

**Datamanagementplan  
Kustgenese 2.0 (versie 1.0)**

Programma 2017-2019





# **Datamanagementplan Kustgenese 2.0 (versie 1.0)**

**Programma 2017-2019**

dr. ir. K. den Heijer  
ir. G. Santinelli

Met medewerking van ir. C.H. Clemens (Witteveen+Bos)

1220339-015



**Titel**  
Datamanagementplan Kustgenese 2.0 (versie 1.0)

<b>Opdrachtgever</b>	<b>Project</b>	<b>Kenmerk</b>	<b>Pagina's</b>
Rijkswaterstaat WVL	1220339-015	1220339-015-ZKS-0001	25




**Trefwoorden**  
Kustgenese 2.0, datamanagement, morfologie, Ameland, zeegaten, diepe vooroever

**Samenvatting**  
Als onderdeel van de Kustgenese 2.0 onderzoekslijn wordt monitoring en modellering uitgevoerd. Het is van belang dat de data van de monitoringsprogramma's in het Amelandse Zeegat en op de diepe vooroever bij Noordwijk, Ameland en Terschelling in een beheerde omgeving wordt geborgd en ontsloten. Hetzelfde geldt voor door modellen geproduceerde data.

Dit datamanagementplan beschrijft de rollen en verantwoordelijkheden van betrokken partijen (Rijkswaterstaat, SEAWAD en Deltares), de technische structuur van het datamanagementsysteem en de werkwijze, procedures en afspraken.

Dit is versie 1.0 van het datamanagementplan. Het plan wordt twee keer per jaar geëvalueerd en zo nodig geüpdatet.

**Referenties**  
Kustgenese 2.0 – zaaknummer 31123135

Versie	Datum	Auteur	Paraaf	Review	Paraaf	Goedkeuring	Paraaf
1.0	maart 2018	K. den Heijer G. Santinelli		W. Stolte		D. Walstra	

**Status**  
definitief



## Inhoud

<b>1</b>	<b>Introductie</b>	<b>1</b>
1.1	Inleiding Kustgenese 2.0	1
1.2	Doelstellingen datamanagement Rijkswaterstaat	2
1.3	Doelstelling datamanagementplan	2
1.4	Welke data in het Kustgenese 2.0 datamanagementsysteem?	3
1.5	Huidige versie en wijzigingen ten opzichte van vorige versie	3
1.5.1	Huidige versie	3
1.5.2	Wijzigingen ten opzichte van vorige versie	3
1.6	Leeswijzer	4
<b>2</b>	<b>Achtergrond en definities datamanagement(systeem)</b>	<b>5</b>
<b>3</b>	<b>Taken en Verantwoordelijkheden</b>	<b>7</b>
3.1	Betrokken partijen	7
3.2	Beschrijving taken en verantwoordelijkheden	7
3.2.1	Product eigenaar - Rijkswaterstaat	7
3.2.2	Dataleveranciers – diverse partijen	7
3.2.3	Datagebruiker – diverse partijen	8
3.2.4	Databeheerder - Deltares	8
<b>4</b>	<b>Inwinning, verwerking en gebruik data</b>	<b>9</b>
4.1	Introductie	9
4.2	Uitgangspunten	10
4.3	Vorbereiding	10
4.3.1	Verkrijgen van toegang tot de Subversion repository Kustgenese 2.0 - voor opslag en verwerking ruwe data	10
4.3.2	Invullen standaard datamanagement formulier	10
4.3.3	Opslag in repository	11
4.4	Uitlezen ruwe data	11
4.5	Opslag ruwe data	11
4.6	Opwerking ruwe data, validatie verwerking en opslag verwerkte data	11
4.6.1	Opwerkingscripts	11
4.6.2	Standaarden	12
4.6.3	Review/test scripts	12
4.6.4	Validatie	13
4.6.5	Opslag en ontsluiting van standaard data	13
4.7	Datagebruik	14
4.8	Overige data	14
4.8.1	Verwerking overige data	14
4.8.2	Modeltoepassing	14
4.9	Trainingen	15
<b>5</b>	<b>Datamanagementsysteem</b>	<b>17</b>
5.1	Introductie	17
5.2	Onderdelen datamanagementsysteem (en Minimal Viable Product)	17
5.2.1	Basis bouwstenen (generiek toepasbaar)	18
5.2.2	Extra bouwstenen	19

5.3	Aandachtspunten	19
5.4	Beheer en onderhoud datamanagementsysteem	20
5.5	Eisen en wensen	20
<b>6</b>	<b>Administratieve en Juridische afspraken</b>	<b>21</b>
6.1	Introductie	21
6.2	Afspraken betreffende verzamelen, managen en delen van data (dataproducent)	21
6.3	Afspraken betreffende data uitwisseling (data gebruiker)	21
6.4	Eigendom	21
<b>7</b>	<b>Archivering, delen en publiceren van data</b>	<b>23</b>
7.1	Introductie	23
7.2	Functionaliteiten Archief	23
7.3	Vragen RWS beleid	23
7.3.1	Bestaande archief van RWS is leidend	23
<b>8</b>	<b>Referenties</b>	<b>25</b>
 <b>Bijlage(n)</b>		
<b>A</b>	<b>Dataset status overview</b>	<b>A-1</b>
<b>B</b>	<b>Manual on using SubVersion (SVN)</b>	<b>B-1</b>
<b>C</b>	<b>Standard Data Management Template</b>	<b>C-1</b>
<b>D</b>	<b>Repository Kustgenese 2.0 directorystructuur</b>	<b>D-1</b>



# 1 Introductie

## 1.1 Inleiding Kustgenese 2.0

Kustgenese 2.0 is een onderzoekslijn van het Deltaprogramma Kust en komt voort uit de strategische beslissing Zand (Deltaprogramma 2015). Kustgenese 2.0 maakt ook onderdeel uit van het Nationaal Kennis- en innovatieprogramma Water en Klimaat (NKWK; [Link](#)<sup>1</sup>).

Kustgenese 2.0 richt zich op zes thema's:

1. Langetermijn-kustontwikkeling: aanvullende monitoring en modelontwikkeling om meer inzicht krijgen in de werking van het zeegatsysteem, uitwisseling op dieper water en andere factoren die het benodigde suppletievolume bepalen.
2. Pilotsuppletie buitendelta Amelander Zeegat: aanleggen van een pilotsuppletie om te leren over de morfologie en ecologie van zeegaten.
3. Ecologie: effecten en kansen van veranderende suppletiehoeveelheden en suppletielocaties voor de ecologie.
4. Lerend werken: ervaring opdoen met mogelijke strategieën door de inzet van proeftuinen en proefsuppleties.
5. Datamanagement: alle data die worden gebruikt of zijn ingewonnen moeten op een veilige en toegankelijke manier worden beheerd.
6. Opkomende onderwerpen: zoals de optimalisatie van zandwinning, ruimtelijke ordening en economie.

Als onderdeel van thema 5 Datamanagement heeft Rijkswaterstaat Water, Verkeer en Leefomgeving (WVL) aan Deltares gevraagd om een datamanagementsysteem op te zetten en te onderhouden. Het datamanagementsysteem ondersteunt alle thema's.

Tegelijkertijd voert Deltares voor de langetermijn-kustontwikkeling (thema 1) in opdracht van Rijkswaterstaat WVL een 3-jarig onderzoeksprogramma uit. Voor de langetermijn-kustontwikkeling zijn reeds diverse beschrijvende documenten opgesteld over de informatiebehoefte [ref. 1]. Het onderzoeksprogramma is een belangrijke gebruiker van het datamanagementsysteem, maar dus niet de enige gebruiker.

De looptijd van het onderhoud van het datamanagementsysteem door Deltares is gelijk aan de looptijd van het 3-jarig onderzoeksprogramma, dat wil zeggen tot 1 oktober 2019. Het is de verwachting dat het systeem na deze datum aan Rijkswaterstaat Centrale Informatievoorziening (CIV) of marktpartij wordt overgedragen. Dit besluit wordt medio 2019 genomen.

