



Referentie Architectuur

Stelsel Digitale Tweeling Fysieke Leefomgeving

Versie 0.9

datum

11 februari 2022

Opstellers

Werkgroep stelsel architectuur: Guido Bayens, Jan Buijn, Eric van Capelleveen, Jan van Gelder, Michel Grothe, Bart de Lathouwer, Walter Lohman, Marcel Reuvers, Frans de Waal, Ton Zijlstra

Inhoudsopgave

Samenvatting	3
1 Inleiding	5
1.1 Begrippen	5
1.2 Toepassingsgebied	7
1.3 Doelgroep	8
1.4 Werken met Bouwblokken	8
2 Klantperspectief	10
2.1 Een gemeenschappelijke informatiebasis	11
3 Leidende principes	13
3.1 Doelgericht op het algemeen welzijn	14
3.2 Actief open en transparant	14
3.3 Eerlijk kwalitatief en inclusief	15
3.4 Robuust en veilig	15
3.5 Onder controle en toezicht van mensen	16
3.6 Belanghebbenden intensief betrokken	17
4 Architecturale ontwerpprincipes	18
4.1 Inleiding	18
4.2 De architecturale ontwerpprincipes	19
4.3 Data als een product	23
4.4 Selfservice data-infrastructuur als platform	23
4.5 Data soevereiniteit	25
4.6 Service oriëntatie.	26
4.7 Open uitwisselingsstandaarden	26
4.8 Kwaliteit data	26
5 Constructieprincipes	27
5.1 Inleiding	27
5.2 Specificatie van constructieprincipes	27
6 Stelselafspraken	29
6.1 Inleiding	29
6.2 Wat verstaan we onder 'het stelsel'?	29
6.3 Betrokkenen, afspraken en rollen	29
6.4 Overleggen	30
6.5 Stelselafspraken	30
7 Verantwoording	33
7.1 Over de opstellers	33
8 bijlagen	34
8.1 Digitale tweeling en het data mesh concept	34
8.2 Het data mesh concept	34

Samenvatting

Dit document beschrijft de referentie-architectuur voor het stelsel van digital twinning voor de fysieke leefomgeving. Werken met digitale tweelingen (digital twins) begint een grote vlucht te nemen bij het ontwerpen en besturen van onze fysieke leefomgeving. Digitale Tweelingen bieden onder meer virtuele stadsreplica's die het gemakkelijk maken om de complexe onderlinge relatie tussen verkeer, luchtkwaliteit, lawaai en andere stedelijke factoren te begrijpen. Krachtige analyses leiden tot modellen van de verwachte impact van potentiële veranderingen. Zij helpen om betere data gedreven operationele besturing en beleidskeuzes voor de langere termijn te maken.

Digitale tweelingen maken gebruik van zeer uiteenlopende typen data en complexe rekenmodellen, welke door een digitale tweeling gecombineerd worden tot vooral drie dimensionale modellen en statische of dynamische gegevens die iets zeggen over de voormalige, huidige of toekomstige situatie.

Om de diversiteit aan gegevens afkomstig van een grote variatie aan bronnen uit te kunnen wisselen is het gewenst afspraken te maken over de wijze waarop dit plaatsvindt. Deze afspraken hebben betrekking op de aard en vorm van data, de wijze waarop deze via netwerken tussen partijen kunnen worden uitgewisseld, onderdelen van de uitwisselingsinfrastructuur en de governance van dit geheel.

Dit document beschrijft niet de architectuur van een individuele digitale tweeling, noch de wijze waarop modellen en schermen aan gebruikers worden gepresenteerd. Er wordt niet ingegaan op de werking van digitale tweelingen zelf. Dit document bevat ook niet een blauwdruk voor een "digitale tweeling van Nederland". Dat zou ondoenlijk zijn. Digitale tweelingen worden ontwikkeld voor een specifiek doel, een stad, een wijk, een verkeerssysteem of een landbouwgebied. Dit document is dus alleen gericht op de uitwisseling van data om daarmee de interoperabiliteit tussen de vele afzonderlijke digitale tweelingen te stroomlijnen.

De architectuur van het stelsel is gebaseerd op het idee van een 'system-of-systems'; systemen die met systemen samenwerken. Daarbij wordt uitgegaan van twee belangrijke architecturale principes:

- Data blijft bij de bron¹;
- Er is een eenvoudige federatieve structuur die de uitwisseling van data mogelijk maakt.

Met andere woorden: De architectuur is gebaseerd op een combinatie van decentrale dataopslag, -verwerking en -toepassing, aangevuld met een federatieve structuur om de uitwisseling van data en dataproducten mogelijk te maken. De belangrijkste federatieve componenten van het stelsel zijn: datanetwerk(en), centrale indexen voor het vinden van de juiste dataverstreckers (*registry*) en voorzieningen voor identiteits- en toegangsmanagement (*dentitie & access control*). In aanvulling hierop worden principes en standaarden beschreven die betrekking hebben op onder meer de aard van de dataproducten, metadatering en het uitwisselingsformat en -protocol (gebaseerd op API's).

De beschreven architectuur is "open". Dit document en de daarin vastgelegde principes en zoveel mogelijk van de genoemde standaarden kunnen beschouwd worden als publiek goed. Toepassing van deze principes en standaarden zijn niet aan restricties met betrekking tot intellectueel eigendom gebonden. Daarom wordt ook deze referentie architectuur gepubliceerd onder de voorwaarden van de Creative Commons Attribution 4.0 International Public License (CC-BY)².

Het ontwerp van het stelsel is gebaseerd op de wens om te komen tot het optimale inzet van digitale tweelingen voor allereerst de "publieke zaak". Daarbij spelen publieke waarden een belangrijke rol. Hieronder vallen zaken als het honoreren van ethische normen, inclusiviteit en bescherming van privacy en veiligheid. Van daaruit zijn ontwerpprincipes opgesteld voor de inrichting van het stelsel.

Bij het opstellen van deze referentie architectuur is gebruik gemaakt van inzichten die zijn ontwikkeld in andere programma's en samenwerkingsverbanden op mondiaal-, EU- en nationaal niveau. Belangrijk voorbeelden hiervan zijn het EU-programma Inspire, het EU-programma Gaia-X, het Digital Urban European

¹ Strikt genomen is deze formulering niet altijd houdbaar. We komen hierop later in dit document terug.

² <https://creativecommons.org/licenses/by-sa/3.0/>

Twins initiatief (DUET), International Data Spaces (IDS), het Digital Twin Consortium en de Nederlandse Overheid Referentie Architectuur.

- De Inspire-richtlijn heeft tot doel een infrastructuur voor ruimtelijke gegevens van de Europese Unie te creëren ten behoeve van het milieubeleid en het milieubeleid of de activiteiten van de EU die gevolgen kunnen hebben voor het milieu.
- DUET is een Europees innovatie-initiatief dat gebruik maakt van de geavanceerde mogelijkheden van cloud-, sensorgegevens en analyses in de vorm van Digitale tweelingen, om de besluitvorming in de publieke sector democratischer en effectiever te maken³.
- IDS, een internationaal samenwerkingsverband van ruim honderdtwintig bedrijven en instellingen, streeft naar samenwerking en data-uitwisseling op domeinniveau, *data spaces* genoemd. Het programma Digitale Twin voor de Fysieke Leefomgeving, is gericht om precies dat te doen voor het fysieke leefdomein in Nederland.
- Gaia-X, een internationaal samenwerkingsverband met ruim 300 deelnemende organisaties, heeft zich als doel gesteld om de uitwisseling van data tussen 'data spaces' te faciliteren. Een term als interoperabiliteit staat hierbij centraal. De bedoeling is dat er mechanismes ontwikkeld worden die diensten van deelnemende leveranciers vinden, verbinden en combineren

Een belangrijk architecturale keuze betreft het zogenaamde 'data mesh principe'. Data mesh kent een viertal hoofdprincipes zoals het eerdergenoemde uitgangspunt van het combineren van decentrale dataopslag en verwerking en een federale uitwisselingsinfrastructuur. Hieraan wordt het idee van 'data als een product' toegevoegd. Hierbij zien we het al lang bestaande idee van service oriented architecture terug, maar dan in de vorm van het aanbieden van min of meer kant-en-klare, herbruikbare datasets, inclusief een uitgebreide set van meta-data, die door geautoriseerde afnemers kunnen worden afgenomen. Dit impliceert dat geen toegang wordt verleend tot interne datasystemen van aanbieders van data, maar dat deze hun dataproducten als het ware in de etalage zetten, geschikt voor potentiële afnemers ervan. Het resultaat van deze uitgangspunten is een stelsel waarbij deelnemers via federatieve indexen ('wegwijzers') hun weg kunnen vinden naar aanbieders van dataproducten, welke zij onder bepaalde condities mogen afnemen. Hiermee is de kern van het stelsel beschreven.

Naast technische voorzieningen en afspraken over standaarden, dienen voor een goede werking van het stelsel eveneens afspraken gemaakt te worden over de governance van het stelsel en de 'spelregels' om er deel van te kunnen uitmaken. Deze stelselafspraken worden in dit document benoemd, maar niet in detail uitgewerkt.

N.B.: In *deze* versie van dit document ontbreekt de concrete invulling van hoofdstuk 5, waarin de constructieprincipes zullen worden opgenomen.

³ DUET: D5.1 System Architecture & Implementation Plan

1 Inleiding

Deze referentie architectuur bevat afspraken en principes voor een digitaal twinstelsel voor de fysieke leefomgeving. De fysieke leefomgeving kent drie invalshoeken: de gebouwde omgeving, de natuurlijke omgeving en de sociale omgeving. Deze referentie architectuur is bedoeld voor iedereen die een rol speelt bij het tot stand brengen van een digitale tweeling of digitale tweeling voor de fysieke leefomgeving. De afspraken, uitgangspunten, principes en standaarden die in dit document worden gepresenteerd, dragen bij aan het tot stand brengen van een stelsel dat bestaat uit vele toepassingen van het idee van digitale tweelingen.

Dit document bevat een "open" architectuur. Dat wil zeggen: de architectuur kan door iedereen, overheid en bedrijfsleven, toegepast worden. Met de afspraken, principes en standaarden uit dit document beogen we een bijdrage te leveren aan het delen van data en modellen afkomstig uit uiteenlopende sectoren en kennisgebieden. Daarmee komen allerlei zinvolle toepassingen in de vorm van driedimensionale beelden, beslismodellen, simulaties of realtime besturingen van operationele processen binnen handbereik.

U leest de eerste versie van deze referentie architectuur. Het onderwerp 'digital twinning' is relatief nieuw⁴. Op veel fronten wordt gewerkt aan de verdere ontwikkeling van het idee van digital twinning. Daardoor bestaan momenteel vaak nog vele oplossingen naast elkaar en is er nog geen dominante aanpak of zijn er nog geen dominante standaarden die de interoperabiliteit eenduidig mogelijk maken. Anderzijds zijn er ook nog onopgeloste vraagstukken. Tegen deze achtergrond is deze referentie architectuur opgesteld; zoekend naar zo breed mogelijk gedragen uitgangspunten, principes en standaarden voor het ontwikkelen van digitale tweelingen. Zoekend naar een samenhangend pakket van principes en standaarden en daarmee soms ook bepaalde principes en standaarden uitsluitend. De keuzes die in dit document gemaakt worden, kunnen dan ook discussie uitlokken. Dat past bij de fase waarin deze ontwikkeling zich bevindt. Na deze eerste versie zullen dan ook nieuwe versies volgen, die ons weer een stap dichterbij een compleet stelsel voor digital twinning van de fysieke leefomgeving zullen brengen.

1.1 Begrippen

Voor een goed begrip van dit document, beginnen we met een nadere toelichting op enkele centrale begrippen. We beginnen met een definitie van het idee van een digitale tweeling.

Digital Twin Consortium heeft de volgende definitie geformuleerd⁵:

Een digital twin of digitale tweeling is een virtuele weergave van entiteiten en processen uit de echte wereld, gesynchroniseerd met een bepaalde frequentie en betrouwbaarheid.

- *Digital twin-systemen transformeren bedrijven door holistisch begrip, optimale besluitvorming en effectieve actie te versnellen.*
- *Digitale tweelingen gebruiken real-time en historische gegevens om het verleden en het heden weer te geven en toekomstige situaties te simuleren.*
- *Digitale tweelingen zijn gericht op het behalen van resultaten die voortvloeien uit use cases door gebruik te maken van gegevens vanuit specifieke domeinen, welke worden geïntegreerd door middel van IT-systemen.*

De ontwikkeling van digitale tweelingen verloopt snel. Wereldwijd zijn er vele honderden initiatieven waarin digitale tweelingen worden ontwikkeld of reeds in gebruik zijn. Kennisinstituten, bedrijven en overheidsinstellingen ontwikkelen de benodigde technologie en streven naar standaardisatie van belangrijke

⁴ In dit document hanteren we het begrip 'digital twinning' om het werken met digitale tweelingen binnen een stelsel aan te duiden.

⁵ Het Digital Twin Consortium is een internationaal gezelschap van bedrijven, kennisinstituten en overheidsorganisaties. Zie: <https://www.digitaltwinconsortium.org/index.htm>

componenten en aspecten van digital twinning. Ook in Nederland zijn alleen al binnen de overheid ruim 100 initiatieven⁶.

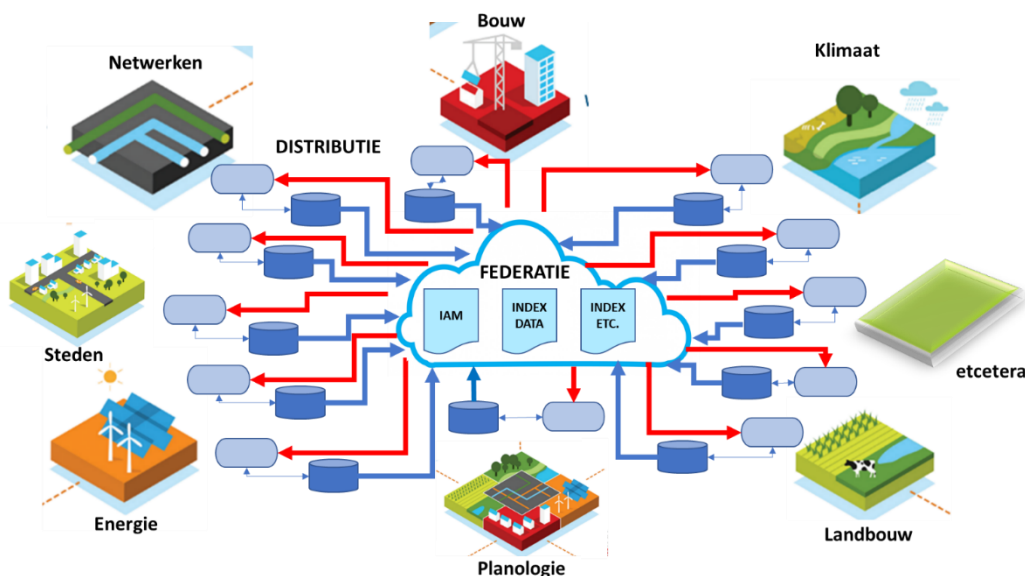
Digitale tweelingen worden ontwikkeld voor uiteenlopende toepassingen en domeinen, zoals industriële productie (SMART Industry of Industry 4.0⁷), ruimtevaart, defensie, gezondheidszorg, veiligheid enzovoorts. Deze referentie architectuur beperkt zich tot digitale tweelingen voor de fysieke leefomgeving, waarop we later in deze inleiding dieper ingaan.

Bij het inzetten van digitale tweeling technologie voor de fysieke leefomgeving komen uiteenlopende typen informatie samen, zoals geo-informatie en informatie over milieu, infrastructuur, gebouwen en dynamische informatie zoals verkeer, waterlopen, energiestromen, etc. Deze informatie wordt door uiteenlopende organisaties verzameld en bewerkt, zoals het Kadaster, gemeenten, provincies, waterschappen en private en publieke infra-beheerders voor verkeer, vervoer, energie, riolering en dergelijke. Zij allen hebben behoefte om gegevens van elkaar te gebruiken om meer holistische modellen en visualisaties samen te stellen. Digitale tweelingen richten zich niet alleen op de vastlegging van de fysieke leefomgeving, maar met name op processen in die fysieke leefomgeving: denk bijvoorbeeld aan verkeersstromen, stikstof distributie, geluidsbelasting of de drukte in de stad. Daarvoor is niet alleen statische data over de fysieke leefomgeving nodig, maar juist ook dynamische data (waarnemingen uit sensoren) over hoe die processen zich afspelen in de tijd en ruimte.

Om gegevens efficiënt en effectief te kunnen delen dient bij voorkeur een goed uitwisselingsstelsel tot stand te komen. Digitale tweelingen zijn tegelijk 'consumenten' en 'producenten' van data. Het geheel van afspraken en voorzieningen die hiervoor nodig zijn noemen we het uitwisselingsstelsel of kortweg stelsel. Deze referentie architectuur bevat ontwerpprincipes, standaarden en afspraken voor het stelsel voor dit uitwisselingsstelsel. We spreken ook wel over het stelsel voor digital twinning.

Kenmerkend voor dit stelsel is een decentrale verwerving, opslag en bewerking van gegevens in combinatie met een lichte, federatieve structuur voor de uitwisseling van gegevens en modellen.

Figuur 1 geeft een gestileerde weergave van dit uitgangspunt. Diverse deelnemers aan het stelsel verzamelen, beheren en bewerken gegevens en deze worden in de vorm van informatieproducten via netwerken (internet) uitgewisseld. Enkele federatieve voorzieningen, zoals indexen voor het snel vinden van de gewenste gegevens en het zorgen voor een identiteit- en toegangscontrole, maken het uitwisselen van gegevens mogelijk.



Figuur 1 stelselarchitectuur voor Digital Twinning

⁶ Zie het rapport Landschapsanalyse, Digitale Tweelingen van de Fysieke Leefomgeving, Geonovum, december 2021.

⁷ https://nl.wikipedia.org/wiki/Smart_Industry

1.2 Toepassingsgebied

Deze referentie architectuur beperkt zich tot de fysieke leefomgeving (gebouwde, natuurlijke en sociale omgeving). Daarbij wordt onder meer gedacht aan de gebouwde omgeving, ruimtelijke ordening, verkeersstromen, milieu (samenstelling van lucht, grond en water), geluid, temperatuur, openbaar vervoer, watersystemen en infrastructuren voor verkeer (wegen, bruggen, viaducten, tunnels, rail), gas (aardgas, waterstof, industriële gassen), warmte, water, riolering, datacommunicatie, begroeiing en bevolking (aantallen, demografie).

Digitale tweelingen kunnen bijdragen aan het in kaart brengen van de situatie op deze (combinatie van) domeinen in het verleden, het heden of de toekomst. Het kan gaan over statische of dynamische gegevens. Inzet van een digitale tweeling kan onderzoekers en beleidsmakers ondersteunen bij het maken van afwegingen in het ontwerp van aspecten van de fysieke leefomgeving, zoals het ontwerp van nieuwe woonwijken, infrastructuur of het doorrekenen van effecten van maatregelen in het kader van de energietransitie of de levensomstandigheden in een bepaald gebied. Digitale tweelingen kunnen ook een krachtig hulpmiddel zijn bij het dagelijks sturen van processen, zoals verkeersstromen of het opvangen van negatieve gevolgen van voor ons land te verwachten overvloedige regenval.

Binnen de genoemde domeinen zijn vaak al uitgebreide databestanden en modellen op basis van meetgegevens of te evalueren aannames voorhanden. Deze data en modellen worden in beschrijvende zin gebruikt of om directe informatie te krijgen over een bepaalde situatie om daaraan sturing te kunnen geven. Modellen kunnen ook gebruikt worden om "what if" analyses of simulaties uit te kunnen voeren waarmee betrouwbare voorspellingen kunnen worden gedaan.

Er is dus al veel en er liggen vaak verschillende disciplines en uitgangspunten aan ten grondslag. Digital twinning is te zien als een volgende stap in deze ontwikkeling. Het gaat om het bij elkaar brengen van data en modellen om op die wijze een multidimensionaal model van (een deel van) de werkelijkheid samen te stellen. Een voorbeeld: Wat is het effect op de verkeersdoorstroming van een tijdelijke wegafsluiting en wat betekent dit voor de milieubelasting (CO₂-uitstoot, geluid) van de alternatieve routes waarvan het verkeer gebruik zal gaan maken en welke aanpassingen in het verkeersbegeleidingssysteem zullen hierbij nodig zijn om een verkeersinfarct te voorkomen? Wat zijn de tijdelijke gevolgen voor de bewoners van het stadsdeel waarheen het verkeer zich zal verplaatsen? De gecombineerde inzet van rekenmodellen biedt inzicht in dergelijke complexe vraagstukken.

Deze referentie architectuur bevat daarom principes (lees: afspraken) en standaarden die relevant zijn voor de uitwisseling van data en modellen tussen verschillende domeinen. Het mogelijk maken van uitwisseling van data en modellen tussen meerdere organisaties noemen we *digital twinning*. Het geheel van afspraken en voorzieningen om deze uitwisseling mogelijk te maken noemen we het *stelsel voor digital twinning*.

De fysieke leefomgeving, nader uitgewerkt

Naast de digitale tweeling is ook de 'fysieke leefomgeving' één van onze sleutelwoorden. De fysieke leefomgeving bepaalt de werkingssfeer van de digitale tweeling waar wij het over hebben. Dat betekent dat onze digitale tweelingen over onze fysieke leefomgeving gaan. Een duidelijke afgebakende definitie van de fysieke leefomgeving is er niet. De fysieke leefomgeving is de omgeving, waarin wij als mensen leven, wonen, werken, reizen, recreëren, et cetera. Wij voeren onze activiteiten uit in de fysieke leefomgeving, die bestaat uit allerlei onderdelen: bouwwerken, infrastructuur, water, bodem en ondergrond, lucht, natuur, landbouw, landschappen en cultureel erfgoed. Deze opsomming is natuurlijk niet uitputtend. En soms wordt ook wel over de natuurlijke leefomgeving gesproken als het gaat om onderdelen als lucht, water, bodem en natuur (het milieu). En de door de mens gemaakte objecten horen bij de fysieke leefomgeving. Objecten kunnen statisch zijn, omdat ze langdurig op dezelfde plek staan. Zo vallen bouwwerken (onze woningen, kantoren en fabrieken) en infrastructuur (wegen, kanalen en kabels en leiding) onder de fysieke leefomgeving. Maar ook objecten die meer dynamisch zijn, zoals voertuigen. Maar ook de landbouwgewassen zijn meer dynamisch, omdat ze maar enkele maanden aanwezig zijn; ze komen en gaan. Het wordt hoog dynamisch als we het hebben over de veranderlijkheid van het weer, de stroming van het water of onze eigen verplaatsingen in het dagelijks leven voor het wonen, werken

en recreëren. Wij als mensen maken daarmee ook deel uit van de fysieke leefomgeving. Met onze activiteiten in de fysieke leefomgeving maken wij mede diezelfde leefomgeving. Die activiteiten zorgen voor de dynamiek in de leefomgeving.

Onze fysieke leefomgeving is een complex, zich continue aanpassend systeem, dat bestaat uit een aantal deelsystemen die nauw op elkaar inwerken. De fysieke leefomgeving is daarmee een connectie en samenwerking tussen deze deelsystemen. Vaak worden drie deelsystemen benoemd, die te typeren zijn enkele naar gemeenschappelijke inzichten:

- Economische deelsysteem;
- Natuurlijke deelsysteem; en
- Sociale deelsysteem.

Het economische deelsysteem van de fysieke leefomgeving heeft betrekking op de economische activiteiten en processen in de ruimte, zoals industrie, bouw en handel, de land- en tuinbouw, transport(netwerken), waterwinning, de afvalstromen (en circulaire economie) en niet te vergeten de energiesector. In het natuursysteem vinden de natuurlijke verschijnselen en processen plaats, zoals de natuurgebieden en biodiversiteit, het water(systeem), klimaat en lucht, bodem en ondergrond. Ook wel het 'milieu' genoemd. Het sociale systeem, ook wel sociaal-culturele systeem, bestaat uit de sociale aspecten en interacties in de fysieke leefomgeving. Denk daarbij aan huisvesting (de woningen), het zorgsysteem (zorgcentra, ziekenhuizen, verzorgingshuizen), het onderwijs (scholen en universiteiten), de voorzieningen voor sport, cultuur en recreatie en natuurlijk de openbare ruimte. Ook ons ruimtelijk gedrag valt hieronder.

Deze drie deelsystemen tonen een grote overeenkomst met het 3P-model van duurzame ontwikkeling, nl. People-Profit-Planet. De combinatie van mensen, milieu, en winst maken dient in harmonie te zijn. Wanneer dit niet zo is, zullen de andere elementen hieronder lijden. Wanneer bijvoorbeeld winst maken te veel prioriteit krijgt, zullen mens en milieu hieronder leiden, bijvoorbeeld via de uitstoot van gevaarlijke stoffen (effect op de volksgezondheid) of door achteruitgang van biodiversiteit. Het woord 'profit' is inmiddels vervangen door 'prosperity' oftewel welvaart. Naast economisch gewin dient ook de maatschappelijke winst in de afwegingen mee te worden genomen. Zo dienen ook de drie deelsystemen van de fysieke leefomgeving in evenwicht te zijn. En ook hier speelt welvaart een toenemende rol. Het concept van het 'brede welvaart denken' staat daarbij steeds meer centraal; "Brede welvaart gaat over de welvaart en welzijn van individuele mensen. Dat wil zeggen: de kwaliteit van leven in het hier en nu en hoe zich dat verhoudt tot latere generaties of die van mensen elders in de wereld."⁸ Door brede welvaart te hanteren als uitgangspunt voor het afwegen van belangen bij ontwikkelingen en veranderingen in de fysieke leefomgeving, kunnen ook de afwegingen tussen de sociaal-culturele, natuurlijke en economische deelsystemen van de fysieke leefomgeving worden meegenomen.

1.3 Doelgroep

Dit document is gericht op iedereen die betrokken is bij het ontwikkelen van vormen van digital twinning binnen het fysieke domein. Dit betreft uiteraard bedrijfs-, informatie-, data- en infrastructuur-architecten, maar ook model- en dataspecialisten, ontwikkelaars, user interface- en beveiligingsexperts en beheerders. Voor personen die iets verder van de techniek afstaan zijn alleen de hoofdstukken over het klantperspectief, de publieke waarden en de stelselafspraken van belang.

1.4 Werken met Bouwblokken

Naast deze Referentie Architectuur is een document opgesteld dat de principes en standaarden uit deze referentie architectuur vertaalt in 'bouwblokken'. Met een bouwblok wordt een functioneel onderdeel van

⁸ <https://www.scp.nl/actueel/nieuws/2021/02/12/plan-van-aanpak-planbureaus-om-brede-welvaart-te-verankeren-in-nederlandse-begroting>

een digitale tweeling nader ingevuld. Met deze bouwblokken kan een specifieke digitale tweeling-oplossing worden samengesteld. Functies waarin deze bouwblokken kunnen voorzien zijn bijvoorbeeld: informatiebeveiliging, datatoegang, databewerking, metadatering, modellering, simulatie, voorspellen, visualisatie en rekenen.

Deze bouwblokken kunnen meerdere doelen dienen:

- Basis voor het vaststellen van het doel van een specifieke digitale tweeling (= wat wil je ermee doen?);
- Basis voor een ontwerp van een specifieke digitale tweeling;
- Basis voor het maken van een correcte aansluiting op het digitale tweeling stelsel;
- Basis voor het opstellen van een programma van eisen bij eventuele aankoop van (onderdelen van) een digitale tweeling;
- Basis voor het vergelijken van alternatieve invullingen van een bepaalde functie;
- Basis voor de werkverdeling tussen diverse specialisten, die zich elk richten op een bepaalde functie van de digitale tweeling.
- Basis voor het digitale tweeling ecosysteem als geheel. Overkoepelende generieke bouwstenen zijn nodig om de afzonderlijke digitale tweelingen met elkaar te kunnen verbinden, beveiligde data uitwisseling, interoperabiliteit, het kunnen zoeken en toegang voor gebruikers ondersteunen.

Door het Geonovumprogramma "Digital Twin voor de Fysieke Leefomgeving" is een afzonderlijk document opgesteld waarin de bouwblokken worden beschreven.

2 Klantperspectief

Nederland staat voor grote maatschappelijke opgaven die ingrijpende gevolgen hebben voor de inrichting van onze leefomgeving. Deze opgaven hangen vaak onderling nauw samen. We geven wat voorbeelden:

- We willen ons energiesysteem verduurzamen, wat aanpassingen in de energie-infrastructuur vergt.
- Het mobiliteitssysteem van snelwegen en vaarwegen loopt tegen capaciteitsgrenzen aan. Deze infrastructuur is voor een belangrijk deel aan vervanging toe, met effecten op ruimte, luchtkwaliteit, klimaat, verkeersveiligheid en de leefbaarheid in steden.
- Er is behoefte aan de bouw van een miljoen woningen, waarvoor locaties bepaald moeten worden.
- De landbouw en mobiliteit vraagt om een transitie onder andere vanwege de stikstofuitstoot.
- Het klimaat verandert, waardoor rekening gehouden moet worden met periodes van extreme droogte en hevige neerslag, stijging van de zeespiegel en verzilting van de bodem.

Naast hun onderlinge verwevenheid kenmerken deze opgaven zich door urgentie en tijdsdruk. Ze zijn bovendien complex en doen een beroep op datgene wat in Nederland schaars is: de fysieke ruimte. Om deze in onderlinge samenhang en complexiteit aan te pakken zijn vele partijen aan zet: overheden, burgers, kennisinstellingen en het bedrijfsleven.

Voor ieder individueel maatschappelijk vraagstuk is het nodig om de samenwerking tussen de genoemde partijen vroegtijdig te faciliteren met een gemeenschappelijke en betrouwbare informatiebasis. Deze informatie is bij meerdere partijen aanwezig en dient – onder bepaalde voorwaarden – uitgewisseld te kunnen worden tussen de genoemde partijen. Afzonderlijke partijen werken in toenemende mate met digitale tweelingen⁹. Het stelsel voor digital twinning maakt uitwisseling van informatie en daarmee samenwerking mogelijk.

Het werken met en combineren van data vraagt in toenemende mate ook om een wettelijke basis. Veel wetgeving is al tot stand gekomen, zoals de Omgevingswet, de Wet Open Overheid; Wet Hergebruik Overheidsinformatie, de Uitvoeringswet Algemene verordening gegevensbescherming (AVG), de AI-act; AI-Verordening; Wet Elektronische Publicatie, etc. Ongetwijfeld zal deze juridische onderbouwing van het omgaan met gegevens nog verder doorgevoerd worden.

Naast de wettelijke basis zitten er ook ethische kanten aan het gebruik van data en modellen. Digitale tweelingen moeten een zo betrouwbaar mogelijk beeld geven van de werkelijkheid en mogelijke uitkomsten van ingrepen. Het gebruik van data en modellen dient daarom verantwoord te gebeuren. Transparantie in het gebruik van data en modellen is voor alle betrokken partijen van belang om het resultaat te kunnen vertrouwen. In hoofdstuk 3 wordt nader ingegaan op deze materie.

De genoemde maatschappelijke opgaven vragen om een nauwe betrokkenheid van belanghebbende burgers en bedrijven. Het is daarom van groot belang dat organisaties die zich bezighouden met het oplossen van deze vraagstukken ook burgers en bedrijven via betrouwbare informatie betrekken bij het zoeken naar oplossingen en de ontwikkeling van beleid en wetgeving. Digitale tweelingen maken dit mogelijk. Zij zijn gebaseerd op digitale modellen, die onder meer driedimensionaal kunnen worden weergegeven en daarom veel inzicht bieden. Deze **verantwoordelijkheid** ligt bij de organisaties die gebruik maken van digitale tweelingen als instrument en valt daarmee buiten de scope van deze referentie architectuur. Toch willen we hierop in de volgende paragraaf wat dieper ingaan.

⁹ Alleen al in het overheidsdomein werden eind 2021 ruim honderd digitale tweeling ontwikkelingen geïnventariseerd.

2.1 Een gemeenschappelijke informatiebasis

Verschillende coalities van overheden, bedrijven en kennisinstellingen werken reeds met een gemeenschappelijke informatiebasis om de complexiteit en verwevenheid van maatschappelijke opgaven het hoofd te bieden. Ze gebruiken daarbij het instrument van de digitale tweeling. Dit concept is afkomstig uit de industrie waarin een exacte digitale kopie van een product, proces of systeem wordt ingezet voor productie- en de onderhoudscyclus. Met de digitale tweeling van de fysieke leefomgeving (DTFL) bedoelen we een digitale representatie van zowel stedelijke als landelijke omgeving, waarin scenario's gecreëerd kunnen worden op basis van statische en dynamische data, modellen en visualisaties. Een digitale tweeling:

1. verbindt betrokkenen rondom de maatschappelijke uitdaging;
2. helpt de huidige en gewenste fysieke leefomgeving te beschrijven en te verbeelden met alle relevante data middels visualisaties in onder meer twee- en driedimensionale modellen.
3. stelt betrokkenen in staat om de gevolgen van ingrepen door te rekenen, te voorspellen en te simuleren;
4. biedt inzicht zodat men sneller komt tot gedragen oplossingen en besluitvorming komt;
5. ondersteunt de operationele besturing van uiteenlopende processen zoals verkeersstromen, waterstromen, energiestromen en dergelijke.
6. ondersteunt beheerders bij een efficiënt beheer en onderhoud van fysieke objecten door effectieve monitoring.

Digitale tweelingen worden ingezet om de complexiteit van de maatschappelijke opgaven het hoofd te bieden. Het is een veelbelovend instrument, dat op grote schaal toegepast kan worden. Dat vraagt van de overheid om te zorgen dat deze technologie betrouwbaar is en dat navolgbaar welke data en modellen zijn toegepast en volgens welke spelregels bouwstenen worden toegevoegd.

Onze stip op de horizon is dat de technologie van de digitale tweeling op nationale schaal gebruikt wordt, als publiek instrumentarium door overheden, burgers, bedrijven en kennisinstellingen, voor het verkennen van maatschappelijke uitdagingen in de fysieke leefomgeving en het ontwerpen en ontwikkelen van oplossingsscenario's. We noemen dit instrument de Digitale Tweeling voor de Fysieke Leefomgeving.

Het gebruik van de DTFL, 'Digital Twinning' geheten, maakt gebruik van het inzicht dat een digitale tweeling biedt om tot gedragen besluiten te komen en deze te vertalen in oplossingen welke gedurende hun levenscyclus worden gemonitord. Door deze technologie en kennis op nationale schaal toe te passen met een publiek instrument worden oplossingen slimmer, sneller en socialer.

Daarom werken we aan de nationale DTFL. Dit is niet één systeem, maar bestaat uit drie componenten:

1. De verzameling van regionale, thematische dan wel stedelijke DTFL's die ontwikkeld zijn voor een specifiek maatschappelijk vraagstuk in de fysieke leefomgeving. DTFL's kunnen herhaalbaar toegepast worden op een andere locatie en nieuwe DTFL's kunnen gebruik maken van reeds bestaande DTFL's of delen daarvan. Indien relevant kunnen DTFL's ook met elkaar verbonden worden. Zodoende wordt de kennis over de reeds gevonden oplossingen maximaal gedeeld.
2. Het stelsel van digital twinning maakt het delen van brondata, modellen en visualisaties mogelijk. Het stelsel van digital twinning biedt toegangswegen naar allerhande bronnen en applicatieve services en koppelvlakken en zorgt zo voor een betrouwbare nationale ontsluiting van de ingrediënten voor een DTFL.
3. De praktijk van Digital Twinning wordt samengebracht in het DTFL-ecosysteem ofwel het stelsel voor digital twinning. Het DTFL-stelsel biedt de componenten rondom de ontwikkeling en het gebruik van een DTFL zoals een set voorwaardelijke standaarden, keurmerken en instrumenten om betekenisvol, voorspelbaar en transparant gebruik te borgen binnen een context van spelregels voor publieke waarden. Zo wordt Digital Twinning op nationale schaal een feit.

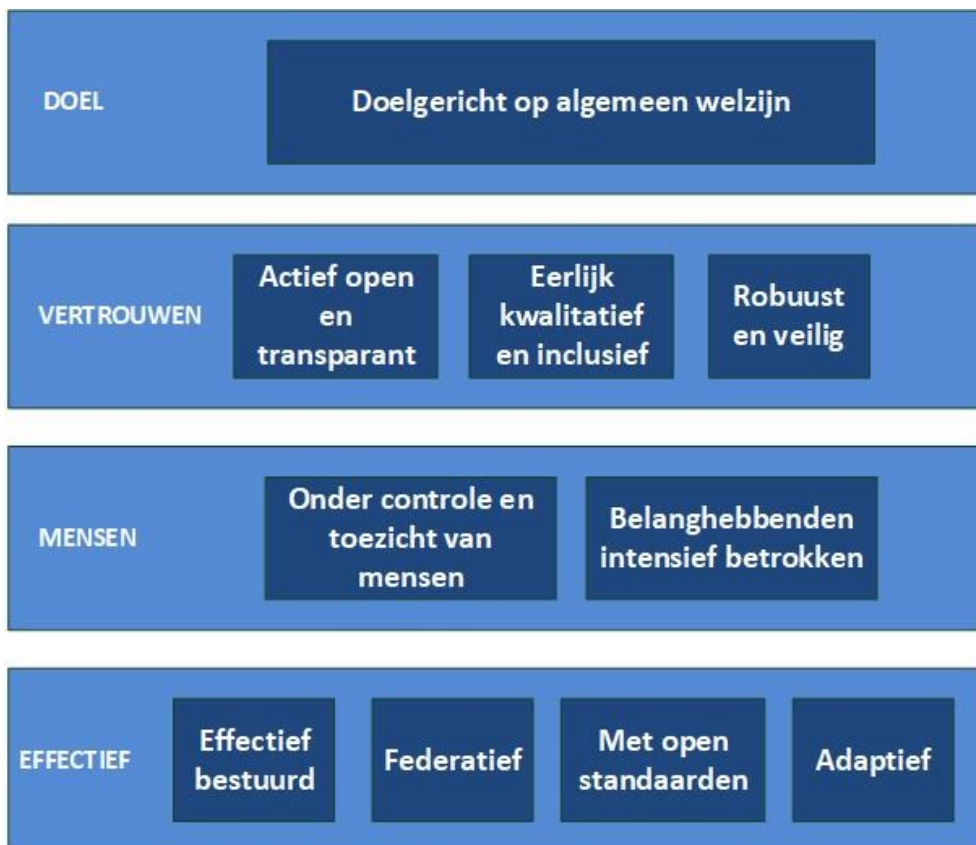
Deze referentie architectuur gaat over het tweede onderdeel van de nationale DTFL: de architectuur voor het stelsel van digital twinning. Het delen en gebruiken van zowel publieke als private data en modellen vraagt om spelregels. Om hiertoe te komen zal de overheid het voortouw nemen om met publieke en private partijen tot afspraken te komen die gebaseerd zijn op publieke waarden. We voorzien dat op dit punt

verschillende data-deel-coalities zullen ontstaan voor de ontsluiting van publieke en private data en modellen rondom bepaalde informatiedomeinen. Met het stelsel voor digital twinning voorzien we in afspraken en enkele fysieke onderdelen, waardoor digitale tweelingen aan elkaar verbonden kunnen worden. Het stelsel staat onder voorwaarden open voor zowel publieke als private partijen. Zij kunnen zowel de rol van leverancier als afnemer van informatie spelen. De voorwaarden waaraan zij moeten voldoen, worden in het laatste hoofdstuk van deze referentie architectuur kort genoemd.

3 Leidende principes

In deze paragraaf staan belangrijke uitgangspunten, die van toepassing zijn op het realiseren en gebruik van digitale tweelingen en het digitale tweeling stelsel. Deze uitgangspunten zijn de leidende principes voor de ontwikkeling en het gebruik van digitale tweelingen en een digitale tweeling stelsel. De 10 principes zijn gecategoriseerd in 4 groepen of pijlers en in onderstaand raamwerk weergegeven (zie figuur 3.1):

1. Doel, een digitale tweeling stelsel dient een gemeenschappelijk en duidelijk doel;
2. Vertrouwen, een digitale tweeling stelsel moeten vertrouwen uitstralen;
3. Mensen, een digitale tweeling stelsel is gemaakt met en voor mensen;
4. Effectiviteit, een digitale tweeling stelsel zal effectief functioneren.



Figuur 3.1 – Raamwerk leidende principes Digital Twin Fysieke Leefomgeving

Er zit geen hiërarchie in de leidende principes: elk principe verdient evenveel aandacht. In deze referentie architectuur zijn de eerste zes principes van de bovenste drie pijlers 'doel', 'vertrouwen' en 'mensen' hieronder kort toegelicht aan de hand van drie aspecten; de uitspraak, verklaring en de implicaties. De vier principes uit de pijler 'Effectief' worden in deze referentie architectuur al uitwerkt in onderliggende architecturale ontwerpprincipes en constructieprincipes. Vandaar dat de toelichting van de principes 'Effectief bestuurd', 'Federatief', 'Met open standaarden' en 'Adaptief' achterwege is gelaten.

3.1 Doelgericht op het algemeen welzijn

Digitale tweelingen en het digitale tweeling stelsel dienen objecten, processen en systemen in de fysieke leefomgeving te verbeteren. Digitale tweelingen zijn nooit het doel, het zijn middelen om het doel mee te bereiken. Verantwoorde digitale tweelingen dienen het algemeen welzijn en creëren waarde voor de Nederlandse maatschappij. Zij werken vanuit een helder, enkelvoudig en bovenal gerechtvaardigd doel.

Verklaring: Een rechtvaardig doel is de noodzakelijke eerste stap om alle andere principes goed na te leven. Het doel bepaalt wie betrokken moet zijn bij de ontwikkeling van de digitale tweeling; welke data minimaal noodzakelijk zijn; hoe inclusief, volledig en accuraat de data moeten zijn, etc. Digitale tweelingen kunnen inzichten geven aan bedrijven, overheden, onderwijs en onderzoek, maatschappelijke organisaties en burgers voor betere planning, ontwikkeling en gebruik van de fysieke leefomgeving, en tevens het beheer en onderhoud van bestaande objecten, processen en systemen verbeteren.

Implicaties: Het opstellen van een rechtvaardig doel vereist voornamelijk begrip. Begrijp en doorleef de herkomst, de context, de dynamiek en het perspectief van het doel, de vraag, het probleem. Samen met eindgebruikers en andere primaire belanghebbenden. Verken of de digital twin de beste weg is om het doel te bereiken. Zijn er nog andere manieren, alternatieve oplossingsrichtingen (onderzocht) waarlangs de vraag beantwoord kan worden? Welke behoefte dient de digitale tweeling en het digitale tweeling stelsel en voldoet de digitale tweeling daaraan?

Digitale tweelingen en het stelsel van digitale tweelingen moeten worden gestructureerd en gedeeld om samenwerking en innovatie in de economie en maatschappij te bevorderen, in overeenstemming met het beginsel van het algemeen welzijn. Daarbij staan de behoeften en wensen van de gebruikers van digitale tweelingen centraal. Algemeen welzijn betekent niet dat digitale tweelingen alleen gefinancierd worden met publieke middelen vanuit de overheid, maar ook investeringen vraagt van het bedrijfsleven voor economische 'waarde creatie'. Digitale tweelingen en het stelsel van digitale tweelingen stimuleren samenwerking en innovatie in de hele economie en maatschappij ten behoeve van het algemeen welzijn.

3.2 Actief open en transparant

Digitale tweelingen en een stelsel van digitale tweelingen moeten zo open mogelijk beschikbaar zijn. Dat wil zeggen dat ze open of gedeelde data bevatten, die toegankelijk zijn via open standaarden om ervoor te zorgen dat ze de meeste waarde creëren voor de hele maatschappij. Transparantie is bedoeld om inzicht, duiding en uitleg te geven. Inzicht in de context en uitvoering van de digitale tweeling en het stelsel, uitleg over de gemaakte keuzes in het proces, duiding over de data, de modellen, visualisatie(tools) en tijdens besluitvorming.

We schrijven: "zo open mogelijk" omdat het van belang kan zijn dat deelnemers aan het stelsel ook informatie delen die niet geschikt is voor een brede, open, verspreiding. Denk hierbij aan privacy- en veiligheidsbeperkingen. Het stelsel dient ook een vertrouwelijke, afgeschermd informatiedeling mogelijk te maken.

Verklaring: Een stelsel van digitale tweelingen zal profiteren van netwerkeffecten: hoe meer mensen of organisaties het stelsel gebruiken, eraan bijdragen en het onderhouden, hoe waardevoller en bruikbaar het wordt.

Implicaties: Openheid is een belangrijk element om digitale tweelingen en een ecosysteem van digitale tweelingen voor Nederland te realiseren en de waarde voor economie en maatschappij te maximaliseren. Een open, op samenwerking gerichte aanpak zal nodig zijn om te zorgen dat het ecosysteem actueel en relevant blijft voor een breed scala aan gebruikers, waardoor de kans op silo's en inefficiëntie wordt verkleind. Openheid vereist ook dat digitale tweelingen vindbaar en toegankelijk zijn, zodat gebruikers

relevante data, rekenmodellen en visualisatietools kunnen vinden en gebruiken. Openheid gaat over het toegankelijk, herbruikbaar en uitwisselbaar maken van data, modellen, visualisatietools en technologieën. Het gaat niet alleen om de technische toegankelijkheid, maar ook de toegankelijkheid voor de belanghebbenden: de gebruikers, burgers en andere primaire belanghebbenden. Leg uit en geef toegang tot de informatie, die belanghebbenden nodig hebben om mee te doen en bij te dragen. Een digitale tweeling stelsel moet het juiste evenwicht weerspiegelen tussen openheid en (op rollen gebaseerde) toegangsbeperkingen, dat gebaseerd is op afgesproken beheers mechanismen.

3.3 Eerlijk kwalitatief en inclusief

De betrouwbaarheid van de digitale tweelingen, een stelsel van digitale tweelingen is afhankelijk van de kwaliteit en inclusiviteit van de data, modellen en technologie. De primaire belanghebbenden moeten deze betrouwbaarheid kunnen beoordelen op aspecten als de relevantie, tijdigheid, nauwkeurigheid, samenhang en consistentie, interpreteerbaarheid en toegankelijkheid van de digitale tweelingen, in relatie tot het beoogde gebruik. Kwaliteitsnormen en kwaliteitsniveaus voor data en modellen moeten worden gedefinieerd en afgesproken om data en modellen te kunnen gebruiken in digitale tweelingen. Het is belangrijk dat de digital twin en het stelsel niet alleen de technologie, maar ook de onderliggende data en de betrokkenen voldoende inclusief en integer zijn. Vooringenomenheid (bias), aannames, het bewust of onbewust bevoordelen of benadelen, onbedoelde discriminatie en profileren van individuen, groepen mensen en locaties is een steeds aanwezig risico bij het gebruik van digitale tweelingen. De maatschappelijke en economische waarde door digitale tweelingen moet aan alle belanghebbenden toekomen

Verklaring: Het verstrekken van de juiste, relevante metadata ondersteunt het verantwoorde gebruik van data en rekenmodellen in digitale tweelingen, waardoor effectieve data-integratie mogelijk wordt om waarde te creëren. Providers van data en rekenmodellen aan een digitale tweeling stelsel dienen de kwaliteit van data en rekenmodellen te beoordelen op basis van de functionaliteit, levensduur en beveiliging, die nodig zijn om het doel te bereiken waarvoor de data en rekenmodellen worden (of zullen worden) gebruikt.

Een digital twin kan nooit een exacte kopie zijn van de werkelijkheid. Juist om die reden is het belangrijk om aan te geven waar de grenzen liggen en in welke aspecten juist wel een realistisch beeld wordt geschetst. Een beeld dat losstaat van vooringenomenheid. Daarbij is het belangrijk om mee te nemen dat zelfs personen die niet in de digital twin worden gerepresenteerd, geraakt kunnen worden door aannames van de digital twin of zijn gebruiker.

Implicaties: Het stelsel van digitale tweelingen moet minimum kwaliteitsnormen voor data en rekenmodellen hanteren. Dit is vooral belangrijk voor gevallen waarin de juiste kwaliteitsnorm moeilijk te bepalen zijn. Kwaliteitsniveaus moeten transparant, gedefinieerd, meetbaar en beheerd worden binnen het stelsel van digitale tweelingen. Bij de verzameling, analyse, interpretatie, visualisatie en geautomatiseerde besluiten van en met data moet gelet en beoordeeld worden op vooringenomenheid om eventuele onbedoelde effecten te voorkomen, waaronder het ontstaan van onbedoelde discriminatie. Deze vooringenomenheid geldt niet alleen voor de betrokkenen, die de digital twin ontwikkelen en gebruiken, maar ook voor de technologie zelf. Technologie is nooit neutraal. Er zijn periodieke controles op vooringenomenheid (bias) in digitale tweelingen nodig om ervoor te zorgen dat de digitale tweelingen eerlijk en niet-discriminerend zijn.

3.4 Robuust en veilig

Een digitale tweeling en het stelsel van digitale tweelingen dient robuust en veilig te zijn. Dat betekent, dat data en rekenmodellen en de technische infrastructuur waarop digitale tweelingen draaien een hoge

beschikbaarheid hebben en goed beveiligd zijn (met evt. toegangsbeperkingen) tegen ongewenst gebruik. Een digital twin mag daarbij alleen de gevoelige informatie (oftewel persoonlijke data) vrijgeven indien voor die informatie uitdrukkelijk toestemming is verkregen van de betrokkenen en het vrijgeven ook wettelijk is toegestaan. Een digitale tweeling moet getoetst door onafhankelijke partijen, om blijvend de veiligheid en gegevensbescherming te kunnen waarborgen.

Verklaring: Data beveiliging en cyber security zijn essentieel om de integriteit van digitale tweelingen en het stelsel van digitale tweelingen te waarborgen. Een robuuste en veilige benadering zal vertrouwen geven aan de diverse gebruikers in de integriteit en het gebruik van digitale tweelingen en het stelsel. Dit principe dient in balans te zijn met de andere principes als openheid en transparantie.

Daarnaast wordt ook een digital twin wettelijk verplicht de privacy van individuen te respecteren. Digitale tweelingen zijn een goede manier om het ruimtelijk gedrag van een synthetische bevolking in de fysieke leefomgeving te bestuderen zonder daarbij inbreuk te maken op de privacy van individuen. Daarbij dienen de ontwikkelaars en gebruikers van de digital twin wel te reflecteren op wat 'persoonlijke data' zijn. De combinatie van verschillende ongevoelige datasets kan er namelijk toe leiden dat gevoelige informatie over personen wordt onthuld.

Implicaties: Beveiliging in digitale tweelingen kan op meerdere manieren technisch ingevuld worden en is een balanceer act tussen gebruiksgemak en geboden zekerheden. Maatregelen zijn mogelijk in de infrastructuur van de data uitwisseling, de gebruikte systemen, procesafspraken rondom digitale tweelingen en in de gebruikte data en rekenmodellen. Met behulp van het kader van de Baseline voor Informatiebeveiliging Rijksdienst kan hier concreet invulling aan gegeven worden. De BIR geeft handvaten die er voor zorg dragen dat data niet-gemanipuleerd (door kwaadwillenden of fouten) van de provider en bron naar de gebruiker komen. Dat ze alleen komen bij degene die er recht op heeft maar vooral ook dat die eindgebruiker kan vaststellen dat de data, rekenmodellen en visualisatietools robuust beschikbaar, betrouwbaar en beveiligd zijn. Pas certificering toe indien mogelijk.

Daarnaast dienen betrokkenen bij het ontwikkelen en gebruiken van digital twins expliciet en heel bewust toestemming te geven voor het (tijdelijk) verzamelen, koppelen, gebruiken en delen van hun data. Daarbij is ook van belang dat heel duidelijk wordt gemaakt voor welk specifiek doel zij hun gegevens beschikbaar stellen. Waarom en hoe juist deze data en de toepassing ervan de voorliggende vraag beantwoorden. Verken daarbij ook 'no go areas' voor de digital twin, ter bescherming van kwetsbare locaties en gemeenschappen.

3.5 Onder controle en toezicht van mensen

Een onafhankelijke vertrouwde derde partij moet als toezichthouder de naleving van ethische en andere relevante afspraken controleren, opdat er verantwoord wordt omgegaan met de digital twin, het stelsel van digitale tweelingen en de onderliggende data. De digital twins en het stelsel moeten onder controle en toezicht zijn van mensen. Toezicht en controle moet plaatsvinden gedurende het ontwikkeltraject van de digitale tweeling.

Verklaring: De verantwoordelijkheid voor ethiek moet niet bij één persoon, stafafdeling of commissie liggen. Dat haalt namelijk de verantwoordelijkheid bij anderen weg en uit de interne processen. Het is daarentegen wel wenselijk dat een onafhankelijke partij toezicht houdt op verschillende onderdelen van de digitale tweeling en het stelsel. Bijvoorbeeld bij de beoordeling van leveranciers en hun dataproducten. De ethische codes zijn daarbij een handvat, maar dienen wel iedere keer opnieuw en gedurende het hele proces met het team te worden doorleefd.

Implicaties: Controle en toezicht moet komen vanuit de mensen: het team. Spreek elkaar aan op de morele verantwoordelijkheid en naleving van de ethische waarden en afspraken. Vraag uitleg van externe partijen over hoe zij aan de ethische afspraken gaan voldoen. Een organisatie kan ook verantwoording afleggen over het gebruik van data en naleving van ethische standaarden in het jaarverslag.

3.6 Belanghebbenden intensief betrokken

Voor maatschappelijke opgaven is solidaire betrokkenheid van primaire belanghebbenden randvoorwaardelijk om te komen tot open, eerlijke en inclusieve beeld-, oordeels-, en besluitvorming. En randvoorwaardelijk om te komen tot een open, eerlijk en inclusief stelsel van digitale tweelingen. Ook voor commerciële toepassingen die geen maatschappelijke opgave dienen is de betrokkenheid van primaire belanghebbenden essentieel. Een duurzame vertrouwensrelatie bevordert de kwaliteit, de vraag naar- en de adoptie van (data- en informatie) producten. Veruit de belangrijkste betrokkenen zijn de mensen, die hun persoonlijke data delen én ieder ander die beïnvloed wordt door toepassing van de data en de digital twin. Het is belangrijk om burgers en andere primaire belanghebbenden al in de ontwerpfase van de digital twin te betrekken. En ook invloed, zeggenschap te geven als 'mede-eigenaar'. Laat de bijdragen van primaire belanghebbenden zwaar wegen.

Verklaring: De aard en omvang van de betrokkenheid van primaire belanghebbenden verschillen per opgave, per vraagstuk, per opdracht, per product, per situatie. Dit wordt onder meer beïnvloed door: de impact op mensen; de duur (eenmalig, tijdelijk of structureel); de risico's (privacy, veiligheid, gelijkheid); de geografische scope (internationaal, landelijk, regionaal, (hyper)lokaal); het domein (publiek, privaat); het specifieke doel van het datagebruik; de fase van ontwikkeling en realisatie. Dit en dat bepaalt ook de (werk)vorm waarlangs primaire belanghebbenden betrokken worden. Dat is maatwerk. Vaak wordt de omgeving pas betrokken bij een project in de testfase: zodra de digitaltwin klaar staat voor gebruik. Feedback wordt dan aangemoedigd en meegenomen, maar dit beperkt de belangrijkste belanghebbenden tot een reactie op wat er al staat. Dat is geen intensieve, solidaire betrokkenheid. De kwaliteit van de ethiek hangt af van de kwaliteit van het gesprek met de juiste mensen, van het stellen en verdiepen van de goede vragen op het juiste moment.

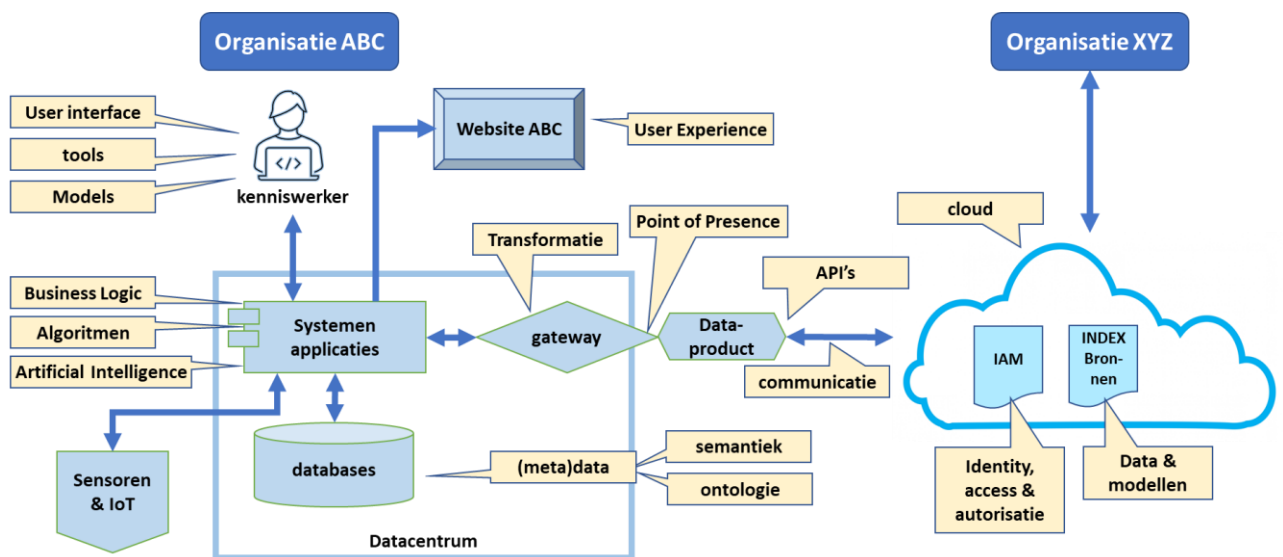
Implicaties: Stel de vraag: Wie zijn de primaire belanghebbenden? Wie is de omgeving? Wie zijn de betrokkenen? Er zijn verschillende soorten partijen die op verschillende manieren betrokken kunnen worden. Werk samen met en leer van (maatschappelijke) partijen die hetzelfde doel nastreven. De overheid is bij maatschappelijke opgaven een belangrijke belanghebbende en vehikel waarlangs burgers bereikt en betrokken kunnen worden.

4 Architecturale ontwerpprincipes

4.1 Inleiding

De maatschappelijke behoefte enerzijds en technische mogelijkheden anderzijds jagen de inzet van digitale tweelingen aan. Maatschappelijke waarden, zoals privacy borging, inclusiviteit, ethiek en informatiebeveiliging, stellen ook eisen aan de inrichting en werking van het digitale tweeling stelsel. De combinatie van deze wensen en eisen dienen tot uitdrukking te komen in zaken als de governance van het stelsel en de organisatorische en technische inrichting ervan.

In dit hoofdstuk gaat het hoofdzakelijk over de technische inrichting: de architectuur van het digitale tweeling stelsel. De architecturale uitgangspunten of principes dienen de maatschappelijke wensen en eisen mogelijk te maken en omgekeerd: maatschappelijke eisen, waarden en wetgeving moeten door het digitale tweeling stelsel worden gerespecteerd. Architecturale principes vormen dus de basis voor de inrichting van het stelsel dat tegemoetkomt aan eisen en wensen en rekening houdt met grenzen die de samenleving stelt. Figuur 2 geeft een vereenvoudigd beeld van aspecten die van belang zijn voor de architectuur van het stelsel. We lichten ze in deze paragraaf kort toe. In hoofdstuk 5 worden deze aspecten dieper uitgewerkt in de vorm van constructieprincipes en standaarden.



Figuur 2 Architecturale componenten digital twinning

We lopen even door het model heen. Organisatie ABC verzamelt, beheert en bewerkt informatie. De gegevens kunnen afkomstig zijn uit interne systemen, zoals bijvoorbeeld een Building Information Modelling system, waarmee afmetingen en kenmerken van een gebouwontwerp worden vastgelegd. Ook kunnen gegevens afkomstig zijn uit metingen, die "handmatig" of via de inzet van sensoren en besturingssoftware van apparaten worden verkregen (bijvoorbeeld Internet of Things). De data worden verzameld, beheerd en bewerkt door de inzet van systemen of applicaties. Deze bevatten bewerkingsregels (business logic of algoritmen) en naarmate de complexiteit ervan toeneemt spreken we meer over artificial intelligence.

Een professional binnen deze organisatie heeft toegang tot deze applicaties en gegevens en beschikt over software (tooling) om deze te bewerken. In ons geval gaat het hierbij natuurlijk om een zodanige bewerking dat het mogelijk wordt uit diverse typen data de gewenste modellen en virtualisaties ervan te realiseren. Het is mogelijk een deel van de aldus ontwikkelde modellen en visualisaties via bijvoorbeeld de website of portal van de organisatie beschikbaar te stellen aan de buitenwereld. Een aardig voorbeeld hiervan is de

app van NS, waarin via de “treinradar” zichtbaar wordt gemaakt welke trein zich op enig moment waar bevindt¹⁰.

Naast deze wijze van informatiepresentatie, bestaat nog een tweede, die zeer belangrijk is voor samenwerkende organisaties binnen het stelsel. Zoals we nog zullen zien gaan we er in het stelsel van uit dat informatie van organisaties via kant-en-klare dataproducten wordt aangeboden aan andere organisaties. Dat betekent dat een organisatie als ABC een aantal informatieproducten “op de plank” heeft liggen, die via een uitleverpunt (gateway) naar andere organisaties verzonden kan worden.

Een afnemende organisatie, XYZ, kan als deelnemer aan het stelsel via een centrale index snel op het spoor komen van de gewenste informatie. Via deze index komt XYZ bij het uitleverpunt van ABC, bestelt het gewenste dataproduct en krijgt dit eenmalig, als abonnement of in de vorm van ‘streaming data’ geleverd. Het dataproduct kent ook een deel waarin relevante meta-informatie is vastgelegd, waardoor de ontvangende partij geïnformeerd wordt over onder meer de herkomst van de data, de context waarbinnen deze is verkregen en wat de kwaliteit ervan is.

Het model biedt een overzicht van architecturale componenten waarover afspraken moeten worden gemaakt, die samen een deugdelijke werking van het stelsel mogelijk maken. Deze referentie architectuur bevat hiervoor de belangrijkste ontwerpprincipes – in dit hoofdstuk –, welke nader uitgewerkt worden in technische constructieprincipes en standaarden (in een volgende versie van dit document).

Hoewel in deze referentie architectuur maximaal rekening is gehouden met actuele inzichten in maatschappelijke wensen, eisen en begrenzingen, de stand van de technologie en architecturale inzichten, zal door middel van versiebeheer het geheel van ontwerp- en constructieprincipes periodiek aangepast moeten worden aan gewijzigde omstandigheden. Architecturale ontwerpprincipes zijn dus kaderstellend en richtinggevend. Zij worden (later) in hoofdstuk 5 verder uitgediept in concrete constructieprincipes en standaarden.

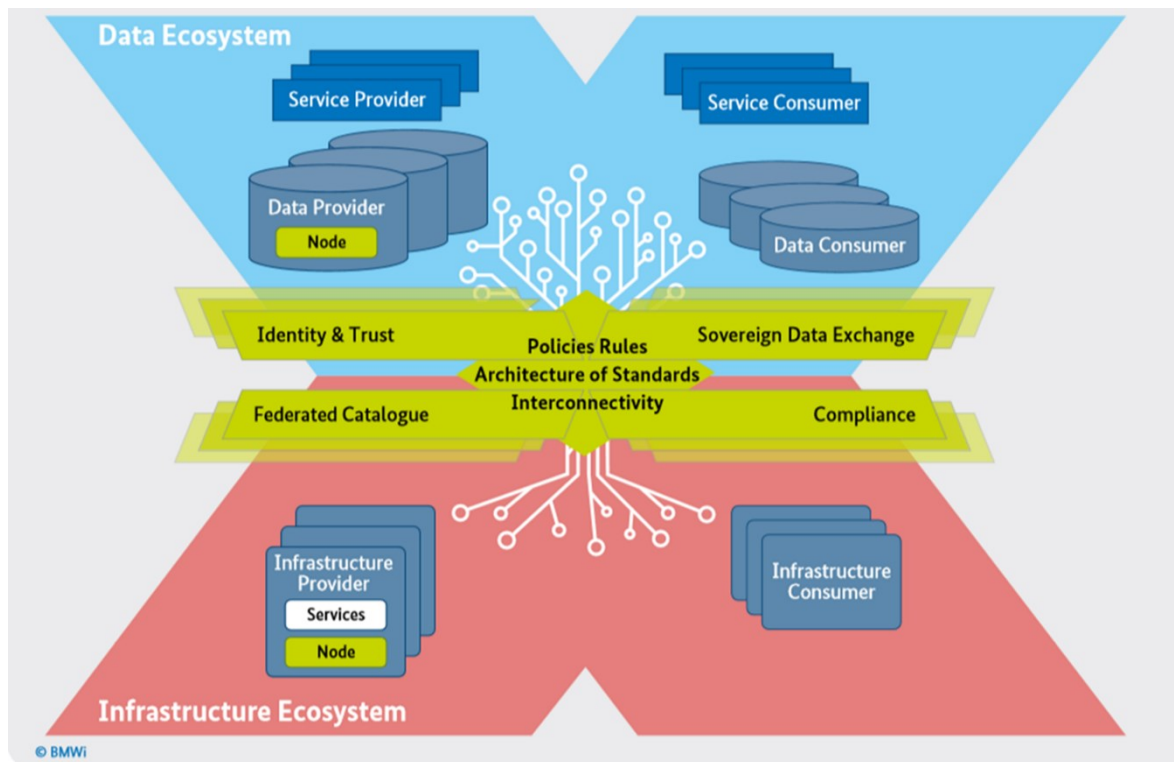
4.2 De architecturale ontwerpprincipes

Bij het ontwerpen van het stelsel wordt uitgegaan van een beperkt aantal, maar fundamentele ontwerpprincipes. We hebben ons daarbij laten leiden door een aantal ontwikkelingen en inzichten, zoals:

- International Data Spaces (IDS), een internationaal samenwerkingsverband van ruim honderdtwintig bedrijven en instellingen, streeft naar samenwerking en data-uitwisseling op domeinniveau, data spaces genoemd. Het programma Digitale Twin voor de Fysieke Leefomgeving, is gericht om precies dat te doen voor het fysieke leefdomein in Nederland. We beschouwen de fysieke leefomgeving in Nederland (gebouwde, natuurlijke en sociale omgeving) dus als een data space.
- De **Inspire**-richtlijn heeft tot doel een infrastructuur voor ruimtelijke gegevens van de Europese Unie te creëren ten behoeve van het milieubeleid en het milieubeleid of de activiteiten van de EU die gevolgen kunnen hebben voor het milieu. Deze Europese infrastructuur voor ruimtelijke gegevens zal het delen van ruimtelijke milieu-informatie tussen organisaties in de publieke sector mogelijk maken, de toegang van het publiek tot ruimtelijke informatie in heel Europa vergemakkelijken en helpen bij het maken van beleid over de grenzen heen.
- **Gaia-X**, is een initiatief waarin ruim 300 vertegenwoordigers van het bedrijfsleven, de politiek, academici en wetenschap uit Europa en de rest van de wereld samenwerken om een federatief en veilig ecosysteem voor data-infrastructuur te creëren¹¹. De architectuur van Gaia-X is gebaseerd op het principe van decentralisatie. Gaia-X is het resultaat van een veelheid aan individuele platforms die allemaal een gemeenschappelijke standaard volgen – de Gaia-X-standaard. Het idee is om samen ontwikkelen een data-infrastructuur te ontwikkelen op basis van de waarden openheid, transparantie en vertrouwen. Wat er dus ontstaat is geen cloud, maar een netwerksysteem dat veel cloudserviceproviders met elkaar verbindt.

¹⁰ <https://treinenradar.nl/>

¹¹ <https://www.gaia-x.eu/>



12

- De **FAIR**-principes¹³. Deze principes vragen aandacht voor zaken als vindbaarheid, metadatering en (her)bruikbaarheid van data.
- **Data mesh**¹⁴. Dit concept draagt onder meer de volgende inzichten aan¹⁵:
 - Data wordt door deelnemende organisaties aan het stelsel decentraal verzameld, opgeslagen en bewerkt en via een lichte federatieve infrastructuur uitgewisseld, die voorzien is van onder meer indexfunctionaliteit om data vindbaar te maken. We zien de federatieve aanpak ook terug in de [interbestuurlijke datastrategie van de overheid](#).
 - Data-uitwisseling gebeurt in de vorm van dataproducten.
 - Het stelsel kent een federatieve governance.

De inzichten van onder meer IDS, Inspire, Gaia-X, FAIR en data mesh vind je terug in de uitgangspunten in dit hoofdstuk.

Ter illustratie van de grondgedachte achter het stelsel op basis van onder meer het data mesh concept, geven we het onderstaande, enigszins fictieve voorbeeld.

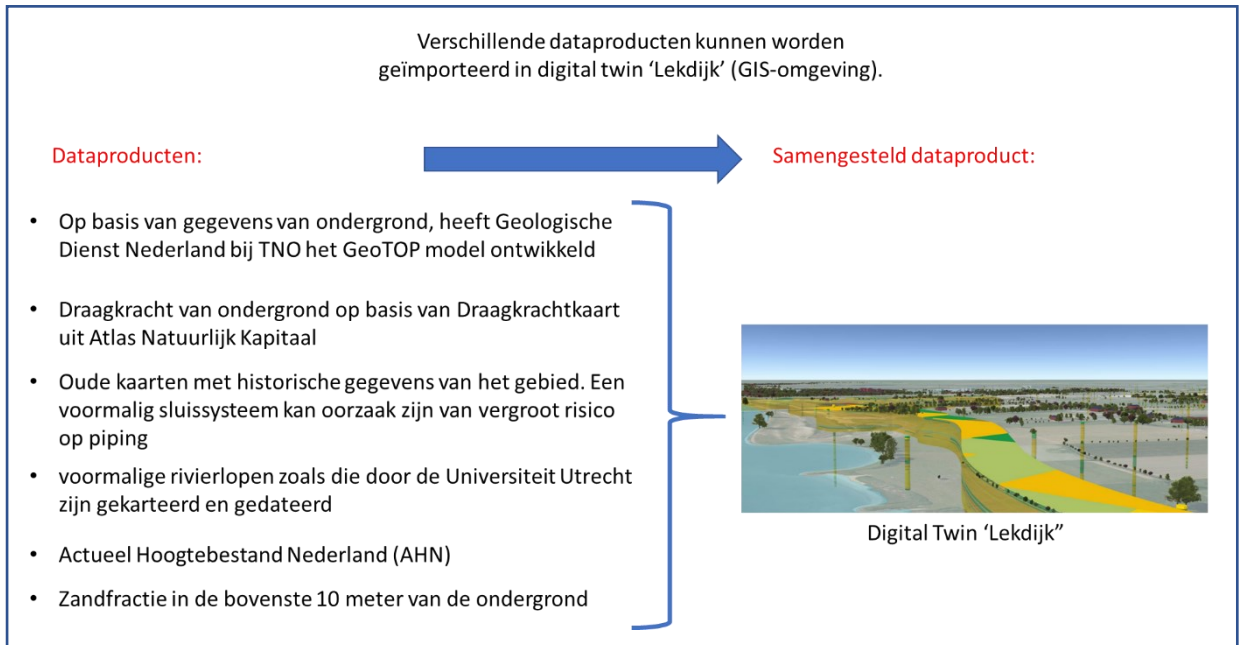
Stel we willen het onderhoud van een deel van de Lekdijk fundamenteel aanpakken. Daarbij maken we graag gebruik van een combinatie van diverse soorten informatie. De onderstaande figuur geeft een beeld van deze verschillende soorten informatie, die elk als dataproduct worden afgenomen van hun leverancier.

¹² Bron: Bundesministerium Für Wirtschaft und Energie

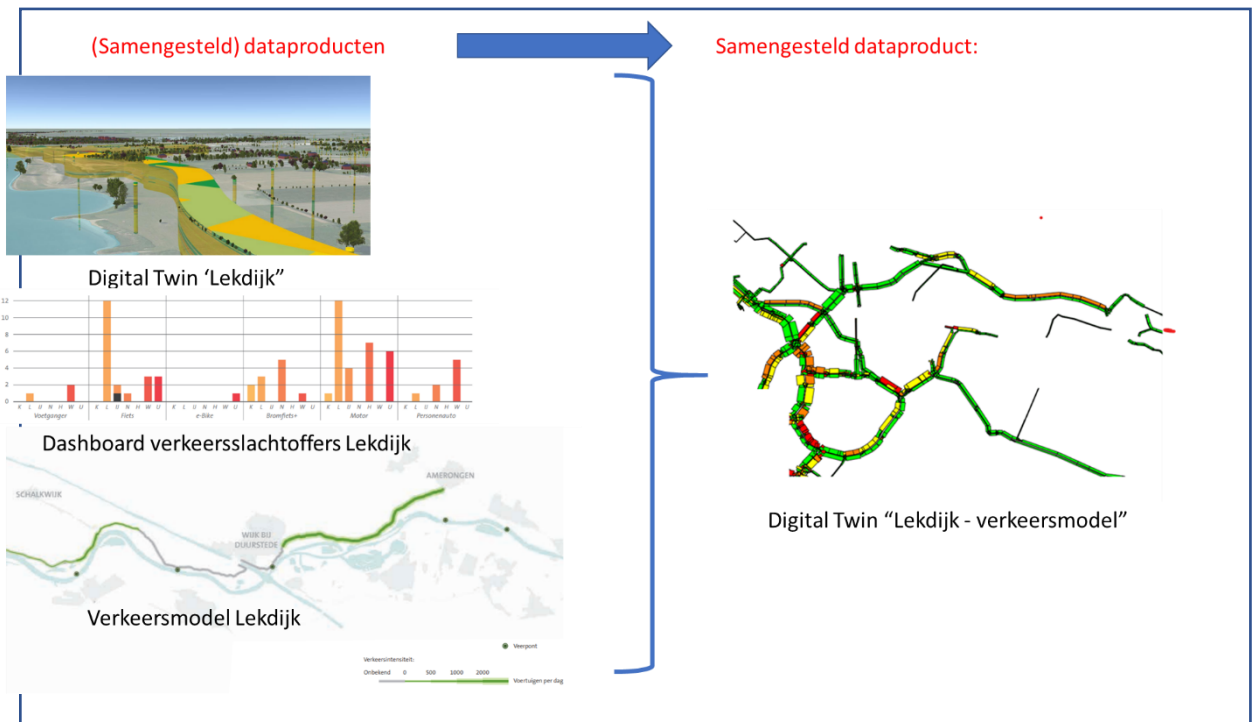
¹³ <https://www.go-fair.org/fair-principles/>

¹⁴ <https://martinfowler.com/articles/data-mesh-principles.html>

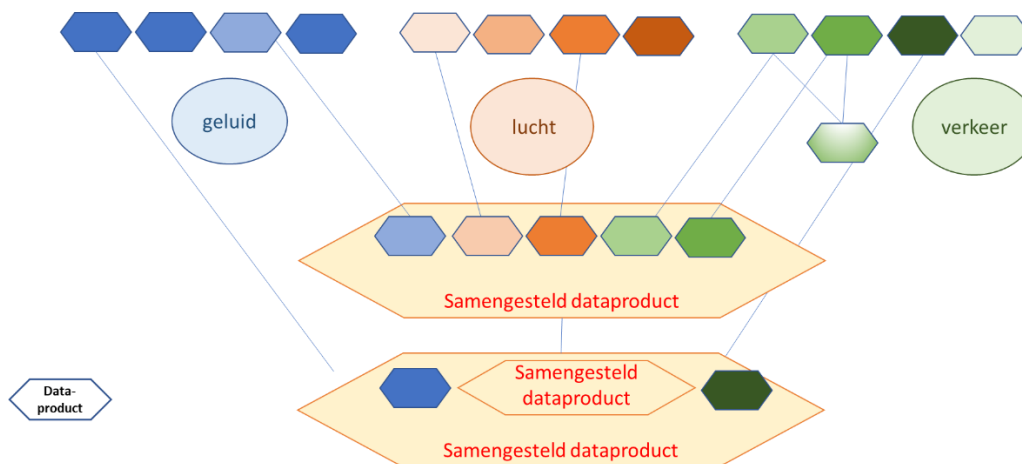
¹⁵ Omdat het data mesh concept zo'n belangrijk rol vervult in deze referentie architectuur, is een nadere toelichting op dit concept opgenomen als bijlage.



De data die gebruikt is voor de digitale tweeling Lekdijk kan ook hergebruikt worden voor het opstellen van een digitale tweeling waarin ook het verkeersmodel, mede gebaseerd op het aantal verkeersslachtoffers in de afgelopen jaren, wordt geïntegreerd.



Wat dit voorbeeld laat zien is de recursieve werking van het data mesh concept. Een meer abstracte weergave is de onderstaande figuur.



Figuur 3 Dataproducten kunnen "opgeteld" worden tot nieuwe dataproducten

We beperken ons in dit hoofdstuk tot architecturale ontwerpprincipes voor het stelsel. Een aantal daarvan heeft uiteraard ook wel betrekking op de wijze waarop data door organisaties beschikbaar wordt gesteld. Hierbij kan gedacht worden aan principes voor het duiden van de semantiek van gegevens, de wijze waarop gegevens en modellen worden voorzien van meta-data, de technische aspecten van het dataproduct en het datacommunicatieprotocol op basis van services en API's. Immers: naarmate organisaties hun data en modellen meer op gemeenschappelijke uitgangspunten en standaarden baseren, wordt het uitwisselen tussen organisaties en het hergebruik van data eenvoudiger. Maar we beginnen met de architecturale ontwerpprincipes.

4.2.1 Aansluiting op omliggende kaders

De architectuur van het digitale tweeling stelsel is mede gebaseerd op architecturale kaders van de overheid en de facto kaders zoals gangbaar binnen bedrijven en organisaties.

Toelichting

De overheid kent een samenhangend kader voor architectuur, zoals de Nederlandse Overheid Referentie Architectuur en daarvan afgeleide architecturen voor onder meer Rijk, Provincies, Gemeenten en Waterschappen. De overheid kent ook andere 'stelsels' voor data-uitwisseling, zoals dat voor de Nederlandse Basisregistraties en in EU-verband bijvoorbeeld Inspire. Tenslotte is er ook veel wetgeving, die mede kaderstellend is voor de inrichting van het digitale tweeling stelsel, zoals de Wet Open Overheid; de Algemene Verordening Gegevensbescherming (AVG), de Omgevingswet, de Wet Hergebruik Overheidsinformatie en de Archiefwet. De architectuur van het digitale tweeling stelsel volgt in beginsel deze wetgeving, architecturale kaders en best practices. Veel van de wensen, eisen en begrenzingen dienen via de principes uit dit document *by design* gerealiseerd te worden.

4.2.2 Domein georiënteerd, gedecentraliseerd data-eigendom/ -architectuur

Data wordt door decentraal verworven, opgeslagen, verwerkt en beschikbaar gesteld.

Toelichting

Dit principe vormt samen met de vier volgende principes de kern van het data mesh concept¹⁶. Uitgangspunt is dat organisaties die deelnemen aan het stelsel zelf zorgen voor het aanleggen van databestanden, dataverwerking en het ter beschikking stellen van data aan derden.

Data spaces

De Europese Commissie wil een interne datamarkt creëren, waar gegevens van overheidsinstanties, bedrijven en burgers veilig en correct kunnen worden gebruikt voor het algemeen belang. De betreffende verordening stelt gebruikers in staat de controle over hun gegevens te behouden en moedigt het totstandbrengen van gemeenschappelijke Europese gegevensruimten in cruciale sectoren aan.

Dit initiatief stelt regels op voor gemeenschappelijke Europese dataruimtes (in sectoren zoals milieu, energie, mobiliteit en landbouw) om:

- beter gebruik te maken van overheidsdata voor onderzoek in het algemeen belang;
- vrijwillige uitwisseling van data door individuele personen te ondersteunen;
- structuren op te zetten waarmee belangrijke organisaties data kunnen uitwisselen.

Deze EU-ontwikkeling staat anno 2022 nog in de kinderschoenen. In latere versies van deze referentie architectuur, zullen mogelijk consequenties op Europees niveau verwerkt worden.

4.3 Data als een product

Data wordt beschikbaar gesteld in de vorm van producten.

Toelichting

Organisaties ontwikkelen op basis van data waarover zij beschikken uitwisselbare dataproducten, die tegen bepaalde condities (aansluitvoorwaarden) aan andere deelnemers van het stelsel ter beschikking worden gesteld. Deze dataproducten kunnen de vorm aannemen van operationele data, analytische data, modellen of een combinatie ervan. Een dataproduct wordt ontwikkeld zodat het vindbaar, begrijpelijk, betrouwbaar, interoperabel, toegankelijk, beveiligd, adresseerbaar is en publieke waarde heeft. Dit zijn de voorwaarden voor deelname aan het stelsel. Het dataproduct informeert afnemers ook over onder meer de relevante context, bruikbaarheid, kwaliteitsaspecten, houdbaarheid en herkomst van de geleverde data. Al deze informatie is bij het dataproduct zelf aanwezig.

4.4 Selfservice data-infrastructuur als platform

Het stelsel berust op federatieve componenten, waaronder publieke en private datanetwerken, op cloudtechnologie gebaseerde datacenters, indexen voor het vindbaar maken van dataproducten en voorzieningen voor identiteits- en toegangscontrole. Via federatieve indexen kunnen vragende partijen geautomatiseerd door hen gewenste dataproducten vinden.

¹⁶ [Data Mesh Principles and Logical Architecture \(martinfowler.com\)](https://martinfowler.com/DataMeshPrinciples.html)

Toelichting

De genoemde componenten zorgen voor een ordentelijke afhandeling van het uitwisselen van dataproducten, met garanties voor identificatie van deelnemende organisaties en een hoge mate van vindbaarheid van dataproducten. Deze componenten kunnen worden gehuisvest in een rekencentrum van een publieke of private serviceprovider, met uitstekende netwerkverbindingen, die onder beheer van publieke of private organisaties staan.

Op federatief niveau zijn enkele voorzieningen of componenten nodig die de uitwisseling van dataproducten faciliteren. Aanvankelijk zal het gaan over de volgende componenten:

- Een federatieve catalogus of index voor het kunnen vinden van dataproducten
- Een identificatie, authenticatie en autorisatievoorziening
- Een voorziening voor het systematisch registreren en vinden van zaken als datadefinities (semantiek), datamodellen en binnen het stelsel toegepaste ontologieën.
- Een voorziening waardoor verschillende gegevensmodellen onderling gelinked kunnen worden.

4.4.1 Federatieve catalogus of index

De federatieve catalogus bevat het aanbod van dataproducten (incl. metadata) van de dataleveranciers in het stelsel. Het is een indexarchief van beschrijvingen dat doorzoekbaar is en informatie biedt van de aangeboden dataproducten. De catalogus is bedoeld om de dataproducten eenvoudig te kunnen vinden, inclusief een beknopte beschrijving van de karakteristieken van het dataproduct en de dataleverancier.

De dataleveranciers zijn zelf verantwoordelijk voor het leveren van de beschrijvingen. Deze beschrijvingen worden ook *self-description data* genoemd.

De "Self Descriptions" bevatten de kenmerken van dataproducten, waaronder:

- aard van de data (domein, semantiek)
- herkomst van data (lineage)
- context
- geldigheid
- vertrouwelijkheid
- aangeboden service die aangeroepen kan worden
- identificatiegegevens
- inloggegevens die zijn uitgegeven en ondertekend door vertrouwde partijen, claims over de provider of asset/resource
- betrouwbare verificatiemechanismen
- zoeken en samenstelling van dataproducten in de catalogus
- bewaking, continue validatie en vertrouwensmonitoring op basis van gebruiksbeleid
- contractuele voorwaarden met betrekking tot het gebruik van het dataproduct

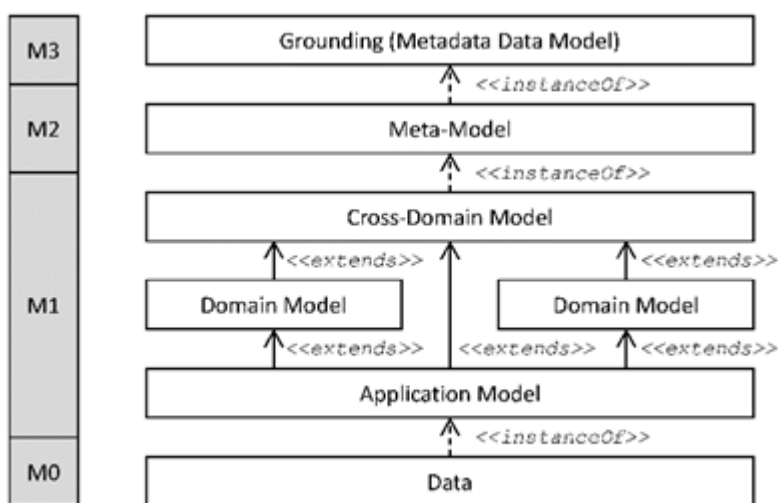
4.4.2 Identificatie, authenticatie en autorisatie

De federatieve componenten zijn alleen toegankelijk voor deelnemers aan het stelsel. Zij dienen voor deelname te voldoen aan een aantal voorwaarden (zie hoofdstuk 6). Daarom is op federatief niveau een voorziening waarmee (werknemers van) deelnemers zich kunnen identificeren. Een federatieve component dient de identificatie te kunnen authentifieren. Mogelijk kan ook een autorisatie noodzakelijk blijken indien dit onvoldoende op het niveau van de aanbieders van dataproducten kan worden geregeld. Mogelijk kan voor deze voorzieningen gebruik gemaakt worden van bestaande voorzieningen, welke worden aangeboden

via Logius¹⁷. Ook eventueel benodigde encryptievoorzieningen worden door Logius binnen de overheid aangeboden.

4.4.3 Ontologie

Elk domein heeft een of meerdere digitale tweelingen conform de vingerende open standaarden, datamodellen en ontologieën in het betreffende domein. Bijvoorbeeld een organisatie werkt conform de BIM (Bouw Informatie Model) standaarden en ontologie en een andere organisatie conform CityGML, smart city ontologie en sensor data. Indien er overkoepelende thema's, vraagstukken zijn waarbij de digitale tweelingen gecombineerd dienen te worden, gebeurt dit door de betreffende datavelden met elkaar in relatie te brengen in een uniforme taxonomie in een knowledge graph of graph network met attributed properties en uit eindelijk op basis van overkoepelend Linked Data Ontologie. Het referentie model hiervoor is:



4.5 Data soevereiniteit

De data-eigenaar bepaalt de gebruiksvoorwaarden van de door hem ter beschikking gestelde data.

Toelichting:

Datasoevereiniteit behandelt hoe gegevens worden uitgewisseld en gedeeld tussen organisaties. Een gegevenseigenaar bepaalt de gebruiksvoorwaarden van zijn gegevens. Met andere woorden: de data-eigenaar bepaalt wat de data-afnemer wel en niet mag doen met de geleverde data.

Om dit goed te kunnen begrijpen moeten we een tweetal begrippen toelichten: data-eigenaar en dataleverancier. Een data-eigenaar is de bron van de data. De data-eigenaar heeft data verzameld of laten ontstaan en vastgelegd¹⁸. Deze data kunnen in de vorm van een of meerdere dataproducten worden aangeboden aan afnemers van data. Hieraan kan de data-eigenaar bepaalde eisen stellen, zoals de mate

¹⁷ <https://www.logius.nl/>

¹⁸ We gaan hier voorbij aan de discussie of het data-object de eigenaar van de data is. Zeker waar het gaat om persoonsgebonden data kan men immers stellen dat het individu de eigenaar van de data is. In het kader van digital twinning beperken we ons tot de verantwoordelijkheid van de organisatie die data heeft vastgelegd.

van vertrouwelijkheid, de 'maximale houdbaarheid' van de data en onder welke voorwaarden 'doorlevering' van data toegestaan is. De afnemer van de data kan echter op zijn beurt de verkregen data verwerken in dataproducten die hij op zijn beurt weer ter beschikking stelt aan derden. Daarmee vervult deze de rol van dataleverancier. In dat geval dienen de eisen van de data-eigenaar weer overgedragen te worden op de volgende afnemer van data in de keten. De dataleverancier is niet de eigenaar van alle data, maar zorgt er wel voor dat de eisen die de eigenaar heeft gesteld ook weer worden gesteld aan afnemers van data 'verderop' in de keten. Omgekeerd kan de afnemer van data "zien" waar de data oorspronkelijk vandaan komt en langs welke bewerkingsstappen in de keten het uiteindelijke dataproduct tot stand is gekomen.

4.6 Service oriëntatie.

Het stelsel is gebaseerd op service oriented architecture als ontwerpprincipe.

Toelichting:

Deelnemers aan het stelsel wisselen data en modellen uit via het beginsel van de service oriëntatie. Dat wil zeggen: Een samenspel van (gestandaardiseerde) vragen en antwoorden via een verbindende infrastructuur. In praktische zin wordt dit ingevuld met services of API's.

4.7 Open uitwisselingsstandaarden

De architectuur van het stelsel is maximaal gebaseerd op de toepassing van open standaarden voor onder meer semantiek, dataproducten, API's, services en communicatieprotocollen.

Toelichting:

Het stelsel staat open voor deelname vanuit overheid en private sector. Federatieve componenten van het stelsel worden via open standaarden vormgegeven. Daarnaast geldt een sterke aanbeveling voor deelnemende partijen om voor dataproducten eveneens gebruik te maken van open standaarden, teneinde de interoperabiliteit van het stelsel te maximaliseren. De architectuur van afzonderlijke digitale tweelingen binnen organisaties, kan uiteraard ook gebaseerd worden op open standaarden, maar om begrijpelijke redenen kunnen commerciële partijen hiervan afwijken.

Door deelnemers aangeboden dataproducten voldoen aan de uitwisselingsstandaarden (semantisch, technisch) zoals beschreven in deze referentie architectuur. Deze eis brengt met zich mee dat – indien een deelnemende organisatie intern afwijkende standaarden hanteert – deze zelf zorgt voor de eventueel benodigde transformatie naar de van toepassing zijnde uitwisselingsstandaard.

4.8 Kwaliteit data

Door deelnemers aan het stelsel aangeboden data (incl. in de vorm van modellen) kan te allen tijde verantwoording worden afgelegd over de herkomst en kwaliteit van data¹⁹.

Toelichting:

Het digitale tweeling stelsel op basis van het data mesh principe leidt tot ketens van combineerbare data. Daarom is het essentieel dat afnemers van dataproducten nauwkeurige, op veldniveau, informatie krijgen over zaken als de herkomst van data (data lineage of data provenance), tijdstip van verwerving van data, betrouwbaarheid van data, geldigheid van data, vertrouwelijkheid van data, gebruikscondities van data en semantische duiding van data.

¹⁹ Zie ook het data mesh framework voor datakwaliteit.

5 Constructieprincipes

5.1 Inleiding

Dit hoofdstuk gaat over het daadwerkelijk construeren van een samenhangende architectuur voor digital twinning. Op basis van de hoofdlijnen van **Ontwerpprincipes** uit het vorige hoofdstuk worden in dit hoofdstuk de constructie principes en de daarbij behorende standaarden beschreven.

5.1.1 Wat is het doel van constructie principes?

Constructieprincipes zijn te beschouwen als richtlijnen voor het ontwikkelen van bijdragen aan het Nederlandse digitale tweeling stelsel voor de fysieke leefomgeving. Veel gespecialiseerde architecten, ontwerpers, ontwikkelaars en beheerders spelen hierbij een rol, zoals user interface experts, solution architects, softwareontwikkelaars, informatie-analisten, data-specialisten, IT-architecten, informatiebeveiligingsdeskundigen, etc. etc. In dit hoofdstuk vinden zij toe te passen ontwerpprincipes en daarbij behorende standaarden.

5.1.2 Hoe worden constructie principes toegepast?

De constructie principes worden toegepast in meerdere typen ontwerpdocumenten, zoals projectstartarchitecturen, solution architecturen, datacommunicatie architecturen, infrastructuur architecturen, etc. etc. Digital twinning kent immers veel componenten, die met elkaar het stelsel vormen. Elk van de hierbinnen toegepaste componenten, worden bij voorkeur in lijn met de constructieprincipes uit dit hoofdstuk ontworpen en gerealiseerd, waardoor zij inpasbaar zijn in het totale stelsel. Toepassing van ontwerpprincipes kan dus ook gezien worden als het zorgen voor een hoge mate van interoperabiliteit binnen het stelsel.

Dit hoofdstuk wordt later verder ingevuld. Elk constructieprincipe wordt op basis van de onderstaande structuur beschreven.

5.2 Specificatie van constructieprincipes

5.2.1 korte naam van het constructieprincipe

Constructieprincipe

Definities

Toelichting

Verwijzing naar ontwerpprincipes

Referenties

Een eerste overzicht van te behandelen constructieprincipes op het niveau van het stelsel is de onderstaande lijst.

- Infrastructuur (netwerken, netwerkcomponenten)
- Identity and access (identiteitsmanagement, authenticatie, autorisatie, single sign on, etc.)
- Beveiliging (o.m. encryptie)
- Dataproducten (typen, samenstelling, metadata, polyglot, kwaliteitsnormen)
- Voorkeursstandaarden voor data en modellen
- Semantische interoperabiliteit (ontologieën)
- Datacommunicatie

- Transformaties
- Vindbaarheid (catalogus, indexering)
- Applicatie-architectuur federatieve componenten
- Beheer federatieve componenten
- Stelselinformatie (technische specificaties voor deelname, overzicht functionaliteiten, release-informatie, FAQ's, etc.)

Sommige constructieprincipes kunnen ook van belang zijn voor individuele organisaties of digitale tweelingen. Daarbij moet gedacht worden aan onder meer:

- Dataopslag
- Data
- Modellen
- Databewerking
- Analytics
- Algoritmen
- Presentatie
- Beveiliging
- Ontwikkeling

6 Stelselafspraken

Dit hoofdstuk is mede gebaseerd op het [stelsel van Nederlandse Basisregistraties](#)

6.1 Inleiding

Het digitale tweeling stelsel bestaat niet alleen uit afspraken over architectuurprincipes en technische standaarden. Het is ook noodzakelijk om afspraken te maken over deelname aan het stelsel en wellicht ook te komen tot wet- en regelgeving op onderdelen van het stelsel. Dit hoofdstuk doet suggesties voor stelafspraken. Deels zijn er in dit hoofdstuk reeds concrete voorstellen voor afspraken geformuleerd, maar er staan ook zaken in waarover later en/of in een ander verband stelafspraken zouden moeten worden gemaakt.

De afspraken variëren van technische standaarden tot en met bestuurlijke verantwoordelijkheid voor de maatschappelijke waarden die grenzen stellen aan deelname aan het stelsel en de integriteit en vertrouwelijkheid van data.

6.2 Wat verstaan we onder 'het stelsel'?

Het digital twinning stelsel van de fysieke leefomgeving is het geheel van afspraken en voorzieningen gericht op het mogelijk maken van de uitwisseling van data en modellen waarin data is vastgelegd. Deze data heeft betrekking op zeer uiteenlopende aspecten van de fysieke leefomgeving, zoals in vorige hoofdstukken van dit document is beschreven.

6.3 Betrokkenen, afspraken en rollen

Bij het stelsel zijn verschillende partijen betrokken. Uiteraard de deelnemende organisaties aan het stelsel, de leveranciers en gebruikers van data, dataproducten en modellen. Voor wat betreft de gefedereerde voorzieningen binnen het stelsel, zoals de identificatie-, authenticatie- en autorisatiefunctie, de verwijzindex en de afspraken over het stelsel als geheel, speelt de overheid een belangrijke rol. Maar het stelsel staat ook open voor niet-overheidsorganisaties. Ook zij zullen daarmee een zekere verantwoordelijkheid hebben. Vervolgens zijn er partijen die een verantwoordelijkheid dragen voor technische voorzieningen, zoals de datacommunicatie-infrastructuur, (cloud)datacenters en tal van voorzieningen die ingezet worden binnen organisaties van deelnemers aan het stelsel, zoals soft- en hardware voorzieningen.

Het gebruik van het woord 'afspraken' betekent dat uitgegaan wordt van een bindend commitment tussen twee of meer partijen. Per type deelnemer aan het stelsel kunnen dan ook verschillende (clusters van) afspraken bestaan. Dit commitment kan bestaan uit privaatrechtelijke overeenkomsten tussen partijen, maar zou ook via wet- en regelgeving bekrachtigd kunnen worden. In hoeverre afspraken ook een wettelijke basis moeten krijgen, laten wij in dit hoofdstuk vooralsnog in het midden. Voor zover nadere wetgeving noodzakelijk wordt geacht, ligt het voor de hand om hiervoor uit te gaan van het ministerie van Binnenlandse Zaken en Koninkrijksrelaties (BZK), gezien de verantwoordelijkheid van dit ministerie voor de informatiehuishouding van de overheid op nationaal niveau. Eventueel wettelijke bepalingen zouden dan onderdeel kunnen zijn van de [Wet Digitale Overheid](#)

6.4 Overleggen

Hoewel uitgegaan wordt van een stelsel dat open staat voor zowel overheid als niet-overheidsorganisaties en een open architectuur, kan het stelsel niet zonder diverse vormen van overleg van diverse stakeholders. Zonder hiervan nu een uitputtende opsomming te geven, kan gedacht worden aan de volgende overleggen:

Een Digital Twinning Governance Board

Dit college kent gemandateerde deelnemers uit de overheid, waarbij gedacht kan worden aan vertegenwoordigers van direct betrokken ministeries en eveneens sterk betrokken uitvoeringsorganisaties en toezichthouders. Deze Board heeft als taak de deugdelijke werking en ontwikkeling van het stelsel te bewaken en bevorderen, alsmede de benodigde randvoorwaarden veilig te stellen en het geheel van stelselafspraken al dan niet via wetgeving te borgen. De Board kan voor zover nodig onderliggende overlegorganen ('commissies') instellen, om aspecten beleidsmatig voor te bereiden voor het overleg in de Board.

Platform Digital Twinning Fysieke Leefomgeving

In dit platform participeren vertegenwoordigers van de belangrijkste stakeholders van het stelsel: overheid, deelnemers aan het stelsel, leveranciers van voorzieningen, kennisinstellingen. Het platform fungeert als het centrale overlegorgaan voor alle aspecten die het functioneren van het stelsel raken, waaronder het functioneren van het stelsel, de verdere ontwikkeling ervan, het doen van voorstellen voor aanvullingen of wijzigingen in de stelselafspraken en de stelselarchitectuur (incl. standaarden), etc. Mogelijk kan dit platform tevens dienen als adviesorgaan voor de Digital Twinning Governance Board. Het platform kan ook tijdelijke of permanente werkgroepen of commissies instellen, die zaken uitwerken ter voorbereiding van het plenaire overleg binnen het platform.

Overleg Beheer centrale stelselvoorzieningen

Dit overleg bestaat uit partijen die verantwoordelijk zijn voor de instandhouding van de centrale voorzieningen binnen het stelsel, zoals infrabeheer, leverancier van centrale datacenterfunctionaliteit, beheer van componenten ten behoeve van identificatie, authenticatie en autorisatie, indexen, etc. De hier gesuggereerde overleggen dragen verantwoordelijkheid voor (een deel van) de stelselafspraken. Zij zien daarbij toe op het handhaven van de gemaakte afspraken en ontwikkelen procedures om dit mogelijk te maken.

6.5 Stelselafspraken

We spreken over een stelsel, omdat het geheel bestaat uit zeer veel partijen en componenten, die met elkaar interacteren. Voor een goede interactie zijn afspraken nodig, evenals voor het onderhouden en ontwikkelen ervan.

Aansluitvoorwaarden

1. Deelname staat open voor in Nederland gevestigde bedrijven, instellingen en overheidsorganisaties.
2. Deelnemers conformeren zich aan de architectuur en standaarden, zoals vastgelegd in de Referentie Architectuur Digital Twinning in de Fysieke Leefomgeving.

Kwaliteitsborging

3. Deelnemers voorzien hun dataproduct van herkomst- en kwaliteitsinformatie, welke doorloopt tot aan de bron van elk informatie-element (ketenkwaliteit).
4. Deelnemers tonen via jaarlijkse auditverklaring aan dat de door hen geleverde dataproducten in overeenstemming zijn met de daarbij aangegeven kwaliteitsaspecten (incl. actuele herkomstinformatie).

Privacy borging

5. Deelnemers zorgen ervoor dat het gebruik van hun data door derden in overeenstemming is met de algemene verordening gegevensbescherming (AVG). Zij maken daartoe bindende en controleerbare afspraken met afnemers van privacygevoelige data.
6. Deelnemers houden een register actueel met de door hen geleverde dataproducten, die privacygevoelige data bevatten. In dit register (of via een gekoppeld bestand) wordt tevens bijgehouden welke van deze dataproducten aan welke afnemer zijn verstrekt. Op vergelijkbare wijze wordt bijgehouden welke contractuele afspraken met de afnemers van deze privacygevoelige dataproducten zijn gemaakt, zodat het gebruik van deze data binnen de AVG-richtlijn blijft. Zie ook onder aansprakelijkheid.
7. Deelnemers die privacygevoelige data ter bewerking aanbieden aan derden (lees: uitbesteding van analysewerkzaamheden), zorgen via contractuele en controleerbare afspraken voor het verwijderen van deze data door de databewerker na afronding van de bewerkingsactiviteiten.

Informatiebeveiliging

8. Deelnemers die dataproducten leveren die de informatiepositie van criminele organisaties of landen buiten de EU zodanig kan versterken, dat risico's ontstaan met betrekking tot maatschappelijke ontwrichting (terrorisme, cyber crime, fysieke bedreiging of afpersing), zorgen door middel van controleerbare en handhaafbare contractuele afspraken met hun afnemers, dat deze gegevens slechts op grond van een strikte doelbinding kunnen worden ingezet.

Autorisatie

10. Deelnemers aan het stelsel dienen zich via E-Herkenning bij de leverende deelnemer te identificeren.

Doorleverregulering

11. Deelnemers zorgen via "algemene leveringsvoorwaarden" voor sluitende, transparante en handhaafbare afspraken over de afname van dataproducten. Onderdeel hiervan zijn afspraken over het doorleveren van data aan afnemers en afnemers van afnemers in het stelsel (ketenverantwoordelijkheid).

Aansprakelijkheid

12. Deelnemers zijn aansprakelijk voor overtredingen van de AVG door directe en indirecte afnemers van privacygevoelige dataproducten (ketenaansprakelijkheid).
13. Deelnemers zijn aansprakelijk voor het gebruik van hun geleverde data door afnemers en afnemers van afnemers (ketenaansprakelijkheid). Via contractering (incl. algemene leveringsvoorwaarden) kan deze verantwoordelijkheid deels bij afnemers gelegd worden.

Contractering

14. Deelnemers sluiten (raam)contracten af met de afnemers van hun dataproducten, waarbij zaken als privacy borging, informatiebeveiliging en doorlevering aan derden conform de stelselafspraken en ter zake doende wetgeving worden geregeld.

Correctieplicht

15. Deelnemers informeren leveranciers van dataproducten over (vermeende) onjuistheden in de kwaliteit van de geleverde data. Normstellend hiervoor zijn de door de leverancier van een dataproduct meegeleverde kwaliteitskenmerken.

Eigenaarschap van data

16. Data die om niet is verkregen en niet wordt beperkt door eisen met betrekking tot privacy borging en informatiebeveiliging, geldt als openbaar en kan niet door een derde als eigendom worden beschouwd.

Financiële aspecten

17. Deelnemers zijn vrij in het beprijzen van hun informatieproducten.

7 Verantwoording

Dit document is opgesteld in de periode oktober 2021 tot en met januari 2022 tot stand gekomen door samenwerking van architecten van onder meer Geonovum, TNO, Kadaster, Gemeente Utrecht en Gemeente Rotterdam. Deze werkgroep maakt deel van het programma Nederlandse Digitale tweeling voor de Fysieke Leefomgeving.

Er is gebruik gemaakt van een grote hoeveelheid recente literatuur over dit onderwerp. Daarnaast is gebruik gemaakt van voorbeelden uit de praktijk, aangevuld met de vaak jarenlange betrokkenheid van de leden van de werkgroep op het gebied van digital twinning, geo-informatie, modelleringstechnieken, data-analyse, referentie architecturen, etc. Via personele unies, gezamenlijke besprekingen en reviewwerkzaamheden hebben deskundigen uit andere werkgroepen binnen het programma Digital Twin van de Fysieke Leefomgeving invloed uitgeoefend op dit voorlopige eindresultaat. Dit geldt ook voor leden van de werkgroep met vertegenwoordigers uit het bedrijfsleven.

De begeleiding van de werkgroep was in handen van een onafhankelijk deskundige op het gebied van bedrijfs- en informatie-architectuur, drs. Guido Bayens MBA.

7.1 Over de opstellers

De werkgroep waardoor dit document is opgesteld bestond uit de volgende deelnemers:

Guido Bayens	Procesbegeleiding en eindredactie
Jan Bruijn	Programmamanager
Eric van Capelleveen	Geonovum
Jan van Gelder	Architect Geonovum
Michel Grothe	Geonovum
Bart de Lathouwer	Architect Rotterdam
Walter Lohman	TNO
Marcel Reuvers	Kadaster
Frans de Waal	Architect Gemeente Utrecht
Ton Zijlstra	Geonovum

8 bijlagen

8.1 Digitale tweeling en het data mesh concept

Digitale tweeling is een instrument en produceert en consumeert dataproducten. Digitale tweelingen van de Fysieke Leefomgeving consumeren dataproducten van meerdere organisaties en omgekeerd: stellen dataproducten ter beschikking aan andere organisaties.

Om het stelsel van Digitale tweelingen van de fysieke leefomgeving te laten functioneren, zijn naast federatieve technische voorzieningen ook stelselafspraken nodig (hoofdstuk 6 geeft hiervoor een eerste aanzet).

Dataprodukt zijn bijvoorbeeld:

- Operationele data
- Analytische data
- Rekenmodel, algoritme
- Samenstelling, combinatie van 1 of meerdere van de bovenstaande dataproducten

Een dataprodukt kan op zichzelf bestaan of is een dataprodukt met een resultaat dat is ontstaan uit een verzameling van (samengestelde) dataproducten.

De dataproducten worden aangeboden aan het stelsel gebaseerd op open standaard services én dataformaten. De (open) standaarden en dataformaten zijn onmisbaar in de aansluitvoorwaarden. Het is 'data', en dat kan voorkomen als objectenservice, bestand(service), maar ook (near) realtime data streams.

De eigenaar/dataprovider heeft de keuze of men het samengestelde dataprodukt als resultaat van een aantal dataproducten en/of de onderliggende dataproducten aan het stelsel aanbiedt. Het dataprodukt dat aangeboden wordt aan het stelsel voldoet (continu) aan alle ontwerpprincipes, aansluitvoorwaarden waar je aan moet voldoen om van de federatieve componenten gebruik te mogen maken.

Het stelsel is aanbod gedreven. Een organisatie die dataproducten ontwikkelt sluit aan op het stelsel, zorgt voor een goede data uitwisseling van het (samengesteld) dataprodukt met het stelsel waarin andere (samengestelde) dataproducten zich bevinden op basis van open standaard services en dataformaten. Men kan een dataprodukt 'open' beschikbaar stellen voor elke deelnemer van het stelsel, of 'gesloten' uitsluitend voor specifieke gerechtigde gebruikers beschikbaar stellen. Het stelsel maakt zoveel als mogelijk gebruik van Europese, landelijke datamodellen, ontologieën die met elkaar verbonden zijn daar waar zinvol op basis van afgesproken referentieobjecten die behoren tot de referentiegrondslag/fundament (denk aan de basisregistraties).

8.2 Het data mesh concept

De data mesh architectuur gaat uit van een viertal primaire uitgangspunten, die een belangrijke rol spelen in deze referentie architectuur (zie hoofdstukken 4, 5 en 6). De vier uitgangspunten van data mesh zijn:

1. Domein gericht decentraal data-eigenaarschap en architectuur

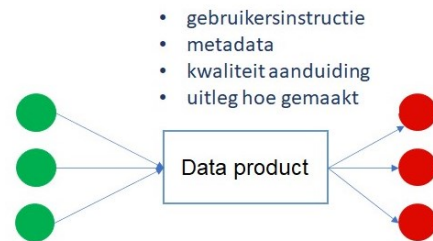
Het concept gaat ervan uit dat organisaties – gegeven hun businessfunctie – verantwoordelijk zijn voor bepaalde data. Bijvoorbeeld: Het Kadaster is verantwoordelijk voor (een deel van de) geo-data, de Kamer van Koophandel voor bedrijfsdata en het Planbureau voor de Leefomgeving voor (een deel van

de) data op het gebied van natuur, milieu en ruimte. Data mesh respecteert deze decentraal belegde verantwoordelijkheden en het daaraan verbonden data-eigenaarschap.

Dit domeingericht denken komt overeen met het idee van de International Data Spaces. We zouden kunnen zeggen dat de fysieke leefomgeving, zoals beschreven in deze referentie architectuur te beschouwen is als een data space, waarbinnen uiteenlopende partijen samenwerken. Er is geen eigenaar van de data space. Het data mesh concept gaat uit van een federatie van samenwerkende partijen.

2. Data als een product

Het dataproduct is een op zichzelf staande eenheid (atomair) en bevat materiedata (operationele data, analytische data of modellen) inclusief metadata, semantiek, ontstaansgeschiedenis, kwaliteit, 'houdbaarheid', e.d. Het dataproduct kan door afnemers – onder bepaalde voorwaarden – eenmalig of meerdere keren worden afgenomen en voor eigen doelen worden ingezet. Data kunnen betrekking hebben op het verleden, het heden of een mogelijke toekomst. Data kan eenmalig verstrekt worden in de vorm van een dataproduct, maar ook realtime, streaming data kunnen als dataproduct geleverd worden.



Het concept Gaia-X is vergelijkbaar op dit punt. In Gaia-X wordt data als een "asset" beschouwd.

3. Federatieve componenten in het stelsel

Het data mesh concept gaat uit van een lichte federatieve infrastructuur. Een of meerdere verbonden netwerk meshtopologieën vormen de technische basis voor deze infrastructuur. Enkele federatieve componenten zijn nodig voor het vinden van de gewenste dataproducten en deze uit te wisselen. Tevens kunnen de federatieve componenten gebruikt worden voor het reguleren van de toegang tot dataproducten door onder meer een identityservice voor het stelsel. De data-infrastructuur is een sterk beveiligde omgeving en maakt een geautomatiseerde zoek- en vindfunctie mogelijk.

Het concept GAIA-X heeft in haar referentie architectuur federatieve componenten ontwikkeld voor de infrastructuur, zoals:

- Federatieve catalogus
- Soevereine data uitwisseling
- Identiteit, authenticatie en autorisatie
- Vertrouwde data uitwisseling
- compliance

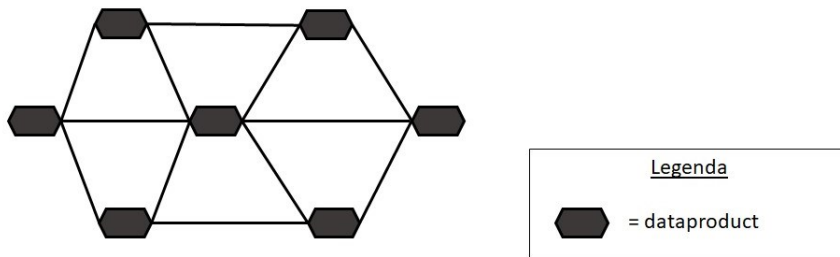
Deze federatieve componenten kunnen we gebruiken in het stelsel.

4. Federatieve governance

Om het stelsel te kunnen laten werken zijn een aantal afspraken en een beheerorganisatie voor de federatieve componenten nodig. Zo zijn onder meer afspraken nodig over bij voorkeur toe te passen ontologieën, de semantiek, de vormgeving van dataproducten en datacommunicatie, zoals vastgelegd in dit document in de hoofdstukken 4 en 5. Daarnaast zijn afspraken nodig over de werking en het beheer van de federatieve componenten binnen het stelsel, zoals kort beschreven in hoofdstuk 6.

Het concept Gaia-X is sterk vergelijkbaar op dit punt, in Gaia-X wordt een "federated governance" voorgesteld waarin de ontwikkeling van dataproducten, "asset" door gedecentraliseerde autonome organisaties gebeurt.

Met het data mesh concept passen we een mesh topologie toe. Dit is een vermaasd netwerk waarin de dataproducten uitgewisseld worden tussen aanbieder en afnemers van data.



Om onderdeel te kunnen zijn van het stelsel dienen de volgende principes gevolgd te worden.

- **Kwaliteitsgarantie en leverbetrouwbaarheid**

De eigenaar van dataproduct heeft de zorgplicht dat het dataproduct geleverd wordt conform de meegeleverde specificaties. Deze kunnen onder meer zaken betreffen als vertrouwelijkheid, doorlevermogelijkheden, houdbaarheid, continuïteit, etc. De federatieve componenten in het stelsel hebben functionaliteit waarmee de kwaliteit, beveiliging, performance en continuïteit van de dataproducten gemonitord wordt.

Tussen of binnen een domein wordt brondata niet samengebracht in een datawarehouse, datalake of andere omgeving. De data worden op bovenstaande wijze (1 t/m 3) verbonden op basis van open standaard services zodat een verbonden stelsel van data ontstaat gericht op herbruikbaarheid.

- **Er bestaan meerdere contexten van de werkelijkheid**

De data hebben binnen het domein waarin ze zijn ontstaan een specifieke betekenis. Domeinen gebruiken, vanwege het taak-specifieke karakter vaak om goede redenen, verschillende begrippen in een andere betekenis of dezelfde begrippen met een verschillende betekenis. Zo kan een 'klant' in het ene domein vergelijkbaar zijn met wat in een ander domein 'cliënt' wordt genoemd. Andersom kan het begrip 'klant' in meerdere domeinen worden gebruikt maar een, ietwat of sterk, verschillend betekenis hebben.

- **Het nDTFL stelsel hanteert de basisregistraties en relaties tussen de basisregistraties als fundament**

Binnen de nDTFL maken we gebruik van dezelfde ondergrond, dit zijn de dataproducten van de basisregistraties. Basisregistraties zijn context onafhankelijk en zijn het fundament onder de data uit de domeinen. Zonder fundament hanteren de domeinen verschillende referenties en kunnen we de resultaten niet meer met elkaar vergelijken. Denk aan het gebruik van topografie die op een andere plek ligt dan de Basisregistratie Grootchalige Topografie (BGT).

De (basis)registraties bestaan uit de Basisregistratie Grootchalige Topografie (BGT), Basisregistratie Topografie (BRT), Basisregistratie Adressen en Gebouwen (BAG), Basisregistratie Ondergrond (BRO), Basisregistratie Kadaster (BRK), Landelijke Voorziening Beeldmateriaal, 3D Basisvoorziening²⁰.

Het onderscheid dat we van oudsher hebben tussen operationele (veelal transactie-gebaseerde) en analytische data zal steeds meer vervallen. Analytische data maken direct gebruik van de operationele data en levert vaak ook weer feedback op de operationele data. Analytische data kunnen weer operationele data zijn voor een ander nDTFL proces. We brengen data naar de modellen en modellen naar de data. Dit organisch stromen van data faciliteren we door de domein-structuur. We hanteren hier het concept van data als een product, zie volgende bullet.

²⁰ Tijdens de uitvoering zullen de referentieregistraties aangepast worden op behoeften

- **Data als een product**

Data als een product lijkt vanzelfsprekend omdat we dit naar ons gevoel al doen. Echter in de huidige infrastructuur beschouwen we data en techniek als 2 aparte dingen terwijl bij data als een product dit een is. Data als een product heeft daarnaast nog een subtiel maar essentieel verschil. Nu leveren we databestanden waar de gegevensgebruiker zelf zijn relevante gegevens maar moet uithalen en vaak zelf moet omzetten naar werkbaar formaat. Met data als een product voegen we product denken (product thinking) toe. De gebruikers moeten als klant worden behandeld en wel op zo'n manier dat de klant gelukkig wordt van de wijze waarop de data geconsumeerd kan worden. Succesvolle producten zijn waardevol, eenvoudig bruikbaar en haalbaar om te maken. Het concept van succesvolle producten is ook toegepast op data als een product.

De data als een product is een atomaire eenheid. Het is niet zo dat op andere plekken 'metadata' wordt opgeslagen van het dataproduct maar bij data mesh is al deze informatie aanwezig bij het dataproduct zelf. Dit past bij de federatieve aanpak.

Dataprodukten bestaan uit verschillende typen informatie. Operationele data of een model, het informatiemodel, begrippen of andere typen zoals ontstaansgeschiedenis, tijdslijnen, kwaliteit, etc. De typen variëren op basis van de aard van het dataproduct.

Een dataproduct kan bestaan uit:

- transactie data – data in transactionele werkprocessen zoals bijvoorbeeld topografie transacties, maar ook sensor data of events;
- analytische data – dit is transactionele data die gecombineerd, bewerkt, verrijkt en/of geaggregeerd is uit verschillende bronsystemen. Analytische data bieden nieuwe kennis en inzichten;
- rekenmodellen, AI of andere modellen
- samenstelling, combinatie van 1 of meerdere van bovenstaande dataprodukten

Dataprodukten bieden data aan gebruikers of afnemende applicaties zoals bijvoorbeeld analyse software, een digitale tweeling, 2/3D-viewers, dashboards, en is softwareproduct (tool) onafhankelijk.

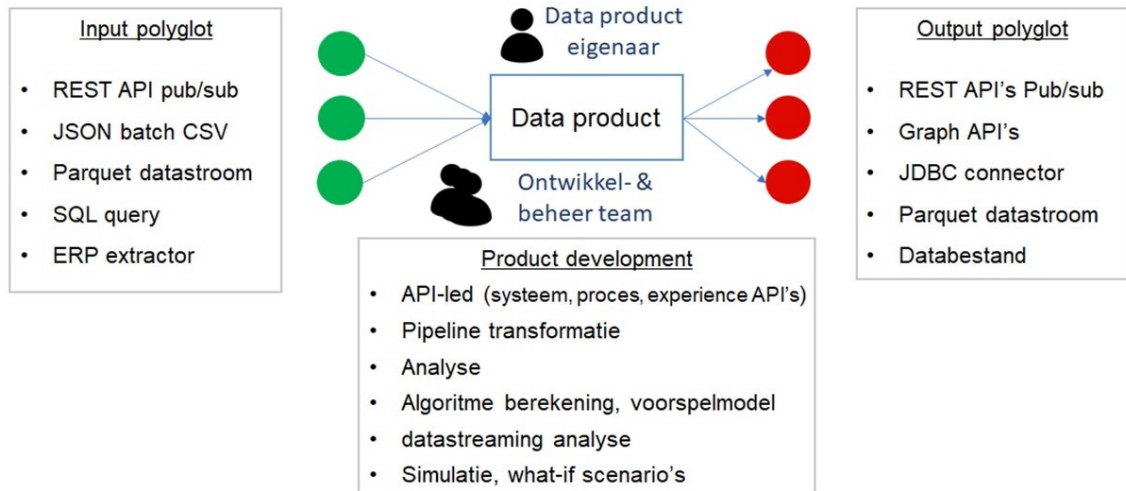
Polyglot input- en output data

- De input en output is in verschillende open standaard dataformaten (bijvoorbeeld graph's, JSON, XML) waarbij de semantiek hetzelfde is uitgedrukt (dit is in hoofdstuk 5 constructie principes verder uitgewerkt)
- De input en output is op basis van open standaard mechanisme zoals: batch; streaming data; API's; geo-webservices en query.
- De input en output is transactionele- en/of analytische data

Binnen het stelsel is het geen eis dat input en output data door mensen gelezen moet kunnen worden. We spreken immers over een 'system of systems'. Daarom dient data wel altijd door machines/ computers gelezen te kunnen worden. Voor de uitwisseling van data tussen machines kunnen ook 'binaire' standaarden gebruikt worden, zoals bijvoorbeeld Google Protobuf. Standaarden als JSON en XML zijn weliswaar ook voor mensen leesbaar, maar voor machine – machine communicatie vaak uiterst inefficiënt.

Dataprodukt ontwikkeling

- Op basis van de input wordt door deelnemende organisaties output data ontwikkeld op basis van diverse (moderne) technieken zoals pipelines voor data transformatie, algoritmen voor data transformatie, streaming data-analyse, etc.; Op basis van de technieken worden dataprodukten ontwikkeld en beheerd.



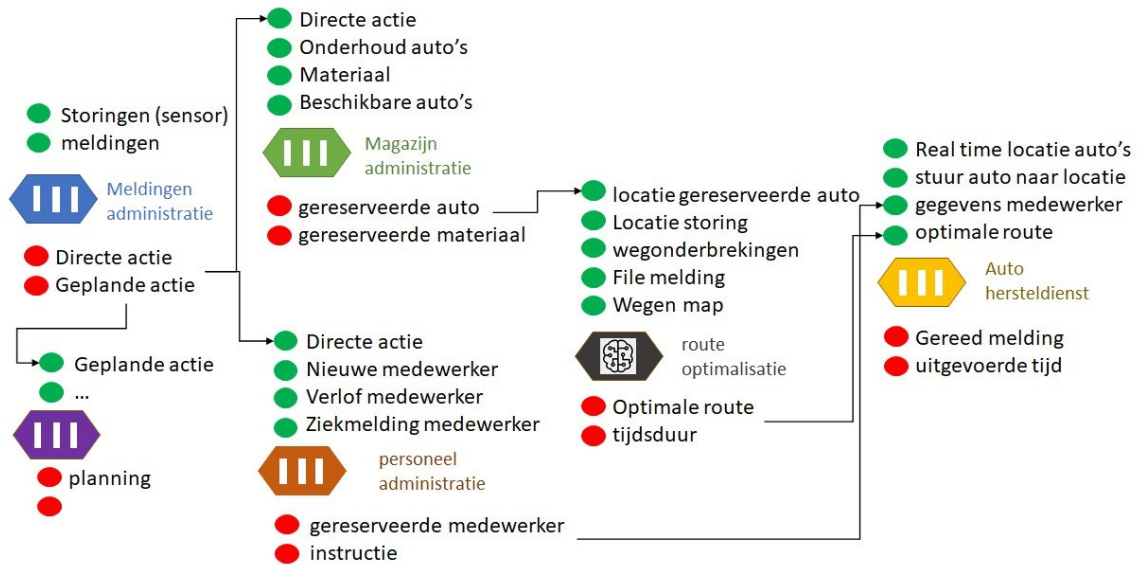
Dataproduct typen

Van een of meerdere dataproducten kan ook een nieuw dataproduct gemaakt worden. Dit kan een samenstelling zijn of het resultaat van een algoritme zoals een tijdelijke analyse, voorspelmodel, e.d. Dit nieuwe product kan ook weer een dataproduct worden dat door andere stakeholders gebruikt kan worden in hun eigen organisatie. Een dataproduct wordt ontwikkeld door een afzonderlijke organisatie en heeft daarmee één producteigenaar. Voor dataproducten geldt dat ze moeten voldoen aan constructieprincipes van het stelsel.



Dataproducten domeinen

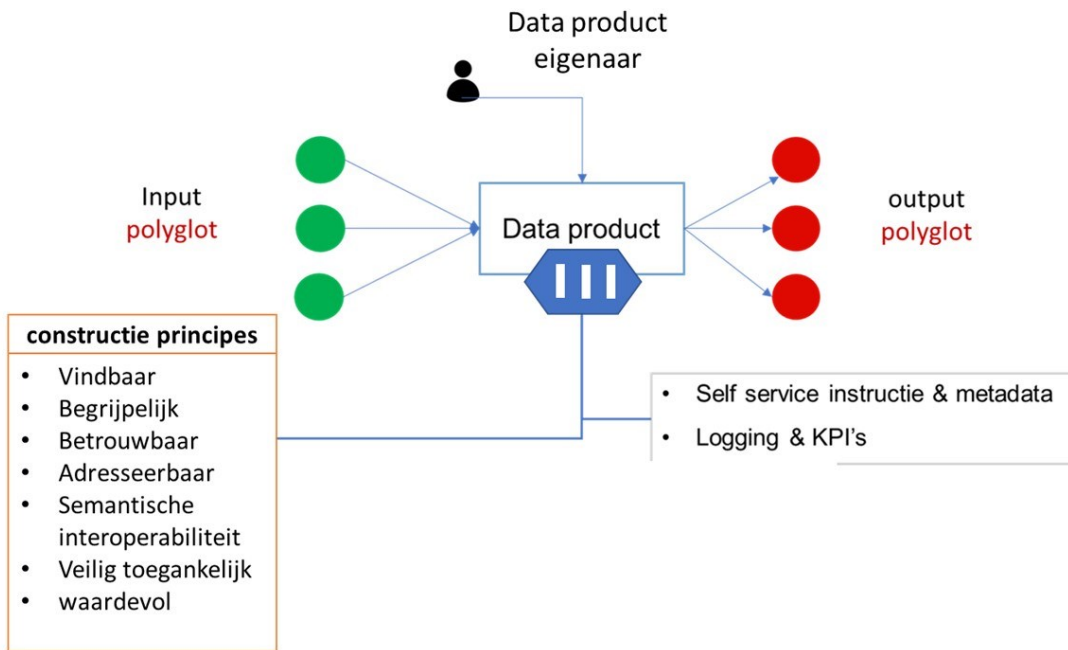
Informatie vraagstukken kunnen betrekkingen hebben op meerdere dataproducten. Ter illustratie een voorbeeld voor verkeersregel installaties. Hierin worden dataproducten in een keten gebruikt waarin een dataproduct een algoritme als transformatie technologie inzet.



Dataproduct keten

Dit kunnen we ook toepassen op een overkoepelend vraagstuk waarbij verschillende dataproducten zijn gebruikt. Bijvoorbeeld het vraagstuk of Nederland voldoende beschermd is tegen het water in geval van enorme regenval en smeltwater in de grote rivieren zoals de Maas en de Rijn. De dataproducten worden dan gebruikt in een informatievraagstuk over water risico's.

Dataproducten hebben kenmerken oftewel ontwerpprincipes die toegepast moeten worden in de ontwikkeling van het dataproduct om te kunnen (mogen) aansluiten op het landelijke stelsel voor digitale tweelingen van de fysieke leefomgeving.



Dataprodukt ontwerpprincipes
Ontwerpprincipes dataprodukt:

- **Vindbaar**

Vindbaarheidsinformatie wordt verstrekt door het dataprodukt zelf.

Vindbaarheid van dataprodukten binnen het nDTFL-stelsel gebeurt aan de bron. Het is aan de individuele dataproducenten om hun vindbaarheidsinformatie op een standaard manier te delen. Elk dataprodukt deelt continue zijn oorsprong, eigenaren, runtime-informatie zoals tijdigheid, kwaliteitsstatistieken, voorbeeldgegevenssets en vooral informatie die is bijgedragen door hun consumenten, zoals de belangrijkste gebruiksscenario's en toepassingen die door de gegevens mogelijk worden gemaakt. Via een centrale mogelijkheid worden deze federatieve catalogus en index bevraagt voor een totaaloverzicht.

- **Begrijpelijk**

Het dataprodukt is zelfbeschrijvend.

Uit welke begrippen, definities, relaties, waardelijsten, e.d. – de semantiek - bestaat het dataprodukt? Naast begrip van de semantiek, moeten datagebruikers begrijpen hoe de gegevens aan hen worden gepresenteerd. Hoe hebben ze toegang? Hoe kunnen ze de data syntactisch opvragen? Wat voor soort SQL-query's kunnen ze uitvoeren? Hoe lees je de data-objecten? Dit gaat idealiter vergezeld van voorbeelddatasets en voorbeeldcodes.

Ten slotte is begrijpbaarheid een sociaal proces. Wij leren van elkaar. Het moet eenvoudig zijn om als datagebruiker ervaringen te delen die weer zorgen voor betere ondersteuning op bruikbaarheid.

- **Betrouwbaar en waarheidsgetrouw**

Elk dataprodukt heeft Service Level Objectives (SLO's) die een acceptabel kwaliteitsniveau garanderen en worden nageleefd door de dataprodukt eigenaar.

Een SLO omvat onder meer:

- Interval van verandering
- Actualiteit
- Volledigheid
- Statistische vorm van gegevens
- Ontstaansgeschiedenis – de datareis
- Datakwaliteit (inzichtelijk volgens een generiek (landelijk) datakwaliteit framework)
- Precisie en nauwkeurigheid in de loop van de tijd
- Operationele kwaliteiten

Hoe beter de data-eigenaar deze gegevens bij zijn dataprodukt vastlegt, bijhoudt en beheert, hoe meer vertrouwen de datagebruikers zullen hebben in het dataprodukt zelf. De SLO is op termijn zodanig gestandaardiseerd dat de tooling van een afnemende partij deze als code kan lezen.

- **Adresseerbaar**

Een dataprodukt moet een vast en uniek identificatie hebben voor de datagebruiker om het handmatig of vanuit software te kunnen openen.

- Een uniek *identificatie* moet de dynamische aard van de data en de mesh-topologie ondersteunen. Het moet aspecten omvatten van dataprodukten die continu kunnen veranderen en tegelijkertijd de continuïteit van het gebruik ondersteunen. **Semantische interoperabiliteit**

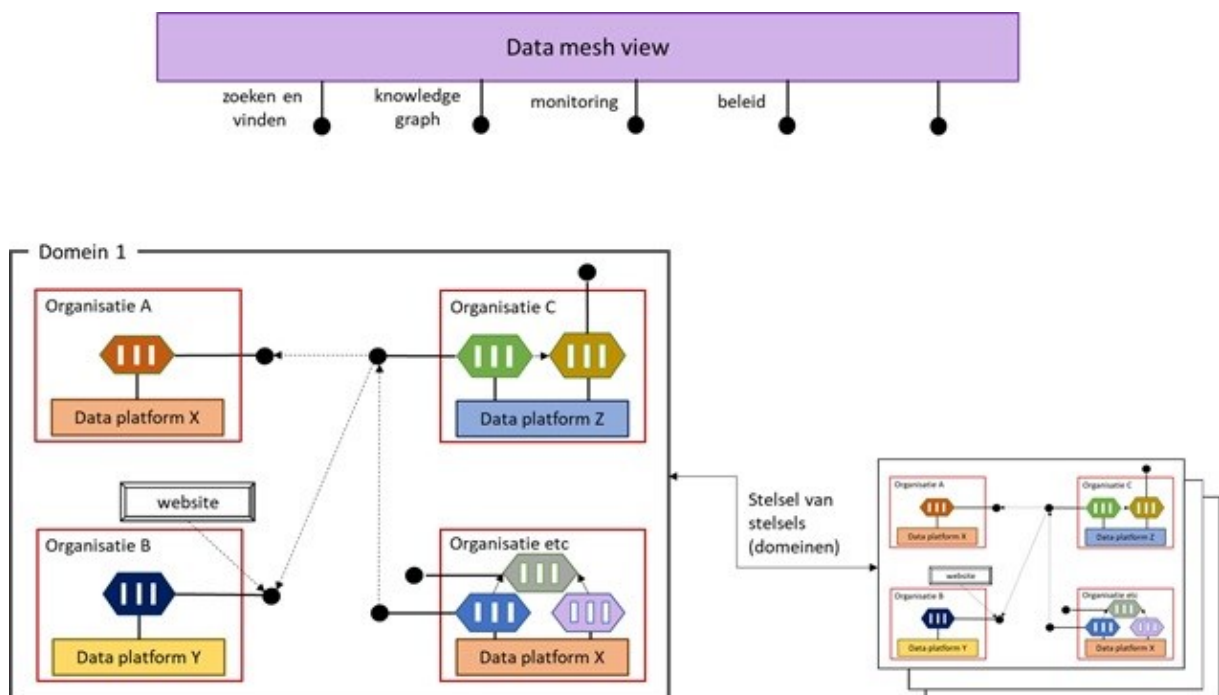
De dataproducten moeten zoveel als mogelijk met elkaar inhoudelijk combineerbaar kunnen worden.

Een van de belangrijkste zorgen in een gedistribueerde data-architectuur is de mogelijkheid om data te correleren/in samenhang over domeinen heen te gebruiken. De sleutel voor semantische interoperabiliteit van data tussen domeinen is het volgen van standaarden en harmonisatieregels, die het mogelijk maken om gegevens over domeinen heen zinvol in samenhang te bevragen. De volgende requirements zijn van belang die gespecificeerd worden in hoofdstuk 5 Constructie principes:

- Volgen van [Metamodel informatiemodellering](#) (MIM) voor [informatiemodellen](#) ,
- Gestandaardiseerde metadata – gebaseerd op wetgeving zoals Wet Open Overheid, Archiefwet (MDTO), etc.
- Geharmoniseerde begrippen en definities – zie bijvoorbeeld ontwikkelingen in programma Omgevingswet
- (Knowledge) graphs – Ontwikkelingen verder uitwerken zoals SAREF <https://saref.etsi.org/>; <https://digital-strategy.ec.europa.eu/en/news/etsi-releases-three-new-saref-ontology-specifications-smart-cities-industry-40-and-smart>
- Mogelijkheid tot hergebruik en koppeling aan begrippen, informatiemodellen, graphs, etc.
- Naar schema's - typen - gedefinieerd door andere dataproducten.

De werking in samenhang beschreven

Door de federatieve opzet zal iedere organisatie zelf moeten zorgen voor het ontsluiten van de dataproducten die ze willen delen met anderen. Dit delen gebeurt met services of API's vanuit de bron. Voorwaarden hierbij zijn dat het vindbaar, begrijpelijk, betrouwbaar en waarheidsgetrouw, adresseerbaar, semantisch interoperabel, toegankelijk en veilig en natuurlijk waardevol is voor de gebruikers. De voorwaarden worden bij een federatieve opzet ingevuld door de bronhouder of aanbieder namens de bronhouder. Het is hierdoor zeer decentraal van opzet. Aan de bron worden de privileges toegekend wie wat mag. Deze federatieve aanpak kan alleen werken als we ons houden aan bepaalde afspraken. Deze worden met elkaar gemaakt en beheerd.



In bovenstaande figuur zien we dat binnen een domein organisatie A het data platform X hanteert en hetzelfde platform ook gebruikt wordt door andere organisaties. Dit kan een samenwerking zijn op basis van open source, closed software of een combinatie. Hier doen we geen uitspraken over, dit is aan het domein. Organisatie A heeft een dataproduct gepubliceerd. Dit kan van alles zijn, van statische gegevens tot bijvoorbeeld een continue stroom van meetgegevens. Het dataproduct voldoet aan de eerder gestelde voorwaarden. Omdat andere organisaties deze gebruiken helpt organisatie A de gebruikers met events op veranderingen op data en het informatiemodel. Ze hebben ook documentatie hoe het dataproduct snel gebruikt kan worden en voor welke doeleinden. De gebruikers komen bij deze informatie door aan te sluiten op de node.

Organisatie B gebruikt een ander data platform en gebruikt het dataproduct in de organisatiewebsite. Organisatie C maakt downstream gebruik van de dataproducten van organisatie A en B. Deze worden met een machine learning (ML) model (groen) verwerkt, en dit leidt tot een nieuw product dat upstream voor hergebruik beschikbaar wordt gesteld. Het ML-model is tevens volgens de randvoorwaarden gepubliceerd. Tenslotte maakt een andere organisatie weer gebruik van het ML-model van organisatie C dat als service beschikbaar is en laat dit los op het dataproduct (blauw). Dit wordt gecombineerd met het paarse interne dataproduct en leidt tot een nieuw samengesteld dataproduct (grijs) dat voor hergebruik beschikbaar wordt gesteld.

Bovenstaande manier van werken is recursief. Het werkt ook zo voor de andere domeinen, organisaties, afdelingen, etc.

Met de data mesh view hebben we een basisinfrastructuur waarmee we binnen een domein en tussen de domeinen het makkelijker maken om dataproducten te kunnen vinden, semantiek samenhang beter kunnen duiden en hergebruiken, op kwaliteit van services wordt gemonitord of op het gebruik van de standaarden. Ook een centrale trustvoorziening kan nodig zijn om toegangsaspecten binnen het stelsel goed te kunnen inrichten of andere dingen waarvan we met elkaar vinden dat deze nodig zijn.

Samenwerkende organisaties binnen het stelsel, of beter nog, samenwerkende computers, dienen te weten bij welke organisatie bepaalde data, modellen en verwerkingscapaciteit aanwezig zijn. Daarvoor hebben we centrale wegwijzers of indexen nodig want bij een federatief stelsel zal de dataproducteigenaar metadata niet opnieuw opvoeren op de basisinfrastructuur. De metadata staat bij het dataproduct en we maken geen onderscheid meer tussen data en metadata. Het is allemaal data geworden die via indexen gevonden wordt. Dergelijke indexen kunnen ook informatie bevatten over de toegankelijkheid van bepaalde dataproducten, modellen of verwerkingscapaciteit. Dit geeft wel aan dat het niet voldoen aan de randvoorwaarden ook zorgt voor het niet vindbaar zijn van het dataproduct. Een federatieve infrastructuur kan niet zonder volwassen eigenaarschap.

Een datacommunicatie-infrastructuur is uiteraard essentieel voor het transporteren van data. Deze infrastructuur bestaat op zijn beurt weer uit verschillende componenten en werkt op basis van veelal mondiale afspraken over toe te passen standaarden. Delen van deze infrastructuur vallen onder verantwoordelijkheid van private partijen, delen onder samenwerkingsverbanden en weer andere delen vallen onder de overheid. Digitale tweelingen maken gebruik van deze infrastructuur.

Hoe werkt dit dan voor een organisatie die een vraagstuk heeft en gebruik wil maken van de nationale digitale tweeling stelsel? Een organisatie beschikt over de nodige hulpmiddelen, tooling, voor het manipuleren van data, het ontwikkelen van (nieuwe) modellen, het uitvoeren van simulaties en het doen van voorspellingen. Zij gebruiken hiervoor een combinatie van dataproducten die ze zelf heeft laten ontstaan of verworven (sensoren, apparaten) en dataproducten die via het digitale tweeling stelsel zijn verkregen. Hiermee wordt het mogelijk om bijvoorbeeld geografische informatie te combineren met resultaten van milieumetingen, verkeersgegevens en meteorologische modellen.

Deze gecombineerde modellen kunnen het resultaat zijn van bepaalde momentopnamen, maar in toenemende mate zijn we in staat om ook met direct aan de werkelijkheid ontleende data (realtime) te werken, waardoor een digitale tweeling van dit deel van de fysieke leefomgeving ontstaat. Met de dataproducten is het dus mogelijk om complexe modellen en visualisaties samen te stellen. Denk hierbij aan het samenstellen van dynamische modellen die het gedrag van verkeersdeelnemers voorspellen in geval van een wegafsluiting of de gevolgen voor fijnstof en geluid bij de verbreding van een weg.

De organisatie kan dit zelf gebruiken en is vrij om het dataproduct weer beschikbaar te stellen voor hergebruik indien andere gebruikers hierom vragen.