelijke scenario's

- Veelal kan voorafgaand aan de uitvoering van het onderzoek worden bedacht of er bevindingen op individueel niveau uit het onderzoek naar voren kunnen komen. De invulling van de functie nevenbevindingen dient echter nader uitgewerkt te worden en afgestemd met datagebruikers, datahouders en burgers / patiënten.

4.2.6. Secundair gebruik van gezondheidsdata, anders dan bij onderzoek

Binnen het Nationale gezondheidsdata-infrastructuur secundair gebruik wordt beoogd om versnippering zoveel mogelijk te voorkomen. De architectuur en bijbehorende processen zijn daarom primair uitgewerkt vanuit het onderzoeksproces, met als doel deze zo generiek mogelijk toe te passen op andere vormen van secundair gebruik van gezondheidsdata, zoals kwaliteitsregistraties en initiatieven als KIK-V.

Veel van de onderliggende processtappen, zoals het lokaliseren van data, het verkrijgen van autorisatie en het beschikbaar stellen van data voor hergebruik, zijn ook in deze context relevant. Zo zal bij de inrichting van een kwaliteitsregistratie eveneens bepaald moeten worden waar de benodigde data zich bevindt en op welke wijze en onder welke voorwaarden deze beschikbaar gesteld mag worden, vaak in een meer structureel en herhaald gebruikspatroon. De verwerkingsomgeving van dergelijke toepassingen kan daarbij worden aangemerkt als een vorm van een BVO.

In samenwerking met deze partijen wordt al verkend in hoeverre de beschreven doelarchitectuur aansluit op deze vormen van secundair datagebruik. In de komende periode wordt deze analyse verder verdiept en geborgd, zodat de toepasbaarheid van de doelarchitectuur waar nodig kan worden versterkt.

4.3. Data-architectuur

Met de doelarchitectuur secundair datagebruik wordt ingezet op de maximale beschikbaarheid van data voor secundair datagebruik. Dit betekent dat data die voor andere doeleinden, zoals dossiervoering of een specifiek onderzoek, is verzameld ook weer beschikbaar komt voor andere toepassingen waaronder onderzoek, beleid en innovatie. De scope van data is breed en omvat zowel bestaande datasets als onderliggende Real World Data (RWD) en andere databronnen waaruit datasets voor specifiek gebruik kunnen worden samengesteld. Zo kan ofwel uit een eerder samengestelde dataset een subselectie worden gemaakt voor een specifiek onderzoek, maar kan ook uit beschikbare data bij een datahouder een dataset worden samengesteld op basis van de beschikbare data-elementen.

Ondersteuning van verschillende vormen van datagebruik

De Nationale gezondheidsdata-infrastructuur secundair gebruik ondersteunt een breed spectrum aan datagebruikers en gebruiksscenario's. Sommige gebruikers hebben behoefte aan toegang tot originele, niet-geïnterpreteerde data om de context, herkomst en betekenis van data zelf te kunnen beoordelen. Andere gebruikers geven de voorkeur aan gestandaardiseerde en geharmoniseerde data, zodat analyses sneller kunnen worden uitgevoerd en data eenvoudiger gecombineerd kan worden over bronnen heen. De doelarchitectuur ondersteunt beide vormen van datagebruik en faciliteert daarmee een groeimodel waarin herbruikbare datamodellen, mappings en harmonisaties stapsgewijs worden opgebouwd. Volledige standaardisatie van alle gezondheidsdata is daarbij niet het uitgangspunt; diversiteit in databronnen en datavormen blijft bestaan en wordt expliciet ondersteund. Daarnaast ondersteunt de Nationale gezondheidsdata-infrastructuur secundair gebruik zowel data access, als data visiting. Hiermee wordt aangesloten op de uiteenlopende

behoeften van onderzoekers, beleidsmakers en innovatoren en blijft de Nationale gezondheidsdata-infrastructuur secundair gebruik breed toepasbaar voor secundair gebruik van gezondheidsdata.

Metadata⁴

Metadata vormt de basis voor het vinden, begrijpen, beoordelen en hergebruiken van data binnen de Nationale gezondheidsdata-infrastructuur secundair gebruik. Door data en databronnen gestandaardiseerd te beschrijven wordt data op een uniforme wijze vindbaar, vergelijkbaar en uitwisselbaar gemaakt. Metadata stelt datagebruikers in staat om inzicht te krijgen in de inhoud, herkomst, kwaliteit, omvang, toegankelijkheid en gebruiksvoorwaarden van beschikbare data en ondersteunt daarmee de selectie van geschikte data. Goede metadata verkleint bovendien de afhankelijkheid van handmatige afstemming tussen datahouder en datagebruiker en draagt bij aan een efficiënter en transparanter proces van dataontdekking en datagebruik. In een federatief landschap waarin data verspreid aanwezig blijft bij de bron, vormt gestandaardiseerde metadata de verbindende laag waarmee databronnen uit verschillende organisaties, domeinen en infrastructuren op een samenhangende wijze kunnen worden ontsloten. Daarmee is metadata een essentiële randvoorwaarde voor interoperabiliteit, FAIR datagebruik en de realisatie van de nationale catalogusfunctie zoals voorzien binnen de EHDS en de Nationale Gezondheidsdata-infrastructuur.

Voor de Nationale gezondheidsdata catalogus is gekozen voor een nationaal metadataschema – het Health-RI metadatamodel - als gemeenschappelijk model voor het beschrijven, publiceren en uitwisselen van metadata. Dit model is gebaseerd op de Europese standaarden DCAT-AP en HealthDCAT-AP, aangevuld met wat de Nederlandse praktijk daarnaast nodig heeft. . Daarmee sluiten we aan op breed geaccepteerde nationale en Europese ontwikkelingen, ondersteunen we de uitwisseling van metadata tussen catalogi, FAIR Data Points en andere metadata-platformen, en ontstaat een schaalbare basis voor interoperabiliteit. Het metadatamodel ondersteunt niet alleen de beschrijving van bestaande onderzoeksdatasets, maar vormt ook de basis voor het ontsluiten van dynamische databronnen, zoals EPD-data, registraties, beelddata, Omics-data en andere gezondheidsgegevens die onder de EHDS voor secundair gebruik beschikbaar moeten komen. Om het model blijvend te laten aansluiten op nieuwe behoeften, wetgeving en ontwikkelingen in het metadata-ecosysteem is een release managementcyclus ingericht voor het gecontroleerd beoordelen, prioriteren en implementeren van wijzigingen aan het metadatamodel.

Datastandaarden en interoperabiliteit

Binnen de Nationale gezondheidsdata-infrastructuur secundair gebruik wordt aangesloten op het landschap van bestaande data- en informatiestandaarden. Onder de noemer Eenheid van Taal wordt in Nederland gewerkt aan verdere standaardisatie in het verlengde van onder andere de Wegiz en de WKZ, waarbij de nadruk vooral ligt op het zorgdomein. Voor secundaire gezondheidsdata-toepassingen worden veelal andere standaarden toegepast. De EHDS onderkent deze diversiteit en gaat er expliciet vanuit dat meerdere standaarden voor secundair datagebruik ondersteund worden. Om databeschikbaarheid over domeinen heen te faciliteren, richt de Nationale gezondheidsdata-infrastructuur secundair gebruik zich daarom op het verbinden van de dominante standaarden uit zowel het gezondheidszorg- als onderzoeksdomein.

Om dit mogelijk te maken worden de meest gebruikte standaarden geïdentificeerd en wordt een nationale selectie ondersteund met uniforme mappingvoorzieningen. Voorbeelden van de meeste gebruikte standaarden zijn openEHR, FHIR, OMOP CDM, DICOM, VCF en ICPC. Een meer volledig overzicht van datacategorieën en datastandaarden is opgesteld door Nictiz: [Verkenning Standaarden voor informatiemodellen v1.01](#). Door standaarden op een uniforme wijze met elkaar te verbinden,

⁴ Binnen de gehele data journey komen andere vormen van metadata voor zoals in "provenance" (wat is er met de data gebeurd, hoe is de data verwerkt ten behoeve van reproduceerbaarheid), proces data, etc. In een volgende versie zal ook aan deze andere vormen van metadata aandacht worden besteed.

wordt datatransformatie tussen verschillende modellen en formaten herbruikbaar en consistent ingericht. De komende jaren ligt de nadruk daarbij op het realiseren van syntactische interoperabiliteit als basis voor verdere ontwikkeling richting semantische interoperabiliteit. Hiermee ontstaat een stapsgewijze aanpak waarin data eerst technisch uitwisselbaar worden gemaakt, waarna verdere harmonisatie en betekenisvolle interpretatie kunnen worden ondersteund.

Het landschap van standaarden en informatiemodellen blijft zich continu ontwikkelen. Ook op Europees niveau wordt gewerkt aan nieuwe standaarden, waaronder EEHRxF en de bijbehorende Logical Information Models (LIM's). De Nationale gezondheidsdata-infrastructuur secundair gebruik wordt daarom zodanig ontworpen dat nieuwe versies en toekomstige ontwikkelingen van standaarden binnen redelijke grenzen ondersteund kunnen worden, zodat gebruik van de Nationale gezondheidsdata-infrastructuur secundair gebruik duurzaam kan aansluiten op nationale en Europese ontwikkelingen.

Databeschikbaarheidsprincipe

Databeschikbaarheid houdt in dat vastgelegde gezondheidsdata idealiter continu beschikbaar, bereikbaar en bruikbaar zijn, zowel voor het primaire als voor het secundaire proces. Het houdt in dat deze data – met waarborgen voor privacy en veiligheid – op ieder moment hergebruikt kunnen worden voor onderzoek, beleid en innovatie, ongeacht waar ze zijn opgeslagen. Naar deze ideale situatie wordt toegewerkt, wetende dat niet voor alle data mogelijk zal zijn.

Voor ieder onderzoek worden specifieke (interoperabiliteits-)oplossingen gerealiseerd om data uit bronsystemen te halen om deze vervolgens te verrijken en daarmee geschikt te maken voor dat specifieke onderzoek maar ook voor toekomstig secundair gebruik waar mogelijk. Door afspraken te maken en tools aan te bieden beoogt de Nationale gezondheidsdata-infrastructuur secundair gebruik tot generieke oplossingen te komen en daarmee het aantal onderzoeksspecifieke oplossingen terug te dringen. Met deze opzet kunnen zowel kosten worden bespaard en kan de doorlooptijd tot de daadwerkelijke start van een onderzoek aanzienlijk worden verkort.

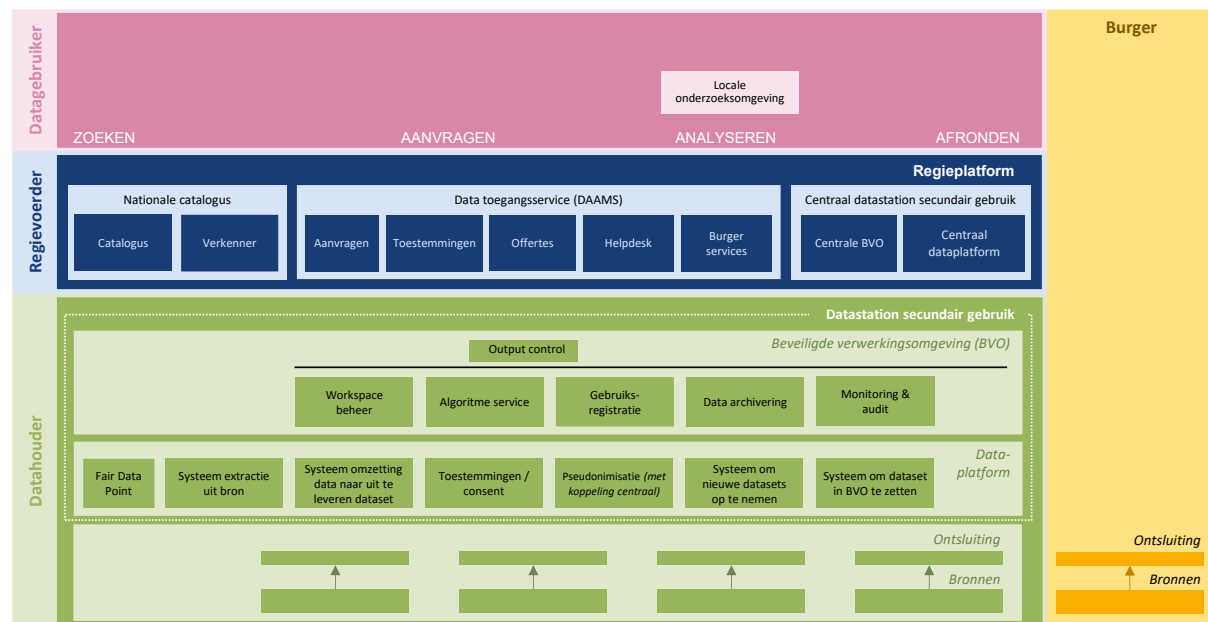
Datakwaliteit

Datakwaliteit vormt een belangrijk onderdeel van de Nationale gezondheidsdata-infrastructuur secundair gebruik en raakt verschillende aspecten van databeschikbaarheid en hergebruik. Om datagebruikers in staat te stellen de geschiktheid van data voor een specifieke toepassing te beoordelen, wordt informatie over datakwaliteit op een gestandaardiseerde wijze vastgelegd en beschikbaar gemaakt via metadata, via een datakwaliteitslabel zoals voorgeschreven door de EHDS. Het Europese project Quantum heeft hiervoor een model ontwikkeld dat handvatten biedt voor het beschrijven en inzichtelijk maken van datakwaliteit. Het is echter de verwachting dat een generiek data kwaliteitslabel niet altijd genoeg is voor het beoordelen van de geschiktheid van de data. Dat vergt additionele functionaliteit waarbij de (prospectieve) datagebruiker en datahouder om te bepalen of de data geschikt is.

Binnen de Nationale gezondheidsdata-infrastructuur secundair gebruik worden verschillende vormen van datatransformatie ondersteund, zoals het geschikt maken van ongestructureerde data voor hergebruik, het omzetten tussen standaarden en het samenvoegen van data uit verschillende bronnen. Omdat deze verwerkingen invloed kunnen hebben op de kwaliteit van data, wordt voorzien in kaders en voorzieningen om de effecten van de verwerkingen inzichtelijk te maken en te monitoren. Daarbij wordt onder meer aandacht besteed aan de invloed van opt-out-voorzieningen op datavolledigheid, de effecten van data-linkage en data-enrichment, en de rol van synthetische data als onderdeel van de dataketen. In dat kader blijft brondata in beginsel beschikbaar als ruwe data, zodat deze opnieuw gebruikt kan worden en kan dienen als referentie voor verificatie van data en uitgevoerde transformaties.

4.4. Applicatie-architectuur

De applicatie-architectuur beschrijft welke applicaties en functies nodig zijn om de beoogde processen en informatievoorziening binnen het secundair gebruik van gezondheidsdata te ondersteunen. Daarbij wordt inzicht gegeven in de verdeling van functionaliteit over applicaties, de onderlinge samenhang tussen deze applicaties en de wijze waarop zij gezamenlijk bijdragen aan het realiseren van de beoogde toekomstige situatie. De focus ligt op de functionele rollen van applicaties en hun interacties, niet op de technische implementatie of concrete productkeuzes. In de volgende figuur zijn deze applicaties en de belangrijkste functies weergegeven:



Figuur 9. Applicaties met belangrijkste functies.

De applicaties en functies worden in deze paragraaf beschreven vanuit het perspectief van de belangrijkste gebruikersgroepen binnen het gezondheidsdata-ecosysteem. Hierbij wordt onderscheid gemaakt tussen:

- applicaties en functies voor datagebruikers (5.4.1), die data willen vinden, aanvragen en gebruiken voor onderzoek, beleid en innovatie, te weten de nationale gezondheidsdata catalogus (1), de nationale gezondheidsdata toegangsservice (2) en de beveiligde verwerkingsomgeving (3);
- applicaties en functies voor datahouders (5.4.2), die data beschikbaar stellen en de toegang daartoe beheren, te weten het dataplatform (4);
- applicaties en functies voor burgers (5.4.3), die hun zeggenschap willen uitoefenen en gebruik maken van hun recht op transparantie, te weten het burgerportaal (5);
- en generieke applicatiefuncties die deze processen ondersteunen (5.2.4), zoals identificatie, authenticatie, autorisatie en adressering. Deze generieke applicatiefuncties die in meerdere applicaties en op meerder niveaus (landelijk, regionaal, lokaal) voorkomen, zijn vanwege de overzichtelijkheid niet opgenomen in de figuur. De generieke applicatiefuncties kunnen, mits op de juiste wijze gespecificeerd en gerealiseerd, generiek worden toegepast.

Op de functies van de bronsystemen en hun ontsluiting wordt verder niet ingegaan. Wel zullen bij een verdere verdieping van de doelarchitectuur eisen worden gesteld aan de ontsluiting van de bronsystemen, waaronder eisen aan de koppelvlakken.

De mate van detaillering van de beschreven functies verschilt. Voor sommige functies zijn de benodigde functionaliteit, verantwoordelijkheden en interacties al relatief goed uitgewerkt, terwijl voor andere functies op dit moment vooral de gewenste richting en rol binnen het stelsel zijn

beschreven. Naarmate de doelarchitectuur verder wordt uitgewerkt en ervaringen worden opgedaan in de praktijk, zullen de beschrijvingen stapsgewijs worden aangescherpt en aangevuld, zodat uiteindelijk voor alle functies een vergelijkbaar niveau van detail ontstaat.

NB: Aan het eind van deze paragraaf wordt het FAIR Data Point (FDP) nader toegelicht (5.2.5), omdat deze toelichting te uitgebreid is om direct op te nemen bij de FDP-functie binnen de Nationale gezondheidsdata catalogus.

4.4.1. Datagebruikers

Applicaties en functies voor datagebruikers, die data willen vinden, aanvragen en gebruiken voor onderzoek, beleid en innovatie, te weten de nationale gezondheidsdata catalogus (1), de nationale gezondheidsdata toegangsservice (2) en de beveiligde verwerkingsomgeving (3).

1. Nationale gezondheidsdata catalogus

Datagebruikers kunnen beschikbare datasets zoeken en vinden in de nationale gezondheidsdata catalogus. Deze catalogus geeft een overzicht van beschikbare datasets op het gebied van health & life sciences in Nederland. De catalogus bevat metadata over de beschikbare datasets, wat betekent dat het de beschrijving van de datasets bevat.

Datasets worden beschreven met een gestandaardiseerd metadataschema voor de nationale gezondheidsdata-infrastructuur secundair gerbuik, zodat ze makkelijk gevonden, doorzocht en vergeleken kunnen worden. Om opgenomen te worden in de catalogus, moeten beschrijvingen aan minimale eisen van dit schema voldoen. Door het schema af te stemmen op de Healthdata@EU-vereisten van de HDAB, zijn datasets uit de Nationale gezondheidsdata catalogus straks ook vindbaar in de Europese HDAB-catalogus.

De functies binnen de catalogus zijn:

- **Zoekfunctie datasets (met beschrijvingen en classificatie):** Functie voor het doorzoeken van aangemelde datasets met metadata zoals beschrijving en classificatie. Om efficiënte en nauwkeurige zoekopdrachten te faciliteren, moet het mogelijk zijn om te zoeken op vrije tekst en gestandaardiseerde metadata-eigenschappen. Ook moet het mogelijk zijn om de datacatalogus te doorzoeken met behulp van een ontologie op basis van een hiërarchische structuur (zoals SNOMED CT), zodat op een concept kan worden gezocht en al zijn afstammelingen in de hiërarchie kunnen worden opgehaald zonder dat de datagebruiker elke mogelijke variant of term hoeft te kennen. Dit stelt gebruikers in staat om gemakkelijk alle relevante informatie met betrekking tot een concept terug te vinden, waardoor de algehele bruikbaarheid en toegankelijkheid van de data catalogus wordt verbeterd. De zoekfunctie dient de datagebruiker te ondersteunen bij het nauwkeurig opgeven van de benodigde dataelementen met bijbehorende in- en exclusiecriteria en vervolgens het zoeken hierop in de beschikbare datasets. De zoekopdracht levert een lijst met datasets op die overeenkomen met de zoekopdracht.
- **Selectie datasets:** De mogelijkheid om één of meerdere datasets te selecteren, om een aanvraag van deze datasets te kunnen initiëren/starten.
- **Lokalisatie van datasets:** Functie die inzicht geeft waar datasets zich bevinden en hoe ze benaderd kunnen worden.
- **Verkenner:** Functie om gevonden datasets / Real World Data (RWD) nader te verkennen zonder directe toegang tot de data. Voordat een eventuele formele data-aanvraag wordt gedaan, kunnen met de inzet van een verkenners databronnen bij datahouders worden onderzocht. Op basis van het zoekresultaat kan vooraf een selectie van data-elementen worden gemaakt, zodat de data-aanvraag zo goed mogelijk kan worden ingediend.
Er zijn verschillende invullingen van deze applicatiefunctie mogelijk:
 - (potentiele) datagebruikers communiceren zelf met de datahouders;

- gebruik maken van bestaande verkenner, zoals Beacon en Locator;
- een generieke verkenner op basis van gefedereerde analyse statistische vragen laten stellen;
NB: Veel datahouders en de door hun gebruikte systemen zijn er momenteel niet op ingericht dat een dergelijke generieke verkenner deze systemen benadert. Het mogelijke maken van het verkennen van data in de systemen vergt dan ook nog een aanzienlijke organisatorische en technische transitie bij de datahouder.
- AI oplossingen.

- **Beheer datacatalogus:** Beheerdersfunctie voor het inrichten, beheren en doorontwikkelen van de Nationale Gezondheidsdatacatalogus. Deze functie draagt zorg voor de kwaliteit, actualiteit en bruikbaarheid van de catalogus en een goede ondersteuning van datahouders en datagebruikers. Ook het beheer van versies van datasets, het beheer van metadata sjablonen, het importeren en exporteren van metadata en het beheer van zoekindexen en zoekmachineoptimalisatie (SEO) maakt onderdeel uit van deze functie. Daarnaast omvat deze functie het beheer van de technische inrichting van de catalogus, het begeleiden van wijzigingen en doorontwikkeling, en het bewaken van aansluiting op relevante standaarden, wet- en regelgeving en gebruikersbehoeften.
- **Monitor beheerders:** Administratief dashboard om gebruikers- en zoekstatistieken weer te geven.
- **Monitor datahouders:** Dashboard dat de datahouder inzichten en managementhulpmiddelen biedt. Het dashboard stelt de datahouder in staat om te volgen hoe zijn datasets presteren in de zoekopdracht en helpt bij het beheren van de metadata van de datasets. Het dashboard biedt ook inzicht in de aanvragen van de datasets van de datahouder.
- **Monitor datagebruikers:** Gebruikersdashboard biedt een gebruiker van de catalogus toegang tot zijn zoekopdrachten en verzoeken om toegang tot datasets.

2. Nationale gezondheidsdata toegangsservice (DAAMS)

De nationale gezondheidsdata toegangsservice heeft als doel het aanvragen, vastleggen en organiseren van autorisatie voor het hergebruik van gezondheidsdata. Deze functie blijft van belang, ongeacht de wijze waarop organisatie van verantwoordelijkheden, verantwoordelijkheden en besluitvorming zijn georganiseerd. De doelarchitectuur gaat daarom uit van een stabiele toegangsfunctie die verschillende varianten van organisatie van verantwoordelijkheden kan ondersteunen.

Binnen de beschouwde tijdshorizon kunnen nieuwe Europese en nationale kaders, waaronder de EHDS en de HDAB-NL, invloed hebben op de verdeling van verantwoordelijkheden voor beoordeling, autorisatie en toezicht. De architectuur anticipeert hierop door functies en voorzieningen zoveel mogelijk los te koppelen van de partij die deze uitvoert. Hierdoor blijft de nationale gezondheidsdata-infrastructuur bruikbaar bij veranderende afspraken over organisatie van verantwoordelijkheden en kan worden voortgebouwd op bestaande investeringen, onafhankelijk van de uiteindelijke organisatorische inrichting.

De nationale gezondheidsdata toegangsservice bevat de volgende functies:

- **Aanvraagformulering:** Functie waarmee een datagebruiker zijn behoefte aan gezondheidsdata gestructureerd en concreet vastlegt in een data-aanvraag, zodat kan worden beoordeeld welke data nodig zijn, waarvoor deze worden gebruikt en onder welke voorwaarden toegang kan worden verleend.
- **Autorisatie en afhandeling aanvragen:** Functie voor beoordelen, goedkeuren van aanvragen voor autorisatie datatoegang. Er zijn twee belangrijke hoofdonderdelen bij de beoordeling: (1) de administratieve, praktische en technische toetsing en (2) de inhoudelijke toetsing:
 - Technische toetsing: De administratieve, praktisch en technische toetsing door nationale gezondheidsdata toegangsservice / HDAB DAAMS.

- Inhoudelijke toetsing: Het beoordelen of een aanvraag voor toegang tot gezondheidsdata voldoet aan de geldende juridische, ethische, organisatorische en inhoudelijke voorwaarden. De verantwoordelijkheid voor deze toetsing ligt, afhankelijk van het toepasselijke kader, bij de HDAB of bij de betreffende datahouder(s).
- **Aanvragenadministratie:** Functie voor administratief beheren van aanvragen. Hierin is ook te vinden waar elke aanvraag over gaat en welke autorisatie geldt.
- **Ondersteuning workflow:** Functie voor faciliteren van het proces van aanvraag tot datagebruik via gestandaardiseerde stappen. Dit kan bijvoorbeeld met een ticketing module waarbij de data aanvraag resulteert in allerlei subtickets (ethische toetsing, vragen richting datahouder, inrichten BVO, aftrappen financiële afrekening etc. etc.).
- **Nationaal coördinatiepunt secundair gebruik:** Het coördinatiepunt zorgt voor een eenduidige toegang, passende doorverwijzing en samenhangende ondersteuning over de gehele gebruikersreis.
- **Contractbeheer:** Functie voor beheren van afspraken rondom datagebruik:
 - **Datagebruikerregister:** Generiek applicatiefunctie voor het vastleggen en beheren van informatie over datagebruikers en relevante kenmerken die nodig zijn binnen het proces van datatoegang en datagebruik.
 - **Offerteservice:** Functie ter ondersteuning van het offerteproces.
NB: Meer eenduidigheid in prijzen en voorwaarden is noodzakelijk om de (financiële) voorwaarden van onderzoeken te faciliteren, die de grenzen van één datahouder overschrijden.
 - **Licensering / data management:** Functie voor gebruiksvoorwaarden, licenties en beheerafspraken rondom datasets.
 - **Registratie gebruik:** Functie voor vastleggen van gebruik toegang tot data en voorzieningen.
 - **Fee management:** Functie voor afhandeling van kosten en vergoedingen.
 - **Betaalservice:** Functie voor de daadwerkelijke betaling, eventueel uitgevoerd door een derde partij.

3. Beveiligde verwerkingsomgeving

Aan BVO's kunnen de datahouders de data aanleveren en binnen deze BVO's kunnen de datagebruikers de data analyseren. Er zijn passende technische en organisatorische maatregelen geïmplementeerd om de data en de privacy te beschermen. Data veilig verwerken kan plaatsvinden in verschillende typen BVO's. Zoals beschreven in paragraaf **Fout! Verwijzingsbron niet gevonden.** **Fout! Verwijzingsbron niet gevonden.**, kan een BVO gebruikt worden voor data access, waarbij gebruikers toegang krijgen tot data binnen een BVO en data visiting, waarbij analyses naar de beschikbare data worden gebracht en uitsluitend gecontroleerde resultaten worden ontsloten.

Binnen Nationale gezondheidsdata-infrastructuur secundair gebruik is gekozen voor een netwerk van BVO's, waarbij zowel datahouders als datagebruikers, als centrale instanties gezamenlijk in de behoeften van datahouders en datagebruikers voor veilig en doelmatig hergebruik van gezondheidsdata voorzien. Uitgangspunt is dat er een netwerk van BVO's blijft bestaan, waarbij er ruimte is voor meerdere oplossingen om de vereisten vanuit onderzoek / analyse te ondersteunen. Als een data-aanvraag wordt goedgekeurd, kan de data veilig uitgegeven worden in een BVO die voldoet aan de vereisten.

Health-RI zal in het netwerk van gecertificeerde BVO's van de nationale gezondheidsdata-infrastructuur secundair gebruik ook een BVO aanbieden om het verwerken van datasets die aangevraagd worden via de nationale gezondheidsdata catalogus en nationale gezondheidsdata toegangsservice te ondersteunen. Deze BVO heeft tot doel behoefte in het landschap in te vullen, namelijk een BVO op "neutrale grond" om aanvragen via de nationale gezondheidsdata catalogus, multicenter studies en in de toekomst ook datahouders zonder BVO beter te ondersteunen.

Datahouders kunnen data beschikbaar maken in de BVO's en hebben door de gemaakte generieke afspraken vertrouwen in het delen van data in de BVO's.

De functies van een BVO zijn:

- **Beheer BVO's:** Functie voor inrichten en beheren van BVO's voor onderzoekers. De BVO's dienen voorafgaand aan opname in het register gecertificeerd te worden.
- **BVO-register:** Functie voor de registratie van erkende beschikbare BVO's.
- **Beheer gereedschappen:** Functie voor beheer van in de BVO aangeboden hulpmiddelen voor het geschikt maken data voor onderzoek, het analyseren van data en het publiceren van onderzoeksresultaten. Hieronder valt onder meer het beheren van de gebruiksvoorwaarden en de afspraken over het niveau van dienstverlening. Ook het contracteren inclusief de licentieverlening, het ondersteunen bij en het trainen in het gebruik van de hulpmiddelen.
- **Gereedschap assemblage:** Functie om veelgebruikte samengestelde, geconfigureerde of aangepaste hulpmiddelen voor datahouders en datagebruikers te selecteren en beschikbaar te stellen.
- **Toegangscontrole:** Functie om te beheren wie toegang heeft tot een BVO en daarmee de toegang tot data.
- **Input control:** Functie om te controleren welke data, hulpmiddelen of algoritmen een BVO binnen mogen.
- **Algoritme services:** Services voor onder meer het uitvoeren, verbeteren en voorspellen van analyses, waaronder gefedereerde analyses.
- **Statistische services:** Services voor onder meer het uitvoeren, verbeteren en voorspellen van statistische analyses op data.
- **Multi Party Computing:** Functie voor gezamenlijke berekeningen over meerdere partijen zonder alle data centraal samen te brengen.
- **Transport:** Functie om datasets of geaggregeerd resultaten beveiligd van de ene omgeving naar de andere te brengen.
- **Output control:** Functie om te controleren of analyseresultaten / algehele datasets de BVO mogen verlaten (privacy/security).
- **Publicatie datasets:** Functie voor het beschikbaar maken van datasets voor veilige publicatie en het publiceren van deze datasets zelf. Waar de datasets veilig worden gepubliceerd dient nog nader onderzocht en bepaald te worden.
- **Publicatie onderzoeksresultaat:** Functie voor het beschikbaar maken van onderzoeksuitkomsten voor publicatie en de daadwerkelijke publicatie zelf.
NB: In veel situaties is het nog niet duidelijk hoe invulling gegeven moet worden aan deze functie. In de komende periode zal dit nader worden uitgewerkt.
- **Teruggave verrijkte data aan datahouder:** Functie voor terugleveren van verrijkte of verbeterde data aan de oorspronkelijke datahouder, mits dit in het contract is afgesproken.
- **Archivering datasets:** Functie voor opslaan en bewaren van datasets voor toekomstig gebruik.
- **Opruiming BVO:** Functie voor het opgeruimd achterlaten van de BVO conform de geldende afspraken.
- **Monitoring & auditing:** Toezicht houden op gebruik en naleving:
 - **Logging:** Functie voor vastlegging van uitgevoerde activiteiten.
 - **Auditing & accounting:** Functie voor controle en verantwoording van gebruik.
 - **Rapportage & compliance:** Functie voor rapporteren over voldoen aan regelgeving.

4.4.2. Datahouders

Applicaties en functies voor datahouders, die data beschikbaar stellen en de toegang daartoe beheren, te weten het dataplatform (4).

4. Dataplatform

Het dataplatform ondersteunt datahouders bij het ontsluiten, beheren en beschikbaar stellen van gezondheidsdata voor secundair gebruik. Binnen het dataplatform worden databronnen beschreven en vindbaar gemaakt, aanvragen behandeld en datasets voorbereid voor gebruik in onderzoek, beleid en innovatie. Daarbij kunnen verschillende bewerkingen worden toegepast, zoals filtering op bezwaar (opt-out), extractie, pseudonimisering, transformatie en optimalisatie van data. Het dataplatform faciliteert zowel het beschikbaar stellen van data in de oorspronkelijke vorm als in gestandaardiseerde en geharmoniseerde vormen, afhankelijk van de behoeften van de datagebruiker. De geprepareerde data kan vervolgens beschikbaar worden gesteld aan een BVO van de datahouder zelf of aan een BVO voor verdere analyse en verwerking. Het is onwenselijk dat datahouders voor iedere specifieke aanvraag een dataset in een specifiek datamodel of formaat moeten aanleveren. Data zou zoveel mogelijk éénmalig in gestandaardiseerde vorm moeten worden omgezet.

De volgende functies maken deel uit van het dataplatform:

- **Data omschrijving service / FAIR Data Point (FDP):** Functie om datasets en metadata machineleesbaar en FAIR te beschrijven en te publiceren. Zie nadere toelichting onder punt
- **Datadefinitieregister:** Generieke applicatiefunctie voor formele data-elementdefinities en datamodellen, als semantische basis voor interpretatie en reproduceerbaarheid.
- **RDF-registratie:** Registratie van metadata volgens het Resource Description Framework (RDF). Hierna kan de metadata worden gepubliceerd via de functie FDP.
- **Coördinatiepunt secundair gebruik:** Eén eenduidige plek waar datagebruikers – en houders terecht kunnen voor vragen over het gebruik, respectievelijk de aanlevering van data. NB: De functie zou ook gebruikt kunnen worden ter ondersteuning van de aanlevering van metadata aan de nationale gezondheidsdata catalogus of gezondheidsdata aan een BVO.
- **Datahouderregister:** Generieke applicatiefunctie voor de registratie van de organisaties die deelnemen aan het gezondheidsdata-ecosysteem in de rol van datahouder. Het register legt vast welke rechtspersonen zijn aangesloten, onder welke voorwaarden zij deelnemen en welke gegevensbronnen, voorzieningen en FAIR Data Points (FDP's) aan deze organisaties zijn gekoppeld. Daarnaast ondersteunt het register de inrichting van verantwoordelijkheden en bevoegdheden door vast te leggen welke vertegenwoordigers namens een datahouder bevoegd zijn om medewerkers te autoriseren voor beheer- en toegangstaken binnen het stelsel. Het datahouderregister vormt daarmee een centrale bron voor de identificatie van datahouders, de koppeling met hun voorzieningen en de afspraken over organisatie van verantwoordelijkheden, die deelname aan het ecosysteem mogelijk maken.
- **Management generieke koppelvlakken:** Functie voor beheren van koppelingen tussen applicaties onderling. Een bijvoorbeeld hiervan is het te beheren koppelvlak dat nodig is om voor een en hetzelfde onderzoek gefedereerde BVO's kunnen benaderen die gebaseerd zijn op verschillende principes. Dat lukt alleen maar als er ook (technische) afspraken zijn gemaakt over de koppelvlakken en deze afspraken ook worden beheerd.
- **Management specifieke koppelvlakken:** Functie voor beheren van koppelingen tussen bronsystemen en de Nationale gezondheidsdata-infrastructuur secundair gebruik. Deze koppelvlakken zijn veelal specifiek omdat leveranciers elk op hun eigen wijze standaarden implementeren in hun systemen.
- **Encryptie:** Functie voor versleuteling van data, waaronder toepassing van het Data Encryption Protocol en Forward Error Correction. Landelijke afspraken zijn nodig om encryptie over datahouders heen te laten werken.
- **Terminologieservices:** Generieke applicatiefunctie ter ondersteuning van uniforme begrippen en coderingen.
- **Datapreparatie bibliotheek:** Bibliotheek met herbruikbare set van regels en hulpmiddelen voor het prepareren van data, zoals standaardiseren en consolideren.

- **Data / File exchange:** Generieke applicatiefunctie voor tijdelijke, veilige uitwisseling van bestanden tussen partijen, zonder bewerking van de data zelf. Faciliteert naadloze integratie en koppeling van datasets binnen diverse systemen en organisaties, waardoor gegevens optimaal beschikbaar en bruikbaar worden.
- **Data minimalisatie:** Functie om bij de extractie data uit een bron de beschikbaar gestelde data te beperken tot wat nodig is. Op kolommen en records. Dit houdt in dat enerzijds geen data wordt verstrekt van burgers die gebruik hebben van hun opt-out recht en anderzijds dat alleen data wordt verstrekt die relevant is voor het doel (doelbinding).
- **Opt-out filter:** Functie om data te filteren op burgers die hun toestemming voor secundair gebruik van gezondheidsdata over henzelf hebben ingetrokken.
NB: Standaard geldt binnen de EHDS dat er sprake is van toestemming door de burger.
- **Consent filter:** Functie om te toetsen of burger toestemming heeft gegeven om data te delen ten behoeve van (een specifiek) onderzoek. Het betreffen onderzoeken die buiten de EHDS vallen.
NB: Mogelijk gaat Nederland ook voor onderzoeken die buiten de EHDS vallen het opt-out principe hanteren.
- **Datakwaliteit validatie:** Functie om te controleren of data volledig, correct en bruikbaar is en op basis van deze controle een kwaliteits- en toepasbaarheidslabel toe te kennen aan de data.
- **Cleansing:** Functie voor opschonen van data voordat deze verder gebruikt, gekoppeld of uitgegeven wordt.
- **Record koppeling:** Generieke applicatie functie om verschillende datasets samen te voegen tot één dataset. Data kan op twee manieren gepartitioneerd zijn:
 - Bij verticale partitionering is record koppeling noodzakelijk om tot een bruikbare dataset te komen. De bron-datasets bevatten dezelfde personen, maar verschillende variabelen. Bron-datasets van verschillende datahouders kunnen gekoppeld worden middels identifiers waardoor een natuurlijke persoon als deelnemer in de verschillende bron-datasets wordt geïdentificeerd. Elke dataset bevat verschillende variabelen. Om de waarden van deze verschillende variabelen aan dezelfde persoon te kunnen koppelen is een persoonsidentificer nodig of een koppeltabel die verschillende persoonsidentifiers aan elkaar koppelt.
 - Bij horizontale partitionering is record koppeling noodzakelijk voor eventuele ontduubeling of stapeling. De datasets bevatten elk dezelfde variabelen maar verschillende deelnemers. Bron-datasets van verschillende datahouders worden aan elkaar gekoppeld, waarbij in elke dataset in principe verschillende natuurlijke personen zijn opgenomen. Dubbelingen (een kopie) en overlap (een kopie met aanvullingen) kunnen ontstaan door verplaatsing van deelnemers naar een ander centrum of bij samengestelde bron-datasets. Om ontduubeling mogelijk te maken is het wenselijk om personen te kunnen onderscheiden op basis van een identifier.
- **Data verrijking:** Functie om data betekenisvoller of beter bruikbaar te maken voor hergebruik.
- **Generieke transformatie:** Generieke applicatie functie voor omzetten van data of concepten van het ene model of formaat naar het andere. Een voorbeeld hiervan is het omzetten van FHIPR naar OMOP.
- **Applicaties specifieke transformatie:** Functie voor omzetten van data naar het afgesproken formaat of standaard. Een voorbeeld hiervan is het omzetten van een leverancier specifieke implementatie van geboortezorgdata naar de afgesproken geboortezorgstandaard.
- **CumuluZ adapter:** Generieke applicatiefunctie die bestaande zorginformatiesystemen (zoals EPD's, HIS'en, ECD's en andere bronsystemen) laat aansluiten op de Nationale gezondheidsdata-infrastructuur secundair gebruik, zonder dat deze systemen zelf volledig hoeven te worden aangepast. De adapter implementeert de generieke functies die nodig zijn voor data-integratie en dataconformiteit.
- **Persistentie platform:** Functie om data op te slaan of beschikbaar te houden op een persistent dataplatform, wanneer de bron zelf onvoldoende performance of standaardisatie biedt. Een persistent dataplatform is een centrale en duurzame opslagomgeving voor data die gedurende

lange tijd beschikbaar moet blijven en ongewijzigd moet worden gehouden. Binnen deze omgeving wordt onderzoeksdata opgeslagen en vanuit deze omgeving wordt deze data beschikbaar gesteld.

- **Pseudonimisatie:** Generieke applicatiefunctie om identificeerbare data te vervangen door pseudoniemen, met landelijke afspraken voor consistent gebruik tussen datahouders. Indien men beschikt over de juiste sleutel(s), zijn pseudoniemen weer terug te leiden tot identificeerbare data. Gepseudonimiseerde gegevens blijven persoonsgegevens onder de AVG, zolang de koppeling met de persoon nog kan worden hersteld. De informatie waarmee die koppeling kan worden gemaakt moet daarom afzonderlijk worden bewaard en beschermd. Pseudonimisatie is met name wenselijk wanneer:
 - gegevens van dezelfde persoon uit verschillende bronnen gekoppeld moeten kunnen worden;
 - aanvullende gegevens over dezelfde persoon verzameld moeten kunnen worden;
 - onderzoeksgegevens nog gecontroleerd moeten kunnen worden tegen de brongegevens;
 - eventueel incidentele bevindingen teruggekoppeld moeten kunnen worden.
- Omdat in veel onderzoek data van burgers bij meerdere datahouders vandaan komt, is het van belang dat bij pseudonimisatie burgers uniek geïdentificeerd kunnen worden over de grenzen van datahouders heen. Dit kan bijvoorbeeld worden gerealiseerd door de unieke identificatie te laten verlopen via een trusted third party.
- **Sleutel lijst:** Generieke applicatiefunctie bestaande uit een lijst waarin de koppeling wordt gelegd tussen gepseudonimiseerde data en de oorspronkelijke identiteit van personen of objecten.
 - **Anonimisatie:** Functie voor verwijderen van herleidbare data zodat personen niet meer identificeerbaar zijn. Het gaat hierbij niet alleen om het verwijderen van directe identificatoren. Ook indirecte identificatie door het combineren of koppelen van data moet worden uitgesloten. Onder meer aggregatie kan worden gebruikt om de kans op identificatie te verkleinen. Een voorbeeld hiervan is het vervangen van een geboortedatum door een leeftijd of leeftijdscategorie.
 - **Synthetisering:** Functie om synthetische of mockdata te genereren, bijvoorbeeld voor discovery of testdoeleinden. Synthetiseren op grond van passend statistisch beveiligde aggregaten heeft de voorkeur boven synthetiseren op grond van persoonsgegevens. Voor het laatste zijn meer waarborgen noodzakelijk. In het algemeen maakt een beperktere distributie van een als zodanig herkenbare synthetische dataset een breder palet van gebruiksscenario's mogelijk. Wanneer persoonsgegevens als input worden gebruikt, moet afzonderlijk worden beoordeeld of de uiteindelijke synthetische dataset nog persoonsgegevens bevat. Die beoordeling is bepalend voor de vraag of de AVG ook op het gebruik en de verdere verstrekking van de synthetische dataset van toepassing blijft.
 - **Publicatie datasets naar BVO's:** Functie voor klaarmaken van datasets voor veilige publicatie naar een BVO en voor het publiceren zelf.

4.4.3. Burgers

Applicaties en functies voor burgers, die hun zeggenschap willen uitoefenen en gebruik maken van hun recht op transparantie, te weten het burgerportaal (5).

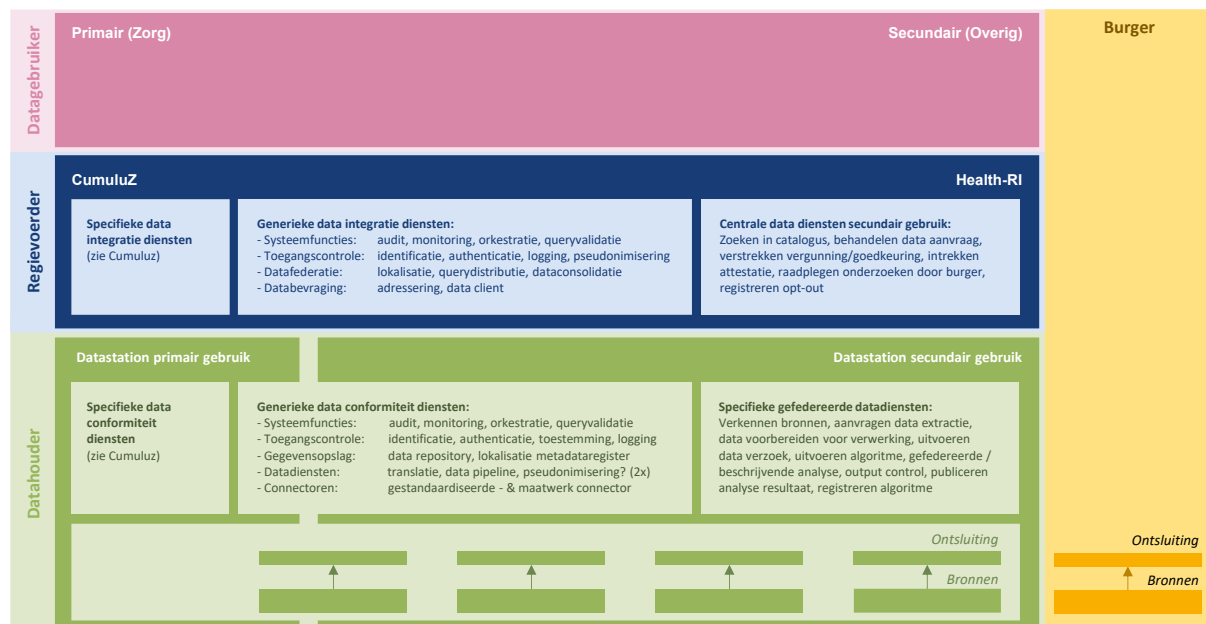
5. Burgerportaal

Functionaliteit gericht op burgers:

- **Registratie zeggenschap:** Functie voor registreren en uitoefenen van zeggenschap van burgers (opt-out / consent) over secundair datagebruik.
- **Transparantie:** Functie voor burger waarmee deze inzicht kan verkrijgen in het gebruik van data over hem / haar voor onderzoek, beleid en innovatie.

4.4.4. Generieke applicatiefuncties

Met de generieke opzet van applicatiefuncties wordt bijgedragen aan het verminderen van versnippering en het bevorderen van interoperabiliteit en herbruikbaarheid van data. Daarom wordt per functionele component beoordeeld in hoeverre deze generiek gespecificeerd en ingezet kan worden voor zowel primair als secundair gebruik van data. Vanuit dit uitgangspunt is zichtbaar dat de functionele componenten binnen de architectuur voor het primaire gezondheidszorgproces op meerdere plaatsen duidelijke overlap vertonen met functionele componenten binnen de architectuur voor secundair datagebruik. Deze overlap is weergegeven in figuur 10:



Figuur 10. Overlap generieke applicatiefuncties primair en secundair hergebruik

De zes (met logging erbij zeven) bekendste generieke applicatiefuncties zijn:

- **Toestemming / zeggenschap:** Functie voor het vastleggen van toestemming (primair gebruik) of het intrekken van toestemming (secundair gebruik).
- **Adressering:** Functie om datahouders, datagebruikers of technische adressen/ontvangers te registreren en te kunnen bereiken.
- **Authenticatie:** Functie om te verifiëren dat een actor daadwerkelijk is wie deze zegt te zijn.
- **Autorisatie:** Functie om te controleren of een actor rechten heeft op data of functionaliteit.
- **Identificatie:** Functie om identiteit van personen, organisaties of applicaties vast te stellen.
- **Lokalisatie data over een persoon:** Functie om vast te stellen waar welke data over een persoon staat.
- **(Logging:** Functie voor het vastleggen van gebeurtenissen, gebruik en acties op verschillende plekken in de Nationale gezondheidsdata-infrastructuur secundair gebruik.)

Deze functies komen voor in meerdere applicaties, zowel bij het primaire, als bij het secundaire gebruik van data. Het ministerie van VWS heeft een regierol bij het ontwikkelen van deze generieke applicatiefuncties. Hierbij heeft VWS ook voor de generieke applicatiefunctie pseudonimisatie (inclusief bijbehorende sleutel lijst) de regie naar zich toegetrokken. Onder deze regie zijn de inspanningen momenteel vooral gericht op het ontwikkelen van voorzieningen, die de behoeften vanuit het primaire gebruik van gezondheidsdata ondersteunen. De Nationale gezondheidsdata-infrastructuur secundair gebruik zal in principe aansluiten bij de keuzes die in dat proces worden gemaakt, waarbij de eisen aan deze generieke functies voor secundair datagebruik ingebracht.

Dat functies dezelfde benaming hebben, betekent overigens niet automatisch dat zij inhoudelijk of technisch gelijk zijn. Verschillen in doelstelling, context of juridische grondslag kunnen ertoe leiden

dat functies voor primair en secundair datagebruik wezenlijk anders moeten worden ingericht. In de komende periode wordt daarom nader onderzocht in hoeverre functionele componenten daadwerkelijk generiek toepasbaar zijn. Daarbij wordt, mede op basis van praktijkbeproevingen en in nauwe samenwerking met onder andere CumuluZ en GIS-A, bepaald welke functies daadwerkelijk gezamenlijk gebruikt kunnen worden en welke componenten functioneel of technisch toch om verschillende implementaties vragen. De uitkomsten hiervan bepalen mede de verdere inrichting, organisatie van verantwoordelijkheden en implementatie van de architectuur.

De overlap tussen primair en secundair datagebruik beperkt zich echter niet tot generieke applicatiefuncties waarover VWS de regierol heeft. Ook functies zoals het datadefinitieregister, pseudonimisatie en ontdebelling komen bij zowel primair, als secundair hergebruik van data terug. Daarnaast worden ook binnen het secundair datagebruik applicatiefuncties onderkend, die bijeen generieke opzet, substantieel kunnen bijdragen aan het verminderen van versnippering en het bevorderen van interoperabiliteit en herbruikbaarheid van data. De volgende generieke applicatiefuncties onderkend:

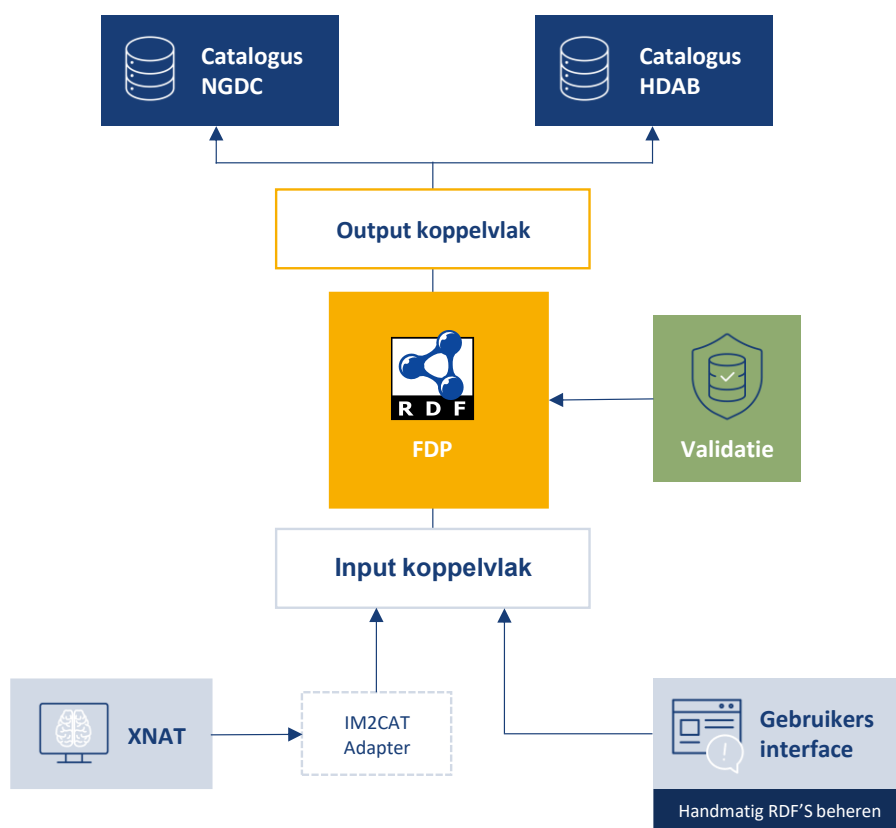
- datadefinitieregister
- datagebruikerregister
- datahouderregister
- terminologie services
- data / file exchange
- record koppeling / ontdebelling
- generieke transformatie
- Cumuluz adapter

Binnen de context van Health-RI werken diverse organisatie momenteel samen om invulling te geven aan een deel van deze generieke applicatiefuncties.

4.4.5. FAIR Data Point

Het FAIR Data Point (FDP) vormt de kernvoorziening voor het beheren, publiceren, valideren en uitwisselen van machine-leesbare metadata binnen de nationale gezondheidsdata-infrastructuur secundair gebruik. Door metadata vast te leggen in RDF-formaat en te beschrijven volgens gestandaardiseerde metadata-schema's, zoals DCAT-AP en daarop gebaseerde community- of domeinprofielen, ontstaat een uniforme en semantisch interoperabele beschrijving van databronnen. Hierdoor kunnen zowel mensen als systemen data op een consistente manier vinden, begrijpen, beoordelen en geautomatiseerd benutten. Het FDP draagt daarmee bij aan interoperabiliteit, herbruikbaarheid en schaalbare samenwerking tussen organisaties, zonder dat hiervoor centrale opslag of centrale regie op metadata noodzakelijk is.

Het beheer van metadata blijft bij de datahouder. Daarmee sluit het FDP aan op het principe van datasoevereiniteit: iedere organisatie bepaalt zelf welke metadata wordt gepubliceerd, welke kwaliteitseisen worden gehanteerd en onder welke voorwaarden informatie beschikbaar wordt gesteld. Doordat metadata inzicht biedt in aspecten zoals herkomst, kwaliteit, licenties, gebruiksvoorwaarden en toepasselijke regelgeving, ondersteunt het FDP niet alleen de vindbaarheid van databronnen, maar ook transparantie, vertrouwen en weloverwogen besluitvorming over het gebruik van data.



Figuur 11: Schematische weergave FDP

Zoals weergegeven in Figuur 11 ondersteunt de voorziening de gehele metadata-levenscyclus. Via een gestandaardiseerd inputkoppelvlak kan metadata handmatig worden beheerd met behulp van een gebruikersinterface of geautomatiseerd worden aangeleverd vanuit bestaande voorzieningen, zoals datarepositories, registraties, databroncatalogi en dataplatformen. Indien nodig kunnen adapters worden toegepast om bestaande systemen aan te sluiten zonder dat de onderliggende inrichting hoeft te worden aangepast. Hiermee wordt de drempel voor deelname aan een dataspace verlaagd en kunnen bestaande investeringen optimaal worden benut.

Onderdeel van de capability is de mogelijkheid om metadata te toetsen aan de binnen een community of dataspace afgesproken standaarden. Door validatie tijdens het publiceren of actualiseren van metadata krijgen datahouders inzicht in de mate waarin beschrijvingen voldoen aan de overeengekomen informatiemodellen, vocabularia en kwaliteitsregels. Dit ondersteunt een consistente toepassing van standaarden en draagt bij aan de interoperabiliteit tussen organisaties en voorzieningen.

Via gestandaardiseerde raadpleeg- en harvestkoppelvlakken kunnen catalogi, andere FDP's en metadata-voorzieningen metadata automatisch ophalen, verwerken en hergebruiken. Hierdoor ontstaat een federatief netwerk van samenwerkende metadata-services waarin metadata éénmaal door de bronhouder wordt beheerd en vervolgens door meerdere catalogi, dataspace's en andere voorzieningen kan worden gebruikt. Een datahouder kan daardoor gelijktijdig deelnemen aan verschillende nationale, Europese of domein specifieke ecosystemen zonder dezelfde metadata op meerdere plaatsen afzonderlijk te hoeven registreren of onderhouden.

Deze federatieve opzet ondersteunt tevens een gecascadeerd model waarin FDP's, catalogi en andere metadata-voorzieningen onderling metadata kunnen uitwisselen. Niet alleen gebruikersinterfaces, maar ook andere metadata-services kunnen via dezelfde gestandaardiseerde koppelvlakken worden aangesloten. Hierdoor ontstaat een flexibel en schaalbaar netwerk waarin

metadata over organisatiegrenzen heen kan worden gepubliceerd, beheerd, geraadpleegd en doorgegeven. Nieuwe catalogi, dataspace of communities kunnen op deze wijze aansluiten zonder dat bestaande voorzieningen fundamenteel hoeven te worden aangepast.

De capability is daarbij metadata-schema-agnostisch. Communities, dataspace en toepassingsdomeinen bepalen zelf welke metadata wordt gestandaardiseerd en welke aanvullende beschrijvingen nodig zijn voor hun specifieke context. Naast een generieke kernset kunnen daardoor ook domein specifieke uitbreidingen worden toegepast. Deze combinatie van standaardisatie en flexibiliteit maakt het mogelijk om interoperabiliteit op nationaal niveau te realiseren, terwijl tegelijkertijd ruimte blijft bestaan voor de informatiebehoeften van individuele domeinen, onderzoeksgebieden en samenwerkingsverbanden.

4.5. Organisatorische architectuur

De organisatorische architectuur legt uit hoe de Nationale gezondheidsdata-infrastructuur secundair gebruik wordt aangestuurd. Het beschrijft de organisatie van regie, besluitvorming, vertrouwen, toezicht, standaardisatie en samenwerking. Daarbij wordt benoemd welke partijen verantwoordelijk zijn, op welke manier afspraken tot stand komen en worden nageleefd, en hoe er gezamenlijk wordt gewerkt aan een betrouwbare, duurzame en goed samenwerkende infrastructuur voor het delen en opnieuw gebruiken van gezondheidsdata.

4.5.1. Uitgangspunten voor organisatie van verantwoordelijkheden

De organisatorische architectuur van de Nationale gezondheidsdata-infrastructuur secundair gebruik is gebaseerd op een aantal leidende uitgangspunten die richting geven aan de inrichting van verantwoordelijkheden, besluitvorming en samenwerking binnen het ecosysteem.

- **Ecosysteem-verantwoordelijkheid;** De Nationale gezondheidsdata-infrastructuur secundair gebruik wordt gerealiseerd binnen een gedistribueerd ecosysteem. Verantwoordelijkheden zijn verdeeld over meerdere partijen, waarbij samenwerking en onderlinge afstemming essentieel zijn.
- **Scheiding tussen richting en implementatie;** De architectuur stuurt op eenduidigheid in functionaliteit, standaarden en koppelvlakken waar uitwisseling plaatsvindt, terwijl ruimte blijft voor variatie in de interne implementatie. Implementatierichtlijnen ondersteunen datahouders en leveranciers bij de concrete invulling.
- **Meervoudige positionering van componenten;** Architectuurcomponenten kunnen afhankelijk van context centraal, regionaal of lokaal worden gerealiseerd. Keuzes hierin worden beïnvloed door onder andere schaalvoordelen, kosten, beschikbaarheid en privacyoverwegingen.
- **Federatief waar mogelijk, centraal waar nodig;** Organisaties behouden zoveel mogelijk hun eigen verantwoordelijkheid en autonomie, terwijl gezamenlijke voorzieningen, standaarden en afspraken centraal worden georganiseerd wanneer dit nodig is voor samenhang, interoperabiliteit of vertrouwen.
- **Verantwoordelijkheid zo dicht mogelijk bij de bron**
Verantwoordelijkheden voor data, metadata en dienstverlening liggen primair bij de partijen die deze beheren of produceren. Hierdoor blijven eigenaarschap, kwaliteit en accountability duidelijk belegd.
- **Scheiding tussen uitvoering, toezicht en besluitvorming**
Rollen voor operationele uitvoering, bestuurlijke besluitvorming en onafhankelijk toezicht zijn van elkaar gescheiden om transparantie, objectiviteit en zorgvuldige afweging van belangen te waarborgen.
- **Transparantie en verantwoording als uitgangspunt**
Besluiten, afspraken, rollen en prestaties zijn inzichtelijk voor betrokken partijen. Organisaties leggen verantwoording af over de naleving van afspraken en de kwaliteit van hun dienstverlening.

• **Europese interoperabiliteit als randvoorwaarde**

De Nationale gezondheidsdata-infrastructuur secundair gebruik sluit aan op Europese wet- en regelgeving, standaarden en infrastructuren, zodat nationale voorzieningen kunnen samenwerken binnen initiatieven zoals de EHDS.

• **Uniforme afspraken, decentrale implementatie**

Binnen de Nationale gezondheidsdata-infrastructuur secundair gebruik gelden gemeenschappelijke afspraken over processen, standaarden, beveiliging en organisatie van verantwoordelijkheden. De implementatie hiervan vindt plaats binnen de eigen context en systemen van deelnemende organisaties.

4.5.2. Ecosysteem

De Nationale gezondheidsdata-infrastructuur secundair gebruik bestaat uit een samenhangend geheel van processen, applicaties en functies die functioneel, technisch, organisatorisch en juridisch lokaal, via knooppunten, landelijk of Europees kunnen worden gepositioneerd. Zo kan bijvoorbeeld uit kostenoverwegingen een applicatie centraal (al dan niet op neutrale grond) worden ingericht vanuit een shared service principe of kan uit privacyoverwegingen gekozen worden voor een lokale / regionale applicatie.

Er wordt in de doelarchitectuur dan ook niet uitgegaan van één vaste organisatorische inrichting, maar van een ecosysteem waarin verschillende partijen vanuit hun eigen rol bijdragen aan databeschikbaarheid en hergebruik. Applicaties en functies kunnen daardoor afhankelijk van de situatie lokaal bij een datahouder, regionaal binnen een samenwerkingsverband, landelijk binnen een voorziening of als onderdeel van een Europese infrastructuur worden gerealiseerd. De gekozen positionering kan in de tijd veranderen onder invloed van technologische ontwikkelingen, organisatorische keuzes, nieuwe wetgeving of Europese ontwikkelingen zoals de EHDS en HealthData@EU infrastructuur.

Stakeholdersamenwerking

De ontwikkeling van Nationale gezondheidsdata-infrastructuur secundair gebruik vraagt om actieve betrokkenheid van de verschillende belanghebbenden binnen het ecosysteem. Daarbij wordt gestreefd naar een balans tussen de autonomie van individuele partijen en de noodzaak tot landelijke en Europese interoperabiliteit, zodat gezondheidsdata veilig, verantwoord en schaalbaar kunnen worden hergebruikt. Om belangen, expertise en maatschappelijke perspectieven structureel te verbinden, worden belanghebbenden georganiseerd in stakeholdergroepen en samenwerkingsverbanden die bijdragen aan de verdere ontwikkeling van de Nationale gezondheidsdata-infrastructuur secundair gebruik. Hierin worden onder meer de perspectieven van veldpartijen, burgers, onderzoeksinstellingen, overheden en leveranciers vertegenwoordigd. Deze stakeholdergroepen dragen bij aan het formuleren van gezamenlijke uitgangspunten, het afstemmen van belangen, het stimuleren van adoptie en het creëren van draagvlak voor landelijke en Europese ontwikkelingen.

Regionale en domeinknooppunten

Een bijzondere vorm van stakeholdersamenwerking binnen de Nationale gezondheidsdata-infrastructuur secundair gebruik wordt gevormd door de regionale en domeinspecifieke knooppunten. Deze knooppunten brengen verschillende actoren en stakeholdergroepen bijeen rondom een gedeelde inhoudelijke, regionale of organisatorische context. Zij vervullen daarmee een verbindende rol tussen de Nationale gezondheidsdata-infrastructuur secundair gebruik en de dagelijkse praktijk van databeschikbaarheid en hergebruik.

Regionale knooppunten richten zich op samenwerking tussen partijen binnen een geografische regio. Zij ondersteunen afstemming tussen datahouders, datagebruikers, leveranciers en andere

betrokken partijen en stimuleren de implementatie van landelijke afspraken en voorzieningen binnen de regionale context.

De regionale knooppunten zijn:

- Amsterdam (vanuit Amsterdam UMC)
- Eindhoven (vanuit TU Eindhoven)
- Groningen (vanuit UMCG)
- Leiden (vanuit LUMC)
- Maastricht (vanuit MUMC+)
- Nijmegen (vanuit Radboud UMC)
- Rotterdam (vanuit Erasmus MC)
- Utrecht (Vanuit UMCU)



Figuur 12. Geografische weergave regionale knooppunten

Domeinknooppunten richten zich op samenwerking binnen een specifiek zorg- of onderzoeksdomein. Hier worden actoren samengebracht die een gezamenlijke inhoudelijke opgave delen, bijvoorbeeld rondom een aandoening, zorgproces of onderzoeksgebied. Domeinknooppunten dragen bij aan harmonisatie van gegevens, afstemming van behoeften en het ontwikkelen van gezamenlijk toepasbare afspraken en voorzieningen. De domeinknooppunten waarmee wordt samengewerkt zijn:

- Oncologie
- Cardiovasculair
- Zeldzame ziekten

Regionale en domeinknooppunten vormen daarmee een belangrijk mechanisme voor kennisdeling, afstemming en adoptie binnen het ecosysteem. De concrete organisatorische inrichting van deze knooppunten kan in de tijd veranderen en valt buiten de scope van deze doelarchitectuur.

4.5.3. Verantwoordelijkheden

In deze paragraaf worden de verantwoordelijkheden van de verschillende partijen binnen het ecosysteem nader toegelicht. De volgende partijen worden daarbij onderkend:

- burgers;
- datahouders;
- regievoerders, zijnde:
 - knooppunten;
 - landelijke partijen;
 - Europese partijen;
- datagebruikers.

Burgers

De burger is als subject van gezondheidsdata en rechthebbende ten aanzien van het gebruik daarvan allereerst belanghebbende bij privacy, transparantie, zeggenschap en maatschappelijk verantwoord gebruik van data. Daarnaast vormen toepassingen (apps, wearables, etc.) die de burger gebruikt zelf een bron van gezondheidsdata. Binnen de doelarchitectuur worden deze toepassingen echter als bronsystemen beschouwd, zonder dat de burger verplicht is om de data uit deze systemen beschikbaar te stellen en herbruikbaar te maken.

Datahouders

De datahouder is primair verantwoordelijk voor:

- het beschikbaar stellen en herbruikbaar maken van brondata en metadata, conform de afspraken in het afsprakenstelsel. De datagebruiker geeft hierbij aan welke BVO de brondata beschikbaar gesteld dient te worden;
- het waarborgen van de kwaliteit, herleidbaarheid en rechtmatigheid van deze data;
- De datahouder kan ervoor kiezen om (delen van) deze verantwoordelijkheid in te vullen via externe of gedeelde (regionale) voorzieningen, mits wordt voldaan aan de geldende afspraken en vereisten.

Knooppunten

Een knooppunt (regio, gezondheidssector, groep samenwerkende partijen, ...), dat samenwerking, afstemming en interoperabiliteit binnen een regio of domein ondersteunt kan een faciliterende rol vervullen voor datahouders door

- applicaties aan te bieden voor het herbruikbaar maken van brondata
- het beschikbaar stellen van (gedeelde) BVO;
- het realiseren en beheren van componenten voor toegang tot data binnen de context.

De keuze om dergelijke voorzieningen via een knooppunt te organiseren is contextafhankelijk en dient aan te sluiten bij de architectuurprincipes en landelijke afspraken.

Landelijke partijen

Op landelijk niveau dragen ook diverse partijen zorg voor het realiseren en beheren van voorzieningen:

- het realiseren en beheren van de nationale gezondheidsdata catalogus en voorzieningen voor dataverkenning, ondergebracht bij Health-RI en de HDAB;
- het realiseren en beheren van componenten voor toegang tot data, ondergebracht bij Health-RI en de HDAB;
- het aanbieden van aanvullende diensten aan datahouders en datagebruikers, zoals het faciliteren van centrale beveiligde verwerkingsomgeving; ondergebracht bij bijvoorbeeld Health-RI en de HDAB;
- het faciliteren van federatieve analyse, gegevensconsolidatie en andere landelijke functies die niet doelmatig lokaal of via knooppunten kunnen worden georganiseerd;
- het ondersteunen van aansluiting op Europese voorzieningen zoals HealthData@EU, ondergebracht bij de HDAB en het Nationaal Contact Punt eHealth).
- het realiseren en beheren van generieke landelijke voorzieningen die door meerdere partijen binnen de Nationale gezondheidsdata-infrastructuur secundair gebruik worden gebruikt, zoals identificatie-, authenticatie-, autorisatie-, registratie- en vertrouwensdiensten, onder regie van het ministerie van VWS en waar mogelijk aangesloten op bestaande nationale voorzieningen

Europese partijen

Europese voorzieningen vervullen een verbindende rol voor grensoverschrijdend gebruik van gezondheidsdata binnen de Europese Unie. Deze voorzieningen ondersteunen de interoperabiliteit en samenwerking tussen nationale infrastructuren, waaronder de Nationale gezondheidsdata-infrastructuur secundair gebruik. Hieronder vallen onder andere:

- het faciliteren van grensoverschrijdende vindbaarheid en toegang tot gezondheidsdata; ondergebracht bij de HDAB;
- het bieden van gemeenschappelijke technische, semantische en organisatorische voorzieningen voor gegevensuitwisseling;
- het ondersteunen van samenwerking tussen nationale HDAB's en andere bevoegde instanties;
- het bevorderen van interoperabiliteit tussen Nationale gezondheidsdata-infrastructuren via Europese standaarden, specificaties en afspraken.

De inrichting en doorontwikkeling van deze voorzieningen vindt plaats binnen Europese initiatieven zoals de EHDS en HealthData@EU. Nationale voorzieningen sluiten hierop aan waar dit nodig is voor interoperabiliteit en grensoverschrijdend gebruik van gezondheidsdata.

1739

1740 **Datagebruikers**

1741 De datagebruiker is primair verantwoordelijk voor:

- 1742 • het opstellen van de onderzoeksvraag;
- 1743 • het indienen van een dataverzoek;
- 1744 • het inrichten en beheren van een beveiligde verwerkingsomgeving of het borgen dat een
- 1745 beveiligde verwerkingsomgeving beschikbaar is voor het uitvoeren van (federatieve)
- 1746 analyses conform de geldende afspraken en voorwaarden. Wanneer de datahouder dit toestaat,
- 1747 kan een datagebruiker ervoor kiezen om data te laten landen in een eigen (lokale) beveiligde
- 1748 verwerkingsomgeving conform de geldende afspraken en voorwaarden;
- 1749 • het uitvoeren van het onderzoek;
- 1750 • het publiceren van het onderzoeksresultaat;
- 1751 • het communiceren van eventuele nevenbevindingen aan de HDAB.
- 1752

5. Beheer doelarchitectuur

De doelarchitectuur is een levend document. Om samenhang, interoperabiliteit en aansluiting op nieuwe inzichten, ontwikkelingen en wet- en regelgeving te borgen, zijn afspraken nodig over het beheer en de doorontwikkeling van de doelarchitectuur.

5.1. Versiebeheer architectuur

De doelarchitectuur is een levend document dat meegroeit met nieuwe inzichten, ontwikkelingen en ervaringen uit de praktijk. Om de kwaliteit, samenhang en toepasbaarheid te borgen, wordt gewerkt met formeel versiebeheer. Wijzigingen worden transparant vastgelegd, zodat inzichtelijk blijft welke onderdelen zijn aangepast en op basis van welke overwegingen.

5.2. Wijzigingsproces architectuur

Nieuwe inzichten, architectuurvraagstukken en wijzigingsvoorstellen worden uitgewerkt in samenwerking met inhoudelijke experts uit het ecosysteem. Op deze wijze wordt gekomen tot een consistente en breed gedragen vastlegging van nieuwe inzichten. Voorgestelde wijzigingen worden beoordeeld en vastgesteld door de architectuurboard en vervolgens verwerkt in de doelarchitectuur.

5.3. Consultatiecyclus

De door experts uitgewerkte en door de architectuurboard (zie hierna) vastgestelde aanvullingen en wijzigingen worden meegenomen in een nieuwe versie van de doelarchitectuur. De doelarchitectuur wordt jaarlijks breed en open geconsulteerd binnen het Nationale gezondheidsdata-infrastructuur secundair gebruik-ecosysteem. Tijdens een openbare consultatieronde in het eerste kwartaal worden stakeholders uitgenodigd om feedback te geven op de actuele versie van de architectuur.

5.4. Architectuurboard

De architectuurboard is verantwoordelijk voor het bewaken van de samenhang, kwaliteit en doorontwikkeling van de doelarchitectuur. Momenteel bestaat de board uit de architectuurcoördinatoren van de regionale knooppunten en een vertegenwoordiging vanuit Surf en CumuluZ.

5.5. Conformance op architectuur

De doelarchitectuur biedt een gemeenschappelijk referentiekader voor voorzieningen, diensten en implementaties binnen de Nationale gezondheidsdata-infrastructuur secundair gebruik. Conformance richt zich op de mate waarin oplossingen, projecten en voorzieningen aansluiten bij de vastgestelde architectuurprincipes, afspraken en doelstellingen. Het doel hiervan is niet om uniformiteit af te dwingen, maar om interoperabiliteit, samenhang en herbruikbaarheid binnen het ecosysteem te bevorderen. Afwijkingen zijn mogelijk, mits deze bewust worden afgewogen, onderbouwd en transparant worden vastgelegd.

Health-RI heeft in afstemming met regionale knooppunten de volgende verbingsprincipes opgesteld om bestaande nieuwe consortia te helpen aansluiten op de Nationale gezondheidsdata-infrastructuur secundair gebruik in Nederland⁵:

1. Consortia committeren zich expliciet om hun infrastructuur te koppelen aan Nationale gezondheidsdata-infrastructuur secundair gebruik en dragen hun kennis en ervaring bij om deze verder te ontwikkelen en te verbeteren. De regionale Health-RI-knooppuntteams kunnen, in afstemming met het centrale Health-RI-team, adviseren hoe dit het beste kan worden georganiseerd.
2. Consortia stemmen hun infrastructuurcomponenten af op de Nationale gezondheidsdata-infrastructuur secundair gebruik.

⁵ Meer informatie over deze principes is te vinden op [Connecting principles for Consortia | Health-RI](#).

- 1803 3. Consortia verbinden zich ertoe nationaal aanbevolen en internationaal erkende (meta)data-
1804 standaarden en/of gemeenschappelijke datamodellen te gebruiken (zie de verwijzingen naar
1805 veelgebruikte standaarden op de Health-RI-website).
- 1806 4. Consortia zorgen ervoor dat hun databronnen vindbaar zijn via de Health-RI-datacatalogus,
1807 volgens de (meta)data-onboardingprocedure zoals gepubliceerd op de Health-RI-website. Dit
1808 helpt bij de voorbereiding op toekomstige eisen vanuit de EHDS.
- 1809 5. Consortia werken expliciet aan het beschikbaar maken van hun middelen (zoals data en
1810 biomaterialen) binnen hun project, met duidelijke toegangsvoorwaarden die aansluiten bij de
1811 algemene richtlijnen van Health-RI en de principes van Open Science en de EHDS.
- 1812 6. Consortia voor infrastructuur wordt geadviseerd om aan te geven hoe hun infrastructuur
1813 aansluit bij bestaande Europese onderzoeksinfrastructuren.
1814

6. Roadmap

Onderzoek, beleid en innovatie maken reeds structureel gebruik van gezondheidsdata. De ambitie van Health-RI is niet om deze toepassingen fundamenteel te vervangen, maar om de bestaande praktijk gefaseerd te verbeteren, zodat zij effectiever, veiliger, meer schaalbaar en betaalbaarder worden. Dit vraagt om een benadering waarin architectuur richtinggevend is, zonder los te raken van de realiteit van gezondheidszorg- en onderzoekspraktijk. Vanuit architectuurdenken bestaat de verleiding om toe te werken naar een technocratisch ideale eindsituatie. In de praktijk blijkt echter dat oplossingen die te ver afstaan van wat het veld op korte termijn kan en wil toepassen, onvoldoende adoptie kennen. Health-RI kiest daarom bewust voor implementatiescenario's die aansluiten bij de behoeften en het absorptievermogen van een substantieel deel van het veld, en die ruimte laten voor iteratieve doorontwikkeling.

Huidige situatie als vertrekpunt

De eerste stappen naar een Nationale gezondheidsdata-infrastructuur secundair gebruik zijn reeds gezet. In de huidige situatie ligt de focus primair op:

- het centraal vindbaar maken van onderzoeksdatasets;
- het aanvraagbaar maken van deze datasets;
- het kunnen analyseren van data binnen een beveiligde verwerkingsomgeving.

Hiermee is het vandaag de dag met name mogelijk om onderzoekdatasets via een catalogus te vinden en aan te vragen. Deze basis vormt het vertrekpunt voor verdere implementatie, niet een tijdelijk stadium dat volledig vervangen moet worden.

Uitbreiding richting gezondheidszorgdata

De eerstvolgende implementatiestap richt zich op het uitbreiden van de huidige functionaliteit naar data uit de gezondheidszorg zelf. Dit betekent dat:

- gezondheidszorgdatasets o.a. uit zorginformatiesystemen, naast onderzoeksdatasets, vindbaar worden in de catalogus;
- gebruikers de mogelijkheid krijgen om de bruikbaarheid en geschiktheid van data te verkennen, voorafgaand aan een aanvraag;
- data zowel in ruwe vorm als getransformeerd naar gestandaardiseerde codering en modellering beschikbaar kan worden gesteld, afhankelijk van de vraagcontext.

Deze stap vraagt nadrukkelijk om nauwe samenwerking met gezondheidszorgorganisaties, omdat de context van zorgregistratie, datakwaliteit en interpretatie bepalend is voor verantwoord hergebruik.

Uitbreiding beproeving in samenwerking met het veld

Health-RI hanteert een implementatiestrategie waarin samen ontwerpen en in de praktijk beproeven centraal staan. Nieuwe functionaliteit en afspraken worden eerst gezamenlijk uitgewerkt, en vervolgens getest in samenwerking met gezondheidszorgorganisaties die beschikken over de benodigde schaal, expertise en middelen. In eerste instantie ligt de focus daarbij op instellingen zoals UMC's en STZ-ziekenhuizen, die als knooppunten een regierol vervullen. Op basis van de daar opgedane ervaringen worden implementatiescenario's aangepast en vereenvoudigd, met als doel deze vervolgens:

- toepasbaar te maken voor kleinere ziekenhuizen;
- door te vertalen naar de huisartsenzorg, GGZ en langdurige zorg;
- en uiteindelijk ook te verkennen in het sociaal domein.

Naast de samenwerking met knooppunten, zal ook worden samengewerkt met (datagebruikers uit) kennisinstellingen die voorzieningen nodig hebben voor het zoeken van data, het aanvragen van data en het uitvoeren van onderzoek. Na beproeving in kleinschalig verband, zal via de knooppunten worden opgeschaald.

Op deze wijze ontstaat zowel voor de datagebruikers, als de datahouders, als de regievoerders een opschaalbaar en adaptief implementatiepad, waarin verschillen in complexiteit, volwassenheid en beschikbare middelen expliciet worden meegenomen.

Samenwerking nationaal

Vanuit VWS wordt gewerkt aan de architectuur van het landelijk gezondheidsinformatiestelsel (GISA). Deze architectuur beschrijft zowel het primair als het secundair gebruik van gezondheidsdata. Health-RI draagt hieraan bij door de architectuur voor secundair gebruik uit te werken, gebaseerd op de principes en keuzes uit de Nationale gezondheidsdata-infrastructuur secundair gebruik-doelarchitectuur.

CumuluZ richt zich op de realisatie van het Landelijk Dekkend Netwerk (LDN) voor databeschikbaarheid. De CumuluZ-doelarchitectuur beschrijft de inrichting van de infrastructuur voor gegevensuitwisseling en ondersteunt zowel primair als secundair gebruik van gezondheidsdata, met een nadruk op het primaire zorgproces. De architectuurkeuzes van CumuluZ worden tevens verwerkt binnen de GISA. Health-RI en CumuluZ stemmen hun doelarchitecturen nauw op elkaar af en werken samen aan generieke functies, standaarden en voorzieningen. Hiermee wordt toegewerkt naar een samenhangende landelijke infrastructuur voor databeschikbaarheid die zowel primair als secundair gebruik van gezondheidsdata ondersteunt. De komende jaren werken beide programma's samen met zorginstellingen aan de verdere realisatie van dit landelijk dekkende netwerk.

Voor een aantal generieke functies binnen het gezondheidsinformatiestelsel heeft het ministerie van VWS de regie genomen op de ontwikkeling van generieke voorzieningen. Het betreft onder meer functies zoals identificatie, authenticatie, autorisatie, adressering, zeggenschap, lokalisatie van data, logging en pseudonimisatie. Deze voorzieningen worden momenteel primair ontwikkeld vanuit de behoeften van het primaire gebruik van gezondheidsdata, maar zijn tegelijkertijd van groot belang voor secundair gebruik van gezondheidsdata.

Op basis van deze doelarchitectuur voor de Nationale gezondheidsdata-infrastructuur secundair gebruik zal de komende periode een nadere softwarearchitectuur worden uitgewerkt voor de generieke applicatiefuncties die relevant zijn voor secundair gebruik. Deze uitwerking beschrijft de functionele en technische vereisten die voortvloeien uit onderzoek, beleid, innovatie en de EHDS. Daarmee wordt ook richting gegeven aan de verdere doorontwikkeling van de generieke voorzieningen onder regie van VWS. De softwarearchitectuur vormt hierbij de brug tussen de behoeften vanuit secundair gebruik en de landelijke implementatie van generieke voorzieningen.

Door deze functies zoveel mogelijk generiek en voor zowel primair als secundair gebruik toepasbaar in te richten, wordt versnippering voorkomen en wordt toegewerkt naar één samenhangend gezondheidsinformatiestelsel. Investerings in generieke voorzieningen dragen daarmee niet alleen bij aan de behoeften van vandaag, maar vormen tevens een duurzaam fundament voor toekomstige nationale en Europese ontwikkelingen, waaronder de EHDS en de HealthData@EU infrastructuur.

EHDS

De Nationale gezondheidsdata-infrastructuur secundair gebruik wordt in de komende jaren verder ontwikkeld in samenhang met de EHDS, de HealthData@EU infrastructuur en de inrichting van de HDAB-NL. Daarbij is het uitgangspunt dat de vanuit het Nationaal Groeifonds gerealiseerde voorzieningen geen tijdelijke tussenoplossing vormen, maar een herbruikbaar fundament voor de toekomstige EHDS/HDAB-situatie. De roadmap richt zich daarom niet op een abrupte vervanging van de huidige inrichting, maar op een geleidelijke overgang waarin bestaande functies, voorzieningen, implementaties en samenwerkingsafspraken zoveel mogelijk worden behouden en doorontwikkeld.

Een laagdrempelige overgang naar de EHDS staat hierbij centraal. Dit betekent dat datagebruikers, datahouders en regie voerende partijen zo veel mogelijk met herkenbare functies blijven werken, zoals catalogus, toegangsservice en beveiligde verwerkingsomgevingen, terwijl de onderliggende verantwoordelijkheden, beoordelingsprocessen en juridische grondslagen kunnen wijzigen. Waar datahouders in de huidige situatie nog een belangrijke rol hebben in aanvraagbeoordeling, autorisatie en outputcontrole, verschuift deze verantwoordelijkheid in de EHDS/HDAB-situatie voor aanvragen die onder de EHDS vallen in belangrijke mate naar de HDAB.

De komende periode wordt de nationale infrastructuur stapsgewijs geschikt gemaakt voor aansluiting op de HealthData@EU infrastructuur en voor uitvoering binnen het HDAB-kader. Dit vraagt onder meer om verdere standaardisatie van metadata, aanvraagprocessen, autorisatie, registratie van gebruik, controle op datagebruik en vrijgave van resultaten vanuit beveiligde verwerkingsomgevingen. Tegelijkertijd blijft ruimte nodig voor nationale keuzes en bestaande onderzoekspraktijken, omdat de EHDS en de nationale invulling daarvan nog verdere uitwerking en aanscherping vragen. De roadmap kiest daarom voor continuïteit over organisatie van verantwoordelijkheid-regimes heen: functies en voorzieningen worden zo ingericht dat zij bruikbaar blijven vóór, tijdens en na de operationalisatie van de HDAB.

Op deze manier ontstaat een groeipad waarin investeringen maximaal worden hergebruikt, implementatielasten voor het veld worden beperkt en stap voor stap wordt toegewerkt naar een nationale infrastructuur die als volwaardig onderdeel van het Europese HealthData@EU ecosysteem kan functioneren. De voorzieningen, standaarden en werkwijzen die in dit groeipad worden gerealiseerd, zijn nadrukkelijk gekozen vanuit de beoogde langetermijnsituatie. Investerings die datahouders vandaag doen in aansluiting op de Nationale gezondheidsdata-infrastructuur secundair gebruik dragen daarmee direct bij aan hun voorbereiding op de EHDS en bouwen voort op eenzelfde architectonische richting. Dit zorgt ervoor dat elke stap niet alleen waarde toevoegt op de korte termijn, maar tegelijkertijd een bouwsteen vormt voor de toekomstige Europese situatie.

Hybride groeimodel

De Nationale gezondheidsdata-infrastructuur secundair gebruik vraagt om een hybride model, waarin zowel ruwe brondata als gestandaardiseerde en geharmoniseerde data beschikbaar zijn voor secundair hergebruik van gezondheidsdata. Beide voorzien in een verschillende behoefte. Ruwe data zijn nodig voor onderzoek, beleid en innovatie waarbij maximale detailinformatie of nieuwe analysemethoden, zoals AI, centraal staan. Transformatie van de ruwe data naar gestandaardiseerde data kan namelijk leiden tot verlies van detailinformatie die voor specifieke onderzoeksvragen relevant is. Gestandaardiseerde data maken daarentegen grootschalig en grensoverschrijdend onderzoek eenvoudiger door betere interoperabiliteit en vergelijkbaarheid van datasets.

Naarmate meer gezondheidszorgdata volgens internationale standaarden wordt vastgelegd en transformatietools zich verder ontwikkelen, zal het aandeel gestandaardiseerde data toenemen. Voor lang niet alle data bestaan echter (internationale) standaarden en het is een utopie dat ooit alle data gestandaardiseerd zullen zijn. Ook ontbreken veelal betrouwbare transformatietools, die nodig zijn om ruwe brondata om te zetten naar gestandaardiseerde data. Het hybride model maakt geleidelijke ontwikkeling naar meer standaardisatie mogelijk, zonder de huidige en ook in de toekomst bestaand blijvende onderzoekspraktijk te beperken.

Toevoegen context

Gezondheidsdata zonder context hebben vaak onvoldoende betekenis voor secundair gebruik. Een dataset kan technisch compleet lijken, terwijl belangrijke informatie over de omstandigheden waaronder een gegeven is ontstaan ontbreekt. Daardoor kan een onderzoeker het gegeven niet goed interpreteren of verkeerd gebruiken. Hoewel de context veelal essentieel is, ontbreekt in veel systemen de mogelijkheid om deze gestructureerd vast te leggen. Gevolg is dat de context ook niet

1970 digitaal beschikbaar worden gesteld voor hergebruik. Aanpassing van systemen om context
 1971 (automatische) vast te leggen, is nodig om dit op te lossen. Het gesprek hierover met leveranciers
 1972 van deze systemen zal de komende periode worden opstart.

1973

1974 **Federatieve analyse als ontwikkelpad**

1975 Federatieve analyse wordt gezien als een veelbelovende privacy-enhanced techniek voor secundair
 1976 gebruik van gezondheidsdata. Tegelijkertijd is onderkend dat deze benadering verdere verkenning
 1977 en beproeving vergt, zowel technisch als organisatorisch. Zo is gestandaardiseerde en
 1978 geharmoniseerde data een voorwaarde voor federatieve analyse. Echter, veel data voldoen
 1979 momenteel nog niet aan deze voorwaarde. Ook laat de kwaliteit van de data te wensen over en
 1980 wordt federatieve analyse bemoeilijkt doordat verschillende organisaties verschillende typen
 1981 gegevens over dezelfde personen beheren (verticale partitionering). In de implementatiescenario's
 1982 wordt federatieve analyse daarom niet als randvoorwaarde, maar als ontwikkelpad gepositioneerd,
 1983 waarin ervaring wordt opgedaan via gerichte pilots en concrete gebruikssituaties.

1984

1985 **AI als katalysator, niet als vertrekpunt**

1986 AI-toepassingen vragen om grote hoeveelheden goed toegankelijke en kwalitatief betrouwbare
 1987 data. Tegelijkertijd laten zij snel de potentiële waarde zien voor gezondheidszorgorganisaties die
 1988 moeten voorzien in de databeschikbaarheid, waardoor het belang van secundair datagebruik voor
 1989 datahouders vaak nadrukkelijker wordt ervaren. AI wordt daarom gezien als katalysator, niet als
 1990 startpunt.

1991

1992 **Gefaseerde introductie van generieke functies**

1993 Parallel aan de uitbreiding van databeschikbaarheid worden functies geïntroduceerd die generiek
 1994 inzetbaar zijn voor secundair datagebruik, ongeacht domein of use case. Het gaat hierbij onder meer
 1995 om:

- 1996 • pseudonimisering;
- 1997 • toepassing van opt-out en dataminimalisatie;
- 1998 • identificatie, authenticatie en autorisatie;
- 1999 • adressering;
- 2000 • logging.

2001 De uitwerking van deze functies geschiedt in nauwe samenwerking met VWS, Nictiz en Cumuluz.

2002

Bijlage 1. Begrippenlijst

Binnen het gezondheidszorginformatielandschap wordt veel terminologie gebruikt, waarbij vaak de aanname gedaan wordt dat men hetzelfde verstaat onder een begrip. Vaak blijken er toch interpretatieverschillen te bestaan. Momenteel gebruikt iedereen zijn of haar eigen definitielijst bij het opstellen van rapporten, contracten en afsprakenstelsels. Dit leidt tot inconsistenties en verwarring, aangezien termen vaak verschillend worden geïnterpreteerd, afhankelijk van de context of het initiatief.

In de hiernavolgende begrippenlijst worden ten behoeve van de leesbaarheid en eenduidige interpretatie van dit document de belangrijkste en meest gebruikte begrippen uit het Health-RI afsprakenstelsel (versie 5) herhaald. Aan deze lijst zijn diverse nieuwe begrippen toegevoegd. Waar nodig zijn definities aangescherpt of geactualiseerd op basis van voortschrijdend inzicht en de positionering ten opzichte van recente en lopende initiatieven, zoals de GISA en de technische vereisten uit THDAS2 en de EHDS-verordening.

Om beter te borgen dat we in Nederland hetzelfde verstaan onder relevante begrippen rondom secundair gebruik van gezondheidsdata, heeft Health-RI een begrippenlijst met breed gedragen begrippen en definities opgesteld. Deze begrippen zullen stapsgewijs worden overgenomen in de [Online Thesaurus Zorg en Welzijn](#).

BEGRIJF	DUIDING
Aanvraagregister	Data-aanvraag- en -uitgifte register, waarin de data-aanvraagdienst de aanvragen beheert.
Administratieve gezondheidszorgdata	Data gegenereerd voor administratieve of factureringsdoeleinden
Administratieve metadata	Data gegenereerd voor beheer en autorisatie
Afsprakenstelsel	Een afsprakenstelsel is een samenhangend geheel van juridische, organisatorische, semantische en technische afspraken, inclusief afspraken over besturing en beheer. Het beschrijft hoe partijen onderling op een veilige, betrouwbare en doelmatige manier data delen.
Anonimiseren	Bewerking van persoonsgegevens die er onomkeerbaar voor zorgt dat de data niet herleidbaar zijn tot een persoon. Zie ook: Online Thesaurus Zorg en Welzijn
Application Programming Interface (API)	Een koppeling tussen verschillende ICT-systemen of softwareprogramma's waarmee data van het ene systeem realtime en zonder tussenkomst van mensen overgebracht kunnen worden naar het andere systeem Bron: Online Thesaurus Zorg en Welzijn
Beschrijvende metadata	Metadata voor identificatie en zoeken (bijv. geslacht van patiënt, type onderzoek)
Betaaldienst	Een dienst die een bedrag kan innen bij een datagebruiker, afhankelijk van afgenomen diensten/datasets
Beveiligde Verwerkings Omgeving (BVO)	De fysieke of virtuele omgeving en organisatorische middelen om te zorgen voor de naleving van het Unierecht, zoals Verordening (EU) 2016/679, met name wat betreft de rechten van datasubjecten, intellectuele-eigendomsrechten, en handels- en statistisch geheim, integriteit en toegankelijkheid, alsook van het toepasselijke nationale recht, en om de entiteit die de beveiligde verwerkingsomgeving biedt in staat te stellen alle gegevensverwerkingsactiviteiten te bepalen en er toezicht op te houden, met inbegrip van het tonen, opslaan, downloaden en exporteren van data en het berekenen van afgeleide data door middel van computer algoritmen Bron: Data Governance Act
Biobanking data	Data verzameld in biobanken die grote collecties biospecimens vertegenwoordigen die zijn gekoppeld aan persoonlijke en gezondheidsinformatie, gebruikt in gezondheids- en medisch onderzoek.

BEGRIJP	DUIDING
Bron	De term bron kan duiden op twee concepten, de plaats waar iets oorspronkelijk vandaan komt, of de plaats waar iets feitelijk verkregen kan worden. Dit onderscheid wordt gemaakt in respectievelijk de begrippen 'herkomst' en 'leveringsbron', zie ook in dit overzicht. Omwille van de leesbaarheid wordt in het GIS ook soms kortweg bron gebruikt, dan is het in de betekenis van leveringsbron. Bron: GIS-A
Bronhouder	De rechtspersoon die de dossierverantwoordelijkheid voor het gezondheidszorgdossier draagt. Bron: NORA, TZW 59379
Bronstelsysteem	De rol die een instantie van een Gezondheidszorginformatiesysteem heeft omdat die is aangesloten op het GIS via één GIS aansluitpunt en daarmee toegang geeft tot de gezondheidsdata die besloten is in het Dossier Zorgaanbieder Bron: GIS-A
Burger	Een individu dat regie voert over diens data in het kader van onderzoek, beleid en innovatie. Zie ook: Online Thesaurus Zorg en Welzijn
Catalogus van datasets	Een verzameling beschrijvingen van datasets die op systematische wijze is geordend en met inbegrip van een gebruikersgericht openbaar gedeelte, waarin informatie over de parameters van individuele datasets langs elektronische weg toegankelijk is via een onlineportaal. Bron: EHDS wetstekst
Centraal verwerken	Bij <i>centraal verwerken</i> wordt de verwerking van data van verschillende <i>datahouders</i> op één centrale plaats verwerkt. De beheerder van de omgeving is hierbij niet noodzakelijkerwijs verbonden aan hetzelfde instituut als de <i>datagebruiker</i> .
Centrale toetsingscommissie	<i>Generieke dienst</i> , zowel voor de gezondheidszorg zelf (primair gebruik) als voor onderzoek, beleid en innovatie (secundair datagebruik), verantwoordelijk voor • Opstellen en beheren generieke gebruiksvoorwaarden • Toetsen of een gewenste uit te voeren analyse voldoet aan de gebruiksvoorwaarden van de aangevraagde dataset(s) • Toetsen (bron overstijgende) data-aanvraag op herleidbaarheid persoonsgegevens en privacyrisico's • Bijhouden jurisprudentie Alternatieve naam: Centrale datatoegangscommissie Zie ook: Online Thesaurus Zorg en Welzijn
Classificatie-standaarden	Voorbeelden: ICD-10 (lijst van ziekten), ICF (classificatie van functioneren), ICPC (classificatiemethode voor eerstelijnszorg), DBC (diagnosebehandelcombinatie)
Clinical trial data	Data verzameld tijdens een RCT om onderzoeksvragen te beantwoorden en publicaties/rapporten te genereren.
Codeboek	Een afsprakenboek om te waarborgen dat alle deelnemers aan het onderzoek dezelfde soorten data op dezelfde wijze verwerken.
Cohort	Een groep personen die bepaalde kenmerken of eigenschappen delen en onderwerp zijn van een studie, dataset of biobank/collectie. Zie ook: Online Thesaurus Zorg en Welzijn
Coördinator gefedereerde analyse	De <i>coördinator gefedereerde analyse</i> zorgt voor de hulpmiddeling om (bij een gefedereerde analyse) de analysevraag in te dienen en zorgt ervoor dat de analysevraag voor de uitvoerders gefedereerde analyse beschikbaar komt en dat de resultaten van de gefedereerde analyse verzameld worden.
Data (ook wel gegevens genoemd)	Een herinterpreteerbare representatie van informatie in een geformaliseerde vorm, geschikt voor communicatie, interpretatie of verwerking door mensen of geautomatiseerde systemen. Bron: ISO 2382
Data Acces	Vorm van toegang tot data waarbij de datagebruiker binnen een beveiligde verwerkingsomgeving toegang krijgen tot data voor analyse.

BEGRIIP	DUIDING
Data Governance Commissie	DGC; commissie die zich richt op het optimaliseren van het dataverwerkingsproces van kwaliteitsregistraties in de medisch-specialistische zorg en het verbeteren van gestructureerde vastlegging aan de bron om zo de toegankelijkheid en beschikbaarheid van deze data te vergroten; de DGC toetst de kwaliteitsregistraties en dataverwerkers op naleving van de landelijke afspraken over de data governance; de DGC stuurt het Shared Service Center Data Governance aan en wordt door het SSC-DG ondersteund; in de DGC zitten vertegenwoordigers van zorgverleners, zorgaanbieders, zorgverzekeraars, burgers en patiënten Zie ook: Online Thesaurus Zorg en Welzijn
Data visiting	Vorm van toegang tot data waarbij een of meer analyses naar de data worden gebracht en uitsluitend het gecontroleerde resultaat aan de datagebruiker wordt getoond.
Databeschikbaarheid	Het beschikbaar, bereikbaar en bruikbaar hebben van vastgelegde gezondheidsdata van de specifieke informatiebehoefte. Dit ten behoeve van alle vormen van (her)gebruik, gezondheidszorgproces- en usecase-onafhankelijk met nadrukkelijke waarborgen voor burger-/patiëntveiligheid en privacy. Bron: NVS van VWS
Datacentrische benadering	Een manier van kijken waarbij de data centraal staat en niet het doel waar het voor is verzameld of wordt gebruikt.
Data-controller	Een data-controller (NL: verwerkingsverantwoordelijke) draagt de hoofdverantwoordelijkheid voor de beveiliging en het rechtmatige gebruik van data. Zie ook: 7.2 Toelichting verwerkingsverantwoordelijkheid
Datadictionary	Een verzameling namen, definities en kenmerken over data-elementen die worden gebruikt of vastgelegd in een database, informatiesysteem of onderdeel van een onderzoeksproject. Het beschrijft de betekenissen en doelen van data-elementen binnen de context van een project en biedt richtlijnen voor interpretatie, geaccepteerde betekenissen en representatie. Een data dictionary biedt ook metadata over data-elementen. De metadata in een data dictionary helpen bij het definiëren van de reikwijdte en kenmerken van data-elementen, evenals de regels voor hun gebruik en toepassing. Bron: Everything You Need to Know About a Data Dictionary
Datagebruiker (ook wel gebruiker van gezondheidsdata) GISA gebruikt hiervoor de term data-afnemer.	Een natuurlijke persoon of rechtspersoon, met inbegrip van instellingen, organen of instanties van de Unie waaraan rechtmatige toegang tot elektronische gezondheidsdata voor secundair datagebruik is verleend op grond van een datavergunning, een goedkeuring van een verzoek om gezondheidsdata of een toegangsgoedkeuring van een gemachtigde deelnemer binnen de HealthData@EU infrastructuur Zie ook: EHDS wetstekst Een datagebruiker kan bijvoorbeeld een onderzoeker zijn die data nodig heeft voor wetenschappelijk onderzoek, een overheidsinstantie die data nodig heeft voor het monitoren van de volksgezondheid of een zorgverlener die toegang nodig heeft tot gezondheidsdata voor een diagnose of behandeling.

BEGRIIP	DUIDING
Datahouder (ook wel houder van gezondheidsdata). GISA gebruikt hiervoor de term data-aanbieder.	Een natuurlijke of rechtspersoon, overheidsinstantie, agentschap of ander orgaan in de gezondheidszorg- of de gezondheidsector, met inbegrip van, indien noodzakelijk, vergoedingsdiensten, alsook elke natuurlijke of rechtspersoon die producten of diensten ontwikkelt die bestemd zijn voor de gezondheids-, de gezondheidszorg- of de gezondheidsector, die wellnessapps ontwikkelt of vervaardigt, die wetenschappelijk onderzoek verricht in verband met de gezondheidszorg- of de gezondheidsector of optreedt als sterfteregister, alsook instellingen, organen of instanties van de Unie die: i) het recht of de verplichting hebben om overeenkomstig het toepasselijke Unierecht of nationaal recht en in hun hoedanigheid van verwerkingsverantwoordelijke of medeverwerkingsverantwoordelijke persoonlijke elektronische gezondheidsdata te verwerken voor de verstrekking van gezondheidszorg of zorg of voor doelstellingen op het gebied van volksgezondheid, vergoedingen, onderzoek, innovatie, beleidsvorming, officiële statistieken of patiëntveiligheid of voor regelgevingsdoeleinden; ofwel ii) in staat zijn om niet-persoonsgebonden elektronische gezondheidsdata ter beschikking te stellen door controle op het technische ontwerp van een product en aanverwante diensten, met inbegrip van het registreren, verstrekken, beperken of uitwisselen van die data; Bron: EHDS wetstekst
Datakwaliteit	De mate waarin de elementen van elektronische gezondheidsdata geschikt zijn voor het beoogde primaire gebruik en secundaire gebruik ervan Zie ook: EHDS wetstekst
Dataminimalisatie	Dataminimalisatie wil zeggen dat een dataset niet meer data bevat dan noodzakelijk is voor het doel waarvoor de dataset is aangevraagd.
Datapunt	Data die in samenhang, in een afgeronde tijdperiode, is geregistreerd met een specifiek doel en betekenis. Het is de kleinste betekenisvolle eenheid van data. Bron: GIS-A
Dataset	Een <i>dataset</i> is een collectie van vergelijkbare data die betrekking hebben op een groep <i>datasubjecten</i> (of entiteiten). De collectie heeft een zekere homogeniteit en een omvang die past bij de vraagstukken waarvoor onderzoekers (en andere geïnteresseerden) data zoeken, vinden en aanvragen in de catalogus, zodat het zin heeft het als samenhangend geheel te zien, waaruit voor hergebruik kan worden geput. Een <i>dataset</i> kan statisch zijn, d.w.z. dat de <i>dataset</i> na oplevering niet meer verandert. Aan de andere kant kan een <i>dataset</i> ook dynamisch zijn: de <i>dataset</i> is dan aan verandering onderhevig en wordt aangevuld. In dat geval kan de lijst <i>datasubjecten</i> (of entiteiten) en/of de lijst met data-elementen die door de <i>dataset</i> wordt beschreven ook veranderen. Zie ook: Online Thesaurus Zorg en Welzijn
Dataset met een grote impact bestemd voor secundair datagebruik	Een dataset waarvan het hergebruik aanzienlijke voordelen met zich meebrengt vanwege de relevantie ervan voor gezondheidsonderzoek Zie ook: EHDS wetstekst
Dataspace	Een omgeving waarin verschillende organisaties of systemen data kunnen delen en gebruiken, terwijl ze controle houden over hun eigen data. Bron: Afgeleid van de beschrijving van de term dataspace door de Europese Commissie.

BEGRIIP	DUIDING
Datastation secundair datagebruik	Een door een <i>bronhouder</i> beheerde omgeving waarin datasets en datadiensten van een dataleverancier worden ontsloten volgens afgesproken standaarden, met expliciete regels voor toegang, autorisatie, logging en gebruik, zodat data betrouwbaar en interoperabel kunnen worden gebruikt en worden uitgewisseld binnen een dataspace. Bron: Wiki Datastation (4.1), waarbij in de eerste zin dienstverlener is vervangen door bronhouder.
Datasubject	Een <i>datasubject</i> (betrokkene) is een identificeerbare natuurlijke persoon, waarvan gezondheidsdata wordt verzameld, verwerkt, opgeslagen of gedeeld binnen het <i>Health-RI ecosysteem</i> voor onderzoek, beleid en innovatie. Dit kan bijvoorbeeld een patiënt zijn waarvan medische data worden opgeslagen en gedeeld tussen verschillende <i>zorgverleners</i> , <i>onderzoekers</i> of overheidsinstanties. N.B.: de term ' <i>datasubject</i> ' hoeft taaltechnisch strikt genomen niet betrekking te hebben op een persoon. Het laat, vertaald naar 'dataonderwerp', ook de mogelijkheid over voor onderwerpen die niet-personen (entiteiten) zijn. Echter, ' <i>datasubject</i> ' is zeer sterk geassocieerd met natuurlijke persoon vanwege artikel 4 lid 1 in de GDPR (Engelstalige versie) waar ' <i>datasubject</i> ' expliciet genoemd wordt. In de AVG wordt gesproken over 'betrokkene'. Zie ook: Online Thesaurus Zorg en Welzijn
Datavergunning	Een administratief besluit dat door een instantie voor toegang tot gezondheidsdata aan een gebruiker van gezondheidsdata wordt afgegeven om bepaalde in de datavergunning gespecificeerde elektronische gezondheidsdata te verwerken voor specifieke secundaire gebruiksdoeleinden, op basis van de in hoofdstuk IV van de EHDS-verordening vastgestelde voorwaarden Zie ook: EHDS wetstekst
Dataverwerker	Een partij die gezondheidsdata verwerkt voor primaire doeleinden (zorgverlening) of secundaire doeleinden (hergebruik zoals onderzoek, beleid) in opdracht van een andere organisatie. Zie ook: Verantwoordelijke en verwerker Autoriteit Persoonsgegevens
Dynamische dataset	Data-elementen die 'ter plekke' worden samengevoegd tot een dataset
Eenheid van Taal	Vermogen van organisaties en hun processen en systemen om effectief en efficiënt informatie te delen met hun omgeving doordat er duidelijkheid is over de betekenissen van de woorden die worden gebruikt; de betekenis van woorden die in een bepaalde context worden vastgelegd, moet zowel voor mensen als voor computers duidelijk en gestandaardiseerd zijn; hier worden terminologiestelsels voor ingezet Zie ook: Online Thesaurus Zorg en Welzijn
Electronic Health Record (EHR)	Elektronisch patiëntendossier met gezondheidsinformatie zoals voorgeschiedenis, medicatie, behandelplannen, vaccinatiedata, allergieën, radiologiebeelden en laboratorium- en testresultaten.
Elektronisch patiëntendossier of EPD	Een verzameling elektronische gezondheidsdata van een natuurlijke persoon die in het gezondheidszorgstelsel worden verzameld en voor het verlenen van gezondheidszorg worden verwerkt Zie ook: EHDS wetstekst
Extract Transform Load (ETL)	Het proces voor het verzamelen en samenvoegen van data uit meerdere bronnen voor het ondersteunen van detectie, rapportage, analyses en besluitvorming.
FAIR data	Data die voldoen aan de richtlijnen voor het vindbaar (Findable), toegankelijk (Accessible), uitwisselbaar (Interoperable), herbruikbaar en duurzaam (Reusable) beschrijven, opslaan en publiceren van data (de FAIR-principes). Zie ook: Online Thesaurus Zorg en Welzijn
FAIR Data point	Een FAIR Data Point is een systeem om datasets op een FAIR manier te beschrijven, met conform DCAT-AP gestandaardiseerde metadata en om de dataset beschrijvingen beschikbaar te maken door middel van simpele web protocollen. Meer info: https://www.fairdatapoint.org/

BEGRIIP	DUIDING
FAIRificatie	Proces waarbij de herbruikbaarheid van een dataset wordt vergroot door het toepassen van de FAIR-principes ; dit omvat onder andere het toewijzen van unieke en persistente identifiers aan de data, het verrijken van metadata en het registreren of indexeren van deze data in doorzoekbare bronnen Zie ook: Online Thesaurus Zorg en Welzijn
Gefedereerd verwerken	Bij <i>gefedereerde verwerking</i> geeft de datagebruiker de opdracht om een verwerking uit te voeren op één of meer omgevingen die worden ingericht onder controle van elk van de betrokken datahouders.
Gefedereerde analyse	Gedecentraliseerde analyse van data waarbij de berekeningen op lokale, gedistribueerde BVO's worden uitgevoerd in plaats van centrale verwerking in één beveiligde verwerkingsomgeving. Bron: Wiki Datastasion.
Gegevens (ook wel data genoemd)	Een herinterpreteerbare representatie van informatie in een geformaliseerde vorm, geschikt voor communicatie, interpretatie of verwerking door mensen of geautomatiseerde systemen. Bron: ISO 2382.
Geharmoniseerd data	Data die zo is aangepast data ze kunnen worden samengevoegd of onderling vergelijkbaar zijn gemaakt.
Generieke functies	Functies die voor meerdere toepassingsgebieden nodig zijn om vindbaarheid, toegankelijkheid, interoperabiliteit of hergebruik van (gezondheids)data te kunnen realiseren; generieke functies zijn door de zorg-ICT-markt in te vullen op basis van een set van afspraken, protocollen, en open (waar mogelijk internationale) standaarden; deze afspraken worden in normen en technische afspraken vastgelegd; voorbeelden van generieke functies zijn adressering, autorisatie, identificatie, authenticatie, logging en toestemming/zeggenschap Zie ook: Online Thesaurus Zorg en Welzijn
Generieke voorziening	Producten of diensten gericht op het ondersteunen van een generieke functie, waarbij de ontwikkeling wordt overgelaten aan private partijen (waarbij marktwerking mogelijk is), maar er één gezamenlijke initiator is die de randvoorwaarden bepaalt (in de meeste gevallen vanuit de overheid); zo is bijvoorbeeld Mitz door de overheid aangewezen als gemeenschappelijke voorziening voor het regelen van toestemming van patiënten tot het delen van hun gezondheidsdata met andere zorgaanbieders Zie ook: Online Thesaurus Zorg en Welzijn
Genomic en proteomic data	Data afgeleid van genomische en proteomische studies.
Gestandaardiseerde data	Data waarbij de datavelden in elk systeem precies hetzelfde worden begrepen en uitgewisseld door toepassing van een vast model.
Gestructureerde data	Data met een eenduidige codering (bijv. diagnosecodes (ICD-10), labuitslagen (getal + eenheid)).
Haalbaarheidsstudie	Een haalbaarheidsstudie is een beoordeling die bepaalt hoe waarschijnlijk het is dat een (of meerdere) dataset(s) een toegevoegde waarde gaat leveren aan een onderzoek. Als een onderzoek gaat over bv ouderen en die zitten niet of beperkt in de dataset dan kan de onderzoeker besluiten om de betreffende dataset niet aan te vragen. Zie ook: Online Thesaurus Zorg en Welzijn
Health Data Access Body (HDAB)	Een juridische entiteit aangewezen door een lidstaat van de Europese Unie om de toegang tot gezondheidsdata te beheren t.b.v. onderzoek, beleid en innovatie. De rol van een HDAB is om toegang tot gezondheidsdata te vergemakkelijken terwijl de bescherming van persoonlijke data en privacy rechten wordt gewaarborgd. De taken van een HDAB omvatten o.a.: • het verwerken van aanvragen (beoordelen aanvragen en vergunningen verlenen) • voorzien in toegang-beheer van verwerkingsomgeving Bron: EHDS wetstekst

BEGRIIP	DUIDING
Health-RI ecosysteem	Geïntegreerde Nationale gezondheidsdata-infrastructuur secundair gebruik. Verwijst naar het geheel van alle betrokken partijen en elementen die een rol spelen bij wetenschappelijk onderzoek en hoe ze met elkaar interageren. Dit omvat niet alleen de technologische middelen en faciliteiten, maar ook de sociale, culturele, economische en politieke omgeving waarin wetenschappelijk onderzoek plaatsvindt. Binnen het ecosysteem kunnen betrokken partijen op een veilige manier data met elkaar uitwisselen.
Horizontaal gepartitioneerde data	Bij horizontaal gepartitioneerde gezondheidsdata bevat iedere bron data van verschillende patiënten/burgers/deelnemers (de rijen) en bevat iedere bron wel dezelfde kenmerken (de kolommen)).
IAA	IAA staat voor Identificatie, Authenticatie en Autorisatie. Dit zijn stappen in het proces van toegangscontrole. Zie de artikelen Identificatie -en authenticatiedienst en Autorisatiedienst voor meer informatie.
Identiteitregister	Verschillende doelgroepen gebruiken het <i>Health-RI ecosysteem</i> . Deze verschillende doelgroepen kennen ieder specifieke rollen en attributen. Vanuit de NEN-normering wordt momenteel naar identificatie en authenticatie van zorgprofessionals gekeken, waarbinnen ook de uitdaging ervaren wordt dat niet alle verschillende groepen zorgprofessionals onder het UZI identificatieregister vallen. Stichting Health-RI pleit voor een nationale aanpak c.q. standaardisatie van <i>identiteitregisters</i> , waarbij domeinoverstijgend en internationaal gekeken wordt naar de inrichting van rollen en attributen. <i>Identiteitregisters</i> vallen onder de generieke diensten
Imaging/beeld data	Een verzameling van data die wordt opgeslagen door middel van beeldvormingstechnieken. Dit kan variëren van digitale foto's en video's tot medische beelden en digitale documenten.
Informatiestandaarden	Modelleringsstandaarden voor causaliteit tussen data-elementen. Voorbeelden: openEHR, OMOP CDM, xtEHR, zorginformatiebouwstenen
Infrastructuur	De <i>Infrastructuur</i> is het geheel van voorzieningen (organisatie, proces, informatie, applicatie en techniek) voor het verwerken, opslaan, beveiligen, beheren en transporteren van digitale data. Zie ook: Online Thesaurus Zorg en Welzijn
Innovator	De <i>Innovator</i> is een bijzondere <i>onderzoeker</i> die herbruikbare onderzoeksresultaten en algoritmen produceert ten behoeve van <i>datagebruikers</i> die vereenvoudigde toegang tot <i>gezondheidsdata</i> willen hebben, zoals zorgprofessionals, beleidsmakers en <i>burgers</i> . Taken die een <i>innovator</i> uitvoert zijn: • Onderzoeken • Certificeren Medical Device Registry (MDR) op Algoritmen • Algoritmen publiceren Zie ook: Online Thesaurus Zorg en Welzijn
Interoperabiliteit	Het vermogen van organisaties, alsook van softwaretoepassingen of apparaten van dezelfde fabrikant of van verschillende fabrikanten, om via de processen die zij ondersteunen te interageren, met inbegrip van de uitwisseling van informatie en kennis – zonder de inhoud van de data te wijzigen – tussen die organisaties, softwaretoepassingen of apparaten. Bron: EHDS wetstekst
Klinisch (patiëntgebonden) onderzoek	Klinisch onderzoek omvat medisch wetenschappelijk onderzoek met (data over) patiënten. Zie ook: Online Thesaurus Zorg en Welzijn
Klinische data	<i>Klinische data</i> zijn data die te maken hebben met de gezondheid of de behandeling van mensen. Ze kunnen afkomstig zijn van klinisch onderzoek, waarbij behandelingen of andere medische interventies worden getest op menselijke proefpersonen, of van klinische praktijk, waarbij <i>zorgverleners</i> of patiënten data verzamelen over hun patiënten of in het geval patiënten over zichzelf.

BEGRIIP	DUIDING
Knooppunt	Een <i>knooppunt</i> vormt de toegang tot het Health-RI Ecosysteem van de rol van <i>Datahouder</i> en/of <i>Datagebruiker</i> . Er zijn verschillende soorten knooppunten: • <i>Regionale knooppunten</i> • <i>Domein knooppunten</i> Meer informatie paragraaf 4.5.2 Ecosysteem in dit document Zie ook: Online Thesaurus Zorg en Welzijn
Koppeldienst	<i>Generieke dienst</i> , zowel voor de gezondheidszorg zelf (primair gebruik) als voor onderzoek, beleid en innovatie (secundair datagebruik), om data van een persoon over verschillende bronnen te kunnen combineren (Data linken) op basis van identifiers, dan wel sleutels.
Kwalitatieve data	Data die beschrijvend is, minder gestructureerd, en inzicht geeft in inhoud, context
Kwaliteitsregistraties	Nationale dataverzameling, meestal aandoeningsspecifiek, om de kwaliteit van gezondheidszorg te verbeteren.
Kwantitatieve data	Meetbare data die in cijfers worden uitgedrukt en gebruikt om feiten, trends en verbanden te analyseren.
Lokale toetsingscommissie data-aanbieden	De <i>Lokale toetsingscommissie data-aanbieden</i> • stelt op voorhand gebruiksvoorwaarden op voor het aanbieden van de originele <i>gezondheidsdata</i> • Voert toetsing uit van een dataonderzoek-aanvraag ten behoeve van de originele databron Alternatieve naam: Lokale datatoegangscommissie Zie ook: Online Thesaurus Zorg en Welzijn
Lokale toetsingscommissie onderzoek	De lokale toetsingscommissie onderzoek is een commissie, verbonden aan de organisatie waaraan de onderzoeker of de data gelieerd is, die de onderzoeken die de onderzoekers willen starten beoordeelt op o.a. ethische en juridische gronden. Zie ook: Online Thesaurus Zorg en Welzijn
MDR-Notificeerder Instantie	<i>Generieke dienst</i> , zowel voor de gezondheidszorg zelf (primair gebruik) als voor onderzoek, beleid en innovatie (secundair datagebruik), die de Medical Device Registry (MDR) certificatie verleent. Zie ook: Online Thesaurus Zorg en Welzijn
Medisch Ethische Toetsings Commissie (METC)	Een Medisch Ethische Toetsingscommissie beoordeelt medisch onderzoek met proefpersonen onder de Wet medisch-wetenschappelijk onderzoek met mensen (WMO). Deze beoordeling heeft als doel de rechten, veiligheid en het welzijn van de deelnemers te waarborgen. Een WMO onderzoek kan pas worden uitgevoerd nadat een goedgekeurde METC een positief advies heeft afgegeven. Zie ook: Online Thesaurus Zorg en Welzijn , CCMO
Meervoudig gebruik	Gebruik zowel binnen de gezondheidszorg (zgn. primair gebruik) als daarbuiten, zoals bijvoorbeeld onderzoek, beleid en innovatie (zgn. secundair datagebruik). De verantwoordelijkheid om data geschikt te maken voor meervoudig gebruik wordt gedeeld tussen de datahouder en de datagebruiker (in de rol van datahouder).
Metadata	Een beschrijving van een dataset en haar kenmerken. De beschrijving van elke dataset bevat informatie over de bron, de reikwijdte, de belangrijkste kenmerken en de aard van de elektronische gezondheidsdata in de dataset alsook de voorwaarden voor het beschikbaar stellen ervan. Bron: EHDS wetstekst
Nationaal contactpunt voor secundair datagebruik	Een organisatorische en technische gateway die het landsgrensoverschrijdende secundaire gebruik van elektronische gezondheidsdata mogelijk maakt, onder verantwoordelijkheid van een lidstaat van de Europese Unie. De NCP-2nd-use is ook verantwoordelijk voor het waarborgen van de interoperabiliteit en compatibiliteit en voor het bieden van begeleiding en ondersteuning. Zie ook: EHDS wetstekst

BEGRIIP	DUIDING
Netwerken	XDS, AORTA, LSP
Niet-persoonsgebonden elektronische gezondheidsdata	Andere elektronische gezondheidsdata dan persoonlijke elektronische gezondheidsdata, met inbegrip van zowel data die geanonimiseerd zijn zodat zij niet langer betrekking hebben op een geïdentificeerde of identificeerbare natuurlijke persoon (de “betrokkene”) als data die nooit betrekking hebben gehad op een betrokkene Zie ook: EHDS wetstekst
Omics data	Grote hoeveelheden data die zijn verzameld door moleculaire biologie met behulp van technologieën zoals genomics, proteomics, en metabolomics. Deze technologieën zijn essentieel voor het begrijpen van biologische systemen en hebben zich sinds de jaren '90 ontwikkeld. Omics data worden vaak geanalyseerd met behulp van bio-informatica en statistische modellen om veranderingen op moleculair niveau te identificeren en te begrijpen, wat bijdraagt aan gepersonaliseerde geneeskunde en het ontwikkelen van nieuwe therapieën.
Onboarden	Het begrip <i>onboarden</i> wordt gebruikt in de context van <i>metadata</i> en omvat het aanmelden van de metadata van een dataset bij een FAIR Data point.
Onderzoeker	Beroepskrachten die zich bezighouden met wetenschappelijk onderzoek; ze houden zich bezig met het ontwerpen van onderzoeksprotocollen, het uitvoeren van experimenten en het analyseren van data Zie ook: Online Thesaurus Zorg en Welzijn
Onderzoeksdata	Data die als input dienen voor een onderzoek, die tijdens een onderzoek worden gegenereerd en/of die het resultaat zijn van een onderzoek Zie ook: Online Thesaurus Zorg en Welzijn
Onderzoeksregister	Register waarin wordt vastgelegd welke onderzoeken er zijn en in welke toestand de onderzoeken zich bevinden Zie ook: Online Thesaurus Zorg en Welzijn
Ongestructureerde data	Data die niet is vastgelegd of georganiseerd volgens een vaste herkenbare structuur (bijv. vrije tekst)
Ontologie	Terminologiebestanden die gekenmerkt worden door aanwezigheid van gespecificeerde relaties tussen de concepten; er zijn niet alleen synoniemrelaties en hiërarchische relaties zoals in een thesaurus maar daarnaast worden alle typen relaties uitgesplitst, bijvoorbeeld tussen een ziekte en een ziekteverwekker, tussen een grondstof en een product, tussen een beroepskracht en een beroep, enz. Zie ook: Online Thesaurus Zorg en Welzijn
Open standaarden	Standaarden die publiekelijk beschikbaar zijn en die ontwikkeld en goedgekeurd worden via een proces dat gebaseerd is op samenwerking en consensus met de betrokken community Zie ook: Online Thesaurus Zorg en Welzijn
Persistent dataplatform	Een persistent dataplatform is een centrale en duurzame opslagomgeving voor data die gedurende lange tijd beschikbaar moet blijven en ongewijzigd moet worden gehouden. Binnen deze omgeving wordt onderzoeksdata opgeslagen en vanuit deze omgeving wordt deze data beschikbaar gesteld.
Person-generated health data (PGHD)	Wearables- en thuismonitoringsdata
Persoonlijke elektronische gezondheidsdata	Data over gezondheid die in elektronische vorm worden verwerkt Zie ook: EHDS wetstekst
Primair gebruik	De verwerking van elektronische gezondheidsdata voor de verlening van gezondheidszorg om de gezondheidstoestand van de natuurlijke persoon op wie die data betrekking hebben te beoordelen, te behouden of te herstellen, waaronder het voorschrijven en het verstrekken van geneesmiddelen en medische hulpmiddelen, alsmede voor relevante sociale, administratieve of vergoedingsdiensten Zie ook: EHDS wetstekst

BEGRIIP	DUIDING
Pseudonimiseren	Het verwerken van persoonsgegevens op een zodanige wijze dat de persoonsgegevens niet meer aan een specifieke betrokkene kunnen worden gekoppeld zonder dat er aanvullende gegevens worden gebruikt; deze aanvullende gegevens moeten apart worden bewaard en er moeten door de organisatie technische en organisatorische maatregelen worden genomen om ervoor te zorgen dat de persoonsgegevens niet aan een geïdentificeerde of identificeerbare natuurlijke persoon (kunnen) worden gekoppeld Zie ook: Online Thesaurus Zorg en Welzijn
Real world data (RWD)	Data gerelateerd aan het gezondheidszorgproces, vastgelegd het EPD, patiënt gegenereerde e/o gerapporteerde data, declaratiedata, public health data
Real world evidence	Evidence die gebaseerd is op real world data
Resource Description Framework (RDF)	Een W3C-standaard voor het beschrijven en uitwisselen van metadata op een machine leesbare manier. RDF legt informatie vast als zogenaamde triples (onderwerp – eigenschap – waarde), waardoor computers de betekenis van gegevens kunnen begrijpen en relaties tussen datasets kunnen leggen.
Secundair gebruik van (gezondheidszorg)data (hergebruik)	Het gebruik van data voor andere doeleinden dan directe dienst-/zorgverlening aan de burger/patiënt. Denk aan wetenschappelijk onderzoek, beleidsontwikkeling, zorginnovatie & kwaliteit
Semantische interoperabiliteit	Het vermogen van verschillende systemen, organisaties of applicaties om data niet alleen uit te wisselen, maar ook op dezelfde manier te interpreteren en te begrijpen. De betekenis van de informatie blijft dus behouden wanneer die van het ene systeem naar het andere gaat.
Sleutelbestand	Een bestand dat de koppeling mogelijk maakt tussen gepseudonimiseerde data en de oorspronkelijke identiteit van personen of objecten.
Social and Behavioural determinants of health (SBDH)	Omgevingsomstandigheden die van invloed zijn op gezondheid, functioneren en kwaliteit van leven, onderverdeeld in vijf categorieën: economische stabiliteit; toegang tot en kwaliteit van onderwijs; sociale en gemeenschapscontext; buurt en gebouwde omgeving; en toegang tot en kwaliteit van gezondheidszorg
Statische dataset	Samengevoegde data-elementen die eerder als dataset zijn verzameld en opgeslagen voor onderzoek. Preprocessing van data is reeds verricht
Structurele transformatie	Het omzetten van data van het ene datamodel, formaat of dataschema naar een ander datamodel, zodat data door een ander systeem kunnen worden verwerkt of gecombineerd.
Structuur- en communicatiestandaarden	Voorbeelden: DICOM (beeld), EDIFACT, CDA, FHIR
Syntactische interoperabiliteit	De mate waarin systemen data kunnen uitwisselen doordat zij dezelfde structuur, codering en uitwisselingsregels voor data hanteren.
Synthetische data	Een verzameling data die is gegenereerd op basis van een model. Dat model is op zijn beurt gebouwd op basis van bijvoorbeeld persoonsgegevens. Een synthetische dataset kan worden gebruikt voor statistische analyse zonder dat er persoonsgegevens mee onthuld worden.
Terminologiestandaarden	Medische terminologie voor eenduidige vastlegging, zoals SNOMED CT, LOINC (laboratoriumobservaties), GMDN (medical devices).
Toestemmingsdata	Data waarin toestemming (opt-in of opt-out) is vastgelegd. Dit kan zijn onderzoek-specifieke toestemming, brede toestemming zoals bij biobanken en registraties.

Bijlage 2. Afkortingenlijst

Binnen het gezondheidszorginformatielandschap wordt veel afkortingen gebruikt. De afkortingen die voorkomen in dit document, worden in onderstaand tabel voluit weergegeven

AFKORTING	VOLUIT
AI	Artificial Intelligence
BVO	Beveiligde verwerkingsomgeving
CDM (OMOP CDM)	OMOP Common Data Model
DAAMS	Data Access Application Management System
DCAT-AP	Data Catalog - Application Profile
DDS	Datastation Description Service
DEP	Data Encryption Protocol
DICOM	Digital Imaging and Communications in Medicine
DIZRA	Duurzaam Informatiestelsel Zorg Referentie Architectuur
DSS	Datastation specificatie
ECD	Elektronisch Cliënten Dossier
EEHRxF	European Electronic Health Record exchange Format
EHDS	European Health Data Space
EPD	Elektronisch Patiënten Dossier
FAIR	Findable Accessible Interoperabel Reusable
FAQ	Frequently Asked Questions
FDP	FAIR Data Point
FDS	Federatie Datastelsel
FEC	Forward Error Correction
FHIR	Fast Healthcare Interoperability Resources
GDPR	General Data Protection Regulation
GISA	Gezondheidsinformatiestelsel Architectuur
HDAB	Health Data Acces Body
HIS	Huisartinformatiesysteem
HL7	Health Level 7
HRI	Health-RI
IAA	Identificatie, authenticatie en autorisatie
IBDS	Interbestuurlijke Datastrategie
ICPC	International Classification of Primary Care
IAC	Industry Advisory Committee
LAS	Landelijk Afspraken Stelsel
LDN	Landelijk Dekkend Netwerk
LIM	Logical Information Model
LOINC	Logical Observation, Identifiers, Names and Codes
MAR	Maatschappelijke Adviesraad (MAR)
NCPeH	National Contact Point eHealth
NGIS	Nationinale Gezondheidsinfrastructuur voor Secundair datagebruik
NVS	Nationale Visie en Strategie
OMOP	Observational Medical Outcomes Partnership
OpenEHR	Open Electronic Health Record
OVT	Obstakel Verwijder Traject

AFKORTING	VOLUIT
RDF	Resource Description Framework
RWD	Real World Data
SNOMED CT	Systematized Nomenclature of Medicine – Clinical Terms
SPE	Secure Processing Environment
SC	Strategic Committee
TEHDAS2	Towards the European Health Data Space 2
TTP	Trusted Third Party
VCF	Virtual Contact File
Wegiz	Wet elektronische gegevensuitwisseling in de zorg
WKZ	Wet Kwaliteitsregistraties Zorg

2027

Bijlage 3. Toelichting architectuurprincipes

In deze bijlage wordt bij ieder specifiek architectuurprincipe voor de Nationale gezondheidsdata-infrastructuur secundair gebruik een nadere toelichting gegeven. Het betreffen de volgende principes:

1. Data nabij de bron waar mogelijk
2. Verantwoordelijkheid en regie bij de datahouder
3. Decentrale autonomie
4. Alleen delen analyseresultaat
5. Gestandaardiseerde ontsluiting
6. Hergebruik als uitgangspunt
7. Generiek boven specifiek
8. Proces- en gebruikersgericht
9. Systeembenadering
10. Internationaal boven nationaal boven lokaal
11. Omgaan met diversiteit
12. Maximale benutting van bestaande data (incl. legacy)
13. Wat kan wel

De toelichting betreft naast een definitie ook de rationale en de implicaties gegeven.

1. Data nabij de bron waar mogelijk

Gezondheidsdata voor secundair gebruik (zoals wetenschappelijk onderzoek, beleid en innovatie) worden bij voorkeur niet permanent gecentraliseerd in één grote landelijke database. In plaats daarvan blijven de gegevens waar mogelijk staan in de bronsystemen van de datahouder of worden ze naar een regionale/lokale voorziening gebracht dicht bij de bron, totdat ze nodig zijn. In de Nationale gezondheidsdata-infrastructuur secundair gebruik wordt er wel rekening mee gehouden dat soms verplaatsing, kopiëren of centrale opslag van data nodig is. Dit vindt echter alleen plaats wanneer dit noodzakelijk is voor de beoogde toepassing, aantoonbare meerwaarde biedt of vanuit wet- en regelgeving vereist is.

Rationale:

- Actuele en kwalitatieve data: Gegevens verouderen en verliezen context zodra ze worden gekopieerd of verplaatst naar centrale (secundaire) dataomgevingen. Door data bij de bron (of zo dicht mogelijk bij de bron) te bevragen, wordt gewerkt met de meest actuele en authentieke gegevensversie.
- Dataminimalisatie en Privacy (Compliance aspect): In het kader van privacywetgeving (GDPR/AVG) en internationale standaarden (zoals de EHDS voor secundair gebruik) geldt het principe van dataminimalisatie. Het bevragen van data aan de bron voorkomt onnodige accumulatie, opslag en duplicatie van gevoelige persoonsgegevens.
- Eenduidig verwerkingsverantwoordelijkheid en beheer: De datahouder is verwerkingsverantwoordelijke (bepaalt doel en middelen van dataverwerking) en blijft daarmee verantwoordelijk voor de kwaliteit, juistheid en toegang van de gegevens. Deze rol geldt tevens bij hergebruik van de data. Dit voorkomt onduidelijkheid over wie de verwerkingsverantwoordelijke is bij hergebruik.
- Vertrouwen en Zeggenschap: burgers houden beter zicht op waar hun gegevens zich bevinden en voor welke secundaire doeleinden ze beschikbaar worden gesteld wanneer data niet onbeperkt naar centrale opslaglocaties zijn gekopieerd.

Implicaties:

- De doelarchitectuur moet federatieve patronen ondersteunen waarbij de analysecode of verkenopdracht naar de data bij de datahouder toe gaat in plaats van dat alle data (eerst) naar de onderzoeker of één centrale plek wordt gebracht.
- Bronnen dan wel netwerken van bronnen moeten data op een gestandaardiseerde wijze ontsluiten zodat deze direct herbruikbaar zijn voor secundaire analyses zonder handmatige transformaties
- Systeemfuncties moeten beschikken, in geval van dat de analyse-code of verkenvraag naar de data toe gaat, over moderne, veilige koppelvlakken waarmee geautoriseerde (secundaire) gebruikers gerichte vragen (bijvoorbeeld query's of algoritmes) kunnen stellen aan het bronsysteem.
- Om met wet- en regelgeving compliant te zijn dient pseudonimisatie, anonimisatie of aggregatie waar mogelijk direct aan de bron plaats te vinden vóórdat de uitkomst/resultaat ter beschikking wordt gesteld aan de geautoriseerde gebruiker (vragensteller).
- De bron (gegevensverwerkingsverantwoordelijke) dient (geautomatiseerd) te kunnen verifiëren of de gebruiker over het juiste mandaat, toestemming of wettelijke grondslag beschikt om de specifieke data te bevragen.
- Vanwege de "waar mogelijk" dient wanneer een verwerking van data (als in beantwoorden van een vraag of het draaien van een algoritme) bijvoorbeeld organisatorisch, financieel, of technisch niet haalbaar is, er een oplossing beschikbaar te zijn om gecontroleerde en gestandaardiseerde "verwerkingswerkplaatsen" (bijvoorbeeld in een knooppunt, een regio dan wel nationaal of internationaal) in te richten dan wel te kunnen gebruiken.

2. Alleen delen analyseresultaat

Wanneer het functioneel, juridisch en technisch mogelijk is, wordt alleen de analyse toegelaten tot de data (data visiting). Als het mogelijk is worden na de analyse de resultaten gedeeld. De resultaten bevatten normaliter alleen statistische gegevens en dus geen persoonsgegevens. Dit sluit aan bij het AVG-principe van privacy by design, waarbij de meest privacy vriendelijke verwerkingsvorm wordt toegepast die de beoogde doelstelling ondersteunt.

Rationale:

- Maximale privacybescherming (privacy by design): Ruwe gezondheidsgegevens verlaten de veilige verwerkingsomgeving van de datahouder niet. Dit verkleint het risico op datalekken, identiteitsherleiding of onrechtmatig hergebruik drastisch.
- Data governance en -soevereiniteit gegarandeerd: Datahouders behouden het fysieke en juridische beheer over hun data. Zij houden zo de volledige controle over wie toegang krijgt tot de gegevens voor analyse.
- Voorkomen van dataduplicatie ('datakerkhoven'): Het centraal verzamelen van (landelijke) gezondheidsdata leidt tot enorme verzamelingen van verouderde kopieën en vormt een aantrekkelijk doelwit voor cybercriminaliteit (het honeypot-effect). Door gegevens direct aan de bron te laten staan, blijven ze altijd actueel en wordt onnodige datavervuiling of -duplicatie voorkomen.
- Naleving van wet- en regelgeving (o.a. AVG/EHDS): Het principe volgt het AVG-beginsel van dataminimalisatie en sluit aan bij de kaders van de EHDS voor secundair gebruik: er wordt niet meer informatie verwerkt of gedeeld dan strikt noodzakelijk is voor het specifieke onderzoeksdoel.

Implicaties:

- Federated / Gedistribueerde architectuur is vereist: Er moet een infrastructuur staan waarbij analytische scripts naar de lokale datastations van datahouders gestuurd kunnen worden.

- Lokale verwerkingscapaciteit: Datahouders moeten beschikken over voldoende rekenkracht (compute) en infrastructuur om de ontvangen analyse-opdrachten lokaal veilig uit te voeren.
- Gestandaardiseerde data (data harmonisatie): Om het script op meerdere locaties reproduceerbare en betrouwbare resultaten te laten genereren, moeten brongegevens lokaal zijn getransformeerd naar een gemeenschappelijk datamodel.
NB: Momenteel is data harmonisatie vaak niet mogelijk, omdat de data van onvoldoende kwaliteit is.
- Geautomatiseerde en/of handmatige uitvoercontrole (Disclosure Control): Er moeten mechanismen/filters zijn ingeregeld die controleren of het uiteindelijk uit de beveiligde verwerkingsomgeving te exporteren analyseresultaat geen verborgen persoonsgegevens bevat.
- Gestandaardiseerde autorisatie & auditing: Er is een centrale of gefedereerde organisatie van verantwoordelijkheden nodig die vooraf bepaalt welke analyselogica toegestaan is, inclusief een onweerlegbare audit trail van alle uitgevoerde scripts en gegenereerde resultaten.

3. Verantwoordelijkheid en regie bij de datahouder

De datahouder is en blijft verantwoordelijk voor de rechtmatigheid, de kwaliteit en de beveiliging van de gezondheidsdata, én beslist over de beschikbaarstelling voor secundair gebruik binnen de kaders van de wet en landelijke afspraken. Wel wordt de regie van de datahouder beperkt op het moment dat de operationele HDAB aan een datagebruiker via vergunningverlening permissie heeft gegeven om data beschikbaar te krijgen en te gebruiken. De datahouder dient in dat geval de data beschikbaar te stellen. Hiertegen kan slechts bezwaar worden aangetekend bij een vermeende overtreding van intellectueel eigendomsrechten.

Rationale:

- Wettelijke, ethische en maatschappelijke borging: Organisaties in de gezondheids-, preventie- en welzijnssector hebben te maken met het medisch beroepsgeheim, het algemeen geheimhoudingsrecht en de AVG. Door de regie expliciet bij de datahouder te laten, blijft de juridische en ethische keten van verantwoordelijkheid helder en intact, ongeacht of het om curatieve zorg of publieke gezondheid gaat.
- Behoud van maatschappelijk vertrouwen: Burgers vertrouwen hun persoonlijke gezondheids- en leefstijlgegevens toe aan hun specifieke professional of instelling, niet aan een anonieme centrale entiteit. Regie bij de datahouder waarborgt dat dit vertrouwen in de gehele gezondheidsketen behouden blijft.
- Kennis over de domein specifieke context: Gezondheids- en preventiedata zijn sterk contextafhankelijk (bijv. meetomstandigheden, populatiekenmerken of specifieke vragenlijsten). De datahouder begrijpt deze context als enige ten volle, wat voorkomt dat gegevens verkeerd geïnterpreteerd of misbruikt worden bij hergebruik/verwerking van gezondheidsgegevens.
- Eenduidige aansprakelijkheid: Bij een datalek, onrechtmatige verstrekking of verkeerd gebruik van brondata is er geen twijfel over de verantwoordelijke entiteit. De datahouder voert de regie en is daarmee het primaire aanspreekpunt voor toezichthouders (zoals de Autoriteit Persoonsgegevens).

Implicaties:

- Inrichten van domein-overstijgende data governance: Datahouders over het hele brede gezondheidsdomein (van ziekenhuizen en huisartsen tot GGD'en, sport- en preventie-initiatieven) moeten interne processen of toetsingscommissies inrichten. Deze beoordelen dataverzoeken/aanvragen snel op rechtmatigheid en geschiktheid, ter ondersteuning van landelijke structuren.
- Beslissingsbevoegdheid en controle (de 'stopknop'): Bij onderzoeksaanvragen die vallen onder EHDS Onder de EHDS vindt de beoordeling door de HDAB plaats en alleen bij vermeende overtreding van intellectueel eigendomsrechten bezwaar mogelijk. Voor onderzoek dat buiten de EHDS plaatsvindt, beslist de datahouder wel over het beschikbaar stellen van data.

- Plicht tot het borgen van datakwaliteit en voldoen aan de FAIR principes: Regie dragen houdt ook de verplichting in om te zorgen dat preventie- en gezondheidsdata vindbaar, toegankelijk, kwalitatief hoogwaardig, begrijpbaar en herbruikbaar (FAIR) worden vastgelegd. Zowel klinische als leefstijl- en gezondheidsdata moeten aan de bron goed beheerd worden.
- Collectieve ondersteunen en ontzorgen van kleinere partijen: Voor kleinere datahouders (bijv. zelfstandige zorgverleners, lokale preventiewerkers of kleine GGZ-praktijken) is de individuele regielast te zwaar. Dit vereist dat verantwoordelijkheid en regie gemandateerd georganiseerd kunnen worden via (gecertificeerde/vertrouwde) knooppunten, netwerkorganisaties of koepels.
- Transparantie via uitgebreide logging: De datahouder moet voor alle vormen van gezondheids- en preventiedata kunnen verantwoorden wie op welk moment voor welk doel toegang heeft gekregen. Dit vereist een auditeerbaar en onveranderbaar logging systeem binnen de beheeromgeving van de datahouder alsook binnen de gehele keten van verwerking/hergebruik/verrijking van die data.

4. Decentrale autonomie over eigen gezondheidsdata-infrastructuur

(Onafhankelijke) organisaties binnen het brede ecosysteem van gezondheid, preventie en zorg behouden het recht om binnen de geldende wet- en regelgeving hun eigen (ICT-)infrastructuur bronsystemen en databeheer in te richten t.b.v. secundair gebruik, mits zij voldoen aan de (inter)nationale vastgestelde afspraken voor interoperabiliteit en gegevensuitwisseling. De architectuur beschrijft dan ook functionaliteit, geen specifieke oplossingen die invulling geven aan functionaliteit. De decentrale autonomie op de inrichting en implementatie wordt gerespecteerd. Afspraken moeten worden gemaakt over de op te leveren resultaten en de onderlinge koppelvlakken.

Rationale:

- Optimaal aansluiten bij lokale data-architecturen voor hergebruik: Niet elke datahouder heeft dezelfde datastructuur, analytische volwassenheid of technische infrastructuur. Decentrale autonomie maakt het mogelijk dat organisaties een secundaire data-infrastructuur kiezen die past bij hun eigen omvang en analytische capaciteiten.
- Wendbaarheid en innovatiesnelheid: Secundair datagebruik (zoals AI-ontwikkeling, wetenschappelijk onderzoek en epidemische surveillance) ontwikkelt zich razendsnel. Door datahouders niet vast te pinnen op één starre, centraal voorgeschreven analyseketen, kunnen zij snel adaptieve analysetechnologieën (zoals Federated Learning of nieuwe privacy-preserving technieken) adopteren zodra deze beschikbaar komen.
- Voorkomen van vendor lock-in: het verplichten/implementeren van een specifieke voorziening van een leverancier kan leiden tot een ongewenste afhankelijkheid van een leverancier. Decentrale autonomie stimuleert een open ecosysteem van IT-leveranciers voor secundaire dataontsluiting op basis van open protocollen/standaarden.
- Verkleinen risico "single point of failure": Een gedistribueerd netwerk waarin elke autonome datahouder controle houdt over de secundaire toegangspunten/databronnen, voorkomt dat het gehele secundaire datanetwerk/infrastructuur platgelegd kan worden door een storing of aanval op één centrale voorziening.

Implicaties:

- Gestandaardiseerde koppelvlakken en protocollen: Autonomie in de interne inrichting geldt alleen mits de interface naar extern gestandaardiseerd is. De infrastructuur van de datahouder moet kunnen praten met het landelijke (secundaire) netwerk via verplichte afspraken voor federatieve bevraging
- Keuzevrijheid in de 'secundaire schil' (lokale node vs. regionaal knooppunt): Datahouders mogen zelf kiezen of ze een eigen lokale beveiligde verwerkingsomgeving beheren, of dat ze zich

aansluiten bij een gedeelde regionale of sectorale 'hub' die de bevestigingen namens hen afhandelt.

- Harmonisatie van de bevestigings- en accreditatie-eisen: Om te voorkomen dat decentrale autonomie leidt tot willekeur bij de goedkeuring van secundaire analyses, moeten de technische eisen aan een node en het proces voor het toelaten van federatieve queries landelijk uniform gecertificeerd en geharmoniseerd zijn (bijvoorbeeld volgens richtlijnen van HDAB-NL)
- Verantwoordelijkheid voor het 'vertalen' van brondata naar secundaire modellen: Omdat de datahouder autonoom is in de inrichting van de secundaire schil, is de datahouder er zelf voor verantwoordelijk dat lokale brongegevens correct en valide worden getransformeerd naar het voor secundair gebruik afgesproken landelijke of internationale datamodel (bijv. transformatie naar OMOP CDM of ZIB-extracten).
- Decentrale uitvoering van Privacy-Preserving Analyses: Datahouders moeten in hun autonome voorzieningen de functionaliteit kunnen ondersteunen om federatieve zoekvragen of algoritmes lokaal uit te voeren (data visiting), waarbij de data binnen de eigen beveiligde (verwerkings-) zone blijft en alleen geaggregeerde of gecensureerde uitkomsten aan de onderzoeker/aanvrager worden teruggegeven.

5. Gestandaardiseerde ontsluiting

Toegang tot gezondheidsdata en de bijbehorende metadata verloopt zo veel mogelijk via gestandaardiseerde, open, internationaal geaccepteerde en goed gedefinieerde koppelvlakken en informatie-architectuurstandaarden.

Rationale:

- Interoperabiliteit en schaalbaarheid: een eenmalige integratie op basis van een open standaard zorgt ervoor dat een datastation herbruikbaar is voor talloze verschillende vragen, zonder dat er telkens maatwerkinterfaces gebouwd hoeven te worden.
- Voorkomen van vendor lock-in: door te eisen dat koppelvlakken open en leveranciersonafhankelijk zijn, kan de infrastructuur onafhankelijk blijven en eigen keuzes blijven maken
- Geautomatiseerde verwerking: gegevens en bijbehorende metadata moeten niet alleen leesbaar en begrijpbaar zijn voor mensen. Juist ook systemen dienen deze geautomatiseerd een eenduidig te kunnen verwerken/interpreteren.
- Naleving van wet- en regelgeving

Implicaties:

- Hanteren van verplichte open en internationaal geaccepteerde standaarden: de doelarchitectuur moet expliciet de toegestane standaarden aanwijzen
- Beschikbare systeemonafhankelijke ontsluitingslaag: Datahouders moeten een koppelvlak of datastation inrichten die de brondata vanuit de bronhouders systemen valideert, transformeert en klaarzet voor (her)gebruik volgens de afgesproken standaarden
- Strikt beheer van de ontsluitingslaag: alle koppelvlakken moeten zijn voorzien van actuele begrijpelijke documentatie en kennen/volgen/hanteren een vastgesteld change en release management.
- Validatie en conformiteitstoetsing (interoperabiliteitstestbed): er is een toetsings- / testmechanisme nodig waarmee kan worden aangetoond dat de datahouder aangeboden koppelvlakken alsook de betrokken gegevens beide voldoen aan afgesproken standaarden (conformance testing).

6. Hergebruik als uitgangspunt

Data, datasets en onderzoeksresultaten worden vanaf het begin ontworpen voor hergebruik.

2279 **Rationale:**

- 2280 • Vermindering van administratieve/registratielast: Gegevens uit het primaire proces hoeven niet
- 2281 dubbel of apart te worden ingevoerd voor secundair gebruik (onderzoek, beleid en innovatie).
- 2282 • Kosten en tijdsbesparing: het benutten van bestaande (inter)nationale infrastructurele
- 2283 bouwstenen voorkomt dat het wiel opnieuw wordt uitgevonden.
- 2284 • Voorkomen van (verdere) versnippering: als iedereen een eigen infrastructuur of eigen data-silo
- 2285 bouwt, ontstaat er een onbeheersbaar landschap. Hergebruik dwingt eenduidigheid en
- 2286 standaardisatie af.
- 2287 • Kwaliteit en interoperabiliteit: het hergebruiken van vastgestelde standaarden verhoogt de
- 2288 kwaliteit en interoperabiliteit van zowel de gegevens zelf als de daarop uitgevoerde analyses.

2289

2290 **Implicaties:**

- 2291 • Aansluiten bij of op de bron: de data-architectuur moet zo zijn/worden ingericht dat secundair
- 2292 gebruik direct de gegevens uit de primaire processen kan aftappen zonder tussentijdse
- 2293 handmatige conversies/transformaties.
- 2294 • Datakwaliteit bij/aan de bron: omdat data meervoudig wordt gebruikt, verschuift het belang van
- 2295 datakwaliteit volledig naar het moment van de eerste vastlegging in het proces.
- 2296 • Comply or explain (of misschien wel beter comply or die): Nieuwe ontwikkelingen/projecten
- 2297 moeten verplicht aanhaken bij/aansluiten op bestaande landelijke afspraken. Afwijken mag
- 2298 alleen met een hele sterke argumentatie (als je comply or die hanteert dan vervalt afwijken,
- 2299 want je moet aansluiten).
- 2300 • Modulariteit en API-first: Nieuw te bouwen functionaliteiten binnen de doelarchitectuur moeten
- 2301 als generieke, herbruikbare componenten (services/microservices) met open API's worden
- 2302 ontworpen, zodat andere partijen er later eenvoudig op kunnen aansluiten.
- 2303 • Centrale vindbaarheid van voorzieningen: er moet een helder overzicht (catalogus/index) zijn
- 2304 van bestaande architectuurcomponenten, standaarden en data(sets) die (direct) herbruikbaar
- 2305 zijn voor nieuwe initiatieven/ontwikkelingen.

2306

2307 **7. Generiek boven specifiek**

2308 Oplossingen en voorzieningen worden generiek ingericht, zodat ze herbruikbaar zijn over
 2309 gezondheidszorg, onderzoek, beleid en innovatie, tenzij een specifieke casus aantoonbaar een
 2310 uitzonderingspositie vereist. De doelarchitectuur stuurt dan ook actief op uniformiteit en
 2311 samenhang om gezamenlijk te convergeren naar landelijke (en Europese) doelarchitectuur, gedeelde
 2312 standaarden en centrale/federatieve voorzieningen.

2313

2314 **Rationale:**

- 2315 • Schaalbaarheid en herbruikbaarheid: Generieke gegevensstructuren maken het mogelijk om
- 2316 data uit uiteenlopende bronnen op een uniforme manier beschikbaar te stellen voor vele
- 2317 verschillende (onderzoeks-)vragen zonder telkens opnieuw de data-infrastructuur in te moeten
- 2318 richten.
- 2319 • Kosten- en beheerbeheersing: maatwerkoplossingen en unieke koppelvlakken vergen veel
- 2320 onderhoud, verhogen de foutgevoeligheid en creëren vaak een vendor lock-in. Generieke
- 2321 voorzieningen verlagen de totale beheerskosten.
- 2322 • Versnelling van ontsluiten data: door gebruik te maken van standaard bouwblokken /
- 2323 componenten (zoals generieke voorzieningen voor pseudonimisering, autorisatie en BVO's)
- 2324 verkort je de doorlooptijd van ideevorming tot analyse aanzienlijk.
- 2325 • Aansluiting op (inter)nationale kaders: De European Health Data Space en de nationale
- 2326 datagovernance eisen uniformiteit in toegang tot en uitwisseling van gegevens. Een generiek
- 2327 uitgangspunt borgt dat de architectuur hier naadloos op aansluit.

2328

Implicaties:

- Transformatie van data bij de bron of bij ingestie: Gegevens moeten vanuit de specifieke bronnen zo vroeg mogelijk getransformeerd worden naar een of meer geaccepteerde generieke standaardmodellen om analyse over meerdere bronnen mogelijk te maken (benodigde gegevens voor een analyse binnen de bronnen hebben dezelfde betekenis en zijn van voldoende kwaliteit).
- Strike uitzonderingsprocedure ('comply or explain'): Een specifiek model of maatwerkoplossing wordt alleen toegestaan als de generieke standaard aantoonbaar niet volstaat voor die specifieke secundaire toepassing en ook de daarmee gepaard gaande extra kosten zijn gedekt.
- Generieke (architectuur)bouwblokken: Functies zoals zeggenschapsregister, pseudonimisering/anonimisering, geautomatiseerde gegevensbeschrijvingen (metadata), maar ook autorisaties worden als generieke, herbruikbare en betaalbare (micro)services/voorzieningen neergezet.
- Vergt initiële investeringen: de overgang naar een generieke structuur vraagt vooraf een grote investering in standaardisatie en harmonisatie die zich pas terugbetaald bij herhaaldelijk en divers secundair gebruik.

8. Proces- en gebruikersgericht

Richt de doelarchitectuur, voorzieningen en organisatie van verantwoordelijkheden zodanig in dat deze optimaal aansluiten op de feitelijk werkprocessen en behoeften van zowel de datahouders als de datagebruikers, zonder onnodige administratieve of technische barrières. Ontwerpbeslissingen worden hierbij gestuurd door de behoeften van gebruikers en werkprocessen, niet door technologie of architectuur als doel op zich.

Rationale:

- Borging van datakwaliteit aan de bron: als het aanleveren of beschikbaar stellen van data voor secundair gebruik haakt aan het reguliere werkproces van diegene die de data verzamelt en vastlegt (zonder extra handmatige registratielast), blijft de datakwaliteit hoog en het draagvlak behouden.
- Reductie van doorlooptijd om data te kunnen gebruiken ('Time-to-data'): aanvragen voor hergebruik van (bestaande) data duren nu vaak maanden tot jaren door complexe goedkeuringsprocessen. Een gebruikersgericht ontwerp stroomlijnt de stappen (van vinden, doorzoeken, aanvragen van toegang/gebruik tot en met het veilig verwerken van data).
- Voorkomen van "shadow-IT": als het officiële proces om gezondheidsdata te hergebruiken te ingewikkeld of traag is, gaan datagebruikers om het systeem heen werken, met grote privacy- en informatiebeveiligingsrisico's tot gevolg.
- Adoptie en acceptatie: een (doel)architectuur kan technisch perfect zijn, maar als de (gebruikers)interfaces niet intuïtief zijn, dan zal de (doel)architectuur niet op grote schaal geadopteerd worden.
- Behoud van aansluiting op de praktijk: architectuurkeuzes die onvoldoende aansluiten op de dagelijkse werkprocessen van datahouders, datagebruikers en burgers leiden tot complexe oplossingen die lastig uitvoerbaar, beheersbaar en uitlegbaar zijn.
- Implicaties:
 - Dualiteit in gebruikersgroepen (perspectief van drie kanten)
 - de datahouders (aanleverproces): het beschikbaar stellen van data moet zo veel mogelijk een geautomatiseerd 'bijproduct' zijn van het primaire (zorg)proces (proces waar de data voor oorspronkelijke doel voor wordt vastgelegd), zonder extra handmatige registratielast.
 - De datagebruiker: Datagebruikers moeten via een centrale/gefedereerde functionaliteit eenvoudig data kunnen vinden, doorzoeken, toegang kunnen aanvragen, en analyses kunnen uitvoeren in voor hen vertrouwde (tools die zij graag gebruiken) en veilige (BVO

geeft bronhouder en burger vertrouwen dat er op de juiste manier met data wordt gewerkt) omgevingen.

- De persoon over wie de gezondheidsgegevens worden hergebruikt: het hergebruik van de gezondheidsgegevens moet voor de burger transparant en controleerbaar zijn. Dit betekent dat er intuïtieve voorzieningen moeten zijn waar burgers welke gegevens voor welke secundaire doelen worden gebruikt, waar burgers eenvoudig hun wettelijke zeggenschap kunnen beheren, en waar de burgers inzicht krijgen in de maatschappelijke opbrengsten van het meervoudige gebruik van de data.

- Geïntegreerde workflow ondersteuning: aanvraag- en goedkeuringsprocedures (waaronder privacy-toetsing, ethische toetsing en data-uitgifte) worden als één digitaal transparant proces ingericht met duidelijke statussen voor de aanvrager.
- Eindgebruikers betrekken bij het ontwerp: architectuurbouwblokken/functionele componenten worden ontworpen op basis van UX-principes (User Experience) en getoetst en getest door feitelijke eindgebruikers (zoals onderzoekers, beleidsmakers, bronhouders en burgers)
- Inrichten van ondersteunende rollen: naast techniek vraagt een gebruikersgerichte aanpak om ondersteunende voorzieningen (denk aan een Health Data Acces Body (HDAB), loketten welke de “no wrong door policy” hanteren) die gebruikers actief begeleiden door het proces.

9. Systeembenadering

Architectuurbeslissingen worden genomen vanuit een integraal perspectief, met expliciete afweging van impact op het totale stelsel. Keuzes moeten dus bijdragen aan het grotere geheel.

Rationale:

- Voorkomen van suboptimalisatie; zorgt ervoor dat decentrale autonomie niet leidt tot fragmentatie. Lokale keuzes en autonomie worden ingevuld binnen gemeenschappelijke stelselafspraken, zodat lokaal gemak niet leidt tot verlies aan doeltreffendheid voor het gehele ecosysteem.
- Ketenafhankelijkheid van datakwaliteit: de geschiktheid van data voor secundair gebruik wordt direct beïnvloed door de manier waar data voor het oorspronkelijke doel (primair) wordt vastgelegd. Zonder systeembenadering ontstaat er een breuk tussen bron en hergebruik.
- Duurzaamheid en schaalbaarheid: de gezondheidszorg heeft te maken met een continu veranderend landschap van secundair gebruik (denk aan nieuwe/veranderende wetgeving, nieuwe analysemethoden, nieuwe databronnen). Alleen met een systeembenadering kan onze architectuur meebewegen zonder telkens opnieuw te moeten worden ontworpen.

Implicaties:

- Integraal architectuurraster (het organisatiebeleid, de processen, de data/informatie, de applicaties en de techniek): keuzes op het technische vlak (bijvoorbeeld een gefedereerd netwerk) mogen niet losstaan van governance, ELSI aspecten en semantische standaarden.
- Ketengerichte governance: er moet ketenbrede en organisatie overstijgende regie en governance worden ingericht voorzien van duidelijke afspraken over eigenaarschap en rollen (bijvoorbeeld RA(S)CI dan wel TBV (Taken Bevoegdheden Verantwoordelijkheden) matrices op alle onderdelen van de keten).
- Impactanalyses bij bronsystemen: Bij veranderingen in de vastlegging van gegevens (primair gebruik, bijvoorbeeld een EPD of nieuw registratiesysteem voor onderzoek, dan wel een aanpassing van een standaard) dient er een impact analyse gedaan te worden op het secundaire landschap.
- End-to-end privacy en security: Vertrouwelijkheid en beveiliging worden niet per koppelvak geregeld, maar via een stelselbreed vertrouwensframework.

10. Internationaal boven nationaal boven lokaal

Internationale standaarden, afspraken en ontwikkelingen zijn leidend boven nationale keuzes ("pas toe of leg uit"). Nationale afspraken zijn op hun beurt leidend boven lokale implementaties. Afwijkingen van internationaal of nationaal vastgestelde standaarden zijn alleen toegestaan wanneer daarvoor een aantoonbare noodzaak bestaat en de afwijking expliciet wordt gemotiveerd. Waar harmonisatie nodig is, wordt deze bij voorkeur georganiseerd op nationaal niveau in plaats van door individuele partijen afzonderlijk.

Rationale:

- Secundair gebruik van gezondheidsdata vindt steeds vaker plaats in een internationale context. Aansluiting op internationale standaarden vergroot de interoperabiliteit en maakt grensoverschrijdende samenwerking, onderzoek en datadeling mogelijk.
- Het volgen van internationaal en nationaal vastgestelde standaarden voorkomt versnippering van oplossingen, begrippen, datamodellen en werkwijzen.
- Gemeenschappelijke nationale voorzieningen, zoals terminologiediensten, mappings of afsprakenstelsels, voorkomen dubbel werk en verminderen beheerlasten voor individuele organisaties.
- Door aan te sluiten op bestaande standaarden kan sneller worden ingespeeld op ontwikkelingen zoals EHDS, THDAS en internationale onderzoeksinitiatieven, zonder telkens lokale aanpassingen of vertaalslagen te hoeven maken.

Implicaties:

- Bij de keuze voor standaarden, databronbeschrijvingen, informatiemodellen, terminologieën en uitwisselingsspecificaties wordt eerst gekeken naar beschikbare internationale standaarden, vervolgens naar nationale afspraken en pas daarna naar lokale invullingen.
- Lokale afwijkingen worden geregistreerd, onderbouwd en periodiek geëvalueerd met als doel convergentie naar de internationale of nationale standaard.
- Harmonisatievraagstukken die meerdere partijen raken, zoals mappings tussen codestelsels of semantische vertalingen, worden bij voorkeur één keer nationaal opgelost en beheerd.
- Nieuwe nationale afspraken sluiten zoveel mogelijk aan op bestaande internationale standaarden en beperken zich tot aanvullingen die noodzakelijk zijn voor de Nederlandse context.
- Leveranciers, datahouders en regievoerende partijen worden geacht actief rekening te houden met ontwikkelingen in relevante internationale standaarden en afspraken.

11. Omgaan met diversiteit

De doelarchitectuur ondersteunt variatie in datatypen, datakwaliteit, volwassenheid en organisatorische context van datahouders, zodat uiteenlopende datahouders kunnen participeren en data beschikbaar kunnen stellen op een wijze die past bij hun mogelijkheden en situatie. Hierbij wordt rekening gehouden met verschillen tussen bijvoorbeeld gecureerde onderzoekscollecties, real world data, ad-hoc dataverzamelingen en verschillende typen zorg- en kennisinstellingen.

Rationale:

- Een nationale infrastructuur kan alleen duurzaam en inclusief functioneren wanneer zij diversiteit in data, context en volwassenheid niet probeert weg te nemen, maar expliciet ondersteunt.
- Het landschap van gezondheidsdata is van nature divers. Data verschillen in herkomst, structuur, kwaliteit, context en mate van standaardisatie.
- Ook datahouders verschillen sterk in schaal, middelen, expertise en technische volwassenheid. Een UMC heeft vaak andere mogelijkheden dan een verloskundige praktijk of zelfstandige zorgaanbieder.

- Een nationale infrastructuur die uitsluitend optimaal werkt voor hoog volwassen organisaties sluit een groot deel van het zorg- en onderzoeksveld uit.
- Door diversiteit als uitgangspunt te nemen wordt een bredere deelname mogelijk, groeit de beschikbare databasis en ontstaat ruimte voor stapsgewijze volwassenwording van deelnemers.

Implicaties:

- De architectuur kent verschillende toetredings- en implementatieniveaus, passend bij de volwassenheid van deelnemers.
- Standaarden en voorzieningen zijn zoveel mogelijk modulair en schaalbaar toepasbaar.
- Minimale eisen voor deelname zijn helder gedefinieerd, waarbij organisaties kunnen doorgroeien naar hogere niveaus van interoperabiliteit en datakwaliteit.
- Metadata en contextinformatie spelen een belangrijke rol om verschillen in kwaliteit, herkomst en betekenis van data transparant te maken.
- Ondersteuning, tooling en richtlijnen worden beschikbaar gesteld om minder volwassen organisaties te helpen aansluiten.
- Architectuurkeuzes worden beoordeeld op hun toepasbaarheid voor zowel grote als kleine datahouders.

12. Maximale benutting van bestaande data (incl. legacy):

Bestaande data, inclusief historische en legacy-data, worden waar mogelijk actief hergebruikt voor onderzoek, beleid en innovatie. De doelarchitectuur stimuleert dat bestaande gegevens niet opnieuw worden verzameld wanneer deze reeds beschikbaar zijn en geschikt kunnen worden gemaakt voor secundair gebruik. Daarbij worden kwaliteit, herkomst, context en juridische randvoorwaarden expliciet meegewogen.

Rationale:

- Versnelling van onderzoek, beleid en innovatie: Het hergebruiken van reeds beschikbare data maakt het mogelijk sneller inzichten te verkrijgen zonder nieuwe dataverzamelingen op te zetten.
- Voorkomen van dubbele registraties en kosten: Het opnieuw verzamelen van bestaande gegevens leidt tot onnodige administratieve lasten, hogere kosten en inefficiënt gebruik van beschikbare middelen.
- Behoud van waardevolle historische informatie: Legacy-data bevatten vaak unieke longitudinale informatie die niet opnieuw kan worden verzameld en essentieel kan zijn voor trends, cohortonderzoek en evaluaties.
- Vergroten van maatschappelijke opbrengst van investeringen: Data die in het verleden met publieke of private middelen zijn verzameld leveren meer maatschappelijke waarde op wanneer deze verantwoord kunnen worden hergebruikt.
- Ondersteuning van het lerend gezondheidszorgsysteem: Het systematisch benutten van bestaande data vergroot de beschikbare kennisbasis en draagt bij aan continue verbetering van zorg, onderzoek en beleid.

Implicaties:

- Hergebruik boven nieuwe dataverzameling: Bij nieuwe toepassingen wordt eerst onderzocht of bestaande databronnen kunnen worden benut voordat nieuwe registraties of datasets worden ontwikkeld.
- Investeren in standaardisatie en harmonisatie: Legacy-data moeten waar nodig technisch, semantisch en organisatorisch geschikt worden gemaakt voor hergebruik.
- Context en herkomst vastleggen: Om bestaande data verantwoord te kunnen hergebruiken moeten metadata, datakwaliteit en contextinformatie expliciet worden vastgelegd en beschikbaar zijn.

- 2531 • Ondersteuning van verschillende volwassenheidsniveaus: De architectuur moet ruimte bieden
2532 aan zowel moderne als oudere gegevensbronnen en een groeipad bieden richting verdere
2533 standaardisatie.
- 2534 • Transparantie over kwaliteit en beperkingen: Datagebruikers moeten inzicht hebben in de
2535 kwaliteit, volledigheid, actualiteit en beperkingen van bestaande datasets voordat deze worden
2536 toegepast.
- 2537 • Voorkomen van verlies van datawaarde: Bij vervanging van systemen of migraties wordt
2538 expliciet rekening gehouden met behoud en herbruikbaarheid van historisch relevante
2539 gegevens.

2540 **13. Wat kan wel**

2541 Richt beslissingen primair op wat mogelijk is binnen de geldende randvoorwaarden, in plaats van op
2542 wat nog niet mogelijk is. Accepteer dat oplossingen of voorzieningen niet vanaf het begin volledig of
2543 perfect hoeven te zijn, mits zij voldoende waarde bieden om verantwoord mee te starten en verder
2544 te ontwikkelen.

2545 **Rationale:**

- 2546 • Versnelt realisatie van maatschappelijke waarde: wachten op een volledig eindbeeld vertraagt
2547 de beschikbaarheid en het gebruik van data.
- 2548 • Complexe veranderingen vragen om stapsgewijze ontwikkeling: niet alle vraagstukken kunnen
2549 vooraf volledig worden opgelost of voorzien.
- 2550 • Leren door te doen: praktijkervaring levert sneller inzicht op in knelpunten, risico's en
2551 verbeterkansen dan uitgebreide theoretische analyses.
- 2552 • Stimuleert samenwerking en betrokkenheid: een oplossingsgerichte houding helpt partijen om
2553 gezamenlijk vooruitgang te boeken, ook wanneer niet alle randvoorwaarden volledig zijn
2554 ingevuld.
- 2555 • Voorkomt dat perfectie een rem wordt op voortgang: een werkbare tussenstap levert vaak meer
2556 waarde dan het uitstellen van besluitvorming tot volledige zekerheid is bereikt.

2557 **Implicaties:**

- 2558 • Voorzieningen, standaarden en processen mogen gefaseerd worden ingevoerd.
- 2559 • Besluitvorming weegt expliciet af welke risico's acceptabel zijn om voortgang mogelijk te maken.
- 2560 • Pilots, proefimplementaties en andere vormen van gecontroleerd experimenteren worden actief
2561 ondersteund.
- 2562 • Groeipaden worden beschreven zodat organisaties kunnen deelnemen zonder direct volledig
2563 aan het eindbeeld te voldoen.
- 2564 • Kwaliteitsverbetering vindt iteratief plaats op basis van ervaringen uit de praktijk.
- 2565 • Belemmeringen worden niet alleen benoemd, maar altijd gecombineerd met een verkenning
2566 van mogelijke oplossingsrichtingen.