---

<sup>1</sup> <https://waterenklimaat.nl/onderzoekslijnen/kustgenese/> d.d. 18 december 2017

## 1.2 Doelstellingen datamanagement Rijkswaterstaat

### *Rijkswaterstaat WVL - overkoepelend*

Rijkswaterstaat WVL heeft in het programma van eisen de volgende voorwaarden geformuleerd:

- Rijkswaterstaat WVL wil de softwareoplossingen die reeds in opdracht door Deltares zijn ontwikkeld, uitwerken tot een datamanagementsysteem, welke bestaat uit generieke onderdelen die ingezet kunnen worden binnen de projecten en programma's die vanuit WVL de komende 2 tot maximaal 5 jaar worden uitgevoerd, waaronder: Kustgenese 2.0, Digitale Delta, Rivercare, DigiShape en de Zandmotor.
- Het datamanagementsysteem heeft tot doel de data op te slaan en te ontsluiten. De data, inclusief metadata, moet inzichtelijk zijn voor alle geïnteresseerde partijen (beschikbaar achter een login). Het datamanagement dient hierin te voorzien en tevens aan te sluiten bij de hiervoor geldende (open) datastandaarden.
- Nadat Kustgenese 2.0 is afgerond is het nodig de data te archiveren. Het is noodzakelijk om hier op voorhand afspraken over te maken. Deze worden uitgewerkt in het datamanagementplan, zodat RWS hier intern nadere afspraken over kan maken.

Een aanvullende wens is om het beheer van data na afloop van een project/programma te stroomlijnen, zodat invulling duidelijk is en indien nodig kan worden aangesloten op de CIV Bouwstenencatalogus. De ontwikkeling en opzet van het datamanagementsysteem Kustgenese 2.0 staat dus niet op zichzelf, maar staat in de bredere context van de ontwikkeling en opzet van datamanagementsystemen binnen Rijkswaterstaat. CIV is samen met WVL verantwoordelijk voor deze ontwikkeling en opzet, waarmee minimaal twee keer per jaar zal worden afgestemd.

*Openstaande vraag te verwerken in volgende versie datamanagementplan: Afspraken maken en vastleggen over verantwoordelijkheden (in overleg met John Schobben), (aan)sturing en beslissingsbevoegdheid. Wat te doen als de KG2 aanpak niet (volledig) in lijn is met de CIV aanpak?*

De genoemde ontwikkelde generieke softwareoplossingen – als onderdelen van de zogenaamde OpenEarth stack – worden in §5.2 beschreven.

### *Kustgenese 2.0 - dit programma*

Voor Kustgenese 2.0 is opdracht gegeven een datamanagementsysteem op te zetten, dat is samengesteld uit generieke softwareoplossingen. Het datamanagementsysteem betreft een concrete toepassing van de generieke onderdelen met specifiek het doel kwaliteit-geborgde data op een transparante en herbruikbare wijze te delen.

Gedurende de looptijd wordt twee keer per jaar het proces rond de totstandkoming, het gebruik van het datamanagementsysteem en het beheer van de data vastgelegd en geëvalueerd. Dit met als doel het proces rond de realisatie en het gebruik van het datamanagementsysteem vast te leggen en te optimaliseren.

## 1.3 Doelstelling datamanagementplan

Dit datamanagementplan beschrijft hoe voor Kustgenese 2.0 het datamanagement wordt ingevuld. Datamanagement omvat zowel het technische platform als de organisatie, afspraken, rollen en verantwoordelijkheden eromheen. Dit plan gaat op beide in.

Datamanagement kan gedurende de looptijd aan verandering onderhevig zijn, bijvoorbeeld door nieuwe inzichten of door praktische keuzes. Het datamanagementplan is daarom een levend document dat minimaal eens per jaar wordt geëvalueerd en zo nodig wordt aangepast. In §1.5 wordt de versie van het datamanagementplan beschreven.

#### 1.4 Welke data in het Kustgenese 2.0 datamanagementsysteem?

Het Kustgenese 2.0 datamanagementsysteem bevat data van:

- De monitoringcampagne in het Ameland Zeegat in augustus en september 2017. De monitoringcampagne is een samenwerking tussen Rijkswaterstaat, kennisinstituut Deltares en de 3 'SEAWAD-universiteiten' TU Delft, Universiteit Utrecht en Universiteit Twente .
- De monitoringcampagne op de diepe vooroever van Noordwijk, Ameland en Terschelling.
- De monitoringcampagnes in relatie tot de pilotsuppletie Ameland Zeegat. De pilotsuppletie wordt naar verwachting in 2018 uitgevoerd. Monitoring vindt plaats voor, tijdens en na de suppletie.

Voor iedere monitoringcampagne is een monitoringplan beschikbaar waaraan een specificatie en overzicht van informatiebehoefte ten grondslag ligt. Het is aan te bevelen om dergelijke plannen als achtergrondmateriaal op te nemen in het datamanagementsysteem.

Het Kustgenese 2.0 datamanagementsysteem bevat niet:

- Data die al in een andere datamanagementsysteem met vergelijkbare eisen en functionaliteit is opgeslagen. Dit om duplicatie te voorkomen.
- Xband radar ruwe data. Vanwege het grote datavolume van de radar data is het opslaan van de ruwe data op dezelfde manier als de andere datasets niet passend. Er wordt naar een andere oplossing gezocht om deze data beschikbaar te maken.
- Vaklodingen (deze worden ontsloten via OpenEarth, samen met de reguliere vaklodingen)
- Beschrijving van boorkernen (deze worden ontsloten via het DINO-loket)

Uitbreiding van de data in het Kustgenese 2.0 datamanagementsysteem is mogelijk na overleg met en toestemming van Rijkswaterstaat WVL.

#### 1.5 Huidige versie en wijzigingen ten opzichte van vorige versie

##### 1.5.1 Huidige versie

Dit is versie 1.0 (concept) van het datamanagementplan. Het beschrijft de eisen van Rijkswaterstaat WVL voor de eerste versie ("Minimum Viable Product"; MVP) van het op te zetten datamanagementsysteem. Onderhevig document bevat de afspraken rond data. De actuele versie van dit document is te vinden in de Kustgenese repository/datamanagement/datamanagementplan.

##### 1.5.2 Wijzigingen ten opzichte van vorige versie

Er is nog geen vorige versie.

## 1.6 Leeswijzer

Van de lezer worden affiniteit met en achtergrond in datamanagement en datamanagementtechnieken verwacht. Delen van dit document gaan in op technische specificaties die voor niet-ingewijden complex kunnen zijn. Hoofdstuk 3 beschrijft de taken en rollen van deelnemende partijen en van hun medewerkers. Hoofdstuk 4 beschrijft de stappen die nodig zijn om te komen tot het gebruiken van data. Hoofdstuk 5 beschrijft de technologische onderdelen van het datamanagementsysteem en de bijbehorende afspraken.

Daaraan voorafgaand voorziet Hoofdstuk 2 in een beknopte introductie in de achtergrond en gebruikte definities. Voor verdere inleiding worden hands-on trainingen en presentaties gegeven. Trainings- en presentatiemateriaal is beschikbaar op de repository.

Dit document bevat op diverse plaatsen nog uit te voeren acties. Deze zullen worden opgepakt en uitkomsten zullen voor zover mogelijk in de volgende versie van dit plan worden verwerkt.

## 2 Achtergrond en definities datamanagement(systeem)

Datamanagement betreft het beheren van data. In het Kustgenese 2.0 project wordt een schat aan informatie verzameld in de vorm van (veld)metingen en modelsimulaties. Om deze data op een veilige en efficiënte manier op te slaan, te delen met betrokkenen en te ontsluiten voor de buitenwereld is datamanagement nodig. Voor goed datamanagement zijn drie zaken essentieel:

1. Afspraken over de werkwijze die haalbaar/werkbaar zijn voor de gebruikers.
2. Een systeem dat opslaan, delen en ontsluiten van data mogelijk maakt.
3. Gebruik van het systeem volgens de gemaakte afspraken.

Het voorliggend datamanagementplan heeft ten doel om de afspraken (1) vast te leggen en het systeem (2) te beschrijven. Na implementatie van het systeem kan, met medewerking van de gebruikers (3), de data op succesvolle manier beheerd en ontsloten worden.

Afspraken behelzen het hele traject vanaf het aanleveren/invoeren van data tot en met de archivering na afloop van het project. Bij het aanleveren/invoeren van nieuwe (ruwe) data is het belangrijk dat de afgesproken structuur gevolgd wordt en dat alle bijbehorende informatie (metadata) en documentatie bijgevoegd wordt om correcte interpretatie van de data mogelijk te maken. Hierbij moet gedacht worden aan gestandaardiseerde vocabulaires, (SI) eenheden en eenduidige tijd- en locatieaanduiding. Daarna dient de data opgewerkt te worden tot een (standaard) vorm die (interdisciplinaire) uitwisseling mogelijk maakt en bevordert. De combinatie van volledige en eenduidige data in een duurzaam bestandsformaat vormt de basis voor goed gebruik tijdens het project en hergebruik in de toekomst.

Definities:

- Dataset: een verzameling van gegevens die in opgewerkte (standaard) vorm uniform samenhangend en eenduidig is. (in praktijk: waarvoor hetzelfde script gebruikt kan worden voor de opwerking vanuit de oorspronkelijke ruwe data)
- Ruwe data: data die rechtstreeks uit een meetinstrument komt.
- Standaard data: data die qua (bestands)formaat en inhoud een afgesproken (internationale) standaard volgt en voorzien is van metadata waardoor passend gebruik en interpretatie mogelijk is



### 3 Taken en Verantwoordelijkheden

#### 3.1 Betrokken partijen

Binnen Kustgenese 2.0 werken de volgende partijen samen [ref. 2]:

- Ministerie van Infrastructuur en Waterstaat
- Rijkswaterstaat (WVL, CIV)
- Deltares
- Ingenieursbureaus: Witteveen en Bos
- SEAWAD: TU Delft, Universiteit Utrecht en Universiteit Twente
- Aannemers: *nu nog niet bekend*

#### 3.2 Beschrijving taken en verantwoordelijkheden

Op hoofdlijnen kunnen de rollen als volgt worden samengevat:

- Data aanlevering
  - Invullen relevante informatie (zie bijlage C)
  - Kwaliteitscontroles
  - Overdracht ruwe data en standaarddata
  - Opwerkingsscripts voor standaardisatie
- Opdrachtgever
  - Akkoord op individuele datasets (na controle)
  - Controle aanlevering (is alles er, reviews aanwezig)
- Beheerder
  - Werking van het systeem
  - Beschikbaarheid
  - Monitoring en overzicht van aanlever voortgang
  - Review aanleveringen (verbetersuggesties)
- Gebruiker
  - Standaard data als basis
  - Feedback over bevindingen
  - Attributie van data eigenaar

##### 3.2.1 Product eigenaar - Rijkswaterstaat

Rijkswaterstaat fungeert als opdrachtgever voor Deltares ten aanzien van de realisatie van het datamanagementsysteem. Als opdrachtgever is Rijkswaterstaat verantwoordelijk voor het formuleren van wensen en eisen, zodat dit in de vorm van functionaliteiten kan worden uitgewerkt in het datamanagementsysteem. Hierbij is het van belang dat er ook rekening gehouden wordt met partners (SEAWAD). Ook is het essentieel dat Rijkswaterstaat eisen en wensen (zoals formaten van standaard data en kwaliteitseisen) helder communiceert richting data leveranciers.

In het proces is Rijkswaterstaat WVL (Carola van Gelder) verantwoordelijk.

##### 3.2.2 Dataleveranciers – diverse partijen

De dataleveranciers hebben tot taak ruwe data aan te leveren, inclusief scripts voor opwerken van de data tot standaard data. Hierbij is ondersteuning beschikbaar vanuit

databaseerder Deltares. Ondersteuning is beschikbaar in de vorm van handleidingen (bijlagen van voorliggend datamanagementplan), training en direct contact. Het project SEAWAD is ook deels leverancier van data.

### 3.2.3 Datagebruiker – diverse partijen

Algemeen:

- Onderzoekers: voor onderzoekers is het essentieel om inzicht te hebben in alle vormen van data (observaties tijdens metingen, ruwe data, standaard data, visualisaties) en de algoritmes om tot deze data vormen te komen.
- Ingenieurs: ook voor ingenieurs is het belangrijk om inzicht te hebben in alle vormen van data, maar afhankelijk van hun rol zullen ze focussen op specifieke delen. Als hun rol is om de data analyse uit te voeren of visualisaties te maken, zal de standaard data het startpunt zijn. Als hun rol is om data opwerking te verzorgen zal er vanuit de ruwe data gewerkt worden.
- Beleidsmakers: voor beleidsmakers is het belangrijk om visualisaties te hebben die bruikbaar zijn om beleidsadviezen/keuzes te onderbouwen.
- Publiek: ook voor het publiek is het visualisatie deel van de data het meest bruikbaar.

Partners:

- SEAWAD: Het project “SEdiment supply At the WAdden Sea ebb-tidal Delta. From system knowledge to mega-nourishments” (<http://www.stw.nl/nl/content/seawad-sediment-supply-wadden-sea-ebb-tidal-delta-system-knowledge-mega-nourishments>) in uitvoering door TU Delft, Universiteit Twente en Universiteit Utrecht doet onderzoek dat sterke raakvlakken heeft met Kustgenese 2.0. Er is afgesproken om zowel data als kennis te delen. SEAWAD maakt gebruik van het Kustgenese 2.0 datamanagementsysteem en is daarin deels gebruiker en deel leverancier van data. Bram van Prooijen is aanspreekpunt betreffende de data namens SEAWAD.

Gebruikers en (programmeer)software:

- Ingenieurs (opgeleid sinds 2013) en mensen met een geospatiele opleiding (bijv. gebruikers van arcgis, qgis, FME) gebruiken doorgaans python als programmeertaal.
- Biologen, ecologen en statistici werken doorgaans in R
- Ingenieurs opgeleid tot 2013 kregen standaard Matlab aangeleerd.

Om de kwaliteit van de data te bevorderen zijn datagebruikers verplicht om hun relevante bevindingen – zoals geconstateerde fouten of verbeterde metadata – terug te koppelen naar de eigenaar/contactpersoon van de data(set). Dit wordt vastgelegd in de voorwaarden voor gebruik.

### 3.2.4 Databaseerder - Deltares

Deltares heeft van Rijkswaterstaat de opdracht gekregen het datamanagementsysteem te implementeren en tot 1 oktober 2019 te voorzien in beheer en onderhoud. Deltares controleert de naleving van de afspraken, zoals beschreven in dit rapport, over data aanlevering, structurering en opwerking tot standaard data. Deltares controleert niet of datagebruikers relevante bevindingen terugkoppelen. Deltares levert periodiek (kwartaal) rapportages (als pdf) met een overzicht van de beschikbare datasets, met informatie over de opwerkingsniveaus en ontsluiting.

Kees den Heijer ([kees.denheijer@deltares.nl](mailto:kees.denheijer@deltares.nl)) is namens Deltares aanspreekpunt en trekker voor het databaseer.



## 4 Inwinning, verwerking en gebruik data

### 4.1 Introductie

Dit hoofdstuk beschrijft het traject vanaf inwinning en verwerking tot aan de opslag van data binnen Kustgenese 2.0. De paragrafen beschrijven de stappen die de onderzoekers en deelnemers aan Kustgenese 2.0 dienen te doorlopen voor het gestructureerd aanleveren en verwerken van de ruwe data, zodat dit na validatie geschikt is voor gebruik. In Figure 4.1 Workflow van instrument naar standaard data. Figure 4.1 is de workflow van ruwe data uit een meetinstrument naar standaard data (netCDF) en catalogus (Geonetwork) gevisualiseerd. De subversion repository vormt de basis/kern van het datamanagementsysteem. Hierin wordt de ruwe data aangeleverd en worden scripts bijgehouden. De technische specificaties van het datamanagementsysteem komen aan de orde in hoofdstuk 5.

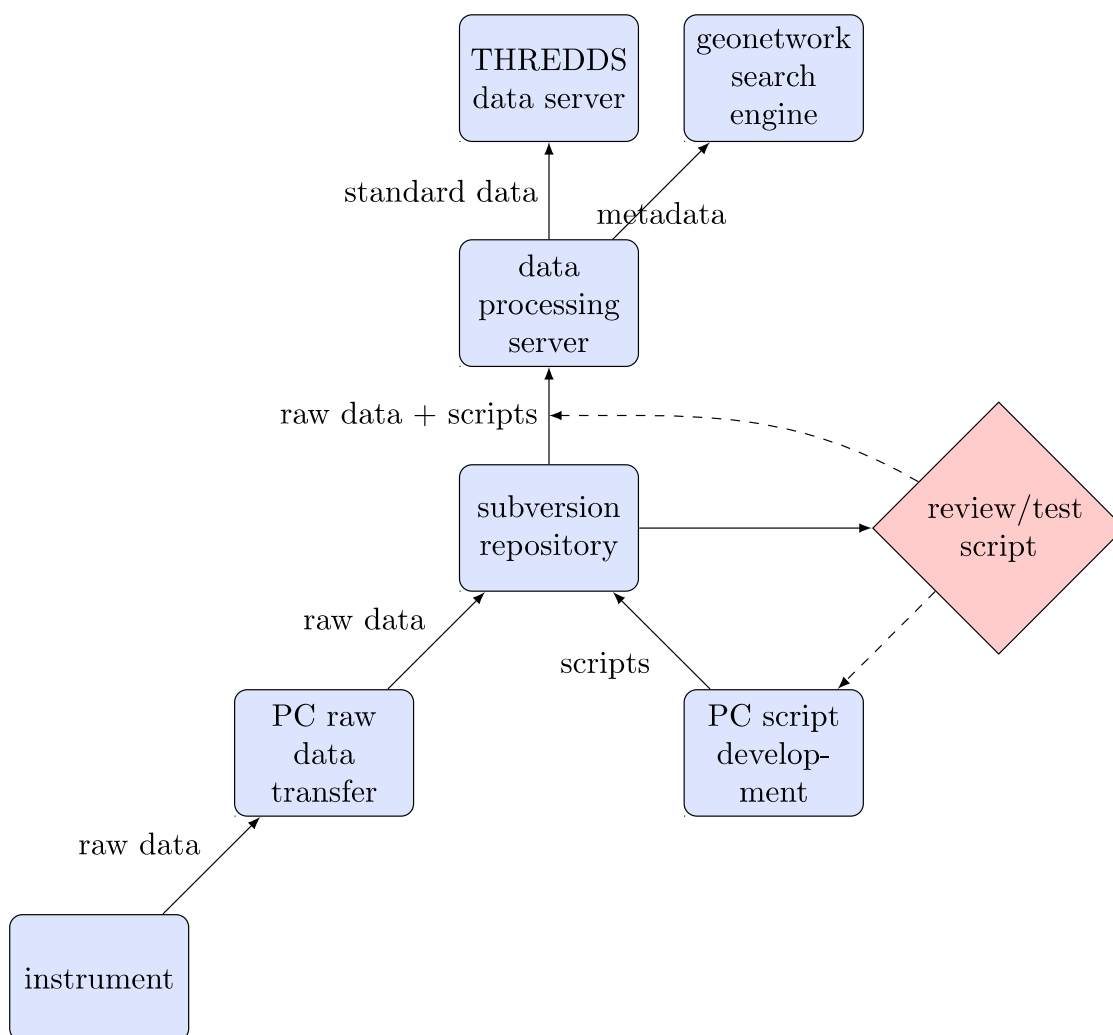


Figure 4.1 Workflow van instrument naar standaard data.

## 4.2 Uitgangspunten

Binnen het datamanagementprogramma zijn de volgende uitgangspunten leidend:

- Voor de naamgeving wordt aangesloten bij het protocol mariene data (van het Informatiehuis Marien). Hierin worden specifieke aanbevelingen gedaan voor semantische standaarden zoals Aquo en CF conventies.
- De Digitale Delta API wordt vooralsnog niet gebruikt voor Kustgenese 2.0. Digitale Delta voorziet momenteel nog niet in een oplossing voor raster data en aansluiten van tijdseries vergt ook nog werk dat niet in Kustgenese 2.0 is voorzien.
- De voertaal voor het werken in de repository is Engels.
- De voorkeur voor programmeertalen is in aflopende volgorde als volgt: Python, R, Matlab (zie ook sectie 3.2.3 voor onderbouwing van volgorde).
- Binnen het programma Kustgenese 2.0 worden bij voorkeur de volgende formats gehanteerd:
  - rasterdata → NetCDF:  
NetCDF heeft als voordeel dat de metadata bij de meetgegevens wordt opgeslagen. Dit heeft als voordeel dat bij de overgang van project naar archief, de bestanden 'self-explaining' zijn. Oftewel mocht een bestand weer worden opgehaald uit archief, dan is deze altijd volledig gevuld.
  - vector data → Geojson
  - overige data → csv, txt, asc, jrk
- UTC (De UTC offset moet altijd vermeld worden)
- als coördinatenstelsel wordt aangesloten bij ETRS89 (is in de praktijk gelijk aan WGS84).
- de WPS is als enige functionaliteit binnen de OpenEarth Stack nu niet opgenomen in het datamanagementsysteem.

## 4.3 Voorbereiding

### 4.3.1 Verkrijgen van toegang tot de Subversion repository Kustgenese 2.0 - voor opslag en verwerking ruwe data

Voor toegang tot de repository van Kustgenese 2.0, dient contact opgenomen te worden met de databeheerder (zie sectie 3.2.4). Instructies voor het gebruik van de subversion repository zijn beschikbaar in bijlage **Error! Reference source not found.**

### 4.3.2 Invullen standaard datamanagement formulier

Om inzichtelijk te krijgen welke data gedurende het programma Kustgenese 2.0 beschikbaar komt, is het nodig om per dataset een datamanagementformulier in te vullen. Dit zorgt ervoor dat aan de voorkant afspraken worden gemaakt over welke data wordt aangeleverd, hoe dit wordt aangeleverd (format), welke nieuwe data wordt gegenereerd, welke data uit andere bronnen wordt gebruikt, en wat aan opslag nodig is. Uit deze formulieren volgt een beeld van de benodigde hoeveelheid opslag en servercapaciteit. De formulieren hebben een dynamisch karakter, wat betekent dat velden naarmate de tijd vordert verder geconcretiseerd en/of aangevuld kunnen worden op basis van voortschrijdend inzicht en ervaring.

Voor Kustgenese 2.0 is een **datamanagementformulier** gemaakt (opgenomen in bijlage **Error! Reference source not found.**), met invulvelden voor onder andere: basisinformatie, beschrijving dataset, metadata en formats, logboek en planning.

### 4.3.3 Opslag in repository

De informatie uit het standaard datamanagementformulier wordt opgeslagen in de repository van Kustgenese 2.0. Zie hiervoor het voorstel voor de mapstructuur, zoals opgenomen in bijlage **Error! Reference source not found.** Het aanmaken van directories voor datasets die nog niet in de mapstructuur voorkomen dient in overleg met de data beheerder gedaan te worden, zodat de consistentie bewaakt kan worden. Indien nodig kan de directorystructuur later worden aangepast (door de beheerder). De relatieve verwijzingen naar de directory's in reken-subroutines wijzigen automatisch mee.

*Merk op: standaard data dient niet in de repository te worden opgeslagen (tenzij anders overeengekomen met data beheerder). Voor bijvoorbeeld ecologische puntdata kan de werkwijze van WOZEP gevolgd worden, namelijk dat standaard data als platte csv in IM metingen (AQUO) worden opgeslagen, alvorens in de database geladen te worden.*

## 4.4 Uitlezen ruwe data

Paragraaf 1.4 geeft het overzicht van de monitoringscampagnes waarvan de data in het datamanagementsysteem opgenomen wordt. Dit datamanagementplan gaat niet over de inwinning van data.

Voor Kustgenese 2.0 wordt een conceptueel datamodel opgesteld met bijbehorende metadata-eisen conform AQUO. Het laat zien wat beschikbaar komt aan ruwe meetgegevens, en geeft inzicht in de eisen aan metadata, zoals wat is er nodig aan metadata en wat is hiervoor het te hanteren format.

Afspraken hierover zorgen voor uniformiteit in naamgeving en te gebruiken eenheden. Tevens legt het datamodel de afspraken vast over standaard headers (kolomkoppen). Het conceptuele datamodel wordt uitgewerkt op basis van de datamanagementformulieren.

## 4.5 Opslag ruwe data

De ruwe data wordt per meetronde uitgelezen en opgeslagen in de Subversion repository van Kustgenese 2.0 (<https://repos.deltares.nl/repos/kustgenese2/trunk>) in de standaard directorystructuur (zie bijlage **Error! Reference source not found.**). Bij de data wordt een 'readme.txt'-file gevoegd, met daarin alle belangrijke informatie over de dataset + eventuele bijzonderheden die een rol spelen bij de interpretatie: bijzonderheden zijn bijvoorbeeld dat het instrument op een bepaald tijdstip is verschoven (een discontinuïteit in de data als gevolg van zo'n verschuiving kan gemakkelijk verklaard worden mits deze informatie goed gedocumenteerd is). Deze informatie dient te worden aangevuld met bevindingen die tijdens de analyse worden opgemerkt.

Elke dataset heeft standaard een raw/ en een scripts/ directory en optioneel een doc/ directory. Meer details over de directorystructuur zijn te vinden in bijlage III ([https://repos.deltares.nl/repos/kustgenese2/trunk/datamanagement/datamanagementplan/annexes/directory\\_structure.pdf](https://repos.deltares.nl/repos/kustgenese2/trunk/datamanagement/datamanagementplan/annexes/directory_structure.pdf))

## 4.6 Opwerking ruwe data, validatie verwerking en opslag verwerkte data

### 4.6.1 Opwerkingscripts

Voor elke dataset dienen opwerkingscripts beschikbaar te zijn om deze volgens afgesproken standaard format beschikbaar te kunnen maken. De data leverancier is daarvoor

verantwoordelijk. De contactpersoon voor de betreffende dataset stelt de opdrachtgever en databeheerder op de hoogte zodra een script gereed is (voor review). Zie ook [https://repos.deltares.nl/repos/kustgenese2/trunk/datamanagement/datamanagementplan/annexes/Procesafspraken\\_dataopwerking\\_incl\\_bijlagen\\_13022018.pdf](https://repos.deltares.nl/repos/kustgenese2/trunk/datamanagement/datamanagementplan/annexes/Procesafspraken_dataopwerking_incl_bijlagen_13022018.pdf) voor meer informatie over de procesafspraken tussen Rijkswaterstaat, Deltares en SEAWAD.

- Voor het opwerken van de ruwe rasterdatabestanden zijn in OpenEarth tools scripts beschikbaar (o.a. door Deltares ontwikkeld) voor het goed vullen van de netCDF bestanden. (<https://svn.oss.deltares.nl/repos/openearthrawdata/trunk/>) en ook (<https://repos.deltares.nl/repos/zandmotor/trunk/>).
- De netCDFkickstarter is beschikbaar op <http://zandmotor.citg.tudelft.nl/netcdfkickstarter> om netCDF bestanden te maken volgens standaard templates en vocabulaires.
- Deltares heeft een ETL (extract-transform-load) tool beschikbaar om data in AQUO standaard in te laden in een PostgreSQL database (<https://svn.oss.deltares.nl/repos/openearthtools/trunk/etl/aquo>). De ETL gaat uit van platte csv data in een van de uitwisselformaten IMWA metingen of IM metingen (<http://www.aquo.nl>)
- Voor ieder type data (sensor) zal een opwerkingscript gemaakt moeten worden door de dataleverancier (zie sectie 3.2.2). Dit script dient opgeslagen te worden naast de ruwe data in de standaard directorystructuur.
- Opwerkingscripts dienen platformonafhankelijk te zijn en eventuele afhankelijkheden van software of toolboxes dienen duidelijk te worden vermeld. Uitgangspunt is dat de opwerking tot standaard data kan plaatsvinden op basis van een “schone” checkout van de subversion repository (uiteraard met de benodigde software geïnstalleerd).

#### 4.6.2 Standaarden

Richtlijnen voor aanlevering van projectdata

([https://www.informatiehuismariene.nl/publish/pages/109375/data\\_delivery\\_for\\_marine\\_monitoring\\_projects.pdf](https://www.informatiehuismariene.nl/publish/pages/109375/data_delivery_for_marine_monitoring_projects.pdf)) is het startpunt met betrekking tot standaarden (let op, in protocol mariene data wordt de term “base data” gebruikt voor wat we in dit rapport “standaard data” noemen). Voor netCDF bestanden dienen de CF conventies (<http://cfconventions.org/>) te worden gevolgd. Deze worden ook ondersteund door de netCDFkickstarter.

#### 4.6.3 Review/test scripts

Opwerkingscripts worden door de data beheerder gereviewed alvorens ze definitief gedraaid worden om de standaard data beschikbaar te maken. De review resulteert in acceptatie of in suggesties voor verbetering. Het resultaat wordt teruggekoppeld naar contactpersoon en opdrachtgever.

Bij de review van scripts wordt gelet op:

- Leesbaarheid/begrijpelijkheid (balans tussen helderheid van code zelf en mate van begeleidend commentaar)
- Documentatie bij data en script
- Metadata (gebruik van vocabulaires)
- Gebruik van relatieve paden (absolute paden werken niet vanzelfsprekend op een andere machine)
- Platformonafhankelijkheid

Bij het testen van scripts is het belangrijkste criterium dat het kan draaien op basis van een “schone” checkout van de subversion repository (uiteraard met de benodigde software geïnstalleerd).

#### 4.6.4 Validatie

Validatie heeft tot doel te controleren of de stap van ruwe data naar standaard data juist is uitgevoerd. Validatie is de verantwoordelijkheid van degene die het script maakt/onderhoud (dataleverancier). Gecontroleerd wordt of data consistent en integer is. Daar waar mogelijk vindt validatie geautomatiseerd plaats. Of validatie is uitgevoerd en succesvol is geweest, wordt bijgehouden en is zichtbaar in het periodieke dataset overzicht (bijlage **Error! Reference source not found.**, NB: nog uitbreiden)

Validatie wordt als volgt uitgevoerd:

- Voor de rasterdata wordt gewerkt met CF conventies: de validatie is hiervoor al beschikbaar. De Thredds Data Server heeft geautomatiseerde services om op de aangeleverde velden een check te doen (Dataset Discovery Report).
- Voor enkele (steekproef) datasets wordt dit uitgevoerd door Deltares (Kees den Heijer).
- De analyse (ofwel het gebruik) van de data is een belangrijke vorm van validatie. Indien “fouten” in de data gevonden worden dient het opwerkingscript hierop aangepast te worden (uitfilteren foutieve data). Indien opmerkelijke zaken in de data worden opgemerkt die niet fout zijn, dan dient dit in de documentatie (readme.txt) toegevoegd te worden.
- De ETL voor opslag van AQUO data in de het RDBMS checkt of the data formeel aan AQUO voldoet (domeintabellen, verplichte velden etc). Daarnaast moet ook nog een inhoudelijke controle plaatsvinden.

*Openstaande actie op te nemen in volgende versie van het datamanagementplan: Voor overige datasets dient validatie nog te worden afgesproken.*

#### 4.6.5 Opslag en ontsluiting van standaard data

De opgewerkte (standaard) data wordt opgeslagen op een projectschijf bij Deltares (p:\1220339-015-data\). NetCDF data wordt via een Thredds Data Server (<https://kustgenese2.deltares.nl/thredds>) ontsloten, welke gekoppeld wordt aan de bovengenoemde projectschijf. De data is voor Deltares medewerkers zowel via het interne netwerk als via the Thredds Data Server beschikbaar. Vector data komt in een PostgreSQL database (Aquo/IM, <http://www.aquo.nl>).

Datasets die aansluiten op bestaande meetreeksen die al elders worden ontsloten, worden in overleg met opdrachtgever daarop aangesloten. Voorbeelden hiervan zijn vaklodingen en JarKus die via OpenEarth (<http://opendap.deltares.nl/thredds/catalog/opendap/catalog.html>) als standaard data worden ontsloten; het is logischer om de gebiedsspecifieke metingen voor Kustgenese 2.0 ook via OpenEarth te ontsluiten. Wel moet er in het Kustgenese 2.0 datamanagementsysteem een verwijzing staan naar de betreffende andere bron.

Mede op verzoek van de NKWK taskforce data management wordt een mirror van geselecteerde datasets bij een of meer cloudproviders overwogen. Dit om gebruikersmogelijkheden op het gebied van rekenen bij de data te vergroten. Dit wordt voor

Kustgenese 2.0 vooralsnog niet uitgevoerd, maar wellicht als het NKWK verder is kan heroverweging plaatsvinden.

## 4.7 Datagebruik

De data is bedoeld zowel voor visualisatie (viewer), onderzoek en (beleids)analyses als voor input- en kalibratiedata van rekenkundige modellen. Verdere specificatie van te onderscheiden gebruikers is te vinden in sectie 3.2.3). Het dataportaal van Kustgenese 2.0 wordt ontsloten via een centrale website (<https://kustgenese.deltares.nl>), welke tevens gebruikt kan worden voor communicatie en PR.

De centrale website is publiek toegankelijk. Via deze website heeft de gebruiker toegang tot:

- Zoekfunctionaliteit
- Viewer
- Standaard data
- Ruwe data (inlog vereist)

## 4.8 Overige data

Naast de data die in het kader van Kustgenese 2.0 wordt ingewonnen, maakt het project ook gebruik van andere (reguliere) datasets. Voor zover deze data op een passende manier ontsloten wordt voor gebruik in Kustgenese 2.0, zal deze data niet opnieuw worden opgeslagen in het Kustgenese 2.0 datamanagementsysteem.

In overleg met de databeheerder kan besloten worden om relevante, specifieke subsets of bewerkingen op bestaande data toch als een 'bevroren' dataset op te slaan. De gebruiker dient hiervoor contact op te nemen met de databeheerder.

### 4.8.1 Verwerking overige data

#### *Beschikbare data van Rijkswaterstaat CIV*

Voor het onderzoeksprogramma Kustgenese 2.0 is data uit bestaande opslaglocaties van CIV relevant, zoals:

- Donar (Data opslag natte Rijkswaterstaat);
- LOL (landelijk opslagsysteem Iodigen).

Deze data is op te vragen via de Servicedesk Data en het waterinfo.rws.nl portaal. Deze data wordt niet opgenomen in het Kustgenese 2.0 datamanagementsysteem.

#### *Data derden*

Bijvoorbeeld data van SOVON over vogels. Vooralsnog wordt deze data niet in het Kustgenese 2.0 datamanagementsysteem opgenomen.

### 4.8.2 Modeltoepassing

#### *Voorbereiden modelinvoer*

Data benodigd voor de modeltoepassing wordt verzameld. Vervolgens wordt deze data ge-upload naar een daarvoor beschikbaar gemaakte directory in de subversion repository.

#### *Opslag bewerkte (model)data*

Het resultaat van de modellering kan apart worden opgeslagen als onderdeel van het datamanagementsysteem als dataproduct. Bijvoorbeeld een NetCDF uitvoer wordt dan opgenomen op Thredds Data Server.

## 4.9 Trainingen

Om het datamanagementprogramma te faciliteren, zijn de volgende trainingen voorzien:

- Training 1  
Doelgroep: dataleveranciers
  - Omgaan met subversion versiebeheer;
  - Invullen datamanagement formulier per onderzoeksonderdeel;
- Training 2  
Doelgroep: dataleveranciers en geïnteresseerde datagebruikers (onderzoekers)
  - Verwerking data (netCDF en CF)
  - Verwerking data (postgresql en AQUO)
- Training 3  
Doelgroep: dataleveranciers en geïnteresseerde datagebruikers (onderzoekers)
  - Visualisatie data (Mapbox)
- Optioneel:  
Doelgroep: datagebruikers
  - Opvragen data via scripting (Python, R)
  - Interactieve en reproduceerbare rekendocumenten (Jupyter notebook)
  - Interactieve en reproduceerbare rekendocumenten (R-shiny en Rmarkdown)

De insteek is dat voor de overige gebruikers geen training vereist is.





## 5 Datamanagementsysteem

### 5.1 Introductie

Het **datamanagementsysteem** heeft tot doel de data op te slaan, te ontsluiten en efficiënt gebruik mogelijk te maken. Het betreft een systeem dat de reproduceerbaarheid van opgewerkte (standaard) dataproducten ondersteunt. De data inclusief metadata moet inzichtelijk zijn voor alle betrokken partijen.

Het datamanagementsysteem moet generiek en robuust genoeg zijn om coherentie en standaardisatie te stimuleren en flexibel genoeg om 'buitensluiting van exotische datatypen' te voorkomen.

### 5.2 Onderdelen datamanagementsysteem (en Minimal Viable Product)

Het datamanagementsysteem wordt opgebouwd uit bouwstenen. Deze bouwstenen zijn gebaseerd op de OpenEarth Stack. Er wordt onderscheid gemaakt tussen de basis bouwstenen die generiek toepasbaar zijn (project overstijgend) en de extra bouwstenen (relevant voor innovatie en ontwikkeling). In de eerste versie van het datamanagementsysteem (Minimal Viable Product) worden de onderstaande basis bouwstenen gecombineerd.

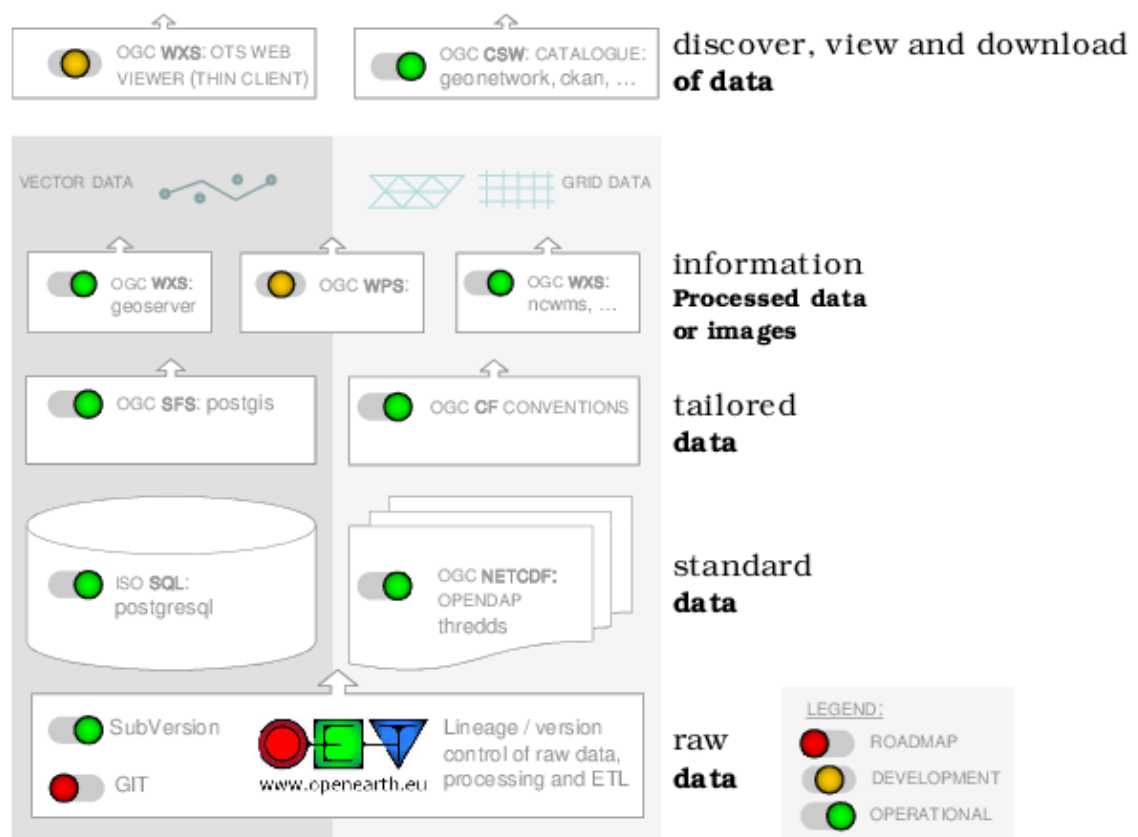


Figure 5.1 OpenEarth stack: weergave van bouwstenen en samenhang.

## 5.2.1 Basis bouwstenen (generiek toepasbaar)

### *Subversion repository*

De Subversion repository is uitsluitend toegankelijk voor betrokkenen en betreft een werkomgeving waarin versiebeheer van toepassing is (<https://repos.deltares.nl/repos/kustgenese2/trunk/>). Het wordt gebruikt voor de opslag van ruwe data en bijbehorende scripts. Voor toegang is een inlog noodzakelijk (zie sectie 4.3.1).

### *Docker host*

De Docker host draait lokaal bij Deltares, en is niet van buiten benaderbaar. Binnen het project wordt deze server gebruikt voor de opwerking van ruwe data tot standaard dataproducten met gebruik van de gevalideerde scripts die door de dataleverancier is aangeleverd.

### *Dataopslag*

Koppeling met servers die opslag nodig hebben voor dataontsluiting (meervoudige opslag voorkomen). De opgewerkte (standaard) data wordt opgeslagen op een projectschijf bij Deltares (p:\1220339-015-data\). Zie ook sectie 4.6.5.

- Thredds Data Server
- Geoserver
- PostgreSQL

### *Thredds Data Server*

De Thredds Data Server stelt de (standaard) data publiek beschikbaar. Binnen het project wordt deze gebruikt voor de opslag van raster data (NetCDF) en voor de ontsluiting van de data via web services.

- OPeNDAP (voor directe toegang tot netcdf data)
- HTTP fileserver (voor download van netcdf en andere data)
- OGC WMS van netcdf data
- OGC WCS van netcdf data

### *PostgreSQL relationele database met postGIS extensie*

- Ontsluiting van vector data

### *Geoserver*

Geoserver is een open source server voor het ontsluiten van vector- en rasterdata via web services. Het maakt gebruik van open standaarden.

- OGC WFS (Web Feature Service: een interface voor het opvragen, aanleveren en bewerken van vector data (maakt gebruik van xml en gml voor dataoverdracht)
- OGC WMS (Web Mapping Service voor vector en rasterdata)
- OGC WCS (Web Coverage Service voor rasterdata)

### *Geonetwork*

- Catalogus/zoek functionaliteit: hiervoor wordt de bestaande instantie <http://dataportal.deltares.nl> gebruikt. Als onderdeel van het dataopwerkingsproces wordt metadata in de database toegevoegd/ge-update.

### *Mapbox*

- Visualisatie van datasets

*Centrale website*

- Startpunt voor de gebruiker. Vanaf hier zijn links naar de overige componenten.

## 5.2.2 Extra bouwstenen

Componenten die voor Kustgenese 2.0 als bruikbare innovaties kunnen worden bestempeld en aansluiten bij de wensen vanuit NKWK taskforce datamanagement, maar die in algemene zin als optioneel kunnen worden beschouwd. Deze maken geen onderdeel uit van de MVP.

*Cloud provider(s)*

- Mirror van data
- Mogelijkheden voor rekenen bij de data

*Pilot Cloud*

Doel is daar waar grote datasets beschikbaar zijn, en de bewerking veel tijd kost een pilot te doen met het rekenen en werken in de cloud en met kleine aanpassingen reproduceerbare data en kaartensets te maken. Het uitvoeren van een dusdanige pilot sluit aan bij de wens van RWS (o.a. NKWK) om live te rekenen met de data en het resultaat direct te visualiseren.

*Jupyter Notebook server*

Interactieve rekendocumenten

*R-shiny en Rmarkdown server*

Interactieve webapplicaties en rekendocumenten

*Pilot interactieve rekendocumenten*

Een voorstel is een pilot uit te voeren met interactieve rekendocumenten (een verzameling packages in een cloud omgeving), zoals Jupyter of R-Shiny. Bijvoorbeeld door de Kustgenese 2.0 data dusdanig te visualiseren, dat het product gebruikt kan worden voor besluitvorming of bijvoorbeeld een managementrapportage. Het kan ook een pilot zijn om de stappen die genomen worden om de kennisvragen te beantwoorden inzichtelijk te maken. Het uitvoeren van een dusdanige pilot sluit aan bij de wens van RWS (o.a. NKWK) om live te rekenen met de data en het resultaat direct te visualiseren.

*WPS (web processing service)*

- Meerdere datasets doorkoppelen of snelle berekeningen doen. Het onderbrengen van data in de cloud (pilot cloud) kan hiervoor een alternatief bieden.

**5.3 Aandachtspunten**

Ontwikkelingen die binnen Kustgenese 2.0 worden nagestreefd, maar die niet hard kunnen worden toegezegd:

- Single Sign On: 1 set credentials voor alle componenten van het datamanagementsysteem en voorkomen van herhaaldelijk inloggen. Momenteel zijn twee inlogs nodig, namelijk voor de repository en voor de website;
- Gemakkelijke uitrol van de softwarecomponenten; momenteel worden de verschillende services handmatig geïnstalleerd en geconfigureerd. Om dit proces te versnellen en reproduceerbaar te maken (voor nieuwe projecten) is het wenselijk om hiervoor een uitrolscript (provisioning) te maken.
- Dichttimmeren van het systeem zodanig dat reproduceerbaarheid (van de dataproducten) gegarandeerd kan worden en transparantie gewaarborgd (in de

huidige opzet wordt enige discipline verwacht van de personen die de dataproducten opwerken en online zetten); in de huidige situatie moet het dataopwerkingsproces handmatig worden gestart en het resultaat gekopieerd naar de definitieve opslag locatie. Door deze handmatige acties te automatiseren kan hier een verbetering opleveren.

- RWS CIV zodanig voorbereiden op de te archiveren data bij beëindiging van het Kustgenese 2.0 project dat alle data een op een (zonder extra transformaties of functionaliteitsverlies) kan worden verhuisd naar CIV (zie sectie 7.3.1).

## 5.4 Beheer en onderhoud datamanagementsysteem

Updates: Alle servers bij Deltares worden maandelijks voorzien van de laatste updates (en in geval van kritische veiligheidsupdates eventueel nog extra tussendoor).

Helpdesk: De databeheerder is tijdens kantooruren beschikbaar voor vragen.

Uptime: Deltares hanteert "best effort" maar geen garantie voor uptime. Server onderhoud vindt elke eerste dinsdag van de maand plaats; gedurende die dag kan het datamanagementsysteem tijdelijk onbereikbaar zijn.

Back-up: van zowel servers als opslag worden minimaal eens per dag back-ups gemaakt welke minstens een half jaar bewaard worden.

## 5.5 Eisen en wensen

- CIV wil graag dat alle mariene projecten een uniforme wijze van dataopslag en -distributie hebben (ref. 2). Vanuit CIV zijn er geen voorschriften voor de Kustgenese 2.0 data omgeving (ref. 2)
- Aansluiten bij Open Data Standaarden en uitwisseling van data mogelijk maken
- CIV heeft de wens dat alle data verwerkt is met de INSPIRE richtlijnen, waarin metadata op strategisch/internationaal niveau wordt opgeslagen (ref. 2). RWS heeft op 18 thema's de INSPIRE standaarden overgenomen.
- Indien gegevens van Kustgenese 2.0 voldoen aan de eisen van het MWTL, dan kunnen ze in de RWS omgeving worden opgeslagen.
- Definitieve producten moeten worden opgeslagen, inclusief bijlages. Tevens beschikbaar stellen van overall managementrapporten (in het Nederlands) en technische rapporten (in het Engels).
- Aansluiten van NKWK en andere projecten/programma's is geen uitgangspunt.

## 6 Administratieve en Juridische afspraken

### 6.1 Introductie

Dit hoofdstuk beschrijft de afspraken rondom het werken met data zoals vastgelegd voor Kustgenese 2.0 (en SEAWAD). Voor alle partners zijn de volgende afspraken van toepassing ten aanzien van opslag, gebruik, publicatie en uitwisseling van data, gedurende de looptijd van het onderzoeksprogramma.

### 6.2 Afspraken betreffende verzamelen, managen en delen van data (dataproductent)

Over het werken met data is afgesproken dat

- Data centraal wordt opgeslagen en
- Data onderling wordt gedeeld.

*Openstaande actie: Standaardtekst Rijkswaterstaat (en eventueel SEAWAD) licentievorm + licentietekst*

### 6.3 Afspraken betreffende data uitwisseling (data gebruiker)

Opgewerkte (standaard) data is openbaar, tenzij gemotiveerd wordt afgeweken. Goedkeuring voor afwijking kan alleen door de product eigenaar (zie sectie 3.2.1) worden gegeven. Validatie op de basisgegevens is van belang. Mogelijk dat vrijgave pas kan nadat de eigenaar er een kwaliteitsmerk op heeft gezet.

De ruwe data repository is in principe alleen toegankelijk voor gebruikers in het kader van Kustgenese 2.0 en SEAWAD. Toegangsverzoeken in bredere context dienen door beheerder te worden afgestemd met product eigenaar (zie sectie 3.2.1) en SEAWAD (zie sectie 3.2.3).

### 6.4 Eigendom

Binnen dit project worden data, modellen en tools aangeleverd door Deltares. Het eigenaarschap van data, software en modellen van Deltares wordt beschreven in de brief 1230103-001-BGS-0008 d.d. 4 april 2016, met name in de volgende paragraaf:

“Deltares ontwikkelt in het KPP-programma en SPA-opdrachten software, schematisaties en toepassingen, zoals voorspellingssystemen. Zoals in de Samenwerkingsovereenkomst tussen de Stichting Deltares en het Ministerie van Infrastructuur en Milieu (KW-31078379) is vastgelegd (artikel 6.1) geldt daarbij: “Tenzij Partijen anders overeenkomen, zijn alle resultaten van projecten en producten die worden geleverd op grond van de overeenkomst openbaar en worden aan een ieder ter beschikking gesteld, tegen betaling van de kosten die gemoeid zijn met het vastleggen van de resultaten op het daarvoor meest geschikte medium, inclusief de eventuele kosten voor instandhouding”. In artikel 6.2 worden in dit verband specifiek computerprogrammatuur, bestanden en modellen genoemd. Onder openbaar moet in dit verband ook “open source” worden begrepen. Software die door Deltares in het kader van de genoemde overeenkomst wordt ontwikkeld is “open source” en op die manier ook voor iedereen beschikbaar.

Voor software is het uitgangspunt dat het intellectueel eigendom (auteursrecht) berust bij Deltares; dat van de schematisaties en toepassingen berust bij IenM. In specifieke gevallen kunnen daar andere afspraken over worden gemaakt.

Software wordt door Deltares ter beschikking gesteld aan derden.

Daarnaast worden er data door derden aangeleverd. Rijkswaterstaat brengt Deltares per dataset op de hoogte van de eigenaar en gebruiksvoorwaarden. De beheerder controleert of deze gegevens bij elke dataset zijn ingevoerd (als onderdeel van de metadata volgens Dublin Core, CF en ISO19115).

## 7 Archivering, delen en publiceren van data

### 7.1 Introductie

Na afronding van het Kustgenese 2.0 is het nodig de data te archiveren (>2020). In dit hoofdstuk is beschreven hoe deze archieffunctie kan worden ingevuld, ervan uitgaande dat het datamanagementsysteem gerealiseerd wordt zoals beschreven in hoofdstuk 5. Doel van dit hoofdstuk is inzicht te geven in de opties die er zijn ten aanzien van archivering (opname in bouwstenencatalogus, dan wel anders), zodat dit nader kan worden uitgewerkt en geconcretiseerd in afspraken.

Let op dat de doorlooptijd van het Kustgenese 2.0 langer is dan het beheer en onderhoud door Deltares. Per 1 oktober 2017 zal overdracht plaatsvinden of zal Deltares beheer en onderhoud voortzetten op basis van nog te maken afspraken.

Meer achtergrond en advies over datamanagement van project naar archieffase is te vinden in [7].

### 7.2 Functionaliteiten Archief

De functionaliteiten van het archief dienen aan te sluiten op het verwachte/beoogde gebruik van de data in de archieffase. Hierbij dienen de doelgroepen (zie paragraaf 3.2.3) en de bijbehorende behoeften beschouwd te worden.

Zorg dat het ontsluiten via de reeds bestaande centrale services (bijv. [geoservices.rijkswaterstaat.nl](http://geoservices.rijkswaterstaat.nl), [opendap.deltares.nl](http://opendap.deltares.nl)) mogelijk is.

De subversion repository met ruwe data en scripts dient na afloop van het project in ieder geval gearchiveerd te worden. Een gearchiveerde subversion repository (svn dump), behelst alle informatie (inclusief versie historie) van de repository, maar is niet gemakkelijk beschikbaar. Als er sprake is van gebruik dat langer doorloopt dan de projectduur (bijvoorbeeld SEAWAD onderzoekers die data nog gebruiken voor hun promotieonderzoek) dan dient grondig afgewogen te worden of het moment van archivering van de subversion repository uitgesteld kan worden tot betreffende gebruikers hun werkzaamheden afgerond hebben. Mocht dit niet mogelijk zijn, dan dienen deze gebruikers tijdig geïnformeerd te worden over de op handen zijnde archiveringsfase, zodat ze daar op in kunnen spelen.

### 7.3 Vragen RWS beleid

Het RWS-beleid is als volgt: zolang het project loopt is de dataopslag de verantwoordelijkheid van het project zelf. Als het project is afgerond wordt de data overgezet naar speciale servers van RWS (of daarbuiten) in een formaat dat aansluit bij de RWS informatie-infrastructuur. Het is wenselijk om vanaf de start van het project zo goed mogelijk aan te sluiten bij de uiteindelijke structuur van RWS.

#### 7.3.1 Bestaande archief van RWS is leidend

*Openstaande actie op te nemen begin 2019: Samen met Rijkswaterstaat (John Schobben) helder krijgen wat de functionaliteiten zijn van het archief (of welke archieffuncties zijn beschikbaar bij Rijkswaterstaat) en hoe dient de data te worden opgewerkt om hier bij aan te sluiten?*

Daarbij is het van belang om twee vragen te beantwoorden:

*Is de opgewerkte data klaar voor het archief?*

Bij de opwerking van de data moet rekening gehouden worden met de functionaliteiten van het archief. Bijvoorbeeld: als er tijdens het project gebruik wordt gemaakt van een PostgreSQL database maar deze niet beschikbaar is in de archieffase, dan kan het een bewuste keuze zijn om uit de database ook bestandsgebaseerde producten (versimpelde csv of shapefile) te creëren die vooral tijdens de archieffase het gebruik vergemakkelijken. Dit kan het beste direct tijdens de opwerkingsfase gedaan worden en niet pas aan het einde van het project.

*Is het archief klaar om de benodigde functionaliteiten van de data te bieden?*

Als er in de archieffase functionaliteiten nodig zijn die nu nog niet geboden worden, moet er iemand de verantwoordelijkheid krijgen om dit tijdig te realiseren. Bijvoorbeeld: als CIV het archief gaat verzorgen moet er gezorgd worden dat netCDF data aangeboden kan worden middels een Thredds Data Server; aangezien deze nu niet in de bouwstenen catalogus staat, moet iemand bij CIV de verantwoordelijkheid krijgen om dit component aan de catalogus toe te voegen.



## 8 Referenties

- 1 Informatie en meetbehoefte ten behoeve van Kustgenese 2.0., Edwin Elias en Claire van Oeveren-Theeuwes d.d. okt. 2016.
- 2 Kustgenese 2, Datamanagementsysteem, Plan van aanpak., Sebastiaan Quirijns, Tinus de Ruiter, concept 01, versie 4.1, d.d. 16 november 2016.
- 3 Doorontwikkeling OpenEarth stack. Implementatie van raster gegevens in OpenEarth Stack, Deltares, definitief, d.d. september 2016.
- 4 Data ontsluiting Rivieren, NKWK-pilot C1, Deltares, 2017, B.
- 5 CIV protocol voor datalevering
- 6 RWS-protocol en voorwaarden opgesteld ism CIV en NKWK
- 7 Datamanagement: van project tot archieffase. Fedor Baart en Willem Stolte, d.d. aug. 2017.



## A Dataset status overview

[https://repos.deltares.nl/repos/kustgenese2/trunk/datamanagement/datamanagementplan/annexes/dataset\\_overview.pdf](https://repos.deltares.nl/repos/kustgenese2/trunk/datamanagement/datamanagementplan/annexes/dataset_overview.pdf)



## **B Manual on using SubVersion (SVN)**

[https://repos.deltares.nl/repos/kustgenese2/trunk/datamanagement/datamanagementplan/annexes/svn\\_manual.pdf](https://repos.deltares.nl/repos/kustgenese2/trunk/datamanagement/datamanagementplan/annexes/svn_manual.pdf)



## **C Standard Data Management Template**

<https://repos.deltares.nl/repos/kustgenese2/trunk/datamanagement/datamanagementplan/annexes/form/data-management-form.pdf>





## D Repository Kustgenese 2.0 directorystructuur

[https://repos.deltares.nl/repos/kustgenese2/trunk/datamanagement/datamanagementplan/annexes/directory\\_structure.pdf](https://repos.deltares.nl/repos/kustgenese2/trunk/datamanagement/datamanagementplan/annexes/directory_structure.pdf